How is Information Valued? Evidence from Framed Field Experiments*

Mitchell Hoffman†

Abstract

Do people buy the right amount of information? In a novel field experiment, businesspeople experts provided guesses about the price and quality of actual websites. Compensation was provided for correct results (high or low). Before answers were revealed, subjects could pay to get a noisy signal. I find that the relationship between subjects' accuracy and their demand for information is much flatter than would be optimal. Subjects underpay for information when signals are valuable and overpay when signals are less valuable. I also find that subjects exhibit significant overconfidence. However, even when the value of information is adjusted to account for subjects' overconfidence or subjects' tendency to sometimes misuse information, subjects underpay when signals are valuable and overpay when signals are less valuable.

JEL Classifications: C93, D03, D83
Keywords: Demand for information; Overconfidence

*I am indebted to John Morgan for his encouragement and guidance on this project. For comments that much improved the paper, I am grateful to David Myatt (the editor) and three anonymous referees, as well as Jason Abaluck, Heski Bar-Isaac, Kevin Bryan, Stefano DellaVigna, Josh Hausman, Teck Ho, Tanjim Hossain, Shachar Kariv, Botond Koszegi, Peter Kriss, Steve Levitt, Ulrike Malmendier, Matt Mitchell, Ismael Mourifie, Barry Nalebuff, Paul Niehaus, Matthew Rabin, Collin Raymond, Emmanuel Saez, Steve Tadelis, Joshua Tasoff, Kenneth Train, Felix Vardy, and various seminar participants. Mahfuz Chy, Ani Goranova, Carlos Inoue, Gayane Markosyan, Palermo Penano, and Jason Tabuzo provided excellent research assistance. I thank the conference organisers Jay Kopita, Aaron Kvitek, Howard Neu, Rick Schwartz, and Connor Young for their assistance. Financial support from the Berkeley Institute of Business and Economic Research, the Russell Sage Foundation Behavioral Small Grants program, the National Science IGERT Fellowship, and the Social Science and Humanities Research Council of Canada is gratefully acknowledged.

†University of Toronto Rotman School of Management, 105 St. George St., Toronto, ON M5S-3E6, Canada; mitchell.hoffman@rotman.utoronto.ca.

In many economic situations, individuals need to decide whether to acquire costly information.[1] While models of endogenous information acquisition usually assume that agents acquire information when the instrumental value of information exceeds its cost, this assumption may be violated when agents have biased beliefs. Consider an individual facing a binary choice {A,B}. With her existing information, she is able to predict the correct choice with probability p. At a cost k, she can do additional research that reveals the right choice with certainty. If the value of the right choice is one and that of the wrong choice is zero, then it is optimal to perform this additional research whenever $k < 1 - p$. However, a large literature shows that people are often overconfident about their ability. Suppose the individual in my example is overconfident about the quality of her pre-existing knowledge, perceiving herself to be correct with probability $p' > p$. In that case, she will become fully informed only when $k < 1 - p'$, and information will be undervalued. On the other hand, belief biases could cause information to be overvalued. For example, when tasks are easy, people often exhibit underconfidence (e.g., Griffin and Tversky, 1992; Moore and Healy, 2008). If $p' < p$, then the individual will overpay for information.

In this paper, I study whether even experts optimally demand information, and examine the sources of any patterns of over- or underpayment. I perform framed field experiments using businesspeople experts in two industries who buy, sell, and operate domain names and websites.[2] In Stage I of the experiments, subjects make guesses about the prices of actually sold domain names or about the quality of actual websites. Subjects make binary guesses (high or low price, or high or low quality), and their confidence for each guess is elicited with an incentive-compatible quadratic scoring rule (QSR). In Stage II, subjects have the opportunity to purchase information to help them possibly guess again on the same tasks. If their willingness to pay (WTP) for information is high enough, they receive the information and guess again. Because the guessing tasks are binary, there is a simple formula for rational acquisition of information. The information takes two forms. In one form, information is a noisy signal of the true price/quality, made with computer-generated

[1] Models with endogenous (costly or costless) information acquisition have been used to study financial markets (e.g., Grossman and Stiglitz, 1980), auctions (e.g., Milgrom and Weber, 1982), voting (e.g., Martinelli, 2006), and medical patient decision-making (e.g., Koszegi, 2003), among many other applications.

[2] Harrison and List (2004) define framed field experiments as experiments using non-standard subjects and that use field context in the task, commodity, or information subjects can use. The term "domain name" refers to an internet property that has not been developed and the term "website" refers to a developed internet property.

random noise. In the second form, information is the guess of another subject from Stage I.

The reason I perform framed field experiments is to combine attractive elements of the lab and of the field for my research question. By using a simple laboratory game with clearly defined rules, I rule out two confounds which may be present in field data. First, I rule out the possibility that subjects are unaware information is available. Second, in my experiment, the costs of information are transparent and are the same for everyone, i.e., no one has an advantage of being able to obtain information more quickly or more cheaply. However, laboratory studies are sometimes criticized on the grounds that they rely on inexperienced undergraduate subjects performing artificial laboratory tasks (Levitt and List, 2007a,b). By using experts performing tasks and purchasing information similarly to how they do so in the real world, I help ameliorate these concerns.[3]

In the main finding of the paper, I find that subjects underpay for information when information is valuable and overpay when information is less valuable. The relationship between empirical accuracy and WTP is far flatter than predicted by theory.

To what extent are my results related to overconfidence? By measuring confidence prior to eliciting WTP, I can examine optimal information acquisition when the value of information is adjusted to account for overconfidence. Empirically, subjects exhibit substantial overconfidence, though interestingly, subjects recognize that other subjects tend to be overconfident. In addition, higher confidence is associated with lower WTP, both across and within subjects. However, after accounting for overconfidence in the value of information, it remains that subjects underpay for highly informative signals, but overpay for relatively uninformative signals.

What else can account for patterns of suboptimal WTP? The preceding discussion assumes that once obtained, information is used optimally in making choices. While the right choice is

[3]For example, I can rule out the criticism of lab experiments that the tasks lacked economic meaning for subjects. It is also likely that biases that exist with experts will also exist with non-experts, but the reverse may not be true. For example, Palacios-Huerta and Volij (2009) and List and Haigh (2005) show that some biases observed among non-experts may be weaker among experts (though other evidence doesn't reach this conclusion (e.g., Cipriani and Guarino, 2009)). One interpretation of the magnitude of sub-optimal information acquisition in my experiment is as a potential lower bound for sub-optimal information acquisition with other subjects in other contexts. In addition to the question of why use businesspeople experts instead of students, there is the question of why I use businesspeople who trade and develop internet real estate and web content. First, the context is one where good information is highly valuable. Second, there is an active domain name appraisal market (where businesspeople purchase information about a domain name from an appraiser), which is a real-world example of costly information acquisition. Third, I had available the generous assistance of the conference organisers in being able to implement business-relevant experimental tasks.

straightforward when signals all point in the same direction, sorting through conflicting signals to come to the right conclusions is much more demanding. I find that information misusage is fairly common. However, even accounting for subjects' tendency to sometimes misuse information, it remains that subjects underpay for valuable signals and overpay for less valuable signals.

Beyond overconfidence and misusage, there are a number of alternative forces that could help explain the results. Conservatism (Edwards, 1968) and base rate neglect (Tversky and Kahneman, 1974) can generate the observed pattern of over- and underpayment if they are present in large enough doses. Other forces may also help explain the results, but seem less successful in fully explaining observed patterns.

The guessing and information acquisition tasks in my experiments are designed to mimic those in the real-world markets from which my subjects are drawn. For example, for domain traders—those involved in buying, selling, and developing domain names, the virtual addresses of the internet—assessing domain prices is a central part of business. Yet, even for professionals, pricing domains is complicated, and there is an active real-world market for domain appraisals, where an expert is paid to provide an estimate of a domain name's value.[4]

Although my paper is one of the first to directly analyse optimal information acquisition over a broad range of information values, several other experimental papers examine costless or costly information acquisition. Kubler and Weizsacker (2004) and Kraemer *et al.* (2006) modify a standard social learning/herding experiment (Anderson and Holt, 1997) so that information is costly. Both papers find evidence of excessive information acquisition. Other recent experiments have examined information acquisition in non-instrumental contexts, e.g., whether people acquire information to feel good about themselves.[5] Information acquisition has also appeared in experimental asset markets and in several experiments on naive advice, but the context in these studies

[4]The subjects are not only highly knowledgeable about the particular items they are guessing about, but are also often familiar with acquiring information about the items. Despite this familiarity, they are probably not experts at information acquisition, and are certainly not experts in experimental economic methods like elicitation of beliefs or WTP. However, using a subject population that is expert at valuing items and familiar with information acquisition is useful for responding to criticisms that participants were totally inexperienced or that the experiment was context-free, and is a novel addition to the literature. I thank two referees for raising this point.

[5]Kubler and Weizsacker (2004), Kraemer *et al.* (2006), and Eliaz and Schotter (2010) also examine aspects of optimal information acquisition. I discuss the relationship between my findings and theirs in Section 4.1. Papers examining non-instrumental information acquisition include Fong and Oberholzer-Gee (2011), Eliaz and Schotter (2010), Burks *et al.* (2013), and Ganguly and Tasoff (2014).

is very different.[6]

My findings are also related to the overconfidence literature. One strand of this literature examines the types and causes of overconfidence.[7] A second strand (Charness *et al.*, 2013; Burks *et al.*, 2013) examines the connection between overconfidence and social signaling. In contrast, my main concern is with the link between overconfidence and optimal information acquisition.

To summarize, there are two main contributions of the paper. First, I develop a simple theory and a novel experiment for analyzing optimal information acquisition, both compared to the rational benchmark and while adjusting the value of information to account for overconfidence and information misusage. Second, I conduct the experiment using expert subjects.

Section 1 presents a basic model of the value of information. Section 2 explains the experimental design. Section 3 provides results. Section 4 discusses the relationship between my findings and others, as well as alternative explanations for my results. Section 5 concludes. Supplementary material is given in the online Appendix.

1 The Model

I present a very simple model of the value of information. The goal is to provide a tight link between theory and my experimental design. As such, I refer to elements of the experimental design, even though they are not fully explained until Section 2.

Subjects wish to guess the value of a binary parameter $\eta \in \{0, 1\}$. In Stage I, in addition to her guess about η, the subject is asked for the confidence of her guess, $p \in [0.5, 1]$, which I elicit in the experiment using a lottery version of the quadratic scoring rule (QSR) that is incentive-compatible under risk aversion. In Stage II, I elicit her WTP for a signal $\omega \in \{0, 1\}$ using the Becker-DeGroot-Marschak (BDM) Procedure. The signal may be a "direct signal," that is, a computer-generated number, or it may be a "subjective signal," that is, the guess of another person. Direct signals are

[6] As in the social learning experiments, asset market experiments (e.g., Sunder, 1992) and naive advice experiments (e.g., Schotter, 2003) often feature situations where information may provide limited instrumental value beyond what should be observed in equilibrium (e.g., other agents in advice games may be unlikely to have superior information). My experiment looks at information acquisition across a whole range of possible values.

