The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself

By DAVID EIL AND JUSTIN M. RAO*

We study processing and acquisition of objective information regarding qualities that people care about, intelligence and beauty. Subjects receiving negative feedback did not respect the strength of these signals, were far less predictable in their updating behavior and exhibited an aversion to new information. In response to good news, inference conformed more closely to Bayes Rule, both in accuracy and precision. Signal direction did not affect updating or acquisition in our neutral control. Unlike past work, our design varied direction and agreement with priors independently. The results indicate that confirmation bias is driven by direction; confirmation alone had no effect.

JEL: D8, D03, A12

Keywords: belief formation, payoff-based updating, self-image

Belief formation is an important aspect of making rational decisions when facing uncertainty. Standard economic theory assumes that decision makers assess the probability of each event with unbiased beliefs formed according to Bayes' Rule. Yet studies have shown that experimental subjects deviate significantly and systematically from Bayes' Rule in their updating behavior (Amos Tversky and Daniel Kahneman 1974, David M. Grether 1980). These experiments typically use underlying quantities whose values do not affect utility directly, such as the color of balls in an urn.

This design ignores an important aspect of the phenomenon, namely the interaction between preferences and beliefs. When judgment concerns a quality of intrinsic importance to the individual, such as his own intelligence, excessive confidence appears to be the norm (Mark D. Alicke and Olesya Govorun 2005, Don A. Moore and Paul J. Healy 2008, Brad M. Barber and Terrance Odean 2001, K. Patricia Cross 2006). The interaction between preferences and inference mistakes has economic consequences. The most important decisions are precisely the ones for which agents care most about the underlying state. Prominent examples include decisions concerning human capital formation, consumption/savings decisions, market entry and exit, asset trading and mate selection.

In the standard model, beliefs matter only through their instrumental value in decision making. Recent theoretical work has relaxed this assumption to give beliefs *direct* importance (Roland Benabou and Jean Tirole 2002, Olivier Compte and Andrew Postlewaite 2004, Markus K. Brunnermeier and Jonathan A. Parker 2005, Botond Köszegi 2006, Guy Mayraz 2009). Although the

* Eil: deil@ucsd.edu, UC San Diego, 9500 Gilman Dr. La Jolla, CA 92093. Rao: jmrao@yahoo-inc.com, Yahoo! Research Labs, 4401 Great America Pkwy Santa Clara, CA 95054. We would like to thank Vincent Crawford and James Andreoni for continued guidance, the Russell Sage Foundation Behavioral Economics Small Grant Program for financial support, Nageeb Ali, Uri Gneezy, Jane Holding, Craig McKenzie, Nikos Nikiforakis, M. Vikram Rao, Michael Schwarz, Joel Sobel and Charles Sprenger for helpful comments and Jessica Huffman for research assistance.

mechanics of each model differ, all explicitly make beliefs (or an updating rule) a choice variable and establish a duality between "rule chooser" and "rule user." Bayesian inference maximizes the instrumental (or indirect) value of beliefs, ignoring any direct benefit. Moving from this point towards optimism involves a second order loss in utility from belief accuracy but a first order gain in direct belief utility. Preference maximization therefore entails a movement away from the Bayesian benchmark towards optimistic updating.

While there is a large experimental literature on information processing, the papers typically fall into one of two categories. Neither is ideally suited to support or reject these models. Papers in the first category use an intentionally neutral quantity.[1] An attractive feature of this protocol is that signals are objective and priors are pinned down by the experimenter, which makes the Bayesian benchmark readily computable. A limitation is that the findings are difficult to apply to an environment in which beliefs have direct importance. These are the situations which often matter the most to the decision maker. Papers in the second category use more realistic underlying quantities, such as personal skill (Gifford W. Bradley 1978), opinions on important issues (Charles Lord, Lee Ross and Mark R. Lepper 1979) and moral judgment (Linda Babcock and George Loewenstein 1997), but employ subjective information structures such as written articles or fictional back-stories. Findings from this literature include confirmatory bias (the tendency to place more weight on signals that confirm one's prior and underweight disconfirming signals) and self-serving bias (conflating self-interest and moral judgment). Using subjective information structures makes it difficult to develop a normative standard for comparison and thereby limits our understanding of the underlying processes at work.[2] It is also difficult to pinpoint the mechanisms through which these biases develop. With our clear and objective information structure, we show that bias exists even when subjects cannot "forget" or "misinterpret" signals. Subjects could observe the full signal history at all times, and the signal generating process was explicit and objective. The persistence of bias in this setting has important policy implications, as giving individuals better tools to retain information, or better understanding of how this information is generated, may not bring their beliefs in line with reality.

This paper bridges the gap between the two protocols by trying to retain many of the attractive features of each. We study how people incorporate objective pieces of new information into their existing beliefs when the underlying quantity has direct belief utility. In this environment signals have intrinsic valence — news is good or bad. The qualities used to generate signal valence were intelligence (*IQ*), as measured by score on an IQ test, and physical attractiveness (*Beauty*), as rated by subjects of the opposite sex. Subjects were ranked one through ten (within gender group for *Beauty*) on a given quality; they updated their beliefs on their position in the ranking in response to objective signals. As a control, subjects also updated their beliefs on a unique randomly assigned integer ranging from one to ten. All subjects participated in one image task (either *IQ* or *Beauty*) and the control task in an experimentally balanced order.

The experiment proceeded as follows. In part one, subjects were ranked for each of their two updating tasks. In part two, subjects performed an information processing exercise for each

[1]This literature started with the broad finding that people update like conservative (or "anchored") Bayesians. Since Kahneman and Tversky (1972) a catalog of systematic and dramatic departures from Bayesian has emerged. Examples include base rate neglect (Grether 1980),the "law of small numbers", and the "representativeness heuristic" (Tversky and Kahneman 1974).

[2]For a recent discussion, see Benoit and Dubra (2007).

task separately and sequentially. At the beginning of each information processing exercise in part two, and before receiving any formal feedback, subjects revealed the distribution of their prior beliefs.[3] Next, they received a series of three signals. After each signal, they updated their beliefs. Each signal was a truthful pair-wise comparison of their rank (out of ten subjects in their session) to a randomly selected anonymous subject in the same session. The signals informed the subject that either they were ranked higher (good news) or lower (bad news) than the comparison subject. Signals were drawn without replacement. Given the priors elicited before any signaling, the objectivity of the signals allowed us to compute the Bayesian response. After completing the updating exercises, we elicited the subjects' willingness-to-pay (WTP) to learn their true rank (on each task) in an incentive compatible manner.

Our primary finding is that subjects incorporated favorable news into their existing beliefs in a fundamentally different manner than unfavorable news. In response to favorable news, subjects tended to respect signal strength and adhered quite closely to the Bayesian benchmark, albeit with an optimistic bias. In contrast, subjects discounted or ignored signal strength in processing unfavorable news leading to noisy posterior beliefs that were nearly uncorrelated with Bayesian inference. This asymmetry was not observed in the control. We call this finding the good news-bad news effect.

The result suggests that bad news has an inherent "sting" that differential processing mitigates. The natural companion hypothesis is that this sting induces an aversion to acquire new information when negative feedback is expected (Köszegi 2006). Consistent with this hypothesis, our second finding is that subjects who believed themselves to be near the top of the distribution in the image tasks were willing to sacrifice some of their earnings to learn their true rank. Subjects who believed they were towards the bottom required a subsidy. In the control, WTP was not affected by beliefs or signal history.

Our experimental findings not only lend support to theoretical work in payoff-based updating but also provide insight into previously identified departures from Bayesian rationality in observed beliefs. These phenomena include bargaining failures associated with self-serving bias (Babcock and Loewenstein 1997) and overconfidence (see Moore and Healy (2008) for a review). Previous explanations fall broadly into two categories: 1) biased processing (Lord, Ross and Lepper 1979, Matthew Rabin and Joel L. Schrag 1999, Bradley 1978, Grether 1980) 2) endogenous recall and acquisition (George A. Akerlof and William T. Dickens 1982, Walter Mischel 1976, Benabou and Tirole 2002). Our results show that valence operates through both channels.

The good news-bad news effect can also be used to explain confirmation bias (CB).[4] In a typical CB experiment, subjects are pre-screened for having strong opinions on a particular issue and then provided with subjective information, such as journal articles, designed to further inform their opinion. A robust finding is that when given the same evidence, both sides of the debate become more convinced of their prior opinions (Lord, Ross and Lepper 1979, Michael J. Mahoney 1977, Scott Plous 1991, Raymond Nickerson 1998). Note that the underlying quantity is not directly

[3] They had received informal feedback by looking around the room, which presumably was informative in *Beauty*. Subjects were chosen for participation randomly through an online database, so it is unlikely they knew each other given the size of the database (4,000+ subjects). As such, looking around the room was probably not very informative in *IQ*.

