DISCUSSION PAPER SERIES

No. 7415

PAYING FOR CONFIDENCE: AN EXPERIMENTAL STUDY OF THE DEMAND FOR NON-INSTRUMENTAL INFORMATION

Kfir Eliaz and Andrew Schotter

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7415.asp**

ISSN 0265-8003

PAYING FOR CONFIDENCE: AN EXPERIMENTAL STUDY OF THE DEMAND FOR NON-INSTRUMENTAL INFORMATION

Kfir Eliaz, Brown University and CEPR
Andrew Schotter, New York University

Discussion Paper No. 7415
August 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7415

August 2009

ABSTRACT

Paying for Confidence: An Experimental Study of the Demand for Non-Instrumental Information

This paper presents experimental evidence that when individuals are about to make a given decision under risk, they are willing to pay for information on the likelihood that this decision is ex-post optimal, even if this information will not affect their decision. Our findings suggest that this demand for non-instrumental information is caused by what we refer to as a "confidence effect": the desire to increase one's posterior belief by ruling out "bad news", even when such news would have no effect on one's decision. We conduct various treatments to show that our subjects' behavior is not likely to be caused by an intrinsic preference for information, failure of backward induction or an attempt to minimize thinking costs.

JEL Classification: C91, D03 and D81
Keywords: anticipatory feelings, disjunction effect, non-instrumental information and thinking costs

Kfir Eliaz
Economics Department
Brown University
64 Waterman St.
Providence, RI 02912
USA

Email: Kfir_Eliaz@Brown.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=153104

Andrew Schotter
Department of Economics
New York University
269 Mercer Street, 7th Floor
New York, NY 10003
USA

Email: as7@nyu.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=113366

Submitted 05 August 2009

The authors would like to thank the Center for Experimental Social Science at New York University for its financial help and Anwar Ruff for programming the experiment. We also thank the Associate Editor, three anonymous referees, Andrew Caplin, Yoram Halevy, Dan Levin, Ariel Rubinstein, and especially Ran Spiegler for their valuable advice. Erkut Özbay provided excellent research assistance. This work was done under grant number SES-0425118 of the National Science Foundation.

1 Introduction

The standard view in Economics is that information is deemed valuable if, and only if, it is instrumental for decision-making. However, a growing number of studies in several fields – Economics, Psychology and Medicine – present evidence to the contrary. In Economics, for example, there are experimental studies of social learning in which a significant proportion of subjects purchase non-informative signals (e.g., Kübler and Weizsäcker (2004), Çelen, Choi, and Hyndman (2005) and Goeree and Yariv (2006)).

The most notable evidence from the Psychology literature is provided in a series of studies by Eldar Shafir (Tversky and Shafir (1992), Shafir and Tversky (1992), Bastardi and Shafir (1998) and Redelmeier, Shafir and Aujla, (2001)), of which the most well-known is a joint experiment with Amos Tversky. In this experiment, students were offered a big discount on a holiday resort, provided it was paid for before the date of an important qualifying exam. A majority of students preferred to forgo the discount and delay their decision until information about the exam arrived. After the results of exam were known, however, a majority of students said they would have gone to the holiday resort regardless of whether they passed or failed. A similar effect was demonstrated in a related experiment where the majority of subjects playing the prisoner's dilemma asked to delay their decision until after they learned the action of their opponent, even though the majority of subjects chose to defect whether their opponent cooperated or defected.

Several studies in Medicine have raised the concern that physicians have the tendency to order too many diagnostic tests beyond the point at which such tests are likely to provide new information (see Allman, Steinberg, Keruly and Dans (1985), Myers and Eisenberg (1985) and Kassirer (1989)) . This concern is best captured by the following quote from the New England Journal of Medicine:

> "We must stop ordering tests that have little chance of changing the scope of diagnostic possibilities. We must become increasingly comfortable with uncertainty; not every diagnosis must be nailed down with the final test...before we embark on a course of therapy." (Putterman and Ben-Chetrit (1995), p.1211)

These studies *seem* to suggest that individuals *may* assign a value to information, which is above and beyond the instrumental benefit it provides. However, it is difficult to asses this hypothesis on the basis of just these studies since there are several factors, not controlled for, which may have contributed towards the decision to acquire "useless" information. For example, subjects in the Economics experiments may not have thought that they were purchasing "useless" information because they either did not update their beliefs according to Bayes rule, or because they did not conform to the equilibrium behavior (and did not expect others to conform). In the Tversky-Shafir experiment subjects might have mistakenly thought that they would choose different actions depending on their grade because they were confused by the fact that there are different reasons to choose the same action – namely, to go on vacation – whether they passed or failed (in one case it's a reward, in another it's a consolation). Finally, there are a host of reasons why physicians may be inclined to subject a patient to more tests than are necessary.

The aim of this paper is, therefore, to investigate whether controlling for all these factors, individuals derive an intrinsic demand for non-instrumental information, and if so, what may be the possible sources of this demand. One possible explanation may be the failure to apply backward induction, or what Tversky and Shafir (1992) (henceforth, TS) refer to as a "disjunction effect". According to TS, this failure may be overcome by essentially forcing individuals to think what they would choose for each piece of information they might receive. A second possible reason for demanding "useless" information may be intrinsic preferences for information, as proposed by Grant, Kajii and Polak (1998,2000) (henceforth, GKP). According to this theory, "information-loving" individuals strictly prefer more accurate signals regardless of whether they make their decision before or after observing the realization of their signal. Finally, individuals may demand information because they wish to simplify their decision and reduce the cost of finding the optimal decision (we adopt the framework proposed in Ortoleva (2009)).

We present experimental evidence on a demand for non-instrumental information for which the above theories do not provide a satisfactory explanation. We propose an alternative explanation, which is based on the notion that individuals may have an intrinsic preference for being "confident" in choosing the right decision. For example, suppose you were faced with a choice between action A and action B, where the outcome depends on an unobservable state of nature. To prefer A over B it is sufficient that, according to your current information, there is more than 50% chance that A is ex-post optimal. However, you may "feel more confident" taking action A if you knew that A is ex-post optimal with a probability much higher than 0.5 (say, 0.8). If this could be verified, then you may be willing to pay for such information in order to "raise your confidence" in choosing A.

To capture this intuition we propose a model of decision-making under risk in which individuals have an intrinsic preference over their posterior beliefs *only at the point of making a decision*. Once a decision is made, an individual's behavior is standard in the sense that he is indifferent among all signals about the outcome of his decision. We interpret this to mean that a decision-maker (DM) cares about *how* he makes a decision, i.e., with what level of confidence that his decision is ex-post optimal. In particular, we assume that decision-makers are mainly driven by a desire to rule out the worst state of nature before making a decision, even if conditional on knowing this state, they would choose the same decision.

In our experiment subjects were faced with the following simple decision problem. A monetary prize X is hidden in one of two boxes, labeled A and B. The probability that each box contains the prize depends on the state of nature. There are two possible states: *high* and *low*. The probability that box A contains the prize is h in the high state and $l < h$ in the low state, where h and l are *both strictly above* $\frac{1}{2}$. The subject's task is to choose a box. If he chooses correctly, he wins the prize. Before he makes his choice, the subject can pay a fee to learn the state of nature. If he chooses not to pay, the subject then must choose a box *without* knowing the state of nature. Whether or not the subject pays the fee, he receives a payment immediately after he makes his choice.

Since a choice of A first-order stochastically dominates a choice of B, knowing the state of nature should not affect one's choice. This is true under any model of decision-making under risk that respects first-order stochastic dominance. It is also independent of the subject's attitude towards risk. Thus, a subject who pays the fee exhibits an intrinsic preferences for non-instrumental information. To determine which of the above three theories best explains these preferences, we ran four experiments, a baseline with two variants and an additional treatment (Treatment 4), to investigate more closely what information subjects are interested in when they pay a fee. We also ran a fifth treatment to investigate the impact of learning. The treatments are briefly described as follows:

Treatment 1 (Baseline). Subjects are first asked if they want to pay a fee to learn the state of nature before making their choice. If they answer yes, they are shown the true state and are then asked to choose a box. If they answer no, they are asked to choose a box without any further information.

Treatment 2. Subjects are first asked which box they would hypothetically choose in each of the states. After they submit their answers, subjects are asked if they would like to pay a fee to learn the state. Subjects who answer yes, learn the state and are then asked to choose a box. Subjects who answer no, are asked to choose a box without ant further information.

Treatment 3. Subjects are first asked to choose a box without knowing the state of nature. *After* they make their choice, but before being paid, they are given the opportunity to pay a fee to learn the state of nature.

Treatment 4. In this treatment there are *three* equally likely states of nature: $h > m > l > \frac{1}{2}$, where $m = \frac{1}{2}(h + l)$. Before choosing a box, subjects are asked if they want to pay a fee for only one of the following pieces of information: (1) whether or not the state is h, (2) whether or not the state is m and (3) whether or not the state is l. If a subject pays the fee for one of the pieces of information, then he is given that information before choosing a box. Otherwise, he chooses a box without any further information. Treatment 4 was run in order to investigate the type of information (i.e., "good news" or "bad news") people are most interested in.

Treatment 5. To investigate the potential impact of learning on subjects' behavior, subjects were repeatedly faced with the same decision problem as in Treatment 1. The question was whether subjects' likelihood of paying diminished over time.

Our findings provide clear evidence that subjects are more than willing to pay for information that will not alter the decision they are about to make. Moreover, our results suggest that this behavior cannot be explained by either the TS disjunction effect nor the GKP theory but is consistent with our confidence hypothesis. For example, under TS, a subject who is willing to pay the fee in the Baseline will not be willing to pay the fee in Treatment 2 while under GKP, a subject would pay the fee in the Baseline if, and only if, he would pay the fee in Treatment 3. We find that a significant proportion of the subjects are willing to pay a fee both in the Baseline and in the second treatment (in contradiction to TS), but hardly any subject is willing to pay a fee in the third treatment (in contradiction to GKP). In particular, for $X = \$20$, $h = 1$ and $l = 0.6$, 18 out of 23 subjects (78%) agreed to pay a fee of $0.5 in the Baseline, 11 out of 21 subjects (52%) agreed to pay this fee in the Treatment 2 and only one out of 21 (4.7%) agreed to pay in the Treatment 3.

Our fourth treatment presents data that indicates that, by and large, subjects are more willing to pay for information about states of nature that decrease their prior rather than those that increase it even though neither information will alter their decision. In this treatment, information acquired by a DM may either be "good news" - in the sense that the posterior probability that the decision is optimal is *higher* than the prior, or it may be "bad news" - in the sense that this posterior probability is *lower* than the prior. We present evidence suggesting that decision-makers give more weight to bad news than to good news. In particular, we show that most of the subjects, who pay for non-instrumental information, prefer to know whether or not the posterior probability is the lowest it can be than to know whether or not it is the highest it can be.

Finally, our fifth treatment provides evidence that our results are robust with respect to learning.

In a small follow up study to this paper, (Eliaz and Schotter (2007)) we examined two extensions to the above treatments. First, we provide evidence that introducing a lag, between the time decisions are made and the time the outcome is revealed, increases the proportion of subjects who are willing to pay a fee, (but this increase is signficant only at the 13% level). Second, we show that very few subjects were willing to pay a fee after they learned their payoffs.

The remainder of the paper is organized as follows. Section 2 introduces alternative theories of a demand for non-intrumental information. In Sections 3 through 6 we present both the design of each treatment and our results. For each treatment we discuss whether its results could be explained by the theories we discussed above. Section 7 offers some concluding remarks.

2 Modeling a demand for non-instrumental information

2.1 Preliminaries

Two-stage lotteries. To simplify the exposition and with an eye towards the experiments, we describe the models in this section in the context of the experimental setting. Consider an environment with only two possible outcomes: one yields a utility of 0 and another yields a strictly positive utility, which we normalize to 1. We will refer to the outcome of receiving a utility of 1 as "winning". With slight abuse of notation, we denote a one-stage lottery over $\{0, 1\}$ by the probability of obtaining 1 .Thus, a lottery that yields x with probability q and 0 with probability $1 - q$ is denoted by q. A two-stage lottery that yields a lottery q_k with probability p_k, where $k = 1, \ldots, K$ and $\sum_{k=1}^{K} p_k = 1$ is denoted $l \equiv (p_k, q_k)_{k=1}^{K}$. Note that a one-stage lottery that yields the outcome 1 with probability q is a special case of a two-stage lottery where $p_k = 1$ for some k.

Decision-problems. We focus on decision-problems involving the choice of an action from the set $D \subseteq \{A, B\}$. Each action generates a state-dependent probability distribution over outcomes. There is a finite set of states Ω, where each state ω is realized with probability $p(\omega)$. In every state $\omega \in \Omega$ the probability of winning when action $d \in D$ is chosen is denoted $q_\omega(d)$. In our experimental design, we shall focus on decision-problems in which the DM needs to guess in which of two locations, A or B, a prize is located. Hence, in each state, the probabilities of winning induced by the two actions are negatively correlated: $q_\omega(B) = 1 - q_\omega(A)$ for each $\omega \in \Omega$. To simplify the exposition, we shall therefore let q_ω denote the probability of winning when A is chosen in state ω.