[7] See Moore and Healy (2008) for a taxonomy of types of overconfidence. Eil and Rao (2011), as well as Mobius *et al.* (2014), explore the link between ego utility and overconfidence, while Cesarini *et al.* (2009) study genetic links. See Ericson (2011) and Grossman and Owens (2012) for links between overconfidence and memory/feedback.

accurate with probability $r \geq 0.5$, that is, $Pr(\omega = \eta|\eta) = r$. For subjective signals, a subject's estimate of the probability that the other person's guess is correct is denoted by s. If the subject receives a signal as a result of the BDM elicitation, she makes a new guess about η to replace her old one. If she is correct, she receives a prize of k (she receives 0 if incorrect).

The following proposition expresses the value of a direct signal. Let $\mathbf{1}(\cdot)$ denote the indicator function, $g(\omega)$ denote the subject's guess given a signal ω, and g denote a guess made without a signal. In my main empirical analyses on optimal WTP, direct signals receive the most attention.

PROPOSITION 1. Optimal WTP for Direct Information

Assume the subject is risk neutral. Suppose her confidence regarding her guess of η is p. Then the optimal WTP for a signal of accuracy r to help win a prize k is $k(r-p) * \mathbf{1}(r > p)$.

Proof. Between the signal and her prior, the subject should rely on the most *ex ante* accurate piece of information. If the signal is less accurate than the subject's prior, her prior should be pivotal (i.e., be used for making decisions), and she gains nothing from obtaining the signal. If the signal is more accurate than the prior, the advantage she gains is the difference in expected payoff between guessing using the signal and guessing using the prior: $b = E(\pi|\omega) - E(\pi) = kPr(g(\omega) = \eta) - kPr(g = \eta) = k(r-p)$. □

This result is straightforward and allows me to see directly whether demand for information is optimal. For example, if a subject is 90% confident in her guess in Stage I, she should bid $WTP = 0$ to obtain an additional signal which is 80% accurate. However, if she is only 50% confident, she should be willing to pay 30% of her winnings to obtain the 80% accurate signal.[8]

Optimal WTP can also be considered for subjective signals, where the signals are in the form of information on the actions of past subjects. The formula for optimal WTP is the same except that the known accuracy of the signal, r, is replaced by the subject's belief about the accuracy of the other subject, s (assuming $s \geq 0.5$). That is, optimal WTP is $k(s-p) * \mathbf{1}(s > p)$.

WTP for subjective information may be affected not only by the behavioral biases mentioned

[8] In many studies on overconfidence, subjects are asked to state a confidence interval (often a 90% interval) regarding their guess on a continuous variable (e.g., Cesarini *et al.*, 2006). The relative advantage of using my binary-guessing design is that allows for a simple and direct test of optimal information acquisition. Optimal information acquisition can be examined relative to subject confidence (as in Proposition 1) or relative to subject accuracy (i.e., letting p represent a subject's chance of guessing correctly).

above, but also by subjects' beliefs about other subjects (e.g., believing that other subjects are overconfident). By asking subjects their beliefs about the accuracy of other players, beliefs about others are incorporated into the value of subjective information. However, I do not model how subjects form beliefs about other subjects.

The formula in Prop. 1 assumes risk neutrality, and will be incorrect if subjects are risk-averse (or loss-averse). With a more general utility function, the optimal WTP for information, b, equates expected utility without information to expected utility when information is purchased:

$$pU(e+k) + (1-p)U(e) = rU(e-b+k) + (1-r)U(e-b), \tag{1}$$

where e is the subject's endowment. The effect of risk aversion on optimal WTP is ambiguous. As seen in the numerical illustration in Appendix Figure D1, information may be more valuable for subjects with moderate risk aversion compared to either subjects who are less risk-averse or more risk-averse. On one hand, buying information may be undesirable to a risk-averse agent since any amount paid lowers her payoff in all states of the world. On the other hand, information also provides "insurance" by increasing the probability that her guess is correct. The effects, though, on the size of optimal WTP in Figure D1 seem relatively small.[9]

One way in experiments to eliminate confounds due to risk aversion is to pay subjects in lottery/probability units (Berg *et al.*, 1986). In theory, this eliminates the effect of risk aversion on optimal WTP. However, making payments in probability units may be less intuitive for businesspeople and would make the experiment less analogous to real-world markets. In addition, in experimental practice, paying in probability units may not eliminate the effect of risk aversion (Walker *et al.*, 1990). In my experiment, instead of paying in probability units, subjects pay for information in dollars and then separately complete a lottery choice task (Holt and Laury, 2002), which measures their risk aversion. I present estimation results later on both under the assumption

[9]Note that when $r = 1$, the value of information is monotonically increasing in risk aversion, as buying information eliminates all risk. Turning from risk aversion to loss aversion, it seems that loss aversion will lower optimal WTP for information if an agent's reference point is her endowment (and $r < 1$). Suppose the agent has a simple kinked utility function $U(x, e) = (x - e) * \mathbf{1}(x \geq e) - L * (e - x) * \mathbf{1}(x < e)$, where $L \geq 1$ is a constant indicating the agent's degree of loss aversion. The value $L = 1$ corresponds to the rational case whereas $L > 1$ means that losses are felt more than gains. For a signal of accuracy r to win a prize k, the optimal WTP equates $r(k - b) - (1 - r)Lb = pk$, which leads to $b = \frac{k(r-p)}{r+(1-r)L} * \mathbf{1}(r > p)$.

of risk neutrality and correcting for risk aversion. Results in both cases are qualitatively similar.

2 Procedure and Background

2.1 Experimental Procedure

Subjects were told they would begin the experiment with \$10. In Stage I, subjects made guesses on between 10-25 tasks. The tasks were guessing whether randomly selected domain names sold for less than \$2,500 or more than \$7,500 or guessing whether randomly selected websites had less than 100 or more than 500 incoming links. Subjects also stated a percentage confidence for their guess from 50 to 100 percent. Subjects were paid according to a risk-independent QSR, very similar to that in McKelvey and Page (1990). If a subject guessed correctly and stated a confidence level c, they received a lottery with a $2c - c^2$ probability of winning \$20 and a $(1 - c)^2$ probability of receiving zero. If they guessed incorrectly, they had a $1 - c^2$ probability of winning \$20 and a c^2 probability of receiving zero. Under these incentives, it is optimal for subjects to accurately report their true confidence level. This method enriches the original QSR of Brier (1950) to be incentive-compatible under risk aversion.[10] Some recent papers assessing (theoretically or empirically) the ability of QSRs and/or alternative scoring rules to accurately elicit beliefs include Offerman *et al.* (2009), Hossain and Okui (2013), and Schlag and van der Weele (2009). Recent papers using risk-invariant scoring rules include Holt and Smith (2009) and Mobius *et al.* (2014).[11]

Stage II consisted of two parts. In one part, subjects had the opportunity to purchase direct information signals on half of the guessing tasks. For each task, I elicited subjects' WTP for direct

[10] To see why the method is incentive compatible, one solves the problem of choosing a confidence level c to maximize expected utility given true believed accuracy q. Letting e be the endowment and k be the prize, the problem is:

$$max_c q[(2c - c^2)U(e + k) + (1 - c)^2 U(e)] + (1 - q)[(1 - c^2)U(e + k) + c^2 U(e)].$$

One finds $c^* = q$, and this does not depend on the level of risk aversion. In contrast, in a standard QSR, the subject's payoff for a reported confidence level c is $2c - c^2$ for a correct guess and $1 - c^2$ for an incorrect one. If subjects in a standard QSR are risk-averse, they may bias reports down toward 0.5.

[11] Controlling risk by paying subjects in lottery units is analysed in the seminal work of Berg *et al.* (1986), though there is some evidence that paying in probability units does not always eliminate the effect of risk aversion (Walker *et al.*, 1990; Selten *et al.*, 1999). However, McKelvey and Page (1990), Hossain and Okui (2013), and Hao and Houser (2010) report experiments where risk-invariant scoring rules seemed reasonably successful in eliciting beliefs, as predicted by theory. While one cannot assume that the risk-invariant QSR used in my experiment is totally foolproof, McKelvey and Page (1990), Hossain and Okui (2013), and Hao and Houser (2010) suggest that risk-invariant scoring rules can be effective in eliciting beliefs.

signals of both 70% and 90% accuracy using the Becker-DeGroot-Marschak (BDM) procedure. Subjects stated WTPs between $0 and $10 for each signal, and it was randomly determined (50-50 chance) whether the 70% or the 90% signal would be available for purchase. If the stated WTP was higher than the random BDM price draw, the subject received a signal, and was told to guess again.[12] In the other part, subjects had the opportunity to purchase subjective information (the guess of another subject) for the remaining tasks from Stage I, again using the BDM procedure.[13] Subjects had no information about a person whose guess they could purchase besides (i) the person's confidence level on the task and (ii) the knowledge that the person previously participated in the experiment at the conference. In addition, for each task, subjects were asked to predict (without incentives for accurate prediction) the probability that the other person's guess was correct; subjects did this after seeing the other person's confidence level, but before stating their WTP.

In Stage III, subjects completed a Holt and Laury (2002) set of choices between lotteries. To eliminate possible wealth effects, subjects were paid according to one decision randomly selected from the entire study.

I performed the experiments at the DOMAINfest and TRAFFIC domain name conferences and the Cybernet Expo online adult conference in 2009. The DOMAINfest and TRAFFIC conferences focused on domain names in general, whereas the Cybernet Expo conference focused on "adult" (i.e., sexually explicit) internet business.[14] At each conference, I set up a table with a briefcase full of cash on it and a sign advertising an "economics experiment." Businesspeople passed by the table, and if interested, participated in the experiment. Across the three conferences, there were 134 subjects.[15] Significant care was taken to ensure that subjects understood the experiment, and

[12]Two remarks are worth making about the BDM. First, the upper limit was chosen to be deliberately high. Even though information should often only be worth several dollars to a rational player, subjects had the option to pay up to $10. This conservative upper limit was chosen to avoid potentially "leading" subjects into stating low WTPs. Second, I used a left-triangular distribution instead of a uniform distribution for the random price so that people would successfully purchase information more frequently (this does not affect the incentive compatibility of the BDM). This allowed me to gather more data on how information is used.

[13]Different people's guesses were randomly selected for different subjective information purchasing decisions (see Appendix E for details).

[14]At DOMAINfest and TRAFFIC, subjects guessed about domain names prices. At Cybernet Expo, subjects guessed about prices for adult domain names and about quality (i.e., number of incoming links) for adult websites.