[4] The standard definition of CB is that agents underweight disconfirming evidence and overweight confirming evidence. Rabin and Schrag (1999) presents a theoretical model of CB that uses this definition as a starting point.

tied to self-esteem, but information still has valence. For example, whether or not the death penalty prevents violent crime does not make one feel good or bad *intrinsically* (unless one is on death row). Rather, the position is important due to the ego utility consequences of "being right." The critical design feature is that confirming signals are always good news and disconfirming signals are always bad news — confirmation and valence are perfectly co-linear.

Our experiment was designed to capture the full signal space; confirming and disconfirming signals can be either good or bad news. Subjects in our experiment did not display CB. For example, a subject who reported believing she had below average beauty tended to put more weight on a signal that informed her she was more attractive than a randomly selected comparison subject (disconfirming good news) and less weight on a negative comparison (confirming bad news). Neither of these types of signals were possible in previous CB experiments since good news was perfectly correlated with priors, leaving CB as the only hypothesis to be rejected. Our results indicate that agreement with prior beliefs is not the underlying cause of CB. We should only expect to observe CB when disconfirming (confirming) signals indeed have negative (positive) valence. In the discussion section we show how the valence-based explanation leads to new policy implications.

We link the good news-bad news effect to a host of economic phenomena that are grounded in non-Bayesian information processing and acquisition. These include decision inaccuracy associated with overconfidence, post-earnings announcement drift in asset prices and pricing bubbles. The application to overconfidence is straightforward. Since strong favorable signals induce a large change in beliefs and strong unfavorable signals are heavily discounted, beliefs tend to err towards the top of the distribution for qualities linked to belief-based utility.[5] Post-earnings announcement drift in small-cap stocks has been shown to be more pronounced for negative, as compared to positive, earnings surprises (Victor Bernard and Jacob Thomas 1989). This asymmetry is consistent with the information processing patterns we observe experimentally. The application to bubbles is germane to the recent financial crisis. While it is well known that rational bubbles can form in macroeconomic models, very optimistic beliefs about future growth in the asset are required to maintain the bubble (Roger Farmer 1999). At a position of extreme optimism, bad news is very informative because current beliefs render negative signals highly unlikely. The good news-bad news effect shows that in this circumstance inference is most biased. Given such a disposition, agents are apt to ignore warning signals and bubbles are more likely to persist than predicted by fully rational models.[6]

The remainder of the paper proceeds as follows. Section 1 details the experimental design, Section II provides results, a discussion follows in Section III and Section IV concludes.

I. Experimental Design

We conducted the experiment at the University of California San Diego Economics Laboratory. There were 7 *IQ* and 4 *Beauty* sessions. *IQ* sessions had 10 subjects each. *Beauty* sessions had 10

[5]Furthermore, the pattern of information acquisition we observe can further fuel overconfidence as people likely to receive good news are the ones seeking out more precise signals.

[6]A recent paper by Chauvin, Laibson and Mollerstrom (2009) shows that the economic cost of bubbles increases quadratically with size. Small bubbles have mild effects but very large ones have can have drastic consequences. The authors argue, *inter alia*, that large bubbles are fueled by market participants ignoring warning signs.

male subjects and 10 females subjects. No subject participated in both *IQ* and *Beauty*. Potential participants were solicited through an online subject database. Prior to entering the lab, they were told only that the experiment would last 1.5 hours and earnings would be $25 on average. Given the sensitive nature of the experiment (receiving information on intelligence and physical attractiveness), we over-subscribed the sessions to allow for the possibility of subjects electing not to participate after reading the IRB consent form.[7] Across all sessions, only 2 subjects elected to leave after learning the nature of the experiment. As such, we randomly selected subjects out of the experiment until the required number remained.

The experiment proceeded in three stages. The first stage of both session types collected the necessary information to rank the subjects on the intelligence or attractiveness. At this juncture subjects, we did not inform subjects why this information was being collected. In *IQ*, subjects took a 25 question IQ test. The questions were selected from a standard Wechsler Adult Intelligence Scale test and involved logic, spatial and verbal reasoning, and general knowledge. Subjects were told the "aims of the experiment depend on you making an honest effort on this IQ test." To incentivize them further, one question was chosen at random ex-post. If they answered this "payment question" correctly $5 was added to earnings.

In *Beauty*, subjects engaged in a speed dating exercise.[8] Each person met 5 subjects of the opposite sex and engaged them in a 4 minute conversation. These meetings were face-to-face with a partition separating each pair. The only restriction placed on the conversations was that they could not reveal identifying information such as their full name or place of residence. To help break the ice, conversation topics were suggested. During the meetings soft music was played to mimic a real speed dating environment. Most subjects engaged in lively conversation and seemed to enjoy the exercise.

After each meeting, subjects filled out a "speed dating questionnaire" rating their conversation partner (scale 1-10) on three dimensions: friendliness, attractiveness and ambition. These forms were kept in a manila envelope during the meetings and were filled out at separate partitioned work stations to maintain anonymity and encourage honest assessment. Ambition and friendliness were included to reduce the anxiety of filling out the questionnaire. The average of the 5 physical attractiveness ratings (one from each "date") was used to give each subject a beauty score. After the speed dating exercise, these sessions proceeded as two separate 10 person sessions of each gender.

While the experimenter read stage 2 instructions to the subject, the scores for the image task in stage 1 were tabulated. The scores ranked the subjects "1" (highest score) through "10" (lowest). Rankings were within gender group for *Beauty*. Ties were broken by a flip of a coin. Subjects were informed of the ranking procedure and understood that the scores gave a strict ordering.

At the beginning of stage 2, subjects were passed a sealed envelope containing a card with a unique number 1-10 written on it. In *Beauty* sessions the numbers were unique within gender group. The number determined rank on the "card task" (*Control*). The subjects were truthfully told that card rank was determined randomly.

[7] Upon entering the lab subjects were told to read carefully the consent form which told them that they would receive feedback on either their intelligence or physical attractiveness. At that point they were given an opportunity to exit discreetly if they no longer wanted to participate.

[8] The design of the speed dating exercise is similar to Fisman et al. (2006) which used speed dating ratings to examine gender differences in mate selection.

At this point we described to the subjects the updating task that they would be performing. Stage 2 consisted of 2 sets (image and control) of 4 rounds each, enumerated 0 through 3. In round 0, subjects entered their prior belief in percentage terms that they occupied each of the 10 ranks on the task. This was done through a computer interface as seen below in Figure 1.

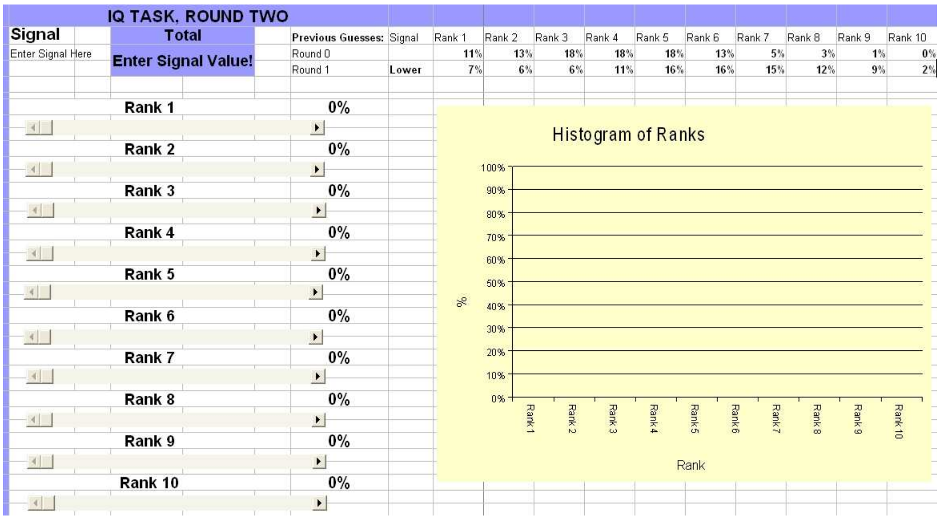

FIGURE 1. INSTRUMENT SCREEN-SHOT.

The percentages always started at 0 for each rank. A graphical representation was provided through an auto-adjusting histogram. The total was displayed at the top of the screen. Beliefs not adding to 100% were not eligible for payment. Subjects made adjustments by clicking the arrows or moving the slider-boxes.[9]

Following rounds 0, 1 and 2, each subject's rank was compared to a randomly selected comparison participant (again these were done within gender group for *Beauty* sessions). This comparison participant was anonymous. Comparisons were bi-lateral and unique each round (drawn without replacement). The result of the comparison was conveyed via a message that read either "You are ranked higher" or "You are ranked lower." The message cards were labeled by round and subject ID letter. As the messages were handed out, subjects were reminded each time that "higher" meant closer to the top rank of 1 and "lower" meant closer to the bottom rank of 10. After receiving the message, subjects entered it into their spreadsheet to ensure the message was accurately received. They then entered their new guesses over their rank on the task. The entire history of rank guesses and signals were provided at the top of the screen. The order of the image and control sets was randomized.