Thus, the choice from D may be viewed as a choice from a *menu* of two-stage lotteries, $M_0 = \{l_A, l_B\}$, where for each $d \in D$,

$$l_d = [p(\omega), q_\omega(d)]_{\omega\in\Omega}$$

I.e., the first stage of l_d consists of drawing a state ω with probability $p(\omega)$, and the second stage consists of drawing a prize with probability $q_\omega(d)$.

Suppose a DM could obtain some information on the state of nature before choosing a decision from D. Obtaining perfect information on the state is akin to choosing a lottery over menus of one stage lotteries in which a menu $M_\omega = \{q_\omega(A)\ , 1 - q_\omega(A)\}$ is drawn with probability $p(\omega)$. We denote such a lottery by $[p(\omega), M_\omega]_{\omega\in\Omega}$. The choice problem of whether or not to obtain information may therefore be viewed as a choice between the sure menu M_0 and the lottery over menus, $[p(\omega), M_\omega]_{\omega\in\Omega}$.

Supppose next that the DM had access only to imperfect information on the state. This means that before making a decision from D, he would observe a realization x from some finite set of realizations, X, where $\lambda(x)$ is the probability of observing realization x and $\eta(\omega \mid x)$ is the posterior probability that the state is ω, conditional on the realization x. Note that for any realization $x \in X$, $\sum_{\omega\in\Omega} \eta(\omega \mid x) = 1$, and for every state ω, $\sum_{x\in X} \lambda(x)\eta(\omega \mid x) = p(\omega)$. We refer to the triplet (X, λ, μ) as a *signal*. For every realization $x \in X$, the DM will face a menu of two-stage lotteries, $M_x = \{l_{x,A}, l_{x,B}\}$, where for each $d \in D$,

$$l_{x,d} = [\eta(\omega \mid x), q_\omega(d)]_{\omega\in\Omega}$$

We therefore view the decision to obtain imperfect information on the state as a choice of the lottery over menus, $[\lambda(x), M_x]_{x\in X}$.

We now present several models that allow a DM to strictly prefer one extended lottery to another even when both induce the same distribution over outcomes.

2.2 Intrinsic preferences for confidence.

The first model attempts to capture a DM who values non-instrumental information because he has intrinsic preferences over *how* he makes a decision. In particular, a DM may intrinsically care about "how confident" he feels in his decision. We model this by assuming that the DM's ranking of signals not only depends on the distributions they induce over final outcomes but also on the induced probability distributions over *posterior beliefs*. Thus, our model follows the growing literature on anticipatory feelings, which proposes to enrich the classical domain of preferences by adding the DM's beliefs as an argument in his utility function (most notably, Akerlof and Dickens (1982), Caplin and Leahy (2001,2004), Brunnermeier and Parker (2005), Eliaz and Spiegler (2006) and Koszegi (2006)).

Define an *extended* two-stage lottery as the pair (l, θ), where θ is the probability that l induces on the outcome 1 (i.e., $\theta = \sum_{k=1}^{K} p_k q_k$). To simplify the exposition, we shall henceforth denote an extended one-stage lottery by q with the understanding that for such a lottery, $\theta = q$. A standard expected utility maximizer (EUM) prefers (l, θ) to (l', θ') if and only if $\theta > \theta'$.Thus, an EUM is indifferent between any pair of two-stage lotteries that induce the same distribution over outcomes. For any extended lottery (l, θ), let $EU(l, \theta) = EU(l) \equiv \theta$. Note that $EU(l)$ is just the first moment of l. An EUM strictly prefers (l, θ) to (l', θ') if and only if $EU(l) > EU(l')$.

We say that a DM has an *intrinsic preferences for confidence* if there exists a function $v : [0, 1] \rightarrow R$ such that the DM's preferences over extended lotteries is represented by a function $V(l, \theta)$ defined as follows:

$$V(l, \theta) \equiv EU(l) + v(\theta) = \theta + v(\theta) \tag{1}$$

We assume that v is at least thrice differentiable with the first two derivative being strictly positive. The interpretation of (1) is that the DM's valuation is composed of two components: the "material" value of the lottery, $EU(l)$, and the "psychological" value of anticipating a reward with probabiltiy θ, i.e., his confidence in winning. Since a higher likelihood of winning a reward should lead to better anticipatory feelings, v is assumed to be increasing in θ. We interpret the convexity of v to mean that a marginal increase in θ has a greater impact on one's confidence, the closer θ gets to certainty.

To describe the DM's attitudes to information we need to define the DM's preferences over lotteries on menus of extended two-stage lotteries. Let M and M' be a two menus of extended two-stage lotteries. We assume the DM prefers M to M' if and only if he prefers the best lottery in M to the best lottery in M',

$$W(M) \equiv \max_{(l,\theta)\in M} V(l,\theta) \geq \max_{(l\prime,\theta\prime)\in M\prime} V(l',\theta') \equiv W(M')$$

We assume that the DM's preferences over lotteries on such menus, are as follows: he prefers $[\lambda(x), M_x]_{x\in X}$ to $[\lambda(y), M_y]_{y\in Y}$ if and only if $\sum_{x\in X} \lambda(x)W(M_x) \geq \sum_{y\in Y} \lambda(y)W(M_y)$.

Example 1. To give a simple illustration of our model, assume $v(\theta) = \theta^n$, where $n \geq 2$. Suppose there are two equally likely states, $\{L, H\}$, where $q_H = 1$ and $q_L = 0.6$ (recall that these denote the probabilities of winning, conditional on choosing A). Consider the decision problem of whether or not to obtain perfect information about the state. The menu corresponding to choosing no information is $M_0 = \{(l_A, 0.8), (l_B, 0.2)\}$, where

$$\begin{aligned} l_A &= [(\frac{1}{2}, 1), (\frac{1}{2}, 0.6)] \\ l_B &= [(\frac{1}{2}, 0), (\frac{1}{2}, 0.4)] \end{aligned}$$

Obtaining information yields a 50-50 lottery between the menu $M_H = \{1, 0\}$ and $M_L = \{0.6, 0.4\}$. Clearly, the DM will choose $(l_A, 0.8)$ from M_0, from M_H he will choose the degenerate lottery 1, and from M_L he will choose the one-stage lottery 0.6. Thus, whether or not he obtains information, the DM will always choose the lottery corresponding to a choice of A. Hence, the information captured by the 50-50 lottery between M_L and M_H is non-instrumental. Consequently, both the lottery over menus and M_0 yield the same expected "material" utility of 0.8, and hence, an EUM would be indifferent between the two options. However, a DM whose preferences are represented by (1) would strictly prefer the lottery over M_L and M_H to M_0. This follows from our assumption on v, which implies,

$$\begin{aligned} W(M_0) &= 0.8 + (0.8)^n \\ \frac{1}{2}W(M_H) + \frac{1}{2}W(M_L) &= 0.8 + \frac{1}{2}(1)^n + \frac{1}{2}(0.6)^n \end{aligned}$$

This example, therefore, demonstrates that a DM with intrinsic preferences for confidence, as modeled here, may exhibit a demand for non-intrumental information. □

For any lottery over menus, $[\lambda(x), M_x]_{x\in X}$, and for every realization $x \in X$, define

(l_x^*, θ_x^*) to be the extended lottery satisfying

$$V(l_x^*, \theta_x^*) = W(M_x)$$

Note that every lottery on menus, $[\lambda(x), M_x]_{x\in X}$, induces a lottery over the *beliefs*, $(\theta_x^*)_{x\in X}$, where with probability $\lambda(x)$ the DM will hold the belief that his likelihood of winning is θ_x^*. Note that the expected utility of this lottery is equal to the expected belief, $\sum_{x\in X} \lambda(x) \cdot \theta_x^*$. Let Θ^* denote the lottery over beliefs $[\lambda(x), \theta_x^*]_{x\in X}$ and let $\mu_m(\Theta^*)$ denote the m-th central moment of Θ^*. For example, $\mu_1(\Theta^*)$ is the mean of Θ^*, $E(\Theta^*)$, $\mu_2(\Theta^*)$ is the variance of Θ^*, $\sigma^2(\Theta^*)$ and $\mu_3(\Theta^*)$ is the skewness of Θ^*, $E[\Theta^* - E(\Theta^*)]^3/\sigma(\Theta^*)$. Our assumptions on v imply that the following:

Lemma 1 *The DM's preferences increase with the variance $\mu_2(\Theta^*)$ and decrease with the skewness $\mu_3(\Theta^*)$.*

Proof. The proof follows from Scott and Horvath (1980) and is given here for the sake of completeness. Let v^n denote the n-th derivative of v. Applying a Taylor expansion to $v(\theta_k^*)$ and taking the expected value of both sides yields

$$E[v(\theta_k^*)] = \sum_{i=1}^{\infty} \frac{\mu_i(\Theta^*)}{i!} v^n[\mu_1(\Theta^*)]$$

From our assumption that $v^2 > 0$ it immediately follows that the DM's preference increase with the variance. Assume that $v^3(\theta) > 0$ for all $\theta \in [0,1]$, or that $v^3(\theta) = 0$ for all θ. The Mean Value Theorem states that for $\theta > \theta'$ there exists $\bar{\theta} \in [\theta, \theta']$ for which $v^1(\theta) = v^1(\theta') - v^2(\bar{\theta})(\theta' - \theta)$. From our assumption on $v^3(\cdot)$, it follows that $v^2(\theta) \leq v^2(\bar{\theta})$ and $v^1(\theta) \leq v^1(\theta') - v^2(\theta)(\theta' - \theta)$. For $\theta' \leq \theta^* = \theta + [v^1(\theta')/v^2(\theta)]$ we have $v^1(\theta) \leq 0$ and $v^1(\theta) < 0$ for $\theta' > \theta^*$. But this contradicts our assumption that $v_1(\cdot) > 0$. ■ ■

Lemma 1 has important behavioral implications. Recall the decision problem described in Example 1 above, where the DM faced the decision of whether or not to obtain perfect information. Note that as long the DM satisfies first-order stochastic dominance, this information is non-instrumental. We argue that Lemma 1 implies that a DM with intrinsic preferences for confidence will strictly prefer to obtain this information. To see this, note that the lottery over $\{M_L, M_H\}$ induces a lottery over the DM's beliefs on winning: with probability $p(H) = \frac{1}{2}$ he will hold a belief of $q_H = 1$ and with probability $p(L) = \frac{1}{2}$, he will have a belief of $q_L = 0.6$. Note that his expected belief of winning when he obtains information, is precisely the same as his prior belief in the absence of information (i.e., the belief induced by the two-stage lottery, l_A). However, when the DM aquires information, he essentially obtains a mean-preserving spread of his prior belief. Lemma 1 implies that this is preferred to having the prior belief $\theta = 0.8$ with certainty. In other words, the DM prefers a situation in which with probability $p(H)$ he would choose A and expect to win with probability q_H and with probability $1 - p(H)$ he would choose A and expect to win with probability q_L, over a situation in which he would choose A and expect to win with probability $p(H)q_H + [1 - p(H)]q_L$. We summarize this conclusion in the following observation,

Observation 1 (demand for non-instrumental information). *Let $[\lambda(x), M_x]_{x \in X}$ be a lottery over menus of one-stage lotteries that induces a lottery over beliefs, Θ^*. Then a DM with intrinsic preferences for confidence prefers $[\lambda(x), M_x]_{x \in X}$ to an extended two-stage lottery (l, θ) that satisfies $EU(l) = \theta = \mu_1(\Theta^*)$.*

Example 2. Assume there are three equally likely states of nature: L, M and H. The probabilities of winning in states L and H, conditional on choosing A, are q_H and q_L, respectively. In state M, the probability of winning, conditional on choosing A (B) is the average probability across states H and L : $\frac{1}{2}q_H + \frac{1}{2}q_L$ ($\frac{1}{2}(1 - q_H) + \frac{1}{2}(1 - q_L)$). Suppose next that the DM was given the option to receive one of two noisy signals, $s_H = (X_H, \lambda_H, \eta_H)$ and $s_L = (X_L, \lambda_L, \eta_L)$, about the state of nature before making his choice. Each signal has two equally likely realizations, h or l (i.e., $X_L = X_H = \{l, h\}$ and $\lambda_H(x) = \lambda_L(x) = \frac{1}{2}$ for each $x \in \{l, h\}$). Signal s_H (s_L) is biased in favor of state H : when the state is H (L) the realization is h (l), otherwise, the realization is l (h). Therefore, $\eta_H(H \mid h) = 1$ and $\eta_H(L \mid l) = \eta_H(M \mid l) = \frac{1}{2}$. Similarly, $\eta_H(L \mid l) = 1$ and $\eta_H(H \mid h) = \eta_H(M \mid h) = \frac{1}{2}$.