[15]Of the 134 subjects, there are 15 subjects with two records per person. These subjects participated at two different conferences (both DOMAINfest and TRAFFIC) or did different tasks at the same conference (Cybernet Expo). I allowed subjects to participate again because they would have no superior information about the particular tasks (they faced different tasks in their second round) and because doing so allowed me to collect more data. I

I was present at experimental sessions to answer questions. Subjects could take as much time as they needed to complete the experiment. Many subjects took around 15-20 minutes, though some took 30 minutes or longer. The experiment was performed using pen and paper. When signals were purchased, they were provided as an "M" ("M" for "More") or "L" ("L" for "Less") written on paper. Subjects were paid immediately afterward in cash. Average earnings were around $23.[16] The experiment provided no direct feedback to subjects on any guesses. (Of course, the signals subjects purchased provided some indirect feedback.)

At the DOMAINfest conference (N=39) in January 2009, there were 10 tasks, and subjects stated WTPs for direct signals on half the tasks and for subjective signals on the other half. The experiment was completed in small groups of around 1-5 subjects. Subjects were seated near one another, but decisions were made privately.[17] At the TRAFFIC conference (N=49) in April 2009 and at the Cybernet Expo conference (N=46) in June 2009, all experiments were administered individually to eliminate any possible group effects. In addition, more tasks were used, thereby allowing more precise measurement of each subject's accuracy and average WTP. At the TRAFFIC conference, 25 tasks were used, and subjects stated WTPs for direct signals only. At the Cybernet Expo conference, 20 tasks were used, and subjects stated WTPs for direct signals on half the tasks and for subjective signals on the other half. For additional details on experimental procedure, see Appendix E.

2.2 Industry Background

In the domain name industry, traders buy domain names generally for the purpose of advertising or re-sale. For example, a domain trader will purchase a name like UsedCarsNYC.com, and then place ads on the website and/or try to sell the name to a car-related firm. In 2007, annual industry size

present the results using all the data and assuming clustering of the error terms with 149 clusters. All the main results are unchanged if I omit the 2nd round of participation for the 15 subjects.

[16] Subject payment was decided after consultation with conference organisers in an effort to make earnings salient for participants. For a subject taking 15-20 minutes to complete the experiment, average earnings of $23 are equivalent to a rate of $69 and $92 per hour, or $138,000 to $184,000 per year (at 40hrs/wk and 50wks/yr). The median income for the subjects is $100,000 to $300,000, though I only have income data for a small share of subjects.

[17] The number of subjects per group at DOMAINfest reflected the number of conference-goers available for the experiment at different points during the conference. Given that decisions were made privately, it seems unlikely that having other subjects present would affect behavior. Moreover, the main results on WTP are robust to excluding data from the DOMAINfest conference.

was estimated at $2 billion (Goldman, 2007). Kesmodel (2008) reports that at least 10,000 people actively invested in domains worldwide in 2007. Early entrants to the industry were sometimes able to directly register valuable names for a small registration fee (Kesmodel, 2008). As good names have become scarcer, they are often acquired through auctions and fixed price sales sites.

Predicting domain prices is difficult. Though prices are correlated with easily observable factors such as length, word count, and the popularity of phrases in search engines (Hoffman, 2007), there are other more subtle factors important for valuation. For example, a seemingly obscure name may be valuable if it is related in some way to categories where advertisers pay high amounts per click (e.g., class-action lawsuits), as may a name with multiple uses (e.g., a name with another meaning in a foreign language). To provide information on domain values, there has emerged a domain name appraisal subindustry, where traders pay to get the opinion of an appraisal expert and/or proprietary computer valuation program. Although the experiment abstracts from real-world domain markets in several ways, the general activities of making predictions about domain values and purchasing information about domain values are familiar to domain trader subjects.[18]

Turning to the online adult industry, adult websites represent a major segment within domain trading and the internet in general. US online adult entertainment revenue was $2.8 billion in 2006 (Edelman, 2009), and Ropelato (2006) estimated that around 12% of websites are pornographic. Online adult industry entrepreneurs often buy domain names. Further, website quality is quite important in the online adult industry, with website quality proxied in the study by the number of incoming links. For websites in categories with fierce competition, understanding the quality of rival websites may be particularly important. Thus, both experimental tasks have an important relation to real-world work even if the experience of doing the tasks is not the same as in the real-world. Appendix C gives further details related to the industries.

2.3 Summary Statistics

Background information was obtained for 96 subject-rounds. Table 1 shows summary statistics. 46% of subjects identify their occupation as industry investor or entrepreneur. 45% of subjects

[18] It should be noted that subjects are not experts at stating confidence levels using QSRs. The possibility that the QSR may have failed in some way to capture beliefs is addressed in Section 4.3.

identify as industry professionals. While they are likely much more familiar with domain prices or website quality than an average person, industry professionals do not likely make a living trading domains or owning websites, and are thus likely less experienced at the experimental tasks than investors / entrepreneurs. The median subject is a white male college graduate making $100,000 to $300,000 per year. About 17% of subjects used a domain appraisal service.[19]

3 Results

The first result concerns overconfidence. Subjects showed significant overconfidence.

Result 1 (Overconfidence) *Internet businesspeople are significantly overconfident. Accuracy and confidence are correlated: Subjects who are more accurate guessers have higher average confidence levels (though only modestly so), and subjects are more confident on easier tasks.*

Subjects answered 62.4% of questions correctly, but expressed average confidence of 77.0%. This difference is highly statistically significant, according to an individual-level two-sided Mann-Whitney test ($z = 9.44, p < 0.01$), my primary statistical test for binary comparisons.[20] These results confirm non-incentivized studies in psychology documenting overconfidence in experts (e.g., Wagenaar and Keren, 1986; Baumann *et al.*, 1991; McKenzie *et al.*, 2008), doing so in a fully incentivized experiment. They also confirm incentivized studies in economics using student subjects documenting overconfidence in different contexts (e.g., Hoelzl and Rustichini, 2005; Camerer and Lovallo,

[19] For mainstream domain traders with more than 5 years of experience, the figure rises to about 1 in 3. Professional appraisals represent an important tool used in researching domain values. One may, perhaps, wonder why the purchase of appraisals is not ubiquitous. One possibility is that domain traders, in effect, vertically integrate and perform the due diligence required for appraisals themselves. A second possibility is that the appraisal market is simply mispriced–unless the domain purchase is very large, appraisal services are cost prohibitive. Indeed, this second explanation is consistent with the survey results where over half of subjects indicated that they would obtain a $400 appraisal on a $1 million domain. A third explanation is that subjects underestimate the value of appraisal information. While the evidence is consistent with all three explanations, a contribution of the experiment is to shed light on whether people may misperceive the value of services like domain appraisals (separate from whether appraisals may be mispriced). Hoffman and Morgan (2015) provides more background on the domain industry, as well as the online adult industry.

[20] Accuracy is also highly significantly different from 50%, or the expected accuracy under random chance ($t = 12.63, p < 0.01$, subject-level t-test, two-sided p-value). Overconfidence is present both in mainstream domain traders, who are overconfident by 15 percentage points (confidence differs from accuracy with $z = 8.19, p < 0.01$), and in adult segment internet businesspeople, who are overconfident by 13 percentage points ($z = 4.91, p < 0.01$). The difference in overconfidence between mainstream and adult segment businesspeople is not significant ($p = 0.40$). Unless otherwise indicated, the tests reported are individual-level two-sided Mann-Whitney tests.

1999; Ericson, 2011).[21] Next, I observe the correlation between accuracy and confidence across persons. As seen in Figure 1, businesspeople who express higher average confidence have higher average accuracy. Further, as seen in Figure 2, the particular domain names and websites on which businesspeople have high confidence levels are the ones for which they are more accurate. These correlations suggest that subjects express confidence levels in a meaningful way, expressing higher confidence on easier tasks and when they are more skillful. It should be noted, however, that the within-person correlation of confidence and accuracy in Figure 1 is fairly weak, with the estimated slope of 0.17 substantially less than one; that is, subjects who express higher confidence are only slightly more accurate on average.

I compare the behavior of subjects identifying as investors or entrepreneurs (N=44) versus other subjects (N=52). Investors and entrepreneurs had an accuracy of 66% and were overconfident by 20 percent (11 percentage points) compared to other industry members who had an accuracy of 61% and were overconfident by 30 percent (16 percentage points). Thus, investors and entrepreneurs are more accurate (z = 2.01, p = 0.04) and less overconfident (z=1.87, p = 0.06) than other industry members, though the differences hover around statistical significance.[22]

While subjects are overconfident, what are their views about the accuracy of others? Do they believe that other subjects are overconfident? This question is important in understanding how businesspeople value obtaining other people's guesses.

Result 2 (Beliefs about Other People's Overconfidence) *When given the confidence level of another paired subject and asked to predict their accuracy, internet businesspeople recognize that other subjects tend to be overconfident.*

In deciding their WTP for a subjective signal (from another subject), subjects were presented with that paired subject's confidence level, and were asked to predict the paired subject's accuracy

[21] Several recent incentivized experiments in economics argue that overconfidence is not always robust (e.g., Cesarini *et al.*, 2006; Blavatskyy, 2009; Benoit *et al.*, 2008). However, Cesarini *et al.* (2006) and Benoit *et al.* (2008) do not focus on absolute overconfidence, as this paper does.

[22] Whether someone is an investor or entrepreneur may be viewed as a measure of industry expertise. The other subjects are comprised of 43 "industry professionals" (i.e., people who work in the two industries, but who are not investors / entrepreneurs) and 9 subjects identifying as "Other." The "Other" category is other individuals at the conferences who likely work in related fields. Excluding subjects from the other group, the differences between investors / entrepreneurs and industry professionals is slightly more pronounced. Investors / entrepreneurs are more accurate ($z = 2.27, p = 0.02$) and are less overconfident ($z = 2.13, p = 0.03$).

on that task. The average belief businesspeople have about other subjects' accuracy is 68%, whereas other subjects' average stated confidence for these situations is 78%. Thus, subjects believe that others are overconfident by about 10 percentage points, and in actuality subjects are overconfident by about 14 percentage points. About 55% of the time, subjects believe that the other subject's accuracy is *less* than their confidence. About 26% of the time, subjects believe that the other subject's accuracy is *more* than their confidence, and 19% of the time, subjects take the other subject's guess at face value.[23]

Thus, though subjects are themselves significantly overconfident, they are able to recognize that other businesspeople are overconfident (though they slightly underestimate the degree of other people's overconfidence).[24] This result is consistent with a large literature in psychology on the "blind spot bias," where subjects believe behavioral biases to be stronger in others than in themselves (Pronin *et al.*, 2007; Pronin and Kugler, 2007; Ehrlinger *et al.*, 2005).

Result 3 (Comparative Statics on the Demand for Information) *The demand for information increases as subject confidence decreases and as signal accuracy increases, as predicted qualitatively by theory. However, the magnitudes of these comparative statics are less than the theoretical ones. I fail to reject the null hypothesis that, other things held equal, subjects value subjective and direct information the same.*

The relationship between confidence and WTP, and between empirical accuracy and WTP, are shown in Figures 3 and 4. In these figures, I plot a nonparametric Fan regression (Fan, 1992) of WTP on confidence or empirical accuracy, with an observation being a person-task. An Epanechnikov kernel (e.g., Li and Racine, 2007) is used. Actual WTP tends to slope downward somewhat, meaning that subjects pay less when they are more confident or more accurate.