Payment for rank guesses was done through the incentive compatible quadratic scoring rule (Reinhard Selten 1998). Following Moore and Healy (2008) subjects were shown the scoring rule formula and told, "While this payoff formula may look complicated, what it means for you

[9]Subjects also provided the percentage chance they thought they would receive a "higher" signal in the following round. However subjects did not complete this task with high enough frequency for this data to be useful.

is simple: you get paid the most on average when you honestly report your best guesses of the probability for each rank. The range of payoff is 0-2 dollars for each round of guesses. The formula rewards accuracy in a way that the way to maximize your average winnings is to report honestly."

Following completion of stage 2, subjects were given an opportunity to learn their true rank on both the control and image tasks. This was done separately for each task through a price list equivalent to a discretized version of the Becker-DeGroot-Marschak (BDM) mechanism (Gordon Becker, Morris DeGroot and Jacob Marschak 1964). This allowed us to capture subjects' willingness to pay for complete information. Each item of the price list had a choice of the following form:

> Do you prefer (A) Receiving \$*x* and learning your true rank on the [IQ/Beauty or Card] task or (B) Receiving \$0 and not learning your true rank?

x ranged from -7.00 to 7.00 in \$1.00 increments. When $x < 0$ the language was changed to "paying $-x$" (e.g. if $x = -3$ it read "paying \$3.00"). There were 15 choices for each task. A random number was drawn for each task to determine which decision would be executed. Subjects were told they should answer each question as if it were going to be executed.[10] If the executed choice led to the subject learning her true rank on a given task, then she was given an "acknowledgment form" on which she initialed by her rank indicating she indeed learned the truth. In *Control*, subjects also got to open the envelope. To ensure choice anonymity, if the executed choice did not lead to a subject learning his true rank then he received a blank acknowledgment form.

Each session lasted about 1.5 hours and subject earnings averaged \$23.33. Full instructions can be found in the appendix.

II. Results

The purpose of the design was to have image and control treatments with objective signals in an identical information structure. This allows for an easy comparison of the patterns observed in the neutral *Control* to *Beauty* and *IQ*. Objective signals have the attractive property that the results of the experiment cannot be explained by subjects simply "inverting" some unfavorable signals, or assuming some different data-generating process for negative signals. The objective signals also allow us to calculate closed form Bayesian posteriors for rounds 1-3, conditional on beliefs in round 0, to establish the normative benchmark. The only assumption required is the truthful reporting of priors in round 0. Even if subjects have higher order beliefs, or "priors over priors," later round beliefs are still measurable with respect to the product topology over round 0 beliefs and the set of signals.[11] It is possible that subjects might strategically understate their priors so as not to be too often "disappointed" by bad news later. If this were the case, what we view as asymmetry in signal response would be a result solely of pessimistic priors. However in

[10]Note that this procedure is equivalent to the BDM but does not require the lengthy and confusing explanation of why subjects have an incentive to report their true willingness to pay.

[11]We discuss this in more detail and show it formally in the Appendix.

the data we find the opposite is the case. We find asymmetric updating in the treatments in which we find the most optimistic beliefs.[12]

Ertac (2006) and Möbius et al. (2010) employ similar designs but with coarser rankings. Ertac elicits subject beliefs over which third of the distribution they occupy, whereas Möbius et al. elicit beliefs over which half of the distribution. This is attractive in the simplicity of beliefs, but no longer allows for computation of a unique Bayesian posterior when the signal is a pairwise comparison, since the likelihood of receiving a ”higher” or ”lower” signal also depends on the distribution of beliefs within each quantile. Both papers therefore use different signal structures than ours. Ertac's signals fully reveal whether or not the subject is in the top third. Möbius et al. use noisy signals of whether or not the subject is in the top half, accompanied by a cover story to make the signals feel more intuitive. Ertac also uses a control, as we do. Möbius et al. do not, instead comparing their data to the Bayesian benchmark. Neither studies search behavior. Möbius et. al finds general conservatism and asymmetry in updating, with stronger reactions to good news, as we do. Their study is generally in agreement with ours. Ertac finds the opposite effect in that bad news is weighted more heavily, however the differences in signal space make it hard to compare the results directly to ours.[13]

There is some worry about the incentive compatibility of the quadratic scoring rule used to determine payoffs to accuracy of beliefs. While it is true that this elicitation mechanism is incentive compatible only for a risk-neutral agent, we believe that the subjects do report their true beliefs. First, the range of payments for each individual round of beliefs is from \$0 to \$2, stakes over which we would expect subjects to be risk-neutral. Second, while subjects are shown the exact formula, the principal thing they are asked to keep in mind is that they make the most on average by reporting their true beliefs. Third, Ertac (2006) shows that when told explicitly the probabilities of various states, subjects report these beliefs back accurately when their responses are paid according to the quadratic scoring rule. Finally and most crucially, the results we focus on here rely on asymmetry in how beliefs change from one round to the next in reaction to positive and negative signals. Any tendency to hedge beliefs due to risk aversion would affect good and bad news equally, and therefore have no bearing on the difference in reaction between the two.

The remainder of this section has three parts. We first present the analysis of the posterior beliefs of the subjects receiving the most "extreme" feedback – all positive or all negative at the time of updating. We then examine the round-to-round changes of beliefs for all subjects in greater detail. Finally, we provide further evidence that beliefs enter into preferences directly by showing how they affect a subject's WTP for full revelation of his rank. Seven of the one-hundred fifty subjects were eliminated from all analysis because they clearly displayed a lack of

[12]One way to measure optimism is the mean of subjects' round 0 beliefs, with 5.5 representing unbiased aggregate beliefs. For *Beauty*, this was 4.26, for *IQ*, 5.24, and for *Control*, 5.36. In other words, the more optimistic subject beliefs, the more asymmetric their updating. This follows a pattern we might expect from asymmetric updating in the field that results in optimistic priors coming into the lab, not a pattern of asymmetric updating in the lab that results from pessimistic reporting of priors.

[13]In Ertac's experiment, signals consist of telling the subjects whether or not they are in the highest group. Ertac's result may seem contradictory to ours at first. However all non-revealing signals in her experiment are uninformative on the relative likelihoods of the lower two ranks. Our finding is that subjects do not respect signal strength for negative signals. This means that they can overweight relatively uninformative negative signals, such as the ones in the Ertac experiment. Therefore we do not consider the results of our experiment and hers to be necessarily inconsistent.

understanding of the experimental protocol.[14]

To clarify exposition, we label "up" messages $s = 1$ and "down" messages $s = 0$. Recall, $s = 1$ indicates the subject's rank is closer to 1 than the comparison participant. In *Beauty* Rank 1 represents the subject whose physical attractiveness rated highest; in *IQ* Rank 1 represents the subject with the highest IQ test score. In *Control* Rank 1 represents the subject whose random "card number" (1-10) is "1."

A. *Posterior beliefs*

The purpose of this section is to examine how subjects incorporated favorable ($s = 1$) and unfavorable ($s = 0$) news into posterior beliefs. Our hypothesis, motivated by the theoretical and empirical work discussed in the introduction, is that for the tasks associated with belief-based utility (*Beauty* and *IQ*), subjects will respond more to favorable news and less to unfavorable news. No difference is expected in the control. For each subject, we calculated their posteriors according to Bayes rule, given the priors they reported in round 0 and the signals observed through that round. We then took the expected rank under this distribution and called this value the "mean Bayesian belief." We compare subjects who have received all good news to subjects who have received all bad news. This comparison retains all observations from round 1 (when the subjects have only received 1 signal) and eliminates some observations from rounds 2 and 3 in order to get the tightest comparison possible.[15] The eliminated observations are analyzed in the following subsection and tell the same basic story.

To begin our analysis, we regress the subjects' observed mean belief on the mean Bayesian belief, grouping observations by whether good or bad news was received.[16] Generalized optimism would be represented by slopes equal to one and negative y-intercepts. This would mean that for every Bayesian posterior, the actual posterior is lower (closer to 1, the "best" rank). A difference in slopes for $s = 1$ and $s = 0$ means that subjects weight signal strength differently depending on if it is a positive or negative signal. A flatter slope means that posteriors are "squished" relative to the Bayesian – posteriors that should be more extreme by Bayes Rule are instead grouped closer

[14]The first reason for elimination was not having beliefs that added up to 100%. The second reason group of subjects eliminated appeared to interpret "up" as meaning closer to 10 (as opposed to closer to 1) but then at some point realized their mistake. This was a function of the somewhat confusing feature of the English language in that "higher" means better in terms of rank in in a distribution but means the opposite in terms of the absolute number. 2 subjects had a computer failure which led to partial data loss.