Suppose the DM acquires signal s_H. Choosing between A and B after a realization of h is equivalent to choosing from the menu of one-stage lotteries, $M_h^H = \{q_H, 1 - q_H\}$, while choosing A or B after a realization of l is equivalent to a choice from the menu of one-stage lotteries, $M_l^H = \{\frac{3q_L + q_H}{4}, 1 - \frac{3q_L + q_H}{4}\}$. Thus, s_H induces the lottery over menus, $[(\frac{1}{3}, M_h^H), (\frac{2}{3}, M_l^H)]$. Similarly, the signal s_L induces the lottery $[(\frac{2}{3}, M_h^L), (\frac{1}{3}, M_l^L)]$, where $M_h^L = \{\frac{q_L + 3q_H}{4}, 1 - \frac{q_L + 3q_H}{4}\}$ and $M_l^L = \{q_L, 1 - q_L\}$. Conditional on choosing s_H, a DM will choose q_H from M_h^H and $(\frac{3q_L + q_H}{4}, \frac{3q_L + q_H}{4})$ from M_l^H. Conditional on choosing s_L, a DM will choose q_L from M_l^L and $(\frac{q_L + 3q_H}{4}, \frac{q_L + 3q_H}{4})$ from M_h^L. Consequently, both signals induce the same expected probability of winning $(\frac{1}{2}(q_L + q_H))$ and the variance of this probability is also identical across the two signals. However, the distribution over the posterior belief of winning, which is induced by s_L (s_H) is *negatively skewed* (positively skewed). Hence, by Lemma 1, the DM would prefer s_L to s_H. $\square$

The preference for negative skewness, as implied by Lemma 1 and illustrated in Example 2, can be stated more generally as follows:

Observation 2 (preference for negative skewness). *Let* $L \equiv [\lambda(x), M_x]_{x\in X}$ *and* $\tilde{L} \equiv [\tilde{\lambda}(y), \tilde{M}_y]_{y\in Y}$ *be a pair of lotteries over menus of one-stage lotteries that induce the lotteries over beliefs,* Θ^* and $\tilde{\Theta}^*$*, respectively. If* $\mu_1(\Theta^*) = \mu_1(\tilde{\Theta}^*)$, $\mu_2(\Theta^*) = \mu_2(\tilde{\Theta}^*)$ *but* $\mu_3(\Theta^*) < 0 < \mu_3(\tilde{\Theta}^*)$, *then a DM with intrinsic preferences for confidence would prefer* L to $\tilde{L}$.

A preference for negative skewness may be interpreted as a preference to rule out "extremely bad news", in the following sense. A distribution that is negatively skewed has a long tail on low values, which in our context represents a very low posterior on the event that the action taken is ex-post optimal. Hence, a DM who chooses such a distribution over posterior may be interpreted as an individual who wants to know whether or not the worst case has occured.

2.3 Intrinsic preferences for information.

One possible reason for demanding non-instrumental information may be sheer "curiosity", or intrinsic preferences for information. GKP propose a model that links such preferences to a failure to reduce compound lotteries. The GKP model generalizes the Kreps and Porteus (1978) model of preferences for the timing of resolution of uncertainty by relaxing both recursivity and expected utility. The key implication of the GKP model is that a DM is sensitive to the timing at which information is revealed. In particular, an information-loving DM is not indifferent among two-stage lotteries that induce the same distribution over outcomes if they differ in the timing of resolution of uncertainty.

To formalize this, we shall adapt a simple version of the GKP model in which a DM reduces a two-stage lottery to a distribution over final outcomes if all information is revealed at the end of the second stage. However, to evaluate two-stage lotteries in which the outcome of the first stage *is* revealed *before* the second stage lottery is carried out, the DM carries out the following computation. First, he computes the certainty equivalents of the second-stage lotteries. Then he computes the expected utility of the first-stage lottey over the second-stage certainty equivalents using a convex utility functional v.

A two-stage lottery l in which the outcome of the first-stage is observed before the second-stage lottery is executed is denoted $(l, 1)$. If only the outcome of l is revealed at the end of the second period, we use the notation $(l, 2)$. We refer to (l, t), $t \in \{1, 2\}$, as a *temporal* lottery. Let $CE(q)$ be the certainty equivalent of a one-stage lottery that yields a utility of 1 with probability q. The DM's preferences over temporal lotteries are represented by the following functional:

$$U(l,t) = \begin{cases} v[CE(\sum_{k=1}^{K} p_k q_k)] & if \quad t = 2 \\ \sum_{k=1}^{K} p_k v[CE(q_k)] & if \quad t = 1 \end{cases}$$

where $v : R \to R$ is increasing and strictly convex. A DM is said to be *information-loving* if he prefers a temporal lottery (l, t) to (l', t') if and only if $U(l, t) > U(l', t')$. The convexity of v implies the following:

Observation 3. *For any two-stage lottery l, an information-loving DM strictly prefers $(l, 1)$ to $(l, 2)$.*

The DM's preferences over menus of temporal lotteries and over lotteries on menus are defined as in the previous model. Let M and M' be a two menus of temporal two-stage lotteries. We assume the DM prefers M to M' if and only if he prefers the best lottery in M to the best lottery in M',

$$W(M) \equiv \max_{(l,t) \in M} U(l,t) \geq \max_{(l',t') \in M'} U(l',t') \equiv W(M')$$

As for lotteries on menus, the DM prefers $[\lambda(x), M_x]_{x \in X}$ to $[\lambda(y), M_y]_{y \in Y}$ if and only if $\sum_{x \in X} \lambda(x) W(M_x) \geq \sum_{y \in Y} \lambda(y) W(M_y)$.

Note that Observation 3 implies a demand for non-instrumental information: when there is no decision to make (i.e., when faced with a singleton menu), any information is non-instrumental. In contrast, a DM with intrinsic preferences for confidence values non-instrumental information only if he faces a *non-singleton menu of choices* at the point of receiving this information. From (1) it follows that when a DM of this type has no action to take (i.e., he faces a singleton menu), he is indifferent between a pair of two-stage lotteries that induce the same probability of winning (i.e., the temporal aspect of a two-stage lottery does not affect his preferences). Hence, we should not observe this DM paying to merely observe the outcome of a given two-stage lottery.

2.4 A disjunction effect.

In order to verify whether or not a signal is instrumental, a DM needs to carry out backward induction. He must ask himself what decision he will take after each realization, and verify that it is different from the action he would take based only on his priors. The disjunction effect refers to the tendency of individuals to substitute simple heuristics that may associate an instrumental value to non-instrumental signals for backward induction.

To formalize the disjunction effect, consider the decision-problem of choosing between A and B, where the winning probabilities associated with each action depend on the state of nature. When the DM is offered the opportunity to learn the state of nature before he makes a choice, assumer he applies the following naïve heuristic to decide whether or not this information is instrumental: if for some pair of states, $\omega, \omega' \in \Omega$, $q_{\omega}(d) \neq q_{\omega'}(d)$ for $d = A, B$, then the information is deemed instrumental. That is, the DM mistakenly interprets changes in the winning probabilities across states as changes in the optimal actions.

Let l_d denote the following two-stage lottery: in the first stage, a state ω is drawn with probability p_{ω}, and in the second stage, a prize is won with probability $q_{\omega}(d)$. Let l_d^{ω} denote a one-stage lottery where a prize is won with probability $q_{\omega}(d)$. Consider the problem of whether or not to obtain information on the state of nature before making a decision $d \in \{A, B\}$. We represent this problem as the choice from a set of menus $\mathcal{M}$, which contains a menu $\{l_A, l_B\}$ and a lottery over menus $(p_{\omega}, \{l_A^{\omega}, l_B^{\omega}\})_{\omega \in \Omega}$. Information on the state ω is non-instrumental for a decision maker, if there exists $d \in \{A, B\}$ such that he would choose l_d from $\{l_A, l_B\}$ and would also choose l_d^{ω} for every realized state ω.

TS interpreted the disjunction effect as a decision error in the sense that it was unlikely to occur when subjects were "forced" to carry out backward induction: before being asked whether or not they wanted access to a signal, they were explicitly asked to imagine what they would choose for each possible realization. This suggests an interpretation of the disjunction effect as an instance of a "choice with frames" a là Salant and Rubinstein (2008). When the decision-problem of choosing between signals is presented with a statement that prompts the DM to think of each realization, the DM applies backwards induction. But when the same decision-problem is presented without such a statement, the DM has a tendency to substitute the heuristic described above for backward induction. We will use this observation as a testable implication of the disjunction effect.

To formalize this interpretation of the disunction effect, we let $\langle \mathcal{M}, F \rangle$ denote the problem of choosing whether or not to obtain information under frame F. Let F^N denote a neutral frame in which the DM is only asked to choose between $\{l_A, l_B\}$ and $(p_{\omega}, \{l_A^{\omega}, l_B^{\omega}\})_{\omega \in \Omega}$. Let F^{BI} denote the frame in which before being asked to choose from M, the DM is asked to imagine which action $d \in D$ he would choose in each state of nature (i.e., the DM is primed to apply backward induction) The following observation is as a testable implication of the disjunction effect.

Observation 4. *If a DM chooses* $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$ *in both* $\langle \mathcal{M}, F^N \rangle$ *and* $\langle \mathcal{M}, F^{BI} \rangle$*, then he does not exhibit a disjunction effect.*

2.5 Costly thinking.

Even when information does not affect the DM's choice of action, it can affect the level of cognitive effort needed to make this choice. For example, suppose there are two equally likely states, ω_L, ω_H, such that in ω_H, action A leads to a sure win, while in ω_L, action A leads to win with probability 0.6. Virtually no effort is required in reaching the conclusion that A should be chosen in state ω_H. On the other hand, some cognitive effort is required to realize that choosing A stochastically dominates choosing B in state ω_L. If a DM acquires information on the state of nature, then there is only a 50% that he would actually need to incur any cognitive effort in figuring out what the optimal action is.[1]

The idea that decision-makers take into account computational costs dates back to Simon (1955). More recently, Ortoleva (2009) proposed an axiomatization of preferences over lotteries of menus such that these preferences may be represented as the difference between an affine evaluation of the content of the menu, and an *anticipated thinking cost* function that assigns to each menu a "thinking cost".[2] In the present set-up, these preferences have the following representation: for any lottery over menus $[\lambda(x), M_x]_{x\in X}$,

$$\mathcal{W}\{[\lambda(x), M_x]_{x\in X}\} = \sum_{x\in X} \lambda(x) \cdot \max_{l\in M_x} EU(l) - C\{[\lambda(x), M_x]_{x\in X}\} \tag{2}$$

For the purpose of our experiment, it suffices to consider only the following key premises of this model. First, the cost of choosing from a singleton menu is zero. This implies that when the DM in this model has already chosen an action in D, he is indifferent between whether or not to obtain information on the state of nature.

[1] Note that cognitive effort here may include psychological costs other than just the computational costs involved in deriving the optimal action. For example, this effort may also include the cost of anticipatory regret, which can be avoided if there are states in which some action leads to almost a sure win.

[2] A related idea, dubbed "costly contemplation", is proposed by Ergin and Sarver (2008). One of the main differences between this paper and Ortoleva's (2009), is that the latter takes as primitive the DM's preferences over lotteries of menus, which is precisely what our experimental design attempts to ellicit.

I.e., for any two-stage lottery l, the DM is indifferent between the temporal lotteries $(l, 1)$ and $(l, 2)$.

The second key premise concerns the separation of the DM's "genuine" ranking of lotteries of menus, the one we would observe in the absence of thinking costs, from the "revealed" rankings, which incorporates thinking costs. We assume that these "genuine" rankings are induced by standard expected utility preferences over lotteries such that any departures from expected utility rankings of lotteries of menus are due to thinking costs. The "genuine" rankings are identified as follows. Let M and M' be some pair of menus of lotteries. Suppose the DM faced the choice between the following pair of lotteries on M and M'. One lottery, denoted $L_{\frac{1}{2}}$, draws each menu with equal probability. A second lottery, denoted $L_{\frac{1}{2}+\varepsilon}$ gives slightly more weight, $\frac{1}{2}+\varepsilon$, to menu M. A DM is said to rank the content of M above that of M' if and only if he prefers $L_{\frac{1}{2}+\varepsilon}$ to $L_{\frac{1}{2}}$. The intuition for this assumption is that both lotteries generate the same thinking costs since both require the DM to consider what he will choose from both M and M'.

We interpret this assumption to mean the following. Suppose the DM is considering a choice from the menus of menus, $\mathcal{M}$, which contains the menu $\{l_A, l_B\}$ and the lottery, $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$. As mentioned above, the menu $\{l_A, l_B\}$ may be viewed as more costly in terms of thinking since it requires the DM to consider the consequence of choosing $d \in D$ in *each* state of nature, while the lottery $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$ requires him to consider the consequence of choosing $d \in D$ *only in the realized state*. However, the discussion in the previous paragraph suggests that if the DM was made to consider the consequence of choosing each $d \in D$ in both states of nature, then both $\{l_A, l_B\}$ and $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$ would generate the *same* thinking costs. This implies that under these circumstances, the DM will rank $\{l_A, l_B\}$ and $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$ according to "genuine" rankings. Hence, if for one of the actions in D, $q_\omega(d) > \frac{1}{2}$ for every $\omega \in \Omega$, then the DM would be indifferent between $\{l_A, l_B\}$ and $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$. Recalling the framings F^N and F^{BI} defined in the previous subsection, we obtain the following observation:

Observation 5. *Assume $q_\omega(A) > \frac{1}{2}$ for every $\omega \in \Omega$, and suppose a DM's preferences are represented by the model of costly thinking. Then in the decision problem with frames, $\langle \mathcal{M}, F^{BI} \rangle$, the DM will be indifferent between $\{l_A, l_B\}$* and $(p_\omega, \{l_A^\omega, l_B^\omega\})_{\omega\in\Omega}$.