To more closely examine these relationships, I consider OLS regressions of the form:

$$y_{in} = \beta_0 + \beta_1 r_{in} + \beta_2 c_{in} + \beta_3 SUBJ_{in} + \beta_4 \hat{p}_i + f_n + \epsilon_{in}$$

[23]For Result 2, I drop subjects from the TRAFFIC conference (since no signals from other subjects were sold at TRAFFIC). The phrase "about 10 percentage points" reflects the difference between two rounded numbers; it is 9.1 percentage points when confidence and accuracy are not rounded.

[24]The difference between predicted accuracy and other subject's confidence is highly significant ($z = 4.51, p < 0.01$). Predicted accuracy is also significantly higher than the other subjects' true accuracy ($z = 2.47, p = 0.01$).

where y_{in} is WTP for a signal by person i on task n, r is the signal accuracy, c is confidence, $SUBJ$ is a dummy for whether the signal is a subjective signal (the guess of another person), $\hat{p}_i$ is a person's empirical accuracy on all tasks, f_n is a task fixed effect (i.e., fixed effects for every domain name or website guessed about), and ϵ_{in} is an error. The task fixed effects also soak up any conference-specific effects (tasks were not re-used across conferences). When the empirical accuracy measure is omitted, subject fixed effects can also be included. Signal accuracy is 70% or 90% for direct signals, and is a subject's estimate of the other person's accuracy for subjective signals.

Table 2 shows that WTP increases in signal accuracy and decreases in confidence. For example, column 1 indicates that increasing the signal accuracy by 10 percentage points is associated with a \$0.38 increase in WTP. Column 2 adds subject fixed effects. Within each person, increasing signal accuracy and decreasing confidence is associated with an increase in WTP. In column 3, WTP decreases in empirical accuracy, but the relation is not statistically significant. Interestingly, conditional on signal accuracy (which is objective for computer signals and subjective for subjective signals), there is no evidence that subjects value subjective information differently from computer-generated information.

In computing the value of information given beliefs (discussed further in Result 6), differences in signal accuracy and confidence will only affect the value of information when the signal accuracy is greater than confidence. Thus, I also consider the following regression:

$$\begin{aligned} y_{in} = \beta_0 + \beta_{1a} r_{in} \mathbf{1}(r_{in} \geq c_{in}) + \beta_{1b} r_{in} \mathbf{1}(r_{in} < c_{in}) + \beta_{2a} c_{in} \mathbf{1}(r_{in} \geq c_{in}) \\ + \beta_{2b} c_{in} \mathbf{1}(r_{in} < c_{in}) + \beta_3 SUBJ_{in} + \beta_4 \hat{p}_i + f_n + \epsilon_{in} \end{aligned}$$

Theory predicts $\beta_{1a} = 20$, $\beta_{1b} = 0$, $\beta_{2a} = -20$, and $\beta_{2b} = 0$. As seen in column 4, the estimated coefficients are 5.1, 1.8, -3.5, and -0.6. These estimates are consistent with theory on a very broad, qualitative level. Using a two-sided t-test, I reject the null hypothesis that the impact of signal accuracy does not depend on whether signal accuracy is greater than or equal to confidence ($\widehat{\beta_{1a}} \neq \widehat{\beta_{1b}}, p < 0.01$), and I reject that the impact of confidence does not depend on whether signal accuracy is higher ($\widehat{\beta_{2a}} \neq \widehat{\beta_{2b}}, p = 0.01$). However, the estimates, $\widehat{\beta_{1a}}$ and $\widehat{\beta_{2a}}$, are far less than

predicted by theory. In column 5, I restrict to direct signals only, and get fairly similar results to column 4. However, the difference between $\widehat{\beta_{1a}}$ and $\widehat{\beta_{1b}}$ is smaller, and the difference between $\widehat{\beta_{2a}}$ and $\widehat{\beta_{2b}}$ is no longer statistically significant ($p = 0.21$). See Appendix C for further discussion.

While Result 3 focuses on comparative statics of actual WTP, Result 4 compares actual WTP to optimal WTP.

Result 4 (Overall Optimality of Demand for Information) *Subjects underpay for information of high value, but overpay for information when the value of information is low. Averaged across all tasks in the experiment, internet businesspeople underpay for information.*

Figure 4 illustrates the main finding of the paper: subjects underpay when information is valuable, but overpay when information is less valuable. Following Section 1, the optimal WTP for information by person i for task n is given by $b_{in} = k(r - p_{in}) * \mathbf{1}(r > p_{in})$ where p_{in} is i's true accuracy on n. To implement this empirically, I instead use $b_{in} = k(r - \hat{p}_i) * \mathbf{1}(r > \hat{p}_i)$, where $\hat{p}_i$ is the share of tasks answered correctly by person i in Stage I. Note that the formula here for optimal WTP uses p to mean accuracy (instead of confidence as in Section 1) so as to analyze optimal WTP relative to a "rational benchmark." I analyze optimal WTP given beliefs later in Result 6. For both the 70% and 90% signals, the dotted line of actual WTP lies below the solid line of optimal WTP at lower levels of empirical accuracy (underpayment), but lies above optimal WTP at higher levels of empirical accuracy (overpayment). Actual WTP line is relatively flat with respect to empirical accuracy. Appendix Figure D4 shows the same result using a simple comparison of means for confidence and accuracy levels in different "bins" (i.e., average WTP when confidence is between 50 and 59 percent; average WTP when confidence is between 60 and 69 percent; etc.).

Why is WTP so flat with respect to empirical accuracy? One explanation is overconfidence. People are likely to be overconfident if tasks are difficult (i.e., if their accuracy is low) and underconfident if tasks are easy (i.e., if their accuracy is high) (Moore and Healy, 2008), and I find this in Figure 2 as well. This could cause people to underpay for valuable signals and overpay for less valuable signals. However, as is shown in Result 6, the relationship between WTP and beliefs is also too flat, suggesting that overconfidence is not the entire story.

A second explanation is a combination of conservatism and base rate neglect. Conservatism is

the psychological bias where people do not update their beliefs to the extent of a rational decision-maker after receiving information, with beliefs biased toward the prior (Edwards, 1968).[25] If subjects anticipate not updating to the extent that a Bayesian would after receiving information, this would make information less valuable for them. For example, if a risk-neutral subject has a 60% accurate prior, a 90% signal should be worth $\$20 * (.9 - .6) = \$6$. However, if the subject anticipates that they will not fully update after receiving the signal, for example, that they will only update to the extent that a rational person would after receiving a 70% signal, then their WTP will only be \$2. Base rate neglect is people's tendency to underweight their priors (Tversky and Kahneman, 1974). For example, if a subject with a 60% prior anticipated suffering from complete base rate neglect, they would be willing to pay \$8 for a 90% accurate signal instead of \$6. Combined together, conservatism and base rate neglect could lead to underpayment for valuable signals and overpayment for less valuable signals.[26] I return to this point again in Section 4.2.[27]

The data can also be used to estimate the degree of overpayment (or underpayment) averaging over all the tasks in the experiment. Before beginning this exercise, it is important to emphasize that the answer may be shaped by which tasks were chosen—had I chosen harder or easier tasks, the answer might be quite different. With this caveat in mind, I examine the below equation:

$$y^*_{in} = b_{in} + \theta + \epsilon_{in} \tag{2}$$

where $b_{in} = k(r - \hat{p}_i) * \mathbf{1}(r > \hat{p}_i)$ is the optimal WTP of person i for information about task n

[25] For some recent evidence on conservatism, see, e.g., Huck and Weizsacker (2002); Eil and Rao (2011); Mobius *et al.* (2014).

[26] For example, consider a subject deciding to purchase a 90% accurate signal. Due to forecasted conservative updating, assume that it is as if the signal will only have accuracy of 70% for him. However, due to base rate neglect, assume that no matter what his accuracy is, it is as if his prior accuracy is only 50%. Thus, his WTP is always \$4 no matter his accuracy. When the subject has accuracy of 50%-70% (that is, when the signal is highly informative), he underpays for information. And when he has accuracy of 70%-100% (that is, when the signal is less valuable), he overpays for information.

[27] An additional possible explanation for the flat relationship between empirical accuracy and WTP is mis-measurement of empirical accuracy, which may occur due to the fact that empirical accuracy is measured using a finite number of tasks per subject. To examine this explanation, I exploit the fact that different subjects faced differing numbers of tasks (10, 20, or 25). Repeating column 3 of Table 2 separately for subjects with 10, 20, or 25 tasks, the relationship between empirical accuracy and WTP is quite flat regardless of whether there are 10, 20, or 25 tasks per person. This suggests that mis-measurement of empirical accuracy seems unlikely to be the sole driver of observed flatness between empirical accuracy and WTP. I cannot rule out, though, that measurement error is a contributing factor. See also Appendix C.

(a particular domain name or website), θ is a systematic taste for information, y^*_{in} is the amount i spends on information about n in a world where the amount spent is unconstrained, and ϵ is a normally-distributed error. The goal is to test the null hypothesis that $\theta = 0$.

In the experiment, WTP is constrained to lie between \$0 and \$10. In cases where optimal WTP is near zero, failing to correct for this constraint may introduce bias.[28] To correct for this, define $y = 1(0 \leq y^* \leq 10) \times y^* + 1(y^* > 10) \times 10$. Then, the equation of interest is: $y_{in} = b_{in} + \theta + \epsilon_{in}$, where I estimate to control for censoring. Plugging in for b_{in}, the equation of interest is:

$$y_{in} - k(r - \hat{p}_i) * \mathbf{1}(r > \hat{p}_i) = \theta + \epsilon_{in} \tag{3}$$

To estimate (3), I first use OLS. However, OLS may suffer from bias, with the direction of bias determined by whether there are more subjects censored from below (this will bias up the estimate) or from above (this will bias down the estimate). My preferred specification is maximum likelihood "tobit," controlling for censoring on the left-hand side at $-b_{in}$ and on the right-hand side $10 - b_{in}$. (More specifically, it is a censored normal regression, as it slightly generalizes the basic tobit by allowing the censoring point to vary by observation, but I refer to it as "tobit.")

Panel 1 of Table 3 shows results estimating under or over-payment in relation to the baseline case of no overconfidence and no misusage. The numbers in the table represent the amount of overpayment (negative numbers represent underpayment), assuming subjects are risk-neutral. In simple OLS regressions, averaged over all tasks, subjects underpay for 70%, 90%, and subjective signals. Under tobit regressions, the coefficients increase in magnitude (more underpayment), reflecting the many instances where subjects are not willing to pay anything for information.