[15]To be precise, this eliminates the round 2 observations for subjects whose round 2 signal was not the same as their round 1 signal (35% in each of the image treatments, 40% in the control), and, of the remaining subjects, the round 3 observations for subjects whose round 3 signal was not the same as their rounds 1 and 2 signals. 52% of all *Beauty*, 49% of all *IQ*, and 49% of all *Control* subjects received either all good or all bad news and therefore are included for all three rounds in this analysis. In all cases, given uniform signals, the split between all good news and all bad news is approximately even.

[16]Since we have split the data here by all good news and all bad news, this represents an inherently between-subjects design. Therefore introducing any subject-specific effects would not allow us to identify our treatment effect. We maintain this model throughout our regression, since with only three rounds of updating, there is little room for within-subject variation of signal valence. We have run regressions using only subjects with mixed signals. The signs of the coefficients are all the same, just with reduced significance due to smaller sample size. The concern may be that subjects who are better Bayesian updaters are the ones more likely to get good news. However such an explanation would seem to indicate a sharper asymmetry in the *IQ* treatment, as IQ scores should correlate better with Bayesian-ness than physical attractiveness. Instead we find that updating is more asymmetric in *Beauty*.

together. This indicates insensitivity to signal strength. Steeper slopes, by contrast, indicate an exaggerated response to signal strength.

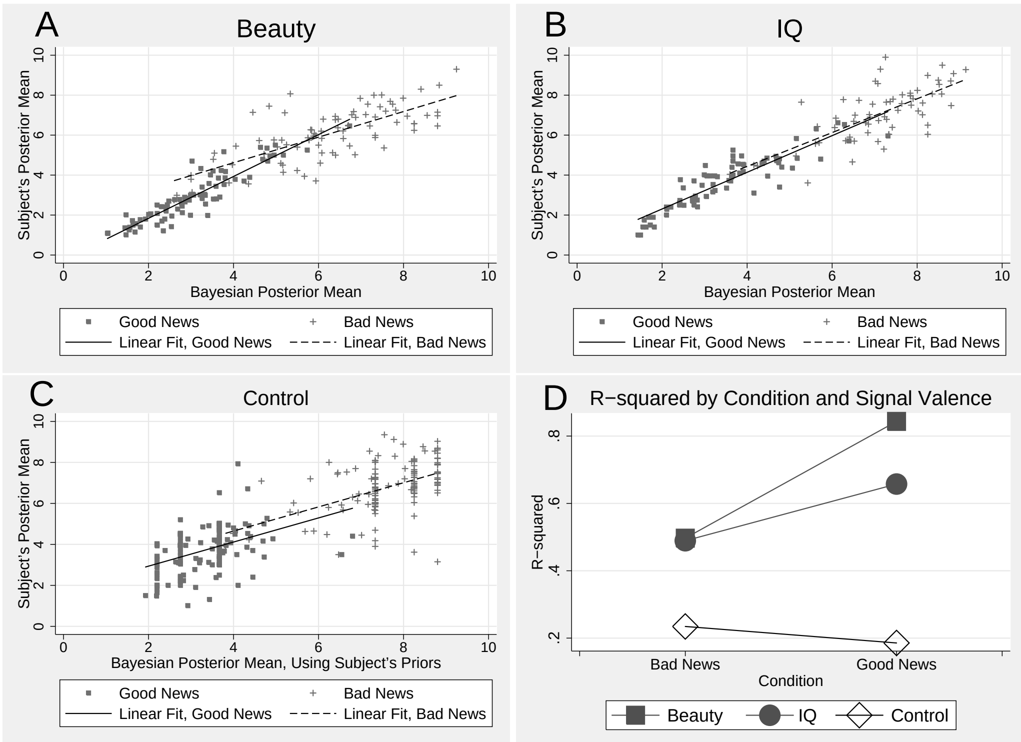

FIGURE 2. BAYESIAN VS. SUBJECTS' POSTERIOR BELIEFS BY SIGNAL VALENCE AND CONDITION

Panels A, B and C of Figure 2 plot subjects' observed mean belief by the mean Bayesian belief for *Beauty*, *IQ* and *Control* respectively. The data is split by signal valence. The fitted lines are from the OLS regressions presented in Table 1. We clearly see asymmetry in the steeper fitted lines for good news as compared to bad news in the image conditions, *Beauty* in particular. This pattern is also indicated by the significant coefficient on the interaction term in Table 1 demonstrating that subjects respect signal strength more for favorable signals than for unfavorable signals. Note that one effect of this is that very few subjects end up with very pessimistic beliefs. The constant slope across signal valence for the control indicates that this asymmetry is driven by belief-based utility inherent in the image treatments.

We see that the slope is indeed constant in *Control* but that it is steeper for good news in both image conditions. The effect is more pronounced in *Beauty*. Table 1 shows that the difference is highly significant, as indicated by the coefficient on the interaction term. Subjects in *Beauty* engaged in asymmetric updating as compared to the control. The difference in slopes across signals for *IQ* is not statistically significant. However we see in panel D that subjects did differentially process information in this condition. Subjects in both image treatments responded more predictably to good news. Their response to bad news was noisier. The following subsection confirms this.

TABLE 1—SUBJECT'S MEAN BELIEF AS A FUNCTION OF BAYESIAN MEAN

	Beauty	IQ	Control
Bayesian μ	0.641***	0.846***	0.589***
	(0.081)	(0.102)	(0.099)
Bayesian μ*1$\{$All $s = 1\}$	0.406***	0.0714	-0.00118
	(0.096)	(0.129)	(0.192)
1$\{$All $s = 1\}$	-2.311***	-0.580	-0.531
	(0.545)	(0.811)	(0.946)
Constant	2.054***	1.051	2.291***
	(0.525)	(0.757)	(0.760)
Observations	163	139	292
R^2	0.867	0.869	0.717

Standard errors clustered at subject level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panels A and B provide visual evidence that beliefs are far more noisy following bad news in the image treatments. No such patterns emerge in panel C (*Control*) – the intercept shift displayed in Panel C is well within a single standard error of 0. Panel D presents the R^2's from the regressions in Table 1, separated by signal valence (i.e. two per condition). The differences in R-squared are statistically significant (as given by a variance ratio test) for *Beauty* ($p = 0.0000$) and *IQ* ($p = 0.0000$) but not for *Control* ($p = 0.6139$). R^2 was 58% higher for good news in *IQ* and 60% higher for good news in *Beauty*. For the image treatments, subjects' beliefs adhered far more tightly to Bayesian rationality when the news was favorable than when it was unfavorable. This is seen through a slope coefficient closer to 1 and much higher R^2. The control shows that this was not simply an artifact of the experimental design unrelated to belief-utility.

B. *Round-to-round changes in beliefs*

The preceding subsection examined how the levels of posterior beliefs compared to the Bayesian predictions. This gives indicative evidence of updating behavior, since the subject's prior informs his own posterior as well as the Bayesian posterior. In this subsection we examine more directly how beliefs change from one round to the next upon the arrival of new information. Here we will use data from all subjects, not just those who received all positive or all negative signals. Recall a negative change is a change towards 1, the highest rank.

Figure 3 presents linear fits of the round-to-round changes in mean beliefs as a function of the Bayesian change. That is, $\Delta\mu_t$, the change in mean subject beliefs for round t, is given by the mean beliefs in round t minus the mean beliefs in round $t-1$. The Bayesian change in round t, $\Delta\mu_{Bayes,t}$, is calculated as the Bayesian posterior in round t given the round 0 prior and the t signals received, minus the Bayesian posterior in round $t-1$, again given the round 0 prior and the $t-1$ signals then received. [17] The associated regression coefficients are in Table 2. [18] Across all conditions and signals, there is a pattern of over-responding to uninformative signals, i.e. the y-intercepts are not zero. We discuss the importance of this result in the following section.

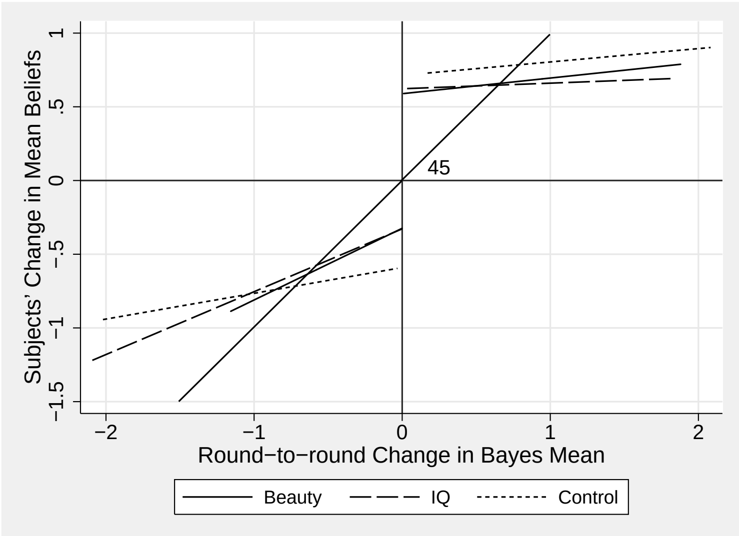

FIGURE 3. LINEAR FITTED VALUES OF ROUND-TO-ROUND CHANGES IN MEAN BELIEF BY CONDITION.