3 Experimental design

The entire experiment was conducted in the laboratory of the Center for Experimental Social Science at New York University. Subjects were recruited from the undergraduate population at New York University. A total of 97 subjects participated. In each treatment subjects were handed written instructions, which they first read on their own and was later read out load by one of the authors. Each subject received a show up fee of \$3, which he could keep regardless of the decisions he made in the experiment. Each subject also received an additional amount of \$4, which he could spend during the experiment and keep any amount that was left. Each treatment lasted around 45 minutes and the average payoff was \$20.20. A post experiment questionnaire was administered after the experiment was over asking subjects to explain their actions.

3.1 The baseline treatment

The instructions for the baseline treatment described a situation in which \$20 was to be placed in one of two boxes, labeled A and B, and the task of the subjects was to choose which box contained the money (detailed instructions of Treatment 1 appear in Appendix A). To determine which box would actually contain the \$20, the computer first chose one of two urns, Urn I or Urn II, with equal probability. Each computerized urn contained 100 balls, some marked with the letter A while others were marked with the letter B (all balls had some mark). In all four treatments subjects were given the composition of the two urns (i.e., how many A balls are in each urn) before being asked to choose a box. The computer then drew at random one ball from the chosen urn and placed the \$20 in the box marked on the ball.

The important feature of our design was that the percentage of A balls in Urns I and II, denoted by α and β respectively, was *strictly above* 50, hence, there was a greater chance that box A contained the \$20 no matter which urn was chosen. This was emphasized in bold letters in the instructions (see Appendix A). Thus, choosing A first-order stochastically dominated choosing B.

Subjects were presented with 15 situations that differed in the composition of each urn (see Table 1).[3] The order of these situations was randomly selected for each subject. In each situation, a subject was asked if he would be willing to pay a fee f to see which urn was selected *before choosing a box*. The size of the fee also varied across the situations as may be seen from Table 1. For example, in one situation $\alpha = 1$, $\beta = 0.6$ and $f = \$2.00$. Here, learning that the ball was drawn from Urn I meant that A is the correct choice with probability 1, while learning that the ball was drawn from Urn II meant that A is still the correct choice but with a lower probability of 0.6.

[3] In Treatment 1, by mistake, Situation 8 that was presented to subjects was identical to Situation 1 (i.e. had the same α, β and f). Hence, there were only 14 non-identical situations presented. Only one subject gave different answers for the two situations, and we did not include him in our data.

After subjects had finished saying yes or no to purchasing information in all 15 situations, one was chosen at random.[4] If for the chosen situation a subject has stated that he would pay to know from which urn the ball was drawn, then that information was revealed to him and he was asked to choose either box A or box B. If the situation chosen for a subject was one for which he decided *not* to pay the fee, then he was asked to choose a box without knowing from which urn the ball was drawn. The outcome - whether or not they chose correctly - was then presented to the subjects and their payoff calculated.

Table 1

Table 1 presents the 15 situations presented to our subjects and exhibits a number of important properties of our experimental design. First, in each of the situations there are more than 50 A balls in each urn, hence, choosing A first-order stochastically dominates choosing B *regardless of which urn is used.* Second, the highest fee was \$4, which is precisely the amount that subjects received at the beginning of the experiment *in addition* to their \$3 show-up fee. Thus, even a subject, who paid a fee of \$4 and failed to choose the box with the prize, still walked out of the experiment with the \$3 show-up fee. For example, if the fee was \$2.00 for a particular situation and the subject decided to pay it and then chose a box that contained the \$20, his or her payoff would be

$$\underbrace{\$20}_{\text{prize}} + (\underbrace{\$4}_{\text{initial endowment}} - \underbrace{\$2}_{\text{fee paid}}) + \underbrace{\$3}_{\text{show-up fee}} = \$25$$

If the chosen box contained no money, the subject's payoff would be \$5.

3.2 Treatment 2: asking subjects to think

We now turn to investigate whether the demand for non-instrumental information that was obtained in the baseline treatment can *all* be attributed to either the disjunction effect or to costly thinking. According to TS, the disjunction effect can be undone by having a person ask himself what he would choose under each contingency (selected urn). Similarly, Observation 5 suggests that when this person answers the question, he effectively incurs the "thinking costs" associated with the decision-problem (since he essentially solves it), hence, no additional information could help reduce these costs. This implies that if our subjects were forced to think what box they would choose for

[4] As mentioned in the Introduction, the information that was offered to subjects in Treatment 4 was different from the information offered in the other three treatments. For more details, see Section 5.

each choice of urn, then they would stop paying for this information. This would not be true if a preference for confidencewas relevant.

To investigate whether this implication would hold in our setup, we modified Treatment 1 as follows. In each of the situations, subjects were asked *three* questions: (1) which box would they choose if they learned that Urn I was used to place the $20? (2) which box would they choose if they learned that it was Urn II? and (3) whether they wanted to pay the fee in this situation. The first two answers were *not* binding, however, they only served to make the subject aware of what he would do in each contingency.[5]

As in Treatment 1, after the subjects answered the above questions in all 15 situations, one situation was drawn at random. If a subject had stated that he wanted to pay the fee for that situation, he was told which urn was selected and then asked to choose a box. If he did not wish to pay the fee for that situation, then he was simply asked to choose a box without knowing the choice of urn. Note that regardless of their answers in the 15 situations, subjects could choose any box they wished for the randomly chosen situation. By Observations 4 and 5, the disjunction effect and the model of costly thinking, cannot explain the behavior of a subject who chooses to pay the fee after answering that he would choose the same box regardless of the choice of urn. Therefore, subjects who exhibit the disjunction effect, or who tried to minimize the cost of thinking, would not pay any fee in Treatment 2, while subjects who exhibit the confidence effect, would behave the same way in Treatments 1 and 2.

3.3 Treatment 3: Intrinsic preferences for information

In this next treatment we investigate whether the behavior of subjects may be explained by GKP's model of intrinsic preference for information per se. To answer this question, we ran the following variation of our baseline treatment. In each of the 15 situations, subjects were asked if they were willing to pay a fee to learn the choice of urn *only after* they have chosen a box. By Observation 3, for a given situation, an information-loving subject would pay a fee in both Treatments 1 and 3. Hence, if our subjects have intrinsic preferences for information, we would expect approximately the same fraction of subjects to pay the fee in Treatment 3 as did in Treatment 1. In contrast, Observation 1 states that a subject with intrinsic preferences for confidence would not

[5]Making the answers binding would mean that subjects essentially guess a box *before* they get the chance to learn which urn was drawn. This effect is investigated in Treatment 3 and in our learning treatment.

be willing to pay in the current treatment.[6]

3.4 Treatment 4: "Good" news vs. "bad" news

An individual who demands non-instrumental information is not indifferent among lotteries on his posterior beliefs even though the posterior beliefs in the support of these lotteries, all induce exactly the same action. In particular, he exhibits a preference towards a mean preserving spread of his prior belief that choosing A yields the prize. The different situations in Treatment 1 allowed us to investigate whether the willingness to pay for non-instrumental information increases with the variance of the lottery on one's posteriors. The current treatment explores the subjects' attitudes towards the *skewness* of lotteries on their posteriors.

Attitudes towards skewness may be interpreted as the reason for preferring a greater variance in one's posterior beliefs. According to the model of intrinsic preferences for confidence, subjects exhibit a preference for *negative* skewness (recall Lemma 1), which may be interpreted as being averse to making a decision with an upward biased belief that this decision is ex-post optimal (i.e., they prefer *not* to make a decision with a "false sense of confidence"). One possible reason for this might be that individuals want to "prepare themselves" for a situation in which there is still a significant chance (e.g., 49%) that the best ex-ante decision may prove to be ex-post wrong. Consequently, subjects are willing to pay for the information offered to them because they would like to know if the "*bad*" realization (a draw of Urn II) has occured. Therefore, the lower the chances are of drawing the A ball from Urn II, the more valuable the lottery is.[7]

A preference towards positive skewness may be interpreted as assigning a sufficiently high value to making the choice of A with the high posterior that they are willing to risk the possibility of obtaining the low posterior (just as, by analogy, individuals may be willing to pay for a lottery that offers a high monetary prize even at the risk of realizing a loss). Thus, the higher the upper posterior is, the more valuable the information is.

Treatments 1-3 do not allow us to address these different interpretations because there are only two states of nature in these treatments and subjects pay to learn the exact state. Consequently, a subject who wants to know whether or not the information is "bad news" is indistinguishable from a subject who wants to know whether or not it is

[6] Our interpretation of GKP is supported by earlier studies of preferences towards the timing of uncertainty resolution, most notably, Chew and Ho (1994), Ahlbrecht and Weber (1996), von Gaudecker, van Soest and Wengström (2008) and the results cited in Barberis, Huang and Thaler (2006).

[7] We use the term "good news" and "bad news" looselt here. By "good (bad) news" we mean a high (low) posterior that the action to be taken (action A) is ex-post optimal.

"good news". We, therefore, ran a fourth treatment where subjects also faced the same 15 situations of Treatments 1-3, but these were presented in a slightly different way (the detailed instructions for this treatment appear in Appendix B). In each situation subjects were told that there were *three* urns used to determine where the \$20 prize was to be placed. The percentage of A balls in Urns I, and III where α and β respectively, and the percentage of A balls in Urn II was $\frac{1}{2}(\alpha + \beta)$. Similar to Treatments 1-3, a computer program randomly selected one of the three urns, and from that urn it randomly selected a ball. The \$20 was then placed in the Box that corresponded to the letter on the drawn ball. Note that given this procedure, the percentage of A balls in urn II is exactly the prior probability that the \$20 will be placed in box A.

For each subject the computer program randomly selected a sequence of the 15 situations. For each situation, subjects were asked if they were willing to pay the fee to learn the answer to one (and only one) of the following questions, and if so, which question they would pay for:

- Question 1 : Was the ball drawn from Urn I ("the high urn")?
- Question 2 : Was the ball drawn from Urn II ("the middle urn")?
- Question 3 : Was the ball drawn from Urn III ("the low urn")?

Note that the answers to both Questions 1 and 3 induce the same mean and variance of the winning probability from A. The difference between the two answers is in the skewness of the distributions over winning probabilities. The distribution induced by Question 1 is positively skewed, while the distribution induced by Question 3 is negatively skewed. By observation 2, if a subject has intrinsic preferences for confidence, then he is expected to pay for an answer to Question 3. Since Question 2 is totally uninformative, paying for an answer to this question is evidence for some confusion on part of the subject.

4 Results

4.1 The baseline treatment

The results of this treatment are presented in Table 2.

Table 2

Note that in some situations, more than 50% of the subjects were willing to pay the fee. In particular, for $\alpha = 1$, $\beta = 0.6$ and $f = \$0.50$, almost 80% of the subjects were willing to do so. Furthermore, when the fee was \$2, or 10% of the prize, about 56% of the subjects agreed to pay the fee when $\alpha = 1$ and $\beta = 0.51$, and about 43% of them agreed to do so when β rose to 0.6.

Table 2 suggests a systematic structure to the data. First, the fraction of subjects who pay the fee decreases with the level of the fee. Consider for example the pair of posteriors $\alpha = 1.0$ and $\beta = 0.51$. While 74% of the subjects paid to find out which of these two posteriors was picked when the fee was \$0.50, about 56% of the subjects paid when the fee was \$2.00 and only 26% paid when the fee was \$4.00. Figure 1 gives a graphical representation of the negative relation between willingness-to-pay and the size of the fee.

Figure 1

Second, fixing the fee and the probability β, the fraction of paying subjects *increases* with α. This may be interpreted as evidence in favor of the model of intrinsic preferences for confidence, which asserts a positive relation between an individual's preference over lotteries on his posteriors and the variance of these lotteries. For example, when the fee is \$0.50 and $\beta = 0.51$, about 48% of the subjects paid the fee when $\alpha = 0.6$, roughly 65% paid when $\alpha = 0.8$, and approximately 74% paid when $\alpha = 1.0$. In contrast, if we fix the fee and the probability α, the fraction of subjects who pay the fee typically *decreases* with β. For example, when the fee is \$2.00 and $\alpha = 1.0$, the percentage of subjects who pay the fee decreases from 56% to 43% when β increases from 0.51 to 0.60. Figure 2 depicts the positive relation between willingness-to-pay and the difference $\alpha - \beta$ (referred to as "spread" in the figure).

Figure 2

To investigate further the correlation between the likelihood of paying the fee and each of the three parameters, α, β and the size of the fee, we ran a simple random effects logit regression on the pooled data of the baseline treatment. In this regression we regressed our binary variable ($\{0, 1\}$) indicating a fee payment on the difference between α, and β (diff) and the fee. We use the difference variable because the spread between α and β reflects the potential increase in confidence that could result from paying the fee. Table 3 summarizes the results of this regression.

Table 3

As evident from Table 3, the coefficient on the difference variable (diff) variable is positive indicating that an increasing the difference between the good and bad urns increases the likelihood of paying the fee, while the coefficient on the fee variable is negative, indicating that as the fee increases subjects are less likely to make a purchase. Both variables are significant at less than the 1% level. This lends further support to our argument in the previous paragraph that our data has a systematic structure and was not generated by a totally random process. In particular, the results of the baseline treatment suggest that the willingness to pay is positively related to the distance between the low and high posteriors.