Appendix Table D2 relaxes the assumption of risk neutrality, with results qualitatively similar to those in Table 3. I estimate (3), but use the value of information under risk aversion (as in equation (1)). Subjects are assigned a coefficient of risk aversion corresponding to the number of safe choices they selected in the Holt and Laury (2002) task. Controlling for risk aversion, the

[28] To see why, consider the case where the optimal WTP on all tasks is zero because subjects are very skillful. Assume also that subjects' true demand for information is on average rational, but that on each task subjects state their WTP with a zero-mean error. In this case, the average measured WTP will be positive (some people will state positive WTPs and some will state a WTP of zero, but it will be impossible to observe negative WTPs), even though the optimal WTP is zero and subjects are rational.

magnitudes of average underpayment are slightly lower. Appendix C gives additional robustness checks and discussion.

Result 5 (Information Usage) *Subjects often pay for information they do not use.*

A rational subject should first assess whether the signal is more accurate than her prior, and if so, purchase the information with a plan to use it. However, when my subjects obtained information that was different from their original guess, 34% of the time they did not follow the new information (101 / 301 times), which is highly significantly different from the rational benchmark of zero misusage ($t = 12.3, p < 0.01$, subject-task-level t-test). The rate of not using purchased information is lowest for the 90% accurate signals (20%, 25/127), and higher for the 70% accurate signals (38%, 44/116) and subjective signals (55%, 32/58). Thus, subjects are least likely to misuse the most accurate information, the 90% signals, and most likely to misuse the information they perceive as least accurate, the subjective signals (recall from Result 2 that subjects believe that other subjects are 68% accurate on average).[29]

A question of interest is whether overconfidence and misusage are correlated within subjects. That is, are subjects who exhibit one type of behaviorial bias more likely to exhibit another one? Appendix Figure D5 compares rates of overconfidence and misusage within subject. The line of best fit has a positive slope of 0.27 (standard error of 0.29), implying that moving from a subject who is well-calibrated to one that is overconfident by 10 percentage points is associated with an increase in misusage of 2.7 percentage points. The correlation, however, is not statistically significant.[30]

[29] One reaction to the evidence that subjects misuse information about one-third of the time is that this may reflect subjects not understanding the experiment. If this were true, misusage should be a random event, and rates of misusage should be uncorrelated with signal accuracy. However, this is not the case. Misusage is least common for the 90% signals, more common for the 70% signals, and most common for the subjective signals, which are the least accurate (recall that subjects' overall accuracy was about 62%). Instead, the evidence is consistent with data from many economics experiments where mostly rational agents sometimes make mistakes, but are more likely to make less costly mistakes. Ignoring a 90% accurate signal is more costly than ignoring a 70% accurate signal (or ignoring a subjective one from a population whose accuracy is about 62%), and subjects do it less often. In addition, it should be noted that the optimal strategy of always using the purchased information over one's prior signal may be somewhat counter-intuitive.

[30] The lack of significance arises in part because of a reduced sample size, as over 35% of subjects never receive a signal that is different from one of their guesses, and thus do not have a well-defined rate of misusage. The evidence that overconfidence and misusage are correlated is thus merely suggestive (at best). Appendix Figure D5 also shows that there is substantial heterogeneity in misusage rates across subjects. Some subjects sometimes misuse information, some never misuse information, and some always misuse information. Significant heterogeneity remains even if one restricts to subjects who received a signal different from their guess multiple times.

Another question is whether misusage declines with industry expertise. The answer is no. Among full-time traders or entrepreneurs, the subject-level misusage rate is 36% compared to 28% for others in the industry, but this difference is insignificant ($p = 0.40$ in a Mann-Whitney test).

One way to think about information misusage is as another manifestation of belief conservatism. By misusing information and sticking with their original guess, my subjects are fully ignoring the information they receive. Given the observed rates of information misusage, I now turn to whether subjects' demand for information is optimal given this misusage.

Result 6 (Optimality of Demand for Information Given Subjects' Beliefs and Information Usage) *Taking subject beliefs as given, subjects underpay for information when it is valuable, but overpay when it is less valuable, as in the baseline. The same pattern also holds taking information misusage rates as given. Averaged across all tasks in the experiment, subjects tend to underpay for information either taking beliefs as given or taking information misusage rates as given.*

Besides examining optimal information acquisition relative to the rational baseline, optimal acquisition can also be examined given subject overconfidence or given subjects' tendency to misuse information. Consider a subject who is 60% accurate and 75% confident about their performance on a task. For a 90% signal, the subject should have WTP of \$6 relative to the rational baseline (\$6=(.9-.6)*\$20) and should have WTP of \$3 given their overconfidence (\$3=(.9-.75)*\$20).

Figure 3 shows a slight variant of the paper's main finding: subjects underpay for valuable information and overpay for less valuable information, as in Result 4, but here, *taking subject beliefs as given* (i.e., basing the value of information on confidence instead of accuracy). Letting c be the subject's confidence, the optimal WTP for information on task n given subject i's beliefs is $b_{in} = k(r - c_{in}) * \mathbf{1}(r > c_{in})$. In Figure 3, actual WTP lies below optimal WTP at low confidence levels (underpayment), but lies above optimal WTP at high confidence levels (overpayment).

Appendix Figure D6 shows that a similar pattern also holds, *taking misusage rates as given.* The optimal WTP for information given actual information usage can be written as $b_{in} = k(\hat{r_i} - \hat{p_i}) * \mathbf{1}(\hat{r_i} > \hat{p_i})$ where $\hat{r_i}$ is the subject's empirical accuracy after receiving information. The disadvantage of using b_{in} in this form is that subjects may receive information on zero or a small number of tasks, depending on how much they are willing to pay and on the random numbers from the BDM. Instead,

I measure optimal WTP for information given usage with $b_{in} = (1-m)k(r-\hat{p}_i) * \mathbf{1}(r > \hat{p}_i)$ where m is the observed rate of misusing information, that is, of choosing to stick with one's initial guess when the signal contradicts one's guess. Because many subjects do not receive information that contradicts one of their guesses, I measure m using the overall rate of misusage among all subjects for a given signal type (e.g., to calculate optimal WTP for 70% direct signals, I use the misusage rate among all subjects on 70% direct signals). Appendix Figure D6 shows underpayment for highly informative signals and overpayment for less valuable signals.

Averaging over all tasks in the experiment, the data can also be used to estimate average over- or underpayment conditional on subject beliefs. To estimate overpayment for information, I use the following equation, again estimated with a tobit procedure:

$$y_{in} - k(r - c_{in}) * \mathbf{1}(r > c_{in}) = \theta + \epsilon_{in}, \tag{4}$$

By estimating both (3) and (4), I attempt to disentangle the share of underpayment remaining after controlling for overconfidence. Equation (3) measures average total underpayment. Equation (4) measures average underpayment, given subject beliefs. The percentage of underpayment remaining after controlling for overconfidence is just the ratio between the two estimates of underpayment: $\frac{\hat{\theta}_4}{\hat{\theta}_3}$. $\hat{\theta}_3$ and $\hat{\theta}_4$ denote estimated overpayment for information in equations (3) and (4), respectively. A similar methodology can be used to measure the share of underpayment remaining after controlling for information misusage.[31]

Results on the optimality of the demand for information given beliefs are shown in Panel 2 of Table 3. Once censoring is accounted for, subjects underpay on average for all three types of signals, though it is not statistically different from zero for subjective signals. Further discussion on these results is given in Appendix B.1.

Panel 3 of Table 3 shows estimates of the amount of average overpayment given how subjects

[31] Incorporating misusage into the value of information, the relevant regression equation is then:

$$y_{in} - (1-m)k(r - \hat{p}_i) * \mathbf{1}(r > \hat{p}_i) = \theta + \epsilon_{in}. \tag{5}$$

The share of underpayment remaining after controlling for information misusage is $\frac{\hat{\theta}_5}{\hat{\theta}_3}$. I use the phrase "controlling for information misusage" or "controlling for overconfidence" in the sense of adjusting optimal WTP to account for misusage or overconfidence. This is not "controlling for" as in a regression.

actually use information. Once a censoring correction is applied, subjects underpay on average for all three types of information. Panel 4 of Table 3 shows estimates of the amount of average overpayment given subject beliefs and given how subjects actually use information. In the specifications controlling for censoring, only the average underpayment for 90% signals remains statistically significant.

Finally, as I discuss in Appendix B.2, there is substantial heterogeneity in the demand for information across subjects.

4 Discussion

I find that people underpay for information when it is valuable and overpay when it is less valuable. First, I discuss the relationship of this finding to the literature. Second, I flesh out how conservatism and base rate neglect could lead to my results. Third, I discuss additional alternative explanations beyond overconfidence, misusage, conservatism, and base rate neglect. These additional alternative explanations seem less successful at fully explaining the results.

4.1 Relation to Other Literature on Optimal Information Acquisition

As discussed in the introduction, there are a few other papers on optimal information acquisition, several of which suggest that subjects may purchase "too much" information. Kubler and Weizsacker (2004) and Kraemer *et al.* (2006) show in herding experiments that many agents over-acquire costly signals (relative to what should occur in equilibrium). Eliaz and Schotter (2010) show that agents sometimes purchase information with no instrumental value in a simple guessing task.

Given that these experiments involve situations where information often has little or no instrumental value, these experiments are fully consistent with my main result, that subjects underpay for information when it is relatively valuable and overpay when it is less valuable. In my study too, I also find that subjects pay positive amounts for information with low value. However, in my study, I observe information acquisition across a wide range of tasks and situations, and find that subjects underpay when information is valuable.

There a few other possibilities that could explain differences between our results. A first

possibility, in relation to Kubler and Weizsacker (2004) and Kraemer *et al.* (2006), concerns beliefs about others versus optimality of the demand for information. In these experiments, I might pay for information either (1) Because I overvalue information, e.g., I value it for non-instrumental reasons, or (2) Because I am concerned that the subjects moving before me may not have behaved optimally. It may be possible that subjects do not overvalue information, but that (2) holds.

A second possibility, in relation to Eliaz and Schotter (2010), concerns the nature of expertise. Eliaz and Schotter (2010) argue that their student subjects are acquiring information in the experiment for its psychic value. One reason I obtain different results could be that experts may not need information for its psychic value (e.g., to confirm that one is right in order to feel good about oneself). It might be possible that experts are averse to receiving information that could contravene their self-image (as I discuss in my "ego utility" alternative explanation below).

4.2 Conservatism and Base Rate Neglect

As discussed in Section 3, conservatism and base rate neglect can help explain why WTP is relatively flat with respect to accuracy and confidence. How much conservatism and base rate neglect would you need to explain my results? For this, I use a simple parameterization of conservatism and base rate neglect based on Grether (1980).