The 45 degree line is given for comparison to the Bayesian response. For *Beauty* and *IQ*, when the news is favorable (lower left quadrant) the fitted curve is much closer to the 45 degree line and has significantly steeper slope – signal strength is respected. The same does not hold true for unfavorable news. Table 2 shows that the slope coefficient for unfavorable news is statistically indistinguishable from 0. Informative down signals are treated in a similar fashion to uninformative ones; beliefs are adjusted down by the same amount, no matter the informativeness of the signal.[19] In the introduction we called this pattern the good news-bad news effect. The effect shows up in both the posterior beliefs and the round-to-round changes. For the image tasks,

[17] Both Figure 3 and Table 2 limit the analysis to changes in the correct direction. There was an equal number of changes in the wrong direction by signal valence. The inclusion of these outliers actually strengthens our argument as 2 were informative down signals in which the subject actually updated slightly updated towards 1. However, these points have such high leverage in the regression and represent less than 3% of the sample. As such they are excluded. Median regression is used in Table 2 to further reduce the influence of outliers; the regressions us bootstrapped standard errors.

[18] The inclusion of higher order terms does not significantly improve fit. We therefore proceed with linear analysis for simplicity.

[19] Our results would be considerably strengthened if we applied a kernel smoothing to prior beliefs. This is because Bayesian updating cannot put any weight on ranks that have prior probability equal to 0. For bad news, this tends to lead to an understatement of the optimistic updating because it was fairly common for a subject to have beliefs 50-40-10 on ranks 1, 2 and 3 respectively. The Bayesian posteriors have support of ranks 1-3, as such the mean does not move very much. However, it was exceedingly rare to see such concentrated (and likely mistaken) beliefs on the low end of the distribution. A smoothing procedure which transformed 50-40-10 to, for instance, 45-35-8-4-3-2-1-1-1 would lead to a much larger required downward change. In reporting beliefs, subjects generally left the possibility open that they were at the top of the distribution, so the smoothing procedure would not affect the analysis of up signals significantly.

signal valence fundamentally affects how new information is incorporated into beliefs. There do appear to be similar patterns in *Control*, albeit of a much lower magnitude and lacking statistical significance.

TABLE 2—MEAN BELIEF CHANGES BY SIGNAL DIRECTION AND CONDITION, DEPENDENT VARIABLE: $\Delta\mu$

Condition	Beauty	IQ	Control
$\Delta\mu_{Bayes}$	0.212	0.0256	0.141
	(0.160)	(0.072)	(0.136)
$\Delta\mu_{Bayes} \times 1\{s=1\}$	0.475*	0.540***	0.141
	(0.263)	(0.131)	(0.172)
$1\{s=1\}$	-0.642***	-0.772***	-1.119***
	(0.150)	(0.101)	(0.172)
Constant	0.437***	0.599***	0.625***
	(0.134)	(0.061)	(0.168)
Observations	206	183	385
R^2	0.49	0.55	0.48

Bootstrapped standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In the preceding subsection we showed that good news led to a much tighter adherence to Bayesian rationality than bad news. Figure 4 compares the standard deviation of the residuals for the regressions in Table 2, by signal valence. We can see the same pattern emerges. Updating after down signals (s=0) is far more noisy in *Beauty* and *IQ*, but the essentially the same in *Control*. A variance ratio test gives $p = 0.0000$ for *Beauty*, $p = 0.085$ for *IQ* and $p = 0.347$ for *Control*. Once again the "belief-based utility effects" in *Beauty* appear stronger than in *IQ* and are virtually non-existent in *Control*. It is worth noting that since social comparisons were used in *Beauty* and absolute comparisons in *IQ*, we cannot conclude that physical attractiveness has a stronger image association than intelligence. Furthermore, our subject pool was composed of students at a top tier university; presumably they receive much more formal feedback on their intelligence than their beauty.[20]

C. *Information acquisition*

Our results in the preceding two subsections showed that subjects dealt with the unwelcomeness of bad news by selectively filtering negative valence signals. Our hypothesis is that this is due to beliefs entering the utility function directly, thereby creating an incentive for self-deception, operationalized through asymmetric updating rules. In this subsection we examine how subjects' willingness to seek out new information varies with their current beliefs. A signal revealed one's true rank, allowing little room for misinterpretation. At this stage of the experiment, knowing

[20] We thank an anonymous referee for pointing this out to us.

FIGURE 4. NOISE IN ROUND-TO-ROUND UPDATING BY TREATMENT AND SIGNAL TYPE. BARS GIVE +/- 1.96 STANDARD ERRORS.

one's rank had no instrumental value with regards to earnings. This allows us to isolate the demand for new information as function of direct belief utility.[21] Köszegi (2006) presents a model which shows that conditioning information acquisition on beliefs is the optimal strategy in the presence of direct belief utility.

WTP is defined as the switching point on a price list offering a series of choices between two options: A) receiving $\$x$ *and* learning their true rank on the task B) receiving $0 *and not* learning their rank. x ranged from -$7.00 to $7.00 in $1 increments. As such, subjects could require as much as $7 subsidy to learn their true rank and could pay as much $7.

Table 3 examines the relationship between WTP and the mean and standard deviation of their final round beliefs (i.e. after receiving all 3 signals). *Control* allows us to account for curiosity and other artifacts of the experimental setting unrelated to self-image.

Subjects WTP to learn their true rank statistically significantly increased as their round 3 mean belief moved towards 1 for both *IQ* and *Beauty* – for *Control* there was not any effect. In the image conditions, subjects who believed they were below average (rank 6+) on average required a subsidy to learn their true rank. WTP also increased with the standard deviation of the belief distribution in these conditions, indicating that subjects responded to informational value as well (although the point estimates are very similar, the effect is only significant in *IQ*). [22]

Figure 5 relaxes the assumption of linearity and shows that subjects in the middle of the distribution have a flat WTP function. These subjects are generally not willing to pay but do not require a subsidy to learn their true rank. The function is convex for high ranks and concave for

[21]One could argue that the information would be of value outside the experiment, in dating markets for instance. This is a reasonable assertion. However, it can explain our results only if the outside value information is heterogeneous in specific and unintuitive ways.

[22]One might suspect that the informational effect might differ by mean belief. Regressions available from the authors show that there are not significant differences if the interaction terms by mean belief quartile are added.

TABLE 3—WILLINGNESS-TO-PAY AS A FUNCTION OF FINAL ROUND BELIEFS

Condition	IQ	Beauty	Control
Final Round μ	-0.325***	-0.206**	0.0112
	(0.107)	(0.0855)	(0.0423)
Final Round σ	0.911**	0.917	0.237
	(0.382)	(0.575)	(0.174)
Constant	0.729	0.298	-0.339
	(0.523)	(0.750)	(0.336)
Observations	77	65	142
R^2	0.099	0.180	0.015

Robust standard errors in parentheses

*** p<0.01, ** p<0.05

low ranks. The very best subjects had a high WTP and the very worst required a subsidy.[23] We posit that these differences are due to ego utility. An alternative explanation is that the value of the information outside the experiment varies predictably over the rank distribution. However, mating markets and job markets, where intelligence and attractiveness are relevant, generally have positive assortative matches(David Watson, Eva C. Klohnen, Alex Casillas, Ericka N. Simms, Jeffrey Haig and Diane S. Berry 2004, Gunter J. Hitsch, Ali Hortasu and Dan Ariely 2010). In this case belief precision is equally valuable at all belief levels.

What type of ego utility function explains this data? In Köszegi (2006), which agents exhibit information aversion depends on the exact nature of the ego utility function. He assumes a discontinuity in ego utility at the 50th percentile, which would lead slightly above average subjects to have the lowest WTP. We find that the most pessimistic subjects have the lowest WTP, while the most optimistic have the highest WTP. Consider a subject with most of the probability mass of her beliefs in the top 3 ranks. On average this subject has positive WTP to learn her true rank. In order to risk receiving bad news, the subject must believe the potential ego gains outweigh the losses. This is the case if the ego utility function has increasing marginal utility as we move from the middle ranks towards rank 1. This increasing marginal utility generates "risk-lovingness" in information over this range. There also must be increasing marginal disutility moving towards rank 10 from the middle ranks. This results in a concave function over this range; subjects are risk-averse in information and require a compensating subsidy. Since agents in the center of the distribution have a flat WTP function, we infer ego utility is relatively constant in this "average region." [24]

[23]There is some evidence that male subjects have a higher WTP when they have beliefs near 1 as compared to females and females require a larger subsidy when beliefs are near 10. However, the differences are not significant. A larger sample size would be needed to examine gender differences at the tails of the distribution.