In order to interpret the above findings as a demand for non-instrumental information it is important to verify that our subjects respected first-order stochastic dominance. I.e., we need to check that in those situations that were randomly chosen to be played out, our subjects chose box A. Indeed, we find that *all* the subjects in the baseline treatment chose box A.

Some insights into what made subjects pay may be obtained by examining their answers to the post-experiment questionnaire. In this questionnaire, we asked the following: "Knowing that the chance of the $20 being in box A was always greater than 50%, one could say that Box A was always the 'best' box to choose no matter what you found out after paying the fee. If so, why did you choose to pay the fee when you did? Did you understand that the chances of the money being in Box A was greater than 50%?"

Of course, not all of the answers provided are easily interpreted and some may not truly reflect the subjects' state of mind at the point of decision. Still, a general message that comes out of the subset of coherent answers is that on the one hand, the majority of subjects understood that the information offered to them was non-instrumental, but on the other hand they had an intrinsic preference to know the likelihood in which choosing A was the correct decision. For instance, of the subset of coherent answers, a representative sample of responses includes the following:

"I paid the fee to give me confidence in my decision. That's why I would only usually pay it when it was small.";

"Yes, I understood that the chance of the money being in Box A was greater than 50%. But, in paying the inspection fee, what I bought was a sense of security, no matter how false or unnecessary or superfluous that sense was.";

"I did understand the chance that the money was in Box A was greater than 50%. However, I wanted to be positive of my choice and to not be disappointed at myself for not choosing to pay the fee and getting the answer wrong."; and

"Knowing the probability that Box A always had a higher probability was certainly a factor in choosing it, but sometimes you just need a be a bit more certain."

The particular design that we employ may give rise to other potential explanations of our findings. One potential explanation may be an induced demand effect in which the whole setup of the experiment may suggest to subject that the option to acquire information has to be good for something. Consequently, subjects may try out the option particularly in those cases where it is cheap. However, our results show that (i) a significant proportion of subjects paid the fee when it was 10% of the prize (more than 40% when there were only A balls in Urn I), and (ii) subjects paid the fee in seemingly systematic way: their willingness to pay was positively related to the distance between the low and high posteriors. Further evidence against the demand-effect explanation is provided in Treatment 3, which we discuss below.

4.2 Treatment 2

Table 4 presents the results of this treatment.

Table 4

The first notable feature of the data is that even when subjects were forced to perform backward induction, more than 40% of them paid the \$0.50 fee. In particular, more than 50% paid this fee when $\alpha = 1.0$ and $\beta = 0.6$. Even when the fee was \$2.00, a third of the subjects paid it when the distance between the high and low posterior was the greatest (i.e., when $\alpha = 1.0$ and $\beta = 0.51$). A second feature of the data is the monotonicity with respect to the cost of information: for each (α, β) pair, the fraction of subjects who pay the fee declines as the fee increases. Finally, the positive relation between the fraction of paying subjects and the distance between the low and high posterior is weaker than in the baseline treatment as it is not present across all the situations. For example, consider those situations in which the fee is \$0.50. For $\beta = 0.6$, more subjects pay the fee when $\alpha = 1.0$ than when $\alpha = 0.8$. Similarly, when $\alpha = 0.8$, more subjects pay the fee when $\alpha = 0.51$ than when $\alpha = 0.6$. However, when $\beta = 0.51$, more subjects pay the fee when $\alpha = 0.8$ than when $\alpha = 1.0$. Still, on the aggregate, when the data from all situations is pooled together, there is a positive correlation between the fraction of paying subjects and the distance between the low and high posterior.

Note that Table 4 may be interpreted as some evidence against the “costly thinking” hypothesis. According to that explanation subjects paid the fee because they were interested in reducing the "cognitive cost" associated with identifying the optimal decision. However, in the present treatment, subjects are forced to incur this cost before they get the opportunity to purchase information. Hence, their decision whether or not to acquire information is independent of this cognitive cost (recall Observation 5).

Comparing the data in Tables 4 and 2, we observe that in many situations, the proportion of subjects who paid the fee in Treatment 2 is lower than the corresponding proportion in the baseline treatment. This suggests that some - but not all - of the demand for non-instrumental information may be explained by the disjunction effect. To test this hypothesis, we compared the proportion of subjects paying the fee situation-by-situation across treatments using a test of proportions. That is, we took a given situation and looked at the fraction of subjects paying the fee in Treatment 1, compared that number to the fraction paying the fee in Treatment 2, and tested if these proportions are significantly different. Table 5 presents the results of our two-sided test of proportions.

Table 5

Note that in some situations, the proportion of subjects who paid the fee in Treatment 2 were *not lower* than the corresponding proportions in Treatment 1. In fact, there was only one situation (situation 12) where there was a significant difference (at the 5% level) in the proportion of subjects paying the fee between Treatments 1 and 2. We interpret this to mean that in situation 12, only *some* subjects - *but not all* - paid the fee because of a disjunction effect. Hence, Table 5 suggests that subjects' behavior was surprisingly similar across Treatments 1 and 2, despite the fact that in Treatment 2 subjects were asked what box they would choose for each choice of urn. This suggests that a disjunction effect, *or* an attempt to minimize thinking costs, may not be the sole explanation of our subjects' behavior.Figure 3 compares the findings of the first and second treatment.

Figure 3

Our interpretation of the data from this treatment rests on the observation that most situations have the following properties. First, most of the subjects who pay the fee state that they would choose Box A whether the ball is drawn from Urn I or II. Second, upon learning which urn was drawn, most subjects choose Box A. For instance, in nine of the 15 situations, at most two subjects said they would choose Box B if Urn

II was used. In six of these situations, those who paid, said they would always choose Box A. In particular, when $\alpha = 1.0$, $\beta = 0.6$ and the fee was \$0.50, about 52% of the subjects paid the fee, all of which said they would always choose Box A. In five of the remaining six situations, where more than two subjects said they would choose Box B if Urn II used, at most two such subjects paid the fee. Finally, of the 21 subjects who participated in Treatment 2, only *one* chose box B, and he did so in only *one* situation (Situation 2 where $\alpha = 1$, $\beta = 0.51$ and $f = \$4.00$) where he chose to pay the fee.

4.3 Treatment 3

Table 6 compares the behavior of subjects in Treatments 1 and 3, and Figure 4 compares the findings in all treatments thus far.

Table 6

Figure 4

Note the overwhelming difference in the fraction of subjects paying the fee in Treatments 3 as opposed to Treatments 1 and 2 (see Table 5). For example, while in situations 3, 6 and 12 the proportion of subjects willing to pay the fee in Treatment 3 was 0.047, 0 and 0.047 respectively (out of the 21 subjects who participated in Treatment 3, only one paid in Situations 3 and 12 and none paid in Situation 6), in Treatment 1 these same proportions were 0.782, 0.565 and 0.739, and in Treatment 2 they were 0.524, 0.333 and 0.429 respectively. Applying a test of proportions to the data in comparing behavior in Treatments 1 and 3 , we find that we can reject the null hypothesis of equality in proportions between Treatments 1 and 3 for all situations (see Table 6) in favor of the alternative one-tailed hypothesis that the proportions in Treatments 1 are *greater* than those in Treatment 3 (the same result holds in a comparison of Treatments 2 and 3).

Our results suggest that in contrast to the GKP model, our subjects were rarely willing to pay a fee when they had already made their decisions, despite the fact that paying for the fee would resolve some of their uncertainty early rather than later. We interpret these findings as evidence that our subjects' behavior cannot be explained by GKP preferences.

Treatment 3 provides further evidence against the demand effect hypothesis. To see why, suppose a subject in the baseline treatment suspected that the information offered to him has some value simply because he was being paid money to make 15 choices regarding information aquisition. Then such a subject should have the same suspicion whether he was offered to buy information before or after he made a decision. However, our results clearly indicate the when the information was offered *after* a decision was made, virtually no subject thought that this information may have some value.

We conclude this Subsection by examining whether our subjects respect first-order stochastic dominance. To investigate this, we checked the number of subjects who chose Box B in each of the situations. Of the 21 participants, one chose B in Situations 7 ($\alpha = 0.6, \beta = 0.51, \$2$) and 15 ($\alpha = 0.8, \beta = 0.6, \$4$), two chose B in Situations 9 ($\alpha = 0.8, \beta = 0.6, \$0.5$) and 11 ($\alpha = 0.8, \beta = 0.51, \$0.5$), three chose B in Situation 13 ($\alpha = 0.6, \beta = 0.51, \$0.5$) and four chose B in Situation 8 ($\alpha = 0.6, \beta = 0.51, \$4$). However, *only one of these subjects chose to pay the fee.* That subject's decision does not bias the results in our favor since our theory predicts that *no* subject should pay a fee. However, it is not clear whether the four subjects who violated first-order stochastic dominance (by choosing B), chose not to pay because of our proposed theory.

4.4 Treatment 4

Table 7 summarizes the answers to the following questions: How many subjects in total were willing to pay a fee? Conditional on paying the fee, what information were most subjects most interested in?

Table 7

As evident from this table, subjects were again willing to pay a fee in this treatment although not with the same frequency as they did in Treatment 1. For example, while 73% of the subjects in Treatment 1 were willing to pay a fee in Situation 12 ($\alpha = 1, \beta = 0.51, \$0.5$ fee) only 56% were willing to do so in Treatment 4.[8] This percentage is still significantly above zero. The same is true for Situation 3 ($\alpha = 1, \beta = 0.51, \$0.5$ fee) where the percentages are 78% versus 31.3%, respectively when comparing Treatments 1 and 4. These differences in percentages may be explained by the fact that the information sold in Treatment 1 is more precise than the information sold in Treatment 4. In Treatment 1 subject can find out exactly what the probability is of drawing an A ball. However, in Treatment 4 subject can only learn that this probability has two equally likely values.

The fractions willing to pay the fee in Treatment 4 in each situation is positively correlated with the fraction paying in Treatment 1 with a correlation coefficient of 0.568, which is significant at the 3% level. Hence, we again see that more subjects are

[8] Note that as we mentioned earlier, in Treatment 4, Urn II in this situation had 52 balls marked by A.

willing to pay the bigger the difference between the high and low urns and the smaller the fee being charged.

We now turn to the question of subjects' attitudes towards skewness. Table 7 suggests that more subjects are willing to pay for an answer to Question 3 concerning the low urn rather than to Question 1, which concerns the high urn. There exists no situation for which subjects are more likely to seek the answer to Question 1 than they are to Question 3. For example, in Situation 12, 72.2% of subjects who paid a fee, did so to find out the answer to Question 3 rather than Question 1. Similar differences existed in other situations as well. Only in Situation 14 ($\alpha = 1, \beta = 0.6, \$2$ fee) were subjects as curious to find out the answer to Question 1 as they were to Question 3. Finally, note that only five of the 32 subjects paid for an answer to Question 2. We, therefore, interpret these findings as evidence in favor of intrinsic preferences for confidence.

5 Learning

In all of the treatments discussed so far, subjects were presented with 15 decisions, which they faced only once. Hence, there was no opportunity for them to learn about any particular decision, and possibly, over time, decide not to purchase information even though they did initially. This leads one to ask whether our subjects, if allowed to learn, might eventually refrain from purchasing non-instrumental information. The results of Treatment 2 suggest that such learning would not occur since learning is equivalent to discovering, via backward indiction, that no matter what the state of nature turns out to be it is always optimal to choose A rather than B. Furthermore, in the real world, decisions similar to the ones we analyze here, are often made infrequently with little or no learning opportunity. For example, a patient with a cancer diagnosis (a once or twice in a lifetime event) is often very eager to purchase information that may change his posterior over whether an operation is the correct action to take, even if it will not change his decision to choose the operaton. Finally, we presented evidence from the medical profession where physicians repeatedly opt for expensive diagnostic tests despite the fact that they know these additional tests are unlikely to change their course of treatment.

Despite these reservations we proceeded to test for learning effects in our subject pool by running one more treatment. In this treatment, one situation was presented to a set of 24 subjects 20 times in an effort to see if their purchases of information would decrease over time. The situation we chose was Situation 12 where $\alpha = 1, \beta = 0.6$ and the fee was \$0.50. In this situation, 78.3% of subjects in Treatments 1 chose to pay for information when it was presented to them once, and 52.4 % of subjects chose to pay when it was presented to them once in Treatment 2. In the current new treatment we are interested in what fraction of subjects would choose to pay for information but, more importantly whether that fraction declines over time as the question posed to them is repeated.

In terms of summary statistics, in approximately 33% (158/480) of the 480 decisions made, subjects were willing to pay the fee to receive information. In addition, exactly 50% of subjects offered to pay for information in at least one round. It, therefore, appears that while these percentages are lower than those of Treatment 1, they are quite significant.

More importantly, there appears to be very little learning in the data. The subject pool is split between those subjects who bought information and those that never did. Conditional on paying, subjects did not decrease their willingness to do so as round progressed. To support this claim, consider Figure 5 where we present the person-by-person purchase decisions of our subjects over the 20 rounds of the experiment.

Figure 5

There are several things to notice about Figure 5. First, the set of subjects seems to be clearly divided between those who purchase information and those who do not. Twelve of the 24 subjects in this treatment bought information at least once while the rest never did. For those who bought, there seems to be little decrease in their willingness to do so as the experiment progressed. For example, while there were 40 purchase decisions by subjects over the first five rounds there were 34 such decisions over the last five rounds.