Base rate neglect is the tendency to ignore one's prior. In a binary model, ignoring one's prior means acting as if one had a 50/50 prior. Base rate neglect is parameterized using $\beta \in [0, 1]$, where prior accuracy, p, is replaced by p', with $\frac{p'}{1-p'} = (\frac{p}{1-p})^{\beta}$. The closer β is to 0, the greater the level of base rate neglect. In my context, conservatism can be thought of as not believing that the signal is as accurate as it actually is. Conservatism is parameterized using $\alpha \in [0, 1]$, where actual signal accuracy, r, is replaced by r', with $\frac{r'}{1-r'} = (\frac{r}{1-r})^{\alpha}$. The closer α is to 0, the greater the level of conservatism. The value of information is then defined by $k(r' - p') * \mathbf{1}\,(r' - p' > 0)$. I estimate α and β using non-linear least squares, minimizing the squared distance between actual WTP and optimal WTP. (For cases where $r = 1$ or $p = 1$, the ratios $\frac{r}{1-r}$ and $\frac{p}{1-p}$ are not well-defined; in these cases, I assume that $r' = 1$ or $p' = 1$. Estimates are qualitatively similar and slightly smaller if I drop observations with $r = 1$ or $p = 1$.)

My main finding is you seem to need "a lot" of conservatism and base rate neglect (relative to estimates in the literature) to rationalize the results. Using subjective beliefs for the prior (that is, relative to observed overconfidence), I estimate that $\alpha = 0.32$ (standard error $= 0.03$) and $\beta = 0.09$ ($s.e. = 0.03$). As seen in Appendix Figure D8, the model with conservatism and base rate neglect (the solid line) can reasonably well approximate the relationship between beliefs and WTP, assuming the estimated parameters. However, to match the observed patterns in the data, a substantial amount of conservatism and base rate neglect would be required. To give a sense of magnitudes, $\alpha = 0.32$ implies that a signal accuracy of 0.9 is regarded for information acquisition purposes as having a signal accuracy of $r' = 0.67$. Further, $\beta = 0.09$ means that a true prior of $p = 0.77$ is regarded for information acquisition purposes as a prior of $p' = 0.53$.

As detailed in Appendix A, my estimates of conservatism and base rate neglect are large relative to those in the experimental literature on belief updating. I interpret this as evidence that conservatism and base rate neglect are unlikely to fully explain behavior in my experiment.

4.3 Additional Alternative Explanations for Main Findings

Ego Utility. One potential alternative explanation is ego utility. Experts may value feeling that they are experts on the subject, and may be averse to receiving a signal that could contradict their expertise. This is especially so given that purchased signals were the main source of feedback in the experiment. That subjects are able to recognize that other subjects tend to be overconfident is consistent with an ego utility explanation. One difficulty with ego utility as a full explanation of underpayment for highly informative signals is that, if greater expertise leads subjects to become more averse to receiving contradictory information, one would expect that greater expertise would lower the demand for information. Appendix Table D1 shows correlates of overpayment for information, and shows that this is not the case—subjects who declared their primary occupation as internet investor / entrepreneur (and who may have greater expertise than other subjects) do not suffer from greater underpayment for information. In fact, the coefficients point more in the opposite direction. However, this is only one test with a coarse measure of expertise, so one cannot rule out that ego utility may be important.

Ego utility thus provides a leading alternative explanation for one-half of my main result, that subjects underpay for highly valuable signals. It is not clear, however, how ego utility could help explain the other half, that subjects overpay for less valuable signals.

Loss Aversion. I showed that incorporating risk aversion into the value of information has little effect on the Table 3 results, but what about loss aversion? Suppose the agent has a simple kinked utility function $U(x, e) = (x - e) * \mathbf{1}(x \geq e) - L * (e - x) * \mathbf{1}(x < e)$, where the subject's \$10 endowment ($e = 10$) is her reference point and $L \geq 1$ is a constant indicating the agent's degree of loss aversion. In this case, the optimal WTP for information is $b = \frac{k(r-p)}{r+(1-r)L} * \mathbf{1}(r > p)$. Although I did not measure a subject's degree of loss aversion during the experiment, I can assume some degree of loss aversion and see if/how my results change. In Appendix Table D4, I re-estimate Table 3 assuming that $L = 3$. The estimates change somewhat (with less average underpayment), but the estimates are generally qualitatively similar. Thus, even if it is assumed that subjects are significantly loss averse, feeling a dollar "lost" three times more than a dollar "gained," this cannot explain away the Table 3 results.

Failure of the Quadratic Scoring Rule. Another possibility is that the experiment could have failed to accurately elicit beliefs using the risk-invariant QSR. To guard against this possibility, significant care was taken to ensure that subjects understood the QSR. The instructions stated in bold that it was optimal for subjects to list their true guesses and confidence levels, and the author was on-hand to answer questions. One rough piece of evidence that subjects took the belief elicitation seriously (beyond the evidence in Result 1 above) is that there is a good deal of within-subject variation in confidence levels. For each subject, I calculated the standard deviation of confidence across tasks, and the average of these standard deviations is about 12%. Moreover, to the extent that the QSR failed in some way to elicit beliefs, this would only be relevant for findings that depended on confidence. It would not affect our finding that the relationship between empirical accuracy and WTP, as well as the relationship between signal accuracy and WTP, are far flatter than predicted by theory.

Selection into Participation. As in many experiments, there is potentially a concern here about whether the subjects who chose to participate in the experiment are different from potential

subjects who did not (Levitt and List, 2007a,b). No businessperson was forced to participate in the study. For example, if the most overconfident subjects were more likely to participate in the study, this would cause my results to be unrepresentative of behavior in the broader population of internet businesspeople subjects. Lacking data on non-participants, I am unable to rule out this possibility. However, there is no clear reason to me why this would be the case. The study was advertised as an "economics experiment," so it is not clear how subjects would select in based on the way they valued information. In addition, as discussed above, some of my results hold within subject, e.g., Table 2 shows that an increase in overconfidence lowers the demand for information even after controlling for subject fixed effects. There is less concern about selection for within subject results (except if the selection occurred on how strongly a subject's WTP responds to changes in beliefs).

Mismeasurement of the Misusage Rate. Subjects only have a chance to misuse information when they successfully purchase information and when the information signal purchased differs from their initial guess. Given that this only occurs on roughly 10% of tasks and is not exogenously assigned, it may be the case that the average observed misusage rate is not representative of the true misusage rate. To assess how much this could matter for over- or underpayment given misusage, I consider the possibility that the true misusage rate exceeds the observed one by 25%. As seen in Appendix Figure D7, our main result still holds.[32]

5 Conclusion

I design a novel experiment on optimal information acquisition and conduct it using businesspeople experts. When subjects have low accuracy relative to signal accuracy, they tend to underpay for information, whereas when they have high accuracy, they tend to overpay for information. Experts exhibit significant overconfidence and confidence is associated with a lower demand for information, but the relationship between confidence and WTP is much flatter than is optimal. Even adjusting the value of information for overconfidence, subjects underpay for highly valuable signals, while

[32] Our main result, of underpayment for highly informative signals and overpayment for less informative signals, also holds if the true misusage rate is less than the observed one. Further, as a robustness check, I re-estimated equation (5), assuming that the true misusage rate exceeds the observed one by 25%. Appendix Table D5 shows qualitatively similar results (though average underpayment for subjective signals is no longer statistically significant).

overpaying for less valuable signals. This overpayment and underpayment also holds adjusting the value of information for subjects' tendency to often misuse information, that is, to ignore signals after purchasing them.

Beyond overconfidence and information misusage, I consider a number of forces that could potentially explain the result that subjects underpay for highly valuable signals and overpay for less valuable signals. One explanation is a combination of conservatism and base-rate neglect. While theoretically promising, a large degree of conservatism and base-rate neglect would be required to rationalize the results. Other alternative forces include ego utility and loss aversion. For these other alternative explanations, I cannot rule them out, particularly ego utility, but they seem unlikely to fully explain the results. Ultimately, while I observe a robust empirical pattern of underpayment and overpayment, future work is required to reach a more definitive explanation for this pattern. By using a controlled experiment (as opposed to field data) to study optimal information acquisition, I can rule out various confounds such as people not being aware that information is available or that the cost of acquiring information is heterogeneous. By using an expert subject pool, I help rule out confounds such as that the information-acquiring agents are unsophisticated or unfamiliar with the information they are acquiring.

It should be strongly highlighted, however, that the experts in the experiment showed clear deviations from optimal behavior. WTP did not vary nearly enough with confidence and accuracy as predicted by theory; confidence is only weakly related to accuracy; and information is misused often. Given these stark deviations from optimal behavior, some readers may be tempted to call into question whether the subjects are truly "experts." As discussed earlier, the subjects have expertise regarding websites and domain names, and there is a relevant real-world market for costly information (i.e., domain name appraisals); using such subjects helps overcome critiques that tasks were artificial, unfamiliar, or lacked meaning for subjects. However, the subjects are not experts in experimental economic methods. Probably more importantly, that the subjects are experts does not imply that they should be immune from common biases. Additional research on optimal information acquisition using other expert populations is clearly warranted.

My results are potentially relevant for interpreting findings on optimal information acquisition

in various field settings. On one hand, people seem to over-value certain types of information, e.g., it has been argued that investors overpay for advice from financial managers (e.g., Gennaioli *et al.*, 2015) and that patients and doctors often request medical information that is of little therapeutic value (Abaluck *et al.*, 2014). On the other hand, there is evidence that people may under-invest in acquiring basic information about school quality (Hastings and Weinstein, 2008) and about features of the tax code (Chetty and Saez, 2013). Such evidence need not be at odds, and is potentially consistent with my main finding that people underpay for high-value signals and overpay for low-value signals (though such evidence surely also reflects important institutional factors beyond my experiment).

It is important to note the limitations of the analysis. I focused on a particular set of tasks using a particular set of signal accuracies. Had I focused on signals with lower or higher accuracy, it is possible that different patterns of under- and overpayment may have been observed. Although I use experts performing tasks which are closely related to real-world ones they actually engage in, the setting in my experiment is still artificial. Furthermore, it is unclear whether my results would translate into natural information acquisition situations. In an experimental setting mimicking common business activities, internet businesspeople under-acquired relatively valuable information and over-acquired less valuable information. But this does not necessarily imply that businesspeople would do so in transactions outside the experiment. To get at this and to deal with the alternative explanations discussed above, additional field experiments are needed.

University of Toronto

References

Abaluck, J., Agha, L., Kabrhel, C., Raja, A. and Venkatesh, A. (2014). 'Negative tests and the efficiency of medical care: What determines heterogeneity in imaging behavior?', NBER Working Paper 19956.

Anderson, L. and Holt, C. (1997). 'Information cascades in the laboratory', *American Economic Review*, vol. 87(5), pp. 847–862.

Baumann, A., Deber, R. and Thompson, G. (1991). 'Overconfidence among physicians and nurses: The 'micro-certainty, macro-uncertainty' phenomenon', *Social Science & Medicine*, vol. 32(2), pp. 167–174.

Benjamin, D., Rabin, M. and Raymond, C. (2014). 'A model of non-belief in the law of large numbers', Working paper, Cornell University.