[24]An alternative explanation is based in the psychological concept of "affirmation." Under this view, signals enter the utility function directly, independent the impact on beliefs. Imagine an agent who is 99% sure she is rank 2. Suppose despite her near certainty, she has a high WTP to learn the truth. Positive (negative) feedback that has little informational value is known as affirmation (disaffirmation). Affirmation (disaffirmation) seeking (avoidance) has some explanatory

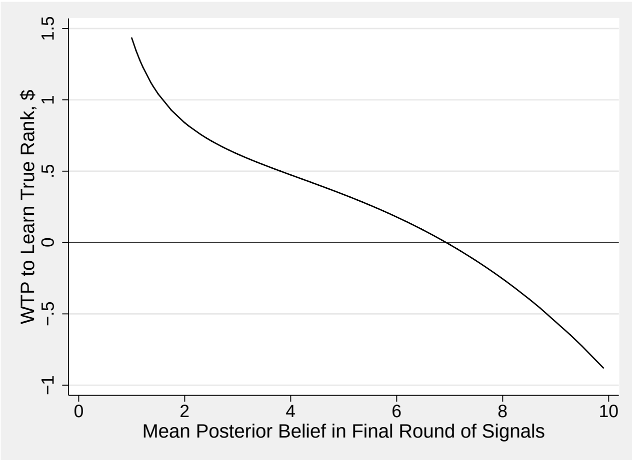

FIGURE 5. FRACTIONAL POLYNOMIAL FIT OF WTP TO LEARN TRUE RANK ON MEAN FINAL ROUND BELIEF IN THE IMAGE TASKS.

The demand for information we observe induces post-realization beliefs that are more precise at the top of distribution, while low types are more likely to have diffuse beliefs. A consequence is that people are not likely to believe with a high degree of confidence that they are in the lower percentiles. No one is certain they are among the worst, but the right amount of people know they are among the best. This pattern of beliefs is not classic overconfidence (mean beliefs above average), but it can have economic consequences for decisions, even those based on only the first moment of beliefs.[25] To see this, consider the following example. Suppose the returns to a college degree are positive only for those in the top 70% of intelligence in the population. If agents below the average have diffuse beliefs with a mean above the 30th percentile, they are likely to enroll, even though many of these agents actually have ability below this point. The matching of types to human capital accumulation will be poor. Another example germane to our study is mate selection. Suppose nobody's mean belief places them in the bottom quartile of mate desirability because they halt the information acquisition process at this point. This halt can potentially disrupt the sorting necessary to reach equilibrium.

III. Discussion

In this section, we first discuss specific economic implications and supportive empirical findings of valence dependent processing and acquisition of information. Next, we highlight key

power in our setting. However, the significant effect of of round 3 belief variance shows that it is not the entire story. This effect shows that subjects do value knowledge more according to the amount of uncertainty it resolves.

[25]The effect on mean overconfidence depends on the signal informativeness, whether priors exhibit over/underconfidence and whether agents update rationally. For instance, if agents are overconfident and signals are perfectly informative, then this search will tend to reduce net overconfidence. If agents are rational and signals are imperfectly informative then this ego utility function will actually induce underconfidence. This results follows from proposition 1 of Köszegi (2006). In section III.C, we discuss how this type of search combined with the inference pattern we observed can lead to overconfidence.

elements of our experimental design which allow us to provide a payoff-based explanation for confirmatory bias. We then illustrate how differential updating and search can complement each other in generating overconfidence. Finally, we discuss the relevant theory that supports our empirical finding and propose avenues of future research.

A. Economic applications

In our experiment signal valence was generated by ranking ten strangers after a short task measuring intelligence or beauty. We chose these qualities and the updating context because in an experimental setting ethical considerations rightly limit "how bad" news can be. In the real-world, news can have much greater significance. In this case signal valence may have an even greater effect on information processing. Here we present empirical evidence consistent with our experimental findings.

Evidence from finance comes from the literature on post-earnings announcement drift (PEAD). Predictable drift in stock prices should not occur in an efficient market, but authors have found that stocks experiencing earnings surprises do predictably drift over a 120 day period following announcement (George Foster, Chris Olsen and Terry Shevlin 1984). Bernard and Thomas (1989) found that PEAD in small cap stocks is significantly more pronounced for bad, as compared to good, news stocks. That is, the initial price response accounts for less of the available information for bad news, which is precisely our primary finding. The asymmetry is only observed for small cap stocks, which are traded primarily by non-institutional investors. This is reassuring for two reasons: 1) it allays concerns that some uncontrolled-for informational asymmetry is driving the results 2) we would not expect institutional investors to respond to belief-based utility. The authors rule out risk-premium-based and other explanations for the observed drift to conclude that PEAD is best explained by incomplete informational response.

The recent housing price bubble provides evidence that people are prone to ignore warning signs that their asset portfolio is priced far above its fundamental value. The existence of a bubble is not a sufficient condition for irrational beliefs; rational agents can fuel bubbles if the expected returns of the "ride to the top" outweigh the expected cost of the burst (Farmer 1999). An agent holding a bubbly asset must be very optimistic about the continued growth; bad news is very informative in this setting because the optimistic beliefs render it highly unlikely. Our experimental findings indicate this is precisely the circumstance where inference is most biased. Informative bad news gets discounted and bubbles persist far longer than they otherwise would have. This has serious economic implications as recent work provides evidence that the net cost of a bubble is quadratically proportional to its size (Kyle Chauvin, David Laibson and Johanna Mollerstrom 2009).

Self-serving bias is the tendency to conflate self-interest with moral judgment. There is strong experimental evidence for self-serving bias (David M. Messick and Keith P. Sentis 1979, Linda Babcock, George Loewenstein, Samuel Issacharoff and Colin Camerer 1995). In a typical experiment subjects are randomly assigned a role in a bargaining game. The role involves a detailed back story. The subject's aim is to negotiate the best deal possible for herself. Despite the fact that roles are randomly assigned, both groups tend to report that they deserve a greater share of the pie *based on the facts of the case*. Self-serving bias has been directly linked to economic and health consequences in the field (Laurie Larwood 1978, Linda Babcock, Xianghong Wang and George Loewenstein 1996). There is also evidence of costly bargaining impasses consistent with self-serving bias (Joseph S. Tracy 1986, Samuel R. Gross and Kent D. Syverud 1991). In bargaining, signals that one will win the case have positive valence – they indicate that a financial windfall is forthcoming or that one did not act wrongly in the past. The good news-bad news effect predicts that these positive signals will receive disproportionate weight as compared to negative signals. Consequently, both sides will exhibit excessive confidence that they are in the right.[26]

The optimism that can result from the good news-bad news effect can be problematic in itself. Many studies have shown that new businesses fail at a very high rate within a few years of their incipience (Timothy Dunne, Mark J. Roberts and Larry Samuelson 1988). Roll (1986) suggests that this is a result of entry mistakes due to overconfidence. Camerer and Lovallo (1999) show in an experimental setting that this is a result of optimistic biases.

An application in medicine is a hospital patient's prediction of his prognosis – here news is literally life-or-death. Studies have found that cancer and HIV+ patients optimistically assess their chances of survival (Jane Weeks, E.Francis Cook, Steven O'Day, Lynn Peterson, Neil Wenger, Douglas Reding, Frank Harrell, Peter Kussin, Neil Dawson, Alfred Connors, Joanne Lynne and Russell Phillips 1998, Richard Eidinger and Schapira Schapira 1984, Jay Bhattacharya, Dana Goldman and Neeraj Sood 2009). It is not that their physicians are misleading or giving them intentionally optimistic forecasts; the physicians' forecasts are shown to be accurate and unbiased. The phenomenon has attracted attention in the medical community because it complicates a physician's ability to apply the treatment she finds most appropriate given the circumstances. An optimistic patient is liable to make decisions that are not in her best interest, typically aggressive and painful treatment. Physicians have adopted strategies to try to alleviate this problem (John Paling 2003).

B. Signal valence and confirmatory bias

The classic account of CB is that people react more strongly to new information that agrees with their prior belief than to information that disagrees with their prior belief.[27] A natural classification in our experiment is that a confirming signal is the one a subject believed was more likely to occur. This means that for a subject whose mean belief places her in the top 5 ranks, above average, $s = 1$ is confirming and $s = 0$ is disconfirming; for a subject whose mean belief places him in the bottom 5 ranks, below average, $s = 1$ is disconfirming and $s = 0$ is confirming.

In Table 4 we look for evidence of CB directly in comparing the subjects' round-to-round change in beliefs to the Bayesian change for confirming and disconfirming signals ($\Delta\mu - \Delta\mu_{Bayes}$). Negative values indicate optimistic updating towards Rank 1.

Examining *Beauty* we see above average subjects did tend to overreact to good news, consistent with CB. Inconsistent with CB, below average subjects receiving $s = 1$ overreacted with

[26]CB has a tough time explaining self-serving bias, especially in controlled experimental settings. With random assignment, it is unreasonable to posit that ex-ante both groups of subjects believe they will be placed in the role that "should" win. Signals cannot be confirming or disconfirming in this setting. Conversely, since winning means higher earnings, news has intrinsic valence.