Figure 6 describes the evolution of purchase decisions over time.

Figure 6

As can be seen, there is no discernible trend downward in the data indicating no significant learning. This is supported by a simple OLS regression regressing the number of subjects choosing to purchase information on the period of the experiment. While the coefficient on the period variable was negative (-0.406) in this regression, it was not significant ($t = -0.67$, $p > 0.514$).

6 Conclusion

This paper attempts to show that people are willing to pay for the opportunity to be more confident in making the right decision. Our findings suggest that decision theory should be enriched to accommodate such preferences. In particular, our findings highlight two behavioral principles that are yet to be captured by any decision-theoretic model: (*i*) individuals may derive an intrinsic benefit from the posterior beliefs they hold, only if they are about to make a decision, and (*ii*) individuals prefer to know whether or not the worst contingency has occured over knowing whether or not the best contingency has occured.

References

[1] Allman, R.M., E.P. Steinberg, J.C. Keruly and P.E. Dans, "Physician Tolerance for Uncertainty. Use of Liver-Spleen Scans to Detect Metastases," *Journal of the American Medical Association*, 254 (1985), 246-248.

[2] Ahlbrecht, M. and M. Weber, "The Resolution of Uncertainty: An Experimental Study," *Journal of Institutional and Theoretical Economics,* 152 (1996), 593-607.

[3] Baron, J., J. Beattie and J.H. Hershey, "Heuristics and Biases in Diagnostic Reasoning," *Organizational Behavior and Human Decision Processes*, 42 (1988), 88-110.

[4] Barberis, N., M. Huang and R.H. Thaler, "Individual Preferences, Monetary Gambles and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96 (2006), 1069-1090.

[5] Bastardi, A., and E. Shafir, "On the Pursuit and Misuse of Useless Information," *Journal of Personality and Social Psychology*, 75 (1998), 19-32

[6] Çelen, B., S. Choi, and K. Hyndman, "Endogenous Network Formation in the Laboratory," New York University, mimeo, (2005).

[7] Chew, S.H. and J.L. Ho, "Hope: An Empirical Investigation of Attitudes Towards Timing of Uncertainty Resolution," *Journal of Risk and Uncertainty*, 8 (1994), 267-288.

[8] Eliaz, K. and A. Schotter, "Experimental Testing of Intrinsic Preferences for Non-Instrumental Information," *American Economic Review (Papers and Proceedings)*, 97 (2007), 166-169.

[9] Ergin, H. and T. Sarver, "A Unique Costly Contemplation Representation," Mimeo, Washington University and Northwestern University (2007).

[10] Goeree, J.K. and L. Yariv, "Conformity in the Lab," Mimeo, Caltech (2006).

[11] Grant, S., A. Kajii and B. Polak, "Temporal Resolution of Uncertainty and Non-Expected Utility Models," *Econometrica*, 68 (2000), 425-434.

[12] Grant, S., A. Kajii and B. Polak, "Intrinsic Preferences for Information," *Journal of Economic Theory*, 83 (1998), 233-259.

[13] Kassirer, J.P., "Our Stubborn Quest for Diagnostic Certainty," *New England Journal of Medicine,* 320 (1989), 1489-1491.

[14] Kreps, D. and E. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica,* 46 (1978), 185-200.

[15] Kübler, D. and G. Weizsäcker, "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory," *Review of Economic Studies*, 71 (2004), 425-441

[16] Myers, L.P. and J.M. Eisenberg, "Do Physicians Order Too Many Lab Tests?" *Internist*, 26 (1985), 14-15.

[17] Ortoleva, P., "The Price of Flexibility: Towards a Theory of Thinking Aversion," Mimeo, New York University (2008).

[18] Putterman, C. and E. Ben-Chetrit, "Testing, Testing, Testing..." *New England Journal of Medicine,* 333 (1995), 1208-1211.

[19] Redelmeier, D., Shafir, E., and P. Aujla, "The Beguiling Pursuit of More Information," *Medical Decision Making*, 21 (2001), 376-81.

[20] Salant, Y. and A. Rubinstein, "(A,f), Choice with Frames," *Review of Economic Studies*, 75 (2008), 1287-1296.

[21] Shafir, E. and A. Tversky, "Thinking Through Uncertainty: Nonconsequential Reasoning and Choice," *Cognitive Psychology,* 24 (1992),449-474.

[22] Simon, H.A., "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69 (1955), 99-118.

[23] Tversky, A. and E. Shafir, "The Disjunction Effect in Choice Under Uncertainty," *Psychological Science,* 3 (1992), 305-309.

[24] von Gaudecker, H-M., A. van Soest and E. Wengström, "Risk Preference in the Small for a Large Population," Netspar Discussion Paper 2008-011, (2008).

Appendix A: **Instructions for Treatment** 1.

This is an experiment in decision making. Various research foundations have supported this research and you will earn money depending upon the decisions you make. You will receive a show up fee of $3, which is yours to keep regardless of the decisions you make in the experiment. When the experiment begins you will also receive an additional amount of $4, which you can spend during the experiment. If you do not spend all of this amount during the experiment, you can keep any amount that is left.

Decision Task:

There are two boxes, labeled A and B. One, and only one, of these boxes contains a prize of $20. The process that determines in which box the $20 will be placed in is as follows. Think of two urns labeled Urn I and Urn II with 100 balls each. In each urn some balls are marked with the letter A and the remaining ones with the letter B so that all balls are marked with one letter. The composition of A and B balls in each urn differs, however.

To determine in which box the $20 is placed, the computer will first choose an urn by flipping a coin with a 50-50 chance of landing heads or tails.

■ If the coin lands *heads* the computer will go to *Urn I* and draw a ball

- If the ball drawn is labeled A, the computer will place $20 in Box A
- If the ball drawn is labeled B, the computer will place $20 in Box B

■ If the coin lands *tails* the computer will go to *Urn II* and draw a ball

- If the ball drawn is labeled A, the computer will place $20 in Box A
- If the ball drawn is labeled B, the computer will place $20 in Box B

For example, suppose that Urn I has 100 A balls and no B balls, while Urn II has 60 A balls and 40 B balls. If the coin lands on heads the $20 will be placed in Box A for sure, while if it lands tails there will only be a 60% chance that the $20 is placed in Box A.

Your task is to choose which of these two boxes to open. If the box you open contains the prize, you will be paid $20.

Before you make your choice we will tell you the composition of the balls in each urn and give you the opportunity to pay a fee to see the result of the coin flip. If you pay the fee, we will tell you whether the coin landed heads or tails so you will know from what urn the ball will be drawn. This means that if you pay the fee you will learn what the chances are that the $20 was placed in Box A.

In the example above, if you pay the fee and learn that the coin landed heads, then you will know that the \$20 is in Box A for sure. If you do not pay the fee, then you will be asked to choose a Box without knowing the outcome of the coin flip (that is, you will only know that there is a 50-50 chance that the ball is either drawn from an urn containing 100 A balls, or from an urn containing only 60 A balls).

In the experiment we will present you with 15 situations of the type described above (their order will be picked at random) and ask you if you want to pay the fee in each one. Each situation will be described by a fee and the compositions of Urn I and Urn II. In every situation each urn will always have a 50-50 chance of being selected

It is important to note that in all of the situations **both Urns I and II will contain more than 50 A balls so that there will always be more than a 50% chance that Box A contains the \$20.** Paying the fee simply allows you to know what the exact chances are.

After you make your choices for the 15 situations, we will then select one of them at random.

■ If for the chosen situation you decided to see the result of the coin flip, we will then tell you which urn was selected and ask you to choose a box (A or B).

■ If for the chosen situation you decided *not* to see the result of the coin flip, we will let you choose a box without telling you from which urn the ball will be drawn.

Payments:

For each situation your payoff will depend on whether you paid a fee to see the outcome of the coin flip and whether there was \$20 in the box you chose. For example, say that fee is $\$X$ and you *choose to pay it.*

■ If you choose a box that has the \$20 in it, you will be paid $\$20+(\$4-\$X)+\$3$: the \$20 prize, the \$4 initial endowment you receive at the beginning of the experiment, minus the \$X for learning which urn was chosen, *plus the \$3 show-up fee* (or in total $\$27-\$X$).

■ If you do not find the \$20, you will be paid $(\$4-\$X)+\$3$: the \$4 initial endowment you receive at the beginning of the experiment minus the \$X for learning which urn was chosen, *plus the \$3 show-up fee* (or in total $\$7-\$X$).

In all the 15 situations you will face, the fee for learning which urn was chosen will not exceed \$4, hence you are guaranteed to keep the \$3 show-up fee.

Suppose you decide *not* to pay the fee and choose a box *without* seeing the result of the coin flip.

■ If you find the \$20 you will be paid $\$20 + \$4 + \$3$: the \$20 prize, plus the \$4 initial endowment, plus the \$3 show-up fee (or in total \$27).

■ If you do not find the \$20, you will be paid $\$7 = \$4 + \$3$: the \$4 initial endowment plus the \$3 show-up fee (or in total \$7).

What we ask you to do:

We shall now present you with a list of situations. As we stated above, one and only one of them will be chosen at random and played out for you to determine your payoff.

Appendix B: **Instructions for Treatment** 4.

This is an experiment in decision making. Various research foundations have supported this research and you will earn money depending upon the decisions you make. You will receive a show up fee of \$3, which is yours to keep regardless of the decisions you make in the experiment. When the experiment begins you will also receive an additional amount of \$4, which you can spend during the experiment. If you do not spend all of this amount during the experiment, you can keep any amount that is left.

Decision Task:

There are two boxes, labeled A and B. One, and only one, of these boxes contains a prize of \$20. The process that determines in which box the \$20 will be placed in is as follows. Think of three urns labeled Urn 1, Urn 2 and Urn 3 with 100 balls each. In each urn some balls are marked with the letter A and the remaining ones with the letter B so that all balls are marked with one letter. The composition of A and B balls in each urn differs, however.

To determine in which box the \$20 is placed, the computer will first choose an urn by randomly choosing a whole number from 1 to 3 such that each of the three numbers, $1, 2$ and 3, has an equal chance of being selected.

■ If the selected number is 1, the computer will go to *Urn 1* and draw a ball.

■ If the selected number is 2, the computer will go to *Urn 2* and draw a ball.

■ If the selected number is 3, the computer will go to *Urn 3* and draw a ball.

After an urn is chosen the computer will draw a ball from it and if the ball has the letter A written on it it will place \$20 in Box A and if it has the letter B written on it it will place \$20 in Box B.

For example, suppose that Urn 1 (which we will call the "High Urn") has 100 A balls and no B balls, Urn 2 (Which we will call the "Middle Urn") has 80 A balls and 20 B balls and Urn 3 (Which we will call the "Low Urn") has 60 A balls and 40 B balls.

If the randomly selected number is 1, the \$20 will be placed in Box A for sure. If the randomly selected number is 2, there will be an 80% chance that the \$20 is placed in Box A. Finally, if the randomly selected number is 3, there will only be a 60% chance that the \$20 is placed in Box A.

Your task is to choose a box (A or B). If the box you open contains the prize, you will be paid \$20. Before you make your choice we will tell you the composition of the balls in each urn and give you the opportunity to pay a fee to learn the answer to one, and only one, of the following questions:

1. Was the ball drawn from Urn 1 (the "High Urn") or was it drawn from either Urn 2 or 3 (the Middle or Low urns)?
2. Was the ball drawn from Urn 2 (the "Middle Urn) or was it drawn from either Urn 1 or 3 (the High or Low urns)?
3. Was the ball drawn from Urn 3 (the "Low Urn) or was it drawn from either Urn 1 or 2 (the High or Middle urns)?

You may also choose *not* to pay for any of the above answers, in which case you would only know that it is equally likely that the ball was drawn from any of the three urns.

In the experiment we will present you with **15** situations, where each situation will be described by a fee and the compositions of Urn 1, Urn 2 and Urn 3 (as in the above example). In every situation each of the three urns will always have an equal chance of being selected. In each situation we shall ask you the following question:

Which of the following options do you want to choose (you may choose only one)?
(*i*) I would like to pay the fee to know whether or not the ball was drawn from Urn 1
(*ii*) I would like to pay the fee to know whether or not the ball was drawn from Urn 2
(*iii*) I would like to pay the fee to know whether or not the ball was drawn from Urn 3
(*vi*) I do not wish to pay for any of the above pieces of information

After you answer the above question for each of the 15 situations, we will then select one of these situations at random.

■ If for the chosen situation you answered that you would pay the fee for one of the three pieces of information (i.e., your answer was either (*i*), (*ii*) or (*iii*)), then we will present you with the information you requested and ask you to choose a box (A or B).

■ If for the chosen situation you answered that you would *not* pay the fee, we will let you choose a box without revealing any additional information.

It is important to note that in all of the situations **all three urns will contain more than 50 A balls so that there will always be more than a 50% chance that**

Box A contains the \$20. Paying the fee simply allows you to know more accurately what the actual chances are.

Payments:

For each situation your payoff will depend on whether you paid a fee and whether there was \$20 in the box you chose. For example, say that fee is $\$X$ and you *choose to pay it.*

■ If you choose a box that has the \$20 in it, you will be paid $\$20+(\$4-\$X)+\$3$: the \$20 prize, the \$4 initial endowment you receive at the beginning of the experiment, minus the \$X fee, *plus the \$3 show-up fee* (or in total $\$27 - \$X$).