Benoit, J.P., Dubra, J. and Moore, D. (2008). 'A proper test of overconfidence', MPRA Paper No. 11954.

Berg, J., Daley, L., Dickhaut, J. and O'Brien, J.R. (1986). 'Controlling preferences for lotteries on units of experimental exchange', *Quarterly Journal of Economics*, vol. 101(2), pp. 281–306.

Blavatskyy, P. (2009). 'Betting on own knowledge: Experimental test of overconfidence', *Journal of Risk and Uncertainty*, vol. 38(1), pp. 39–49.

Brier, G.W. (1950). 'Verification of forecasts expressed in terms of probability', *Monthly Weather Review*, vol. 78(1), pp. 1–3.

Burks, S., Carpenter, J., Goette, L. and Rustichini, A. (2013). 'Overconfidence and social signalling', *Review of Economic Studies*, vol. 80, pp. 949–983.

Camerer, C. and Lovallo, D. (1999). 'Overconfidence and excess entry: An experimental approach', *American Economic Review*, vol. 89(1), pp. 306–318.

Cesarini, D., Johannesson, M., Lichtenstein, P. and Wallace, B. (2009). 'Heritability of overconfidence', *Journal of the European Economic Association*, vol. 7(2-3), pp. 617–627.

Cesarini, D., Sandewall, O. and Johannesson, M. (2006). 'Confidence interval estimation tasks and the economics of overconfidence', *Journal of Economic Behavior & Organization*, vol. 61(3), pp. 453–470.

Charness, G., Rustichini, A. and van de Ven, J. (2013). 'Self confidence and strategic behavior', CESifo Working Paper No. 4517.

Chetty, R. and Saez, E. (2013). 'Teaching the tax code: Earnings responses to an experiment with eitc recipients', *American Economic Journal: Applied Economics*, vol. 5(1), pp. 1–31.

Cipriani, M. and Guarino, A. (2009). 'Herd behavior in financial markets: An experiment with financial market professionals', *Journal of the European Economic Association*, vol. 7(1), pp. 206–233.

Edelman, B. (2009). 'Red light states: Who buys online adult entertainment?', *Journal of Economic Perspectives*, vol. 23(1), pp. 209–220.

Edwards, W. (1968). 'Conservatism in human information processing', in (B. Kleinmuntz, ed.), *Formal Representation of Human Judgement*, pp. 17–52.

Ehrlinger, J., Gilovich, T. and Ross, L. (2005). 'Peering into the bias blind spot: People's assess-

ments of bias in themselves and others', *Personality and Social Psychology Bulletin*, vol. 31(5), pp. 680–692.

Eil, D. and Rao, J.M. (2011). 'The good news-bad news effect: Asymmetric processing of objective information about yourself', *American Economic Journal: Microeconomics*, vol. 3(2), pp. 114–38.

Eliaz, K. and Schotter, A. (2010). 'Paying for confidence: An experimental study of the demand for non-instrumental information', *Games and Economics Behavior*, vol. 70(2), pp. 304–324.

Ericson, K.M.M. (2011). 'Forgetting we forget: Overconfidence and prospective memory', *Journal of the European Economic Association*, vol. 9(1), pp. 43–60, ISSN 1542-4774.

Fan, J. (1992). 'Design-adaptive nonparametric regression', *Journal of the American Statistical Association*, vol. 87(420), pp. 998–1004.

Fong, C. and Oberholzer-Gee, F. (2011). 'Truth in giving: Experimental evidence on the welfare effects of informed giving to the poor', *Journal of Public Economics*, vol. 95(5-6), pp. 436–444.

Ganguly, A. and Tasoff, J. (2014). 'Fantasy and dread: The demand for information and the consumption utility of the future', Working paper.

Gennaioli, N., Shleifer, A. and Vishny, R. (2015). 'Money doctors', *Journal of Finance*, vol. 70(1), pp. 91–114.

Goldman, A. (2007). 'Domain names: 21st century real estate', *Washington Post*, 7-22-2007.

Grether, D.M. (1980). 'Bayes rule as a descriptive model: The representativeness heuristic', *Quarterly Journal of Economics*, vol. 95(3), pp. 537–57.

Griffin, D. and Tversky, A. (1992). 'The weighing of evidence and the determinants of confidence', *Cognitive Psychology*, vol. 24, pp. 411–435.

Grossman, S.J. and Stiglitz, J.E. (1980). 'On the impossibility of informationally efficient markets', *American Economic Review*, vol. 70(3), pp. 393–408.

Grossman, Z. and Owens, D. (2012). 'An unlucky feeling: Overconfidence and noisy feedback', *Journal of Economic Behavior & Organization*, vol. 84(2), pp. 510 – 524.

Hao, L. and Houser, D. (2010). 'Getting it right the first time: Belief elicitation with novice participants', Working paper, George Mason University.

Harrison, G. and List, J. (2004). 'Field experiments', *Journal of Economic Literature*, vol. 42(4), pp. 1013–1059.

Hastings, J. and Weinstein, J. (2008). 'Information, school choice and academic achievement: Evidence from two experiments', *Quarterly Journal of Economics*, vol. 123(4), pp. 1373–1414.

Hoelzl, E. and Rustichini, A. (2005). 'Overconfident: Do you put your money on it?', *Economic Journal*, vol. 115(503), pp. 305–318.

Hoffman, M. (2007). 'The pricing of second-hand internet domain names', Unpublished senior thesis, Yale University.

Hoffman, M. and Morgan, J. (2015). 'Who's naughty? who's nice? experiments on whether prosocial workers are selected out of cutthroat business environments', *Journal of Economic Behavior & Organization*, vol. 109, pp. 173 – 187.

Holt, C. and Laury, S. (2002). 'Risk aversion and incentive effects', *American Economic Review*, vol. 92(5), pp. 1644–1655.

Holt, C.A. and Smith, A.M. (2009). 'An update on bayesian updating', *Journal of Economic Behavior and Organization*, vol. 69(2), pp. 125 – 134.

Hossain, T. and Okui, R. (2013). 'The binarized scoring rule', *Review of Economic Studies*, vol. 80(3), pp. 984–1001.

Huck, S. and Weizsacker, G. (2002). 'Do players correctly estimate what others do? : Evidence of conservatism in beliefs', *Journal of Economic Behavior & Organization*, vol. 47(1), pp. 71–85.

Kesmodel, D. (2008). *The Domain Game*, Xlibris Publishing.

Koszegi, B. (2003). 'Health anxiety and patient behavior', *Journal of Health Economics*, vol. 22(6), pp. 1073 – 1084.

Kraemer, C., Noeth, M. and Weber, M. (2006). 'Information aggregation with costly information and random ordering: Experimental evidence', *Journal of Economic Behavior and Organization*, vol. 529, pp. 432–442.

Kubler, D. and Weizsacker, G. (2004). 'Limited depth of reasoning and failure of cascade formation in the laboratory.', *Review of Economic Studies*, vol. 71, pp. 425–441.

Levitt, S.D. and List, J.A. (2007a). 'Viewpoint: On the generalizability of lab behaviour to the field', *Canadian Journal of Economics*, vol. 40(2), pp. 347–370.

Levitt, S.D. and List, J.A. (2007b). 'What do laboratory experiments measuring social preferences reveal about the real world?', *Journal of Economic Perspectives*, vol. 21(2), pp. 153–174.

Li, Q. and Racine, J.S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton, NJ: Princeton University Press.

List, J.A. and Haigh, M. (2005). 'A simple test of expected utility theory using professional traders', *Proceedings of the National Academy of Science*, vol. 102(3), pp. 945–948.

Martinelli, C. (2006). 'Would rational voters acquire costly information?', *Journal of Economic Theory*, vol. 129(1), pp. 225–251.

McKelvey, R. and Page, T. (1990). 'Public and private information: An experimental study of information pooling', *Econometrica*, vol. 58(6), pp. 1321–1339.

McKenzie, C.R., Liersch, M.J. and Yaniv, I. (2008). 'Overconfidence in interval estimates: What does expertise buy you?', *Organizational Behavior and Human Decision Processes*, vol. 107(2), pp. 179–191.

Milgrom, P. and Weber, R.J. (1982). 'The value of information in a sealed-bid auction', *Journal of Mathematical Economics*, vol. 10(1), pp. 105 – 114.

Mobius, M.M., Niederle, M., Niehaus, P. and Rosenblat, T. (2014). 'Managing self-confidence: Theory and experimental evidence', Working paper, UC San Diego.

Moore, D.A. and Healy, P.J. (2008). 'The trouble with overconfidence', *Psychological Review*, vol. 115(2), pp. 502 – 517.

Offerman, T., Sonnemans, J., Van De Kuilen, G. and Wakker, P.P. (2009). 'A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes', *Review of Economic Studies*, vol. 76(4), pp. 1461–1489.

Palacios-Huerta, I. and Volij, O. (2009). 'Field centipedes', *American Economic Review*, vol. 99(4), pp. 1619–1635.

Pronin, E. and Kugler, M.B. (2007). 'Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot', *Journal of Experimental Social Psychology*, vol. 43(4), pp. 565 – 578.

Pronin, E., Lin, D.Y. and Ross, L. (2007). 'The bias blind spot: Perceptions of bias in self versus others', *Personality and Social Psychology Bulletin*, vol. 28(3), pp. 369–381.

Ropelato, J. (2006). 'Internet pornography statistics', http://internet-filter-review.toptenreviews.com/internet-pornographystatistics.html.

Schlag, K.H. and van der Weele, J. (2009). 'Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality', Working paper, Universitat Pompeu Fabra.

Schotter, A. (2003). 'Decision making with naive advice', *American Economic Review*, vol. 93(2), pp. 196–201.

Selten, R., Sadrieh, A. and Abbink, K. (1999). 'Money does not induce risk neutral behavior, but binary lotteries do even worse', *Theory and Decision*, vol. 46(3), pp. 213–252.

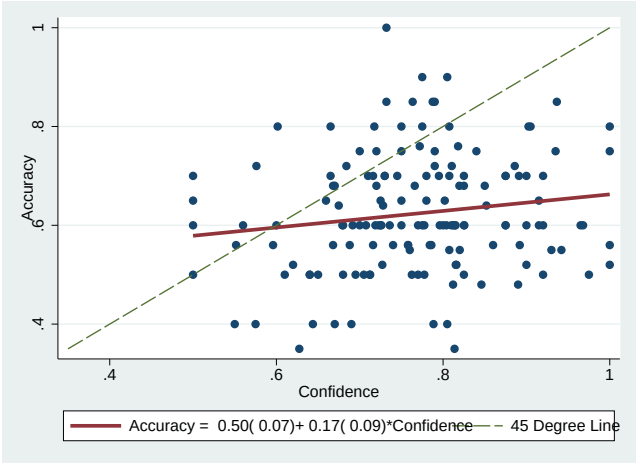
Sunder, S. (1992). 'Market for information: Experimental evidence', *Econometrica*, vol. 60(3), pp. 667–95.