[27]Sometimes confirmatory bias is used to explain seeking information that is unlikely to be disconfirming (failing to test alternative hypotheses). This search-based CB finds support in our data.

TABLE 4—$\Delta\mu - \Delta\mu_{Bayes}$ AS A FUNCTION OF PRIOR GROUP AND SIGNAL VALENCE

	Good news		Bad news	
	Above avg	Below avg	Above avg	Below avg
CB prediction	–	+	–	+
CB in words	Over respond	Under respond	Under respond	Over respond
Beauty	-0.14	-0.13	0.00	-0.13
	(-3.74)	(-1.20)	(-0.03)	(-1.32)
IQ	0.07	-0.08	-0.14	0.02
	(1.46)	(-0.70)	(-0.91)	(0.23)
Control	0.13	0.21	-0.36	-0.37
	(1.69)	(2.11)	(-5.16)	(-3.89)

t statistics ($H_0 = 0$) in parentheses

commensurate magnitude (0.01 less). They did not ignore or under-respond to the disconfirming good news. CB also predicts that below average subjects would over-respond to $s = 0$. But the opposite occurs — there is an under-response. For IQ, the patterns are similar but of a lower magnitude. What appears to matter most in the image conditions is the signal valence, not confirmation with the prior belief. The patterns in the *Control* do not exhibit CB or valence based-processing, consistent with valence-based story.

Our results indicate that agreement with prior beliefs is not the underlying cause of CB. We should expect to observe CB only when disconfirming (confirming) signals have negative (positive) valence. Linking CB to valence allows one to derive it as a consequence of the maximization of preferences that include belief-based utility. The most cited existing model of CB, Rabin and Schrag (1999), uses a mechanical processing bias as a primitive of the model.[28]The model effectively shows how beliefs evolve in the presence of CB, but leaves the underlying causes of the bias unmodeled.

There are two natural links between CB and belief-based utility: 1) self-image associated with "being right" 2) anticipatory utility. The first link is self-explanatory. To understand the second, suppose an agent has made decisions in the past that pay out in a future period, with payoffs increasing with the ex-post accuracy of beliefs. If she changes her mind today, it means past actions have a lower expected payout, which reduces "anticipatory utility" today. Common examples include investment in a technology by a firm, time devoted to one of many potential students and human capital investment in specific skills. Under both links, disconfirming signals are bad news while confirming signals are good news. In other words, valence and confirmation are naturally co-linear in these cases. In our study, valence and confirmation were not always co-linear; in the cases in which they were not, the ego utility of being right was outweighed by the direct concern for self-esteem. By isolating the two effects, we find that valence appears to be the driving force behind observed CB. In past experiments showing CB, confirming signals were always good news. The perfect co-linearity of confirmation and valence rendered it impossible to determine which aspect of the signal was driving CB.

Our results, coupled with payoff-based updating models that show how biased updating rules

[28]Specifically the assumption is that all signals that agree with one's prior are perceived accurately, but a certain portion of disconfirming signals are mistakenly perceived as supporting one's prior belief.

can emerge from preference maximization, remove some of the mystery associated with CB. Perhaps more importantly, our results lead to new policy implications. Consider the example of a division manager making profit forecasts during an economic recession. In this case, a confirming signal is one that indicates profits will remain low. We interpret our results as suggesting that this manager would respond more in line with Bayesian inference for good news signals, as compared to confirming bad news. In many important cases such as this, confirmation and valence are not positively correlated. The importance of valence suggests that forecasts should be made by (more) neutral observers whenever possible. The explanation is also consistent with not observing CB in *Control* – there is very low utility of being right about a random number. Previous studies finding CB used a very specific underlying quantity, although the literature has not focused on this aspect. The quantity is usually not directly tied to belief utility but indirectly linked through identity (e.g. sides of a hot political debate (Plous 1991, Lord, Ross and Lepper 1979) or beliefs which drove costly effort in the past (Mahoney 1977)). Both cases fit within the two links we outlined earlier.

C. The joint effect of the search and inference results

Biased acquisition and processing can lead to path dependence in beliefs. Path dependence in beliefs is stressed in Rabin and Schrag (1999)'s model of CB. Biased acquisition and inference have a similar effect. The bias induced by the good news-bad news effect alone increases with initial overconfidence because bad news becomes increasingly informative. Our search findings suggest these are the signals likely to be actually received. General overconfidence on image tasks will develop when priors are accurate or optimistic, but it might not develop if beliefs evolve to (or begin at) a sufficiently underconfident state. In the latter case, underconfidence can get locked in as people avoid new information.

D. Payoff-based updating and future work

Theoretical papers have explored the consequences of including beliefs directly in the utility function. These papers allow information processing and acquisition rules to be under the agents control. We refer to this class of models as "payoff-based updating" because optimal beliefs are chosen through the maximization of (non-standard) preferences. Processing models generally adopt some form of duality between "rule chooser" and "rule user." A rule is a function mapping informational states to beliefs.[29] The decision problem involves both a benefit to having accurate beliefs about an underlying state of the world, through the instrumental value of improving choice outcomes, and direct belief utility, through a psychological channel such as self-esteem (Akerlof and Dickens 1982, Benabou and Tirole 2002, Köszegi 2006), self-efficacy (Compte and Postlewaite 2004) or anticipatory utility (Brunnermeier and Parker 2005). The models show that Bayes Rule is a special case; it maximizes utility conferred to the instrumental value of beliefs only. It is suboptimal because it entails a first order loss in utility through the direct belief utility channel. The rule that is chosen shades beliefs in a psychologically beneficial direction.[30]

[29] Mayraz (2009) provides a formal axiomatic model in which beliefs are affected by desires.

[30] A model of clinical depression would invert the sign of the direct belief utility to capture the fact that the illness makes one want to stay depressed. In this case the models would predict that depressed agents' beliefs that are more

The good news-bad news asymmetry we find strongly supports these models. This does not mean such a bias is rational to possess, but it does provide an explanation for its existence in the human psyche. In many cases, it very well might be a fit trait. More work is needed to determine if the bias continues to appear in settings in which it is exploitable and agents have sufficient experience.

IV. Conclusion

We have shown that belief-based utility has important implications for information processing and acquisition. Subjects in our study updated their beliefs on physical attractiveness and intelligence, attributes closely tied to self-esteem. New information came in the form of binary signals and was not only confirming or disconfirming but also intrinsically favorable or unfavorable. We found that good news and bad news were incorporated into prior beliefs in entirely different ways. For good news, the inference was less noisy, respected signal strength and conformed much more closely to Bayesian rationality. Conversely for bad news, beliefs were more noisy and signal strength was not respected, which led to discounting bad news on net. By isolating the effects of valence and confirmation, our results indicate that valence is the underlying cause of confirmatory bias; confirmation alone had no effect. Subjects were also significantly more likely to acquire new information if their past feedback had been positive. Neither valence-based processing nor acquisition was observed in our control condition, which used a neutral underlying quantity.

Many economic decisions involve uncertainty and belief formation. Often these beliefs are over quantities of importance to the individual, such as those affecting self-esteem or future monetary payouts. Theoretical work has contended that in these settings beliefs affect utility directly and not simply through an impact on decision making. Such models predict a movement away from Bayesian inference in the direction of optimism. Our experimental results support these models and point to differential processing by valence as an underlying cause of confirmatory and self-serving bias. Valence-based processing can also help explain a host of economic phenomena such as asymmetric post-earnings announcement drift, market bubbles, overly aggressive medical decision making and overconfidence.

REFERENCES

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *The American Economic Review*, 72(3): 307–319.

Alicke, Mark D., and Olesya Govorun. 2005. "The Better-than-Average Effect." In *The Self in Social Judgment*. , ed. Mark Alicke, David Dunning and Joachim Krueger. Psychology Press.

Babcock, Linda, and George Loewenstein. 1997. "Explaining Bargaining Impasse: The Role of Self-serving Biases." *The Journal of Economic Perspectives*, 109–126.

Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer. 1995. "Biased Judgments of Fairness in Bargaining." *The American Economic Review*, 1337–1343.

pessimistic than they ought to be.

Babcock, Linda, Xianghong Wang, and George Loewenstein. 1996. “Choosing the Wrong Pond: Social Comparisons in Negotiations that Reflect a Self-serving Bias.” *The Quarterly Journal of Economics*, 111(1): 1–19.

Barber, Brad M., and Terrance Odean. 2001. “Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment.” *Quarterly Journal of Economics*, 116(1): 261–292.

Becker, Gordon, Morris DeGroot, and Jacob Marschak. 1964. “Measuring Utility by a Single-response Sequential Method.” *Behavioral Science*, 9(3): 226–232.

Benabou, Roland, and Jean Tirole. 2002. “Self-Confidence and Personal Motivation.” *Quarterly Journal of Economics*, 117(3): 871–915.