■ If you do not find the \$20, you will be paid $(\$4-\$X)+\$3$: the \$4 initial endowment you receive at the beginning of the experiment minus the \$X fee, *plus the \$3 show-up fee* (or in total $\$7 - \$X$).

In all the 15 situations you will face, the information fee will not exceed \$4, hence you are guaranteed to keep the \$3 show-up fee.

Suppose you decide *not* to pay the fee.

■ If you find the \$20 you will be paid $\$20 + \$4 + \$3$: the \$20 prize, plus the \$4 initial endowment, plus the \$3 show-up fee (or in total \$27).

■ If you do not find the \$20, you will be paid $\$7 = \$4 + \$3$: the \$4 initial endowment plus the \$3 show-up fee (or in total \$7).

What we ask you to do:

We shall now present you with a list of situations. As we stated above, one and only one of them will be chosen at random and played out for you to determine your payoff.

Table 1

The Situations Used in Treatments 1 and 2

Situation	α	β	Fee
1	1.0	0.60	$4.00
2	1.0	0.51	$4.00
3	1.0	0.60	$0.50
4	0.8	0.60	$2.00
5	0.8	0.51	$4.00
6	1.0	0.51	$2.00
7	0.6	0.51	$2.00
8*	0.6	0.51	$4.00
9	0.8	0.60	$0.50
10	0.8	0.51	$2.00
11	0.8	0.51	$0.50
12	1.0	0.51	$0.50
13	0.6	0.51	$0.50
14	1.0	0.60	$2.00
15	0.8	0.60	$4.00

* This situation was not used in Treatment 1

Table 2:

Baseline Treatment

Fraction of subjects who paid the fee as a function of α, β and the size of the fee[1,2]

Fee:	$0.50		$2.00		$4.00	
	β=0.51	β=0.60	β=0.51	β=0.60	β=0.51	β=0.60
α=1.0	0.739	0.783	0.565	0.435	0.261	0.348
α=0.8	0.652	0.522	0.217	0.217	0.174	0.130
α=0.6	0.478	NA	0.217	NA	NA[2]	NA

[1] A total of 23 subjects participated in this treatment

[2] This is the situation that was not used in Treatment 1.

Table 3

Random Effects Logit: Probability of Paying fee in Treatment 1

Variable	Coefficient	Std. dev.	z	$P > \lvert z \rvert$
fee	-1.095	0.1477	-7.42	0.000
diff (α – β)	7.625	1.4161	5.38	0.000
constant	-0.936	.6205	-1.51	0.131

Observations = 345
Log likelihood = -139.391
Prob > χ^2 = 0.0000
Wald χ^2 (3) = 53.88

Table 4:

Treatment 2

Fraction of subjects who paid the fee as a function of α, β and the size of the fee[1]

Fee:	$0.50		$2.00		$4.00	
	β=0.51	β=0.60	β=0.51	β=0.60	β=0.51	β=0.60
α=1.0	0.429	0.524	0.333	0.285	0.286	0.190
α=0.8	0.524	0.429	0.238	0.286	0.190	0.190
α=0.6	0.429	NA	0.286	NA	0.143	NA

[1] A total of 21 subjects participated in this treatment

Table 5

Test of Proportions: Treatment 1 (T1) *vs.* Treatment 2 (T2)

Situation	T1 proportion	T2 proportion	t	$P > \lvert t \rvert$
1	0.348	0.190	0.8027	0.4270
2	0.261	0.286	-0.1806	0.8575
3	0.783	0.524	1.8369	0.0733
4	0.217	0.286	-0.5123	0.6111
5	0.174	0.190	-0.1390	0.8901
6	0.565	0.333	1.5500	0.1286
7	0.217	0.286	-0.5123	0.6111
8	NA	0.143		
9	0.522	0.429	0.6064	0.5475
10	0.217	0.238	-0.1600	0.8737
11	0.652	0.524	0.8524	0.3988
12	0.739	0.429	2.1547	0.0370
13	0.478	0.429	0.3234	0.7480
14	0.435	0.286	1.0153	0.3158
15	0.130	0.190	-0.5352	0.5967

Table 6: Test of Proportions: Treatment 1 (T1) *vs.* Treatment 3 (T3)

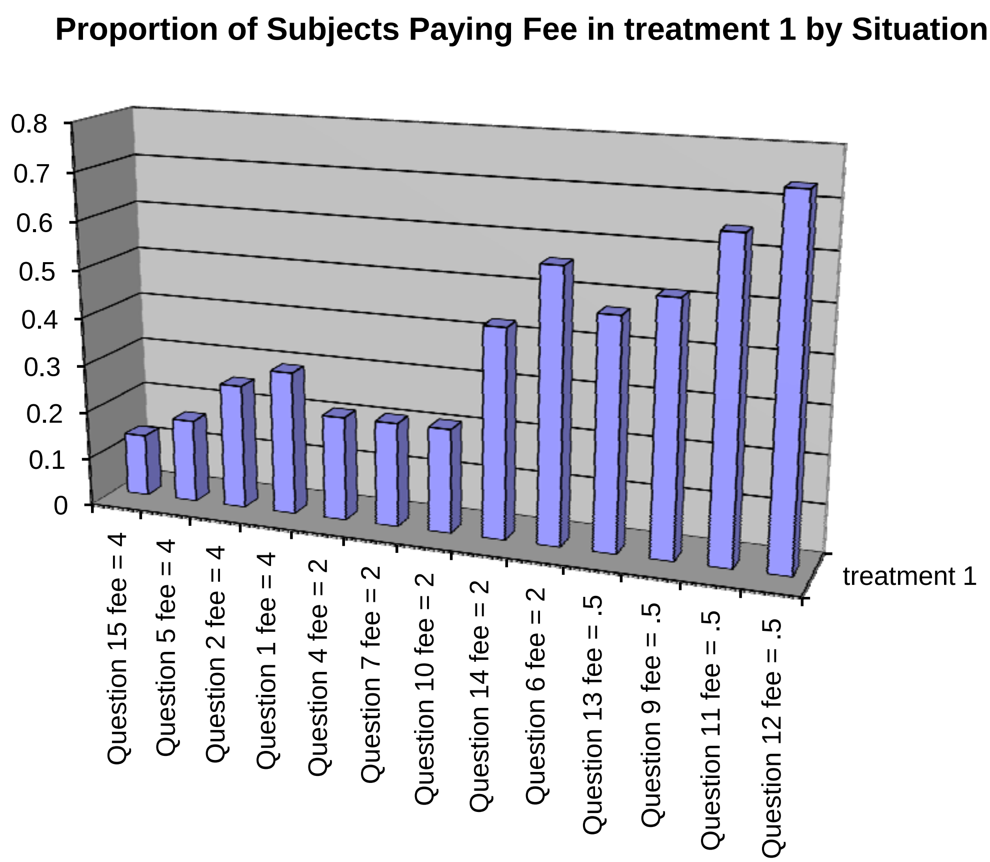
Situation	T1 proportion	T3 proportion	Z*	$P > \lvert z \rvert$
1	0.348	0.000	--	1
2	0.261	0.000	--	1
3	0.783	0.047	15.9	1
4	0.217	0.000	--	1
5	0.174	0.000	--	1
6	0.565	0.000	--	1
7	0.217	0.000	--	1
8	NA	0.047	NA	NA
9	0.522	0.095	6.673	1
10	0.217	0.000	--	1
11	0.652	0.000	--	1
12	0.739	0.047	7.22	1
13	0.478	0.000	--	1
14	0.435	0.000	--	1
15	0.130	0.000	--	1

* Z statistic not calculated when one proportion = 0.

Table 7: Total and Conditional Fraction of Subjects Who Pay Fee for Each Urn[1,2]

Situation	Total (Conditional) Fraction for Urn I		Total (Conditional) Fraction for Urn II		Total (Conditional) Fraction for Urn III		Fraction Paying Fee
1	0.031	(0.250)	0.062	(0.500)	0.031	(0.250)	0.125
2	0.094	(0.500)	0.000	(0.000)	0.094	(0.500)	0.187
3	0.062	(0.200)	0.000	(0.000)	0.250	(0.800)	0.312
4	0.031	(0.125)	0.094	(0.375)	0.125	(0.500)	0.250
5	0.000	(0.000)	0.000	(0.000)	0.125	(1.000)	0.125
6	0.094	(0.333)	0.031	(0.111)	0.156	(0.555)	0.281
7	0.062	(0.286)	0.031	(0.143)	0.125	(0.571)	0.219
8	0.000	(0.000)	0.031	(0.200)	0.125	(0.800)	0.156
9	0.062	(0.167)	0.094	(0.250)	0.219	(0.583)	0.375
10	0.000	(0.000)	0.031	(0.100)	0.281	(0.900)	0.312
11	0.031	(0.071)	0.094	(0.214)	0.312	(0.714)	0.437
12	0.094	(0.167)	0.062	(0.111)	0.406	(0.722)	0.562
13	0.187	(0.461)	0.031	(0.077)	0.187	(0.461)	0.406
14	0.094	(0.333)	0.062	(0.222)	0.125	(0.444)	0.281
15	0.000	(0.000)	0.000	(0.000)	0.094	(1.000)	0.094

[1] Total (Conditional) fraction equals the number of subjects who paid the fee divided by the total number of subjects (number of subjects who paid the fee) in the treatment.
[2] Paying for Urn x (where x=I,II,III) means paying for an answer to Question x.



Proportion of Subjects Paying Fee in treatment 1 by Situation
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Question 15 fee = 4
Question 5 fee = 4
Question 2 fee = 4
Question 1 fee = 4
Question 4 fee = 2
Question 7 fee = 2
Question 10 fee = 2
Question 14 fee = 2
Question 6 fee = 2
Question 13 fee = .5
Question 9 fee = .5
Question 11 fee = .5
Question 12 fee = .5
treatment 1

Figure 1

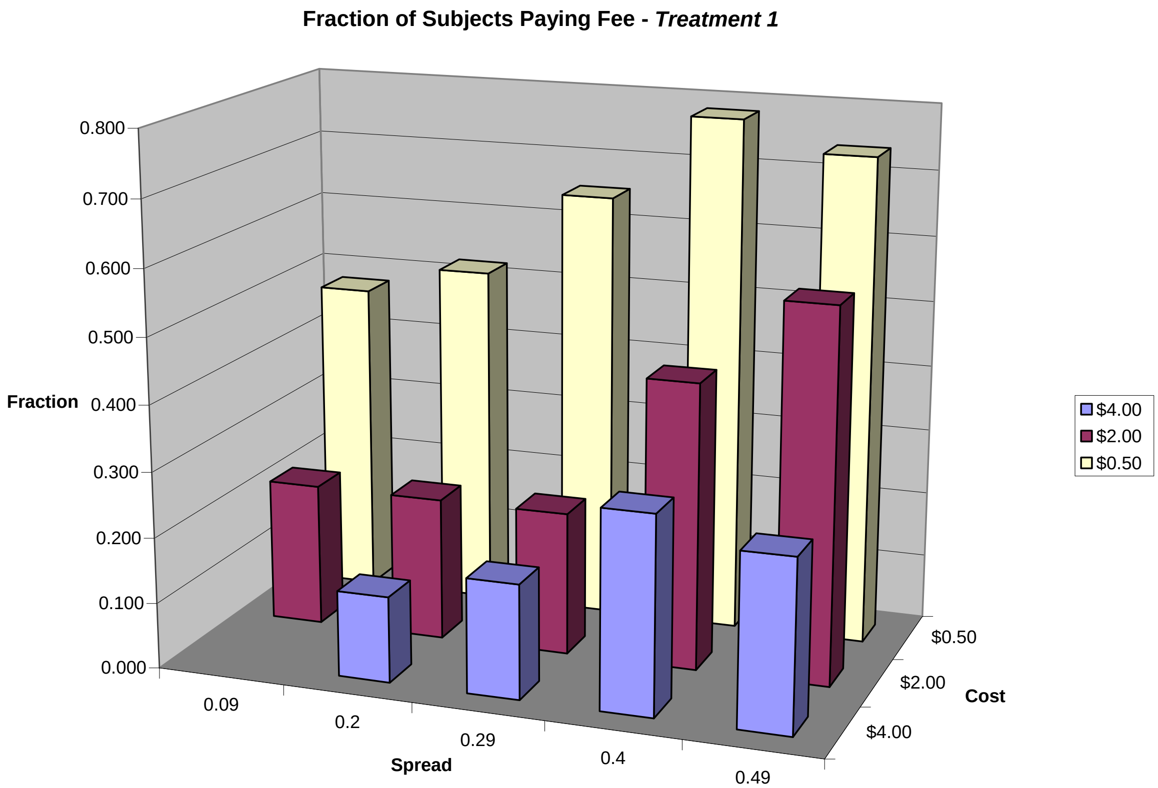

Fraction of Subjects Paying Fee - *Treatment 1*
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Fraction
0.09
0.2
0.29
0.4
0.49
Spread
$0.50
$2.00
$4.00
Cost
$4.00
$2.00
$0.50