Tversky, A. and Kahneman, D. (1974). 'Judgment under uncertainty: Heuristics and biases', *Science*, vol. 185(4157), pp. 1124–1131.

Wagenaar, W. and Keren, G.B. (1986). *Does the Expert Know? The Reliability of Predictions and Confidence Ratings of Experts*, Intelligent Decision Support in Process Environments, Berlin: Springer.

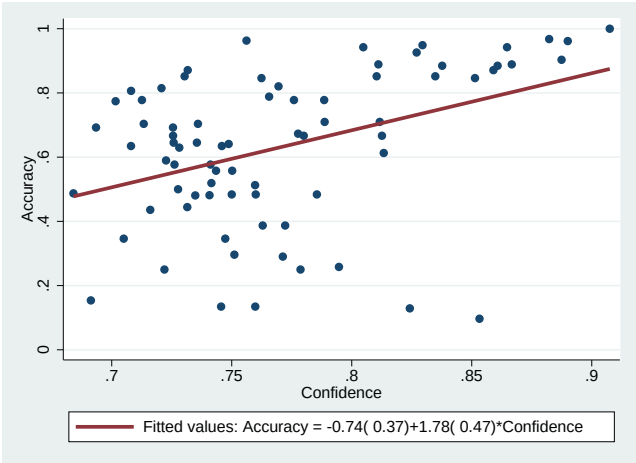
Walker, J.M., Smith, V.L. and Cox, J.C. (1990). 'Inducing risk-neutral preferences: An examination in a controlled market environment', *Journal of Risk and Uncertainty*, vol. 3(1), pp. 5–24.

Figure 1: Confidence and Accuracy Across Subjects



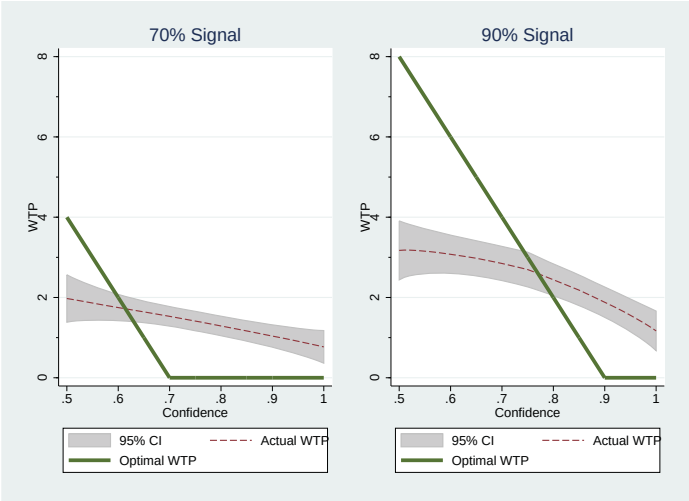
Notes: Each dot represents the average confidence and accuracy per subject. The solid line is based on an unweighted subject-level regression of average accuracy on average confidence. The dotted line is the 45-degree line. Individuals who participated twice are counted separately (one dot per round). The data are from Stage I of the experiment.

Figure 2: Confidence and Accuracy Across Tasks



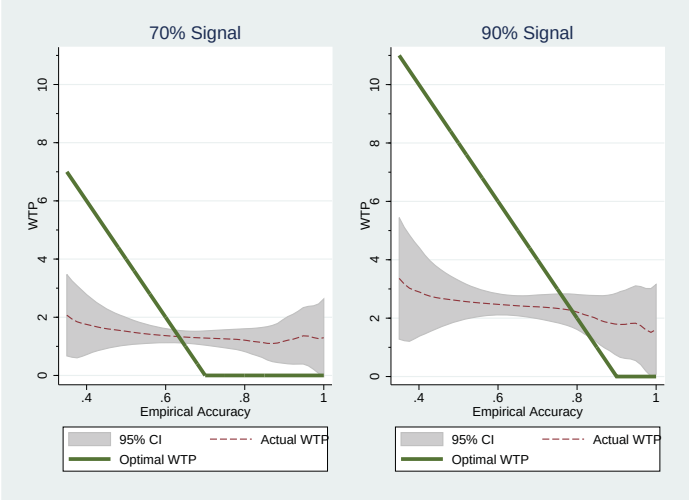
Notes: Each dot represents the average confidence and accuracy per task. The drawn line is based on an unweighted task-level regression of average accuracy on average confidence. The data are from Stage I of the experiment.

Figure 3: Actual WTP and Optimal WTP Given Beliefs



Notes: Actual WTP is plotted using a locally weighted Fan regression with an Epanechnikov kernel (bandwidth=0.25). The 95% confidence intervals account for clustering by subject and are calculated using 50 bootstrap replications. Optimal WTP given beliefs is $20(r - c_{in}) * \mathbf{1}(r > c_{in})$ where r is signal accuracy and c_{in} is subject i's confidence on task n.

Figure 4: Actual WTP and Optimal WTP Given Empirical Accuracy



Notes: Actual WTP is plotted using a locally weighted Fan regression with an Epanechnikov kernel (bandwidth=0.30). The 95% confidence intervals account for clustering by subject and are calculated using 50 bootstrap replications. Optimal WTP is $20(r - \hat{p}_i) * \mathbf{1}(r > \hat{p}_i)$ where r is signal accuracy and $\hat{p}_i$ is subject i's empirical accuracy. Empirical accuracy is a subject's share of tasks guessed correctly in Stage I of the experiment.

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.
Investor / entrepreneur	96	0.46	0.50
Industry professional	96	0.45	0.50
Years involved with industry	70	6.36	3.85
High school or less	80	0.03	0.16
Some college	80	0.23	0.42
College	80	0.40	0.49
Graduate school	80	0.35	0.48
Total income less than 100k	36	0.47	0.51
Total income between 100k and 300k	36	0.39	0.49
Total income more than 300k	36	0.14	0.35
Asian	48	0.10	0.31
Black	48	0.04	0.20
Hispanic	48	0.04	0.20
White	48	0.75	0.44
Male	80	0.78	0.42
Have purchased a domain appraisal	76	0.17	0.38
Would spend $400 to purchase an appraisal if buying a domain name worth around $1 million	35	0.56	0.45

Notes: Survey data were obtained for 96 of 149 subject-rounds. All variables are binary except "Would spend $400 to purchase an appraisal" (where 0="Probably Not", 0.5="Maybe", and 1="Likely") and "Years involved with industry."

Table 2: Determinants of the Willingness to Pay for Information, OLS Regressions

Sample:	All (1)	All (2)	All (3)	All (4)	Direct signals (5)
Signal Accuracy	3.758***	3.575***	3.786***		
	(0.465)	(0.448)	(0.496)		
Confidence (about own accuracy)	-2.886***	-2.693***	-2.862***		
	(0.849)	(0.469)	(0.865)		
Empirical Accuracy			-0.647	-0.716	-0.822
			(1.571)	(1.574)	(1.811)
Subjective Information Dummy			0.030	-0.043	
			(0.202)	(0.211)	
Signal Accuracy * (Signal Acc $\geq$ Conf)				5.108***	5.340***
				(0.992)	(0.806)
Signal Accuracy * (Signal Acc $<$ Conf)				1.787***	3.331***
				(0.614)	(0.643)
Confidence * (Signal Acc $\geq$ Conf)				-3.455**	-2.751**
				(1.348)	(1.365)
Confidence * (Signal Acc $<$ Conf)				-0.571	-1.372*
				(0.728)	(0.777)
Subject fixed effects	No	Yes	No	No	No
Observations	4,809	4,809	4,809	4,809	4,110
R-squared	0.10	0.57	0.10	0.10	0.13

Notes: Standard errors clustered by subject in parentheses. Each column is a regression. The dependent variable is WTP for information in dollars. An observation is a subject-task. Task fixed effects (i.e., fixed effects for every domain name or website guessed about in the experiments) are included in all regressions. Subject fixed effects cannot be included for columns 3-5 because empirical accuracy is measured at the subject level, and is not measured separately for each subject-task. Empirical accuracy is a subject's share of tasks guessed correctly in Stage I of the experiment. Columns 1-4 analyze all payment decisions whereas column 5 restricts to direct signals. For subjective signals, "Signal accuracy" is a subject's perception of the paired subject's accuracy on that task. For the computer-generated (or "direct signals"), signal accuracy is 70% or 90%. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Average Overpayment for Information

	70% Direct Signals		90% Direct Signals		Subjective Signals	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1- Rational Benchmark						
Overpayment (in $)	-0.580***	-2.071***	-3.200***	-4.270***	-0.974***	-2.871***
	(0.198)	(0.394)	(0.259)	(0.397)	(0.316)	(0.650)
Observations	2,073	2,073	2,069	2,069	706	706
Panel 2- Overpayment Given Beliefs						
Overpayment (in $)	0.558***	-0.798**	-0.582**	-1.762***	0.722***	-0.809
	(0.161)	(0.329)	(0.231)	(0.365)	(0.256)	(0.496)
Observations	2,057	2,057	2,053	2,053	699	699
Panel 3- Overpayment Given Misusage						
Overpayment (in $)	0.157	-1.147***	-2.086***	-3.078***	0.523**	-0.857*
	(0.165)	(0.342)	(0.238)	(0.372)	(0.249)	(0.479)
Observations	2,073	2,073	2,069	2,069	706	706
Panel 4- Overpayment Given Beliefs and Given Misusage						
Overpayment (in $)	0.860***	-0.379	0.015	-1.036***	1.278***	-0.052
	(0.151)	(0.304)	(0.215)	(0.340)	(0.240)	(0.437)
Observations	2,057	2,057	2,053	2,053	699	699
OLS	X		X		X	
Tobit MLE, Censoring of WTP at 0 and 10		X		X		X
% Underpayment Remaining After Controlling For Overconfidence	NA	39%	18%	41%	NA	28%
% Underpayment Remaining After Controlling For Misusage	NA	55%	65%	72%	NA	30%
% Underpayment Remaining After Controlling For Overconf & Misusage	NA	18%	NA	24%	NA	2%

Notes: The table reports regressions of overpayment for information in dollars on a constant, where overpayment is actual payment minus the different optimal payments discussed in the text. A positive estimate reflects overpayment and a negative estimate reflects underpayment. An observation is a subject-task. Standard errors clustered by "subject" in parentheses. Individuals with data for two different rounds are counted as two different subjects for clustering purposes. There are 149 clusters for columns 1-4 and 89 clusters in columns 5-6. All regressions are done under the assumption of risk neutrality. In columns marked "Tobit MLE," I perform censored normal regressions (using "cnreg" in Stata) that generalize the basic tobit model by allowing the censoring point to vary by observation. Since actual WTP has censoring points of 0 and 10, overpayment will have censoring points at $-b$ and $10 - b$ where b is optimal WTP. * significant at 10%; ** significant at 5%; *** significant at 1%