Figure 2

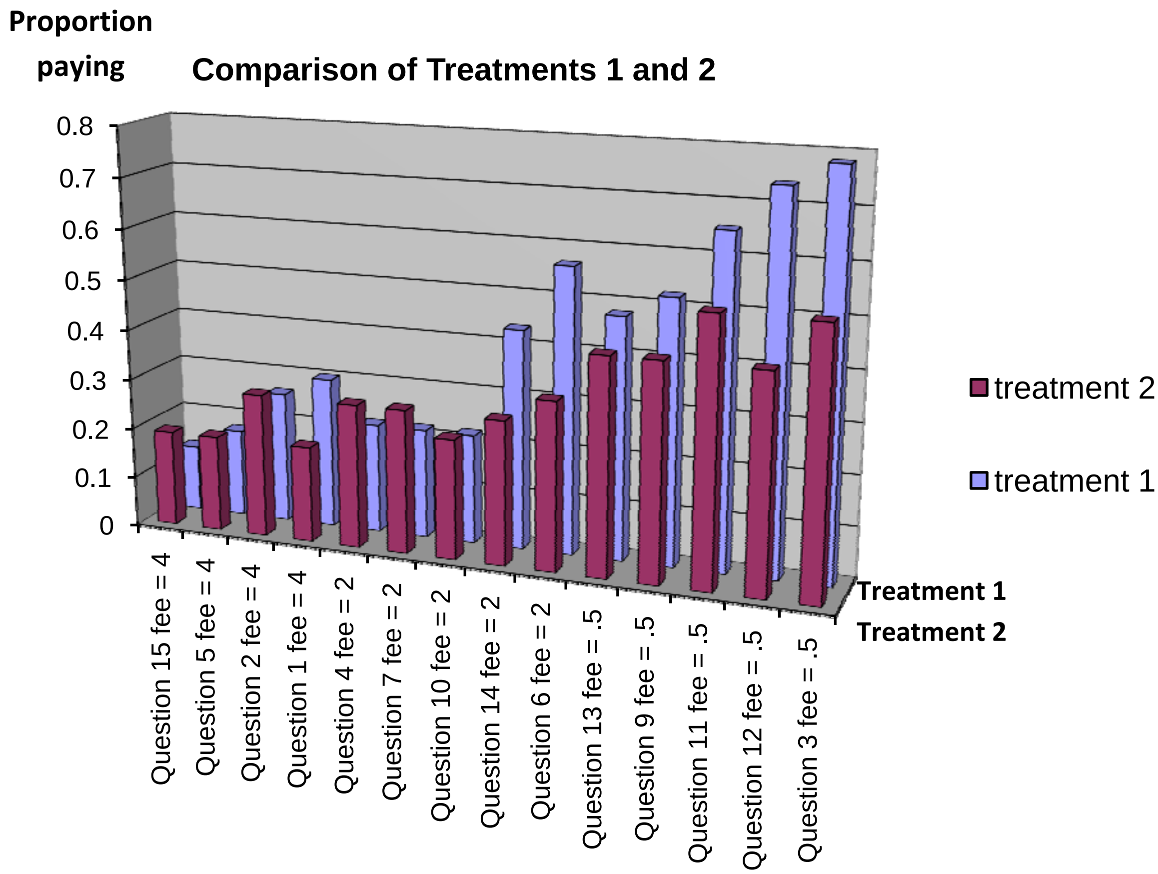
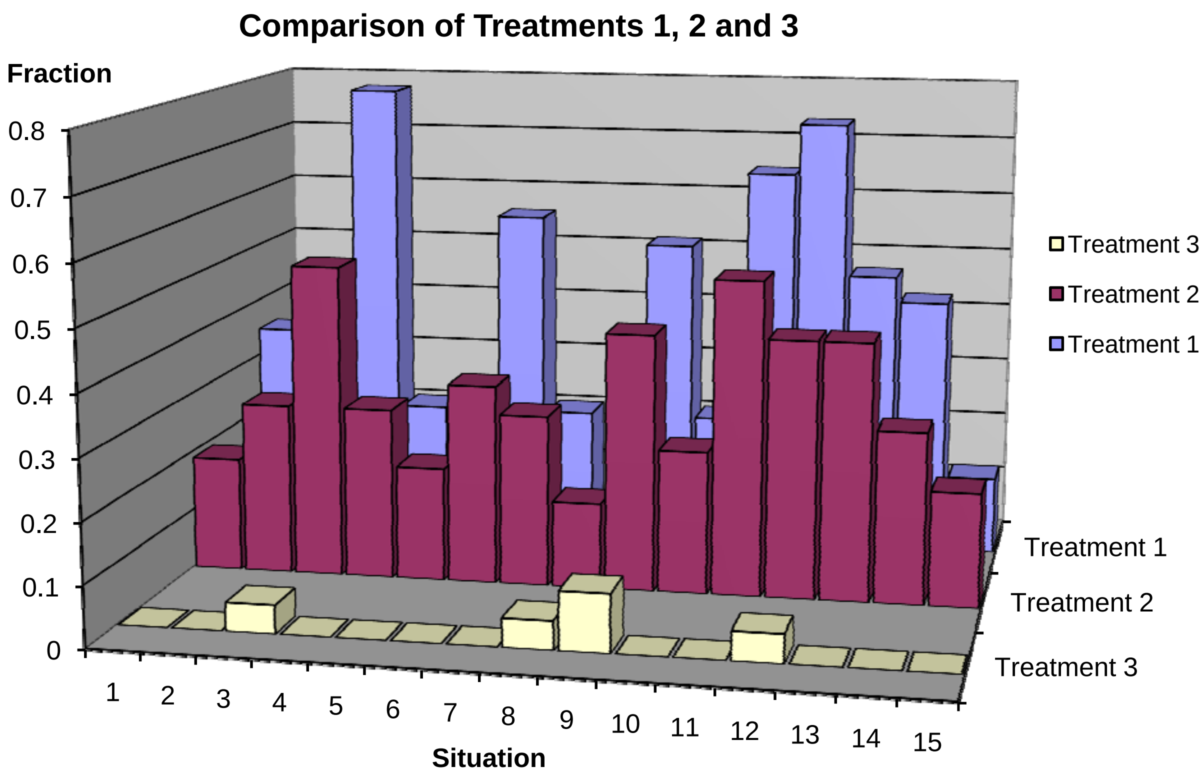

Proportion paying
Comparison of Treatments 1 and 2
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
treatment 2
treatment 1
Treatment 1
Treatment 2
Question 15 fee = 4
Question 5 fee = 4
Question 2 fee = 4
Question 1 fee = 4
Question 4 fee = 2
Question 7 fee = 2
Question 10 fee = 2
Question 14 fee = 2
Question 6 fee = 2
Question 13 fee = .5
Question 9 fee = .5
Question 11 fee = .5
Question 12 fee = .5
Question 3 fee = .5

Figure 3



Comparison of Treatments 1, 2 and 3
Fraction
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Situation
Treatment 1
Treatment 2
Treatment 3
Treatment 3
Treatment 2
Treatment 1

Figure 4

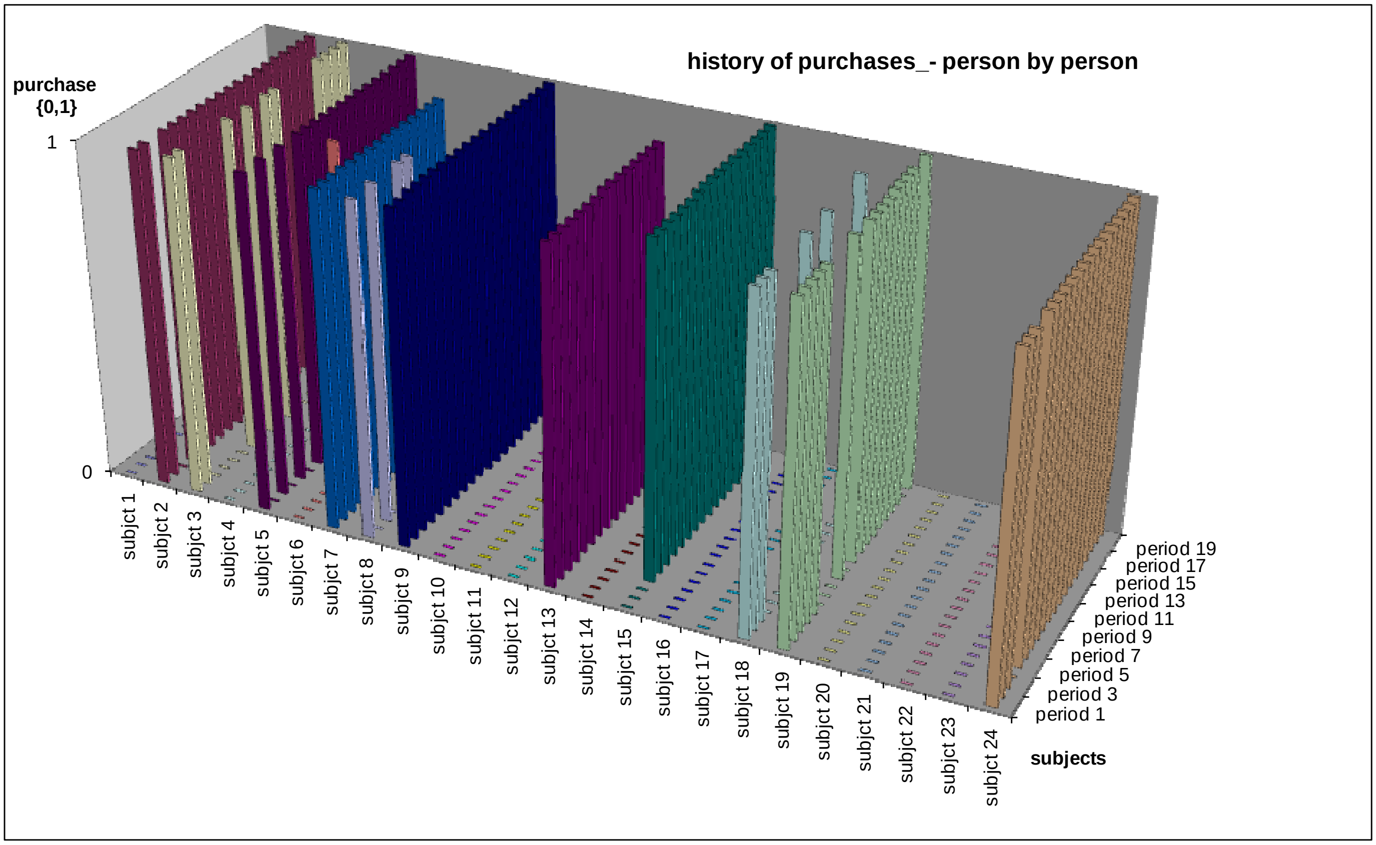

history of purchases_- person by person
purchase
{0,1}
1
0
subjct 1
subjct 2
subjct 3
subjct 4
subjct 5
subjct 6
subjct 7
subjct 8
subjct 9
subjct 10
subjct 11
subjct 12
subjct 13
subjct 14
subjct 15
subjct 16
subjct 17
subjct 18
subjct 19
subjct 20
subjct 21
subjct 22
subjct 23
subjct 24
subjects
period 19
period 17
period 15
period 13
period 11
period 9
period 7
period 5
period 3
period 1

Figure 5

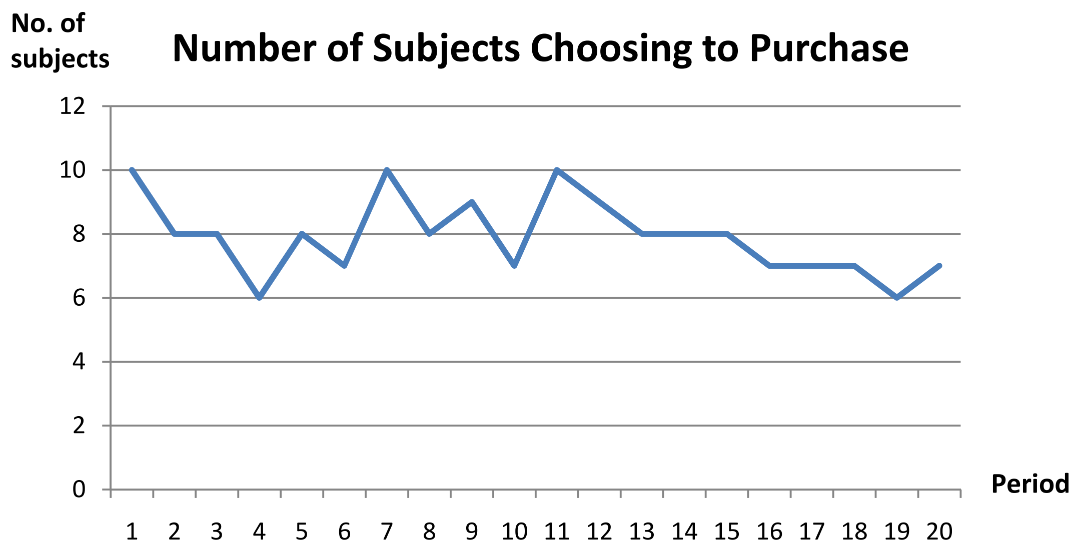

No. of subjects
Number of Subjects Choosing to Purchase
12
10
8
6
4
2
0
Period
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

Figure 6