Benoit, Jean-Pierre, and Juan Dubra. 2007. “Overconfidence?” *MPRA Paper*, 6017.

Bernard, Victor, and Jacob Thomas. 1989. “Post-earnings-announcement Drift: Delayed Price Response or Risk Premium?” *Journal of Accounting Research*, 27: 1–36.

Bhattacharya, Jay, Dana Goldman, and Neeraj Sood. 2009. “Market Evidence of Misperceived Mortality Risk.” *Journal of Economic Behavior and Organization*.

Bradley, Gifford W. 1978. “Self-serving Biases in the Attribution Process: A Reexamination of the Fact or Fiction Question.” *Journal of Personality and Social Psychology*, 36(1): 56–71.

Brunnermeier, Markus K., and Jonathan A. Parker. 2005. “Optimal Expectations.” *The American Economic Review*, 95(4): 1092–1118.

Carrillo, Juan D., and Thomas Mariotti. 2000. “Strategic Ignorance as a Self-disciplining Device.” *Review of Economic Studies*, 529–544.

Chauvin, Kyle, David Laibson, and Johanna Mollerstrom. 2009. “Asset Bubbles and the Cost of Economic Fluctuations.” *MIT Sloan Working Paper*.

Compte, Olivier, and Andrew Postlewaite. 2004. “Confidence-enhanced Performance.” *American Economic Review*, 1536–1557.

Cross, K. Patricia. 2006. “Not Can, But Will College Teaching be Improved?” *New Directions for Higher Education*, 1977(17): 1–15.

Darley, John, and Paget Gross. 1983. “A Hypothesis Confirming Bias in Labeling Effects.” *Journal of Personality and Social Psychology*, XLIV.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1988. “Patterns of Firm Entry and Exit in US Manufacturing Industries.” *The RAND Journal of Economics*, 19(4): 495–515.

Eidinger, Richard, and Schapira Schapira. 1984. “Cancer Patients’ Insight Into Their Treatment, Prognosis, and Unconventional Therapies.” *Cancer*, 53(12).

Ertac, Seda. 2006. “An Experimental Study of Ego-Relevance in Information Processing.” *Cal Tech Doctoral Dissertation*.

Farmer, Roger. 1999. *The Macroeconomics of Self-fulfilling Prophecies.* The MIT Press.

Fisman, Raymond, Sheena S. Iyengar, Emir Kamenica, and Itamar Simonson. 2006. “Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment.” *Quarterly Journal of Economics*, 121(2): 673–697.

Foster, George, Chris Olsen, and Terry Shevlin. 1984. “Earnings Releases, Anomalies, and the Behavior of Security Returns.” *The Accounting Review*, 59(4): 574–603.

Frey, Dieter. 1981. “The Effect of Negative Feedback About Oneself and Cost of Information on Preferences for Information About the Source of this Feedback.” *Journal of Experimental Social Psychology*, 17(1): 42–50.

Grether, David M. 1980. “Bayes Rule as a Descriptive Model: The Representativeness Heuristic.” *Quarterly Journal of Economics*, 95: 537–557.

Gross, Samuel R., and Kent D. Syverud. 1991. “Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial.” *Michigan Law Review*, 90: 319.

Heider, Fritz. 1958. *The Psychology of Interpersonal Relations.* Wiley: New York.

Hitsch, Gunter J., Ali Hortasu, and Dan Ariely. 2010. “Matching and Sorting in Online Dating.” *American Economic Review*, 100(1): 130–63.

Kelley, Harold, and John Michela. 1980. “Attribution Theory and Research.” *Annual Review of Psychology*, 31(1): 457–501.

Köszegi, Botond. 2006. “Ego Utility, Overconfidence, and Task Choice.” *Journal of the European Economic Association*, 4(4): 673–707.

Larwood, Laurie. 1978. “Swine Flu: A Field Study of Self-serving Biases.” *Journal of Applied Social Psychology*, 8(3): 283–289.

Lord, Charles, Lee Ross, and Mark R. Lepper. 1979. “Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence.” *Personality and Social Psychology*, 37(11).

Mahoney, Michael J. 1977. “Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System.” *Cognitive Therapy and Research*, 1(2).

Mayraz, Guy. 2009. “Priors and Desires: A Model of Payoff-Dependent Beliefs.” mimeo.

Messick, David M., and Keith P. Sentis. 1979. “Fairness and Preference.” *Journal of Experimental Social Psychology*, 15(4): 418–434.

Mischel, Walter. 1976. “Determinants of Selective Memory About the Self.” *Journal of Consulting and Clinical Psychology*, 44(1): 92–102.

Möbius, Markus, Muriele Niederle, Paul Niehaus, and Tanya Rosenblatt. 2010. “Self-confidence Management: Theory and Empirical Evidence.” *Working paper.*

Moore, Don A., and Paul J. Healy. 2008. “The Trouble with Overconfidence.” *Psychological Review*, 115(2): 502.

Nickerson, Raymond. 1998. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.” *Review of General Psychology*, 2(2): 175–220.

Paling, John. 2003. “Strategies to Help Patients Understand Risks.” *British Medical Journal*, 327(7417): 745–748.

Plous, Scott. 1991. “Biases in the Assimilation of Technological Breakdowns: Do Accidents Make Us Safer?” *Journal of Applied Social Psychology*, 21.

Rabin, Matthew, and Joel L. Schrag. 1999. “First Impressions Matter: A Model of Confirmatory Bias.” *Quarterly Journal of Economics*, 114(1): 37–82.

Selten, Reinhard. 1998. “Axiomatic Characterization of the Quadratic Scoring Rule.” *Experimental Economics*, 1(1): 43–62.

Svenson, Ola. 1981. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, 47(2): 143–148.

Taylor, Shelley, and Jonathon Brown. 1994. "Positive Illusions and Well-being Revisited: Separating Fact from Fiction." *Psychological Bulletin*, 116: 21–21.

Tracy, Joseph S. 1986. "An Investigation into the Determinants of U.S. Strike Activity." *The American Economic Review*, 76(3): 423–436.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124–1131.

Watson, David, Eva C. Klohnen, Alex Casillas, Ericka N. Simms, Jeffrey Haig, and Diane S. Berry. 2004. "Match Makers and Deal Breakers: Analyses of Assortative Mating in Newlywed Couples." *Journal of Personality*, 72(5): 1029–1068.

Weeks, Jane, E.Francis Cook, Steven O'Day, Lynn Peterson, Neil Wenger, Douglas Reding, Frank Harrell, Peter Kussin, Neil Dawson, Alfred Connors, Joanne Lynne, and Russell Phillips. 1998. "Relationship Between Cancer Patients' Predictions of Prognosis and Their Treatment Preferences." *Journal of the American Medical Association*, 279(21): 1709–1714.

Appendix

Priors over priors

Even if subjects have "priors over priors", all that matters is round 0 beliefs, assuming that they report to us their true belief that they are in rank i. For instance, if a subject believes with certainty that she is in the pth percentile in the population at large, then her beliefs over her rank among subjects in that session, absent any other information, will be binomial. However if she has some uncertainty over her rank in the population, the implied distribution over her rank in her session need not be binomial. Additionally, observation of subjects in the room might inform her judgement in some way that moves her away from a binomial distribution. More generally, assume that a subject has N different priors that have some positive probability. This could be for any number of reasons, such as uncertainty about their place in the population distribution, uncertainty about how members of the opposite sex weight different qualities to arrive at their ranking, uncertainty about which members of the opposite sex are the harsher raters (recall that each subject has "dates" with only half of the subjects of the opposite sex), etc.. Let $f_n(i)$ denote the probability assigned to rank i under prior n. Let $p(n)$ denote the probability assigned to the prior f_n. Our assumption is that subjects report $\sum_{n=1}^{N} p(n)f_n(i)$ as their round 0 belief on rank i. Let this be denoted as $q_0(i)$. Then given some signal realization s_1, the posterior on rank i would be $P(i|s_1) = \sum_{n=1}^{N} \frac{P(s_1|i)f_n(i)}{\sum_{j=1}^{10} P(s_1|j)f_n(j)} P(n|s_1)$. That is, you update on which state is the true one conditional on each prior, and then you also update on which prior is correct. Reducing this yields an expression solely in terms of q_0:

$$
\begin{aligned}
\sum_{n=1}^{N} \frac{P(s_1|i)f_n(i)}{\sum_{j=1}^{10} P(s_1|j)f_n(j)} P(n|s_1) &= \sum_{n=1}^{N} \frac{P(s_1|i)f_n(i)}{\sum_{j=1}^{10} P(s_1|j)f_n(j)} \frac{P(s_1|n)p(n)}{\sum_{m=1}^{N} P(s_1|m)p(m)} \\
&= \frac{\sum_{n=1}^{N} P(s_1|i)f_n(i)p(n)}{\sum_{n=1}^{N} P(s_1|n)p(n)} \\
&= \frac{P(s_1|i)q_0(i)}{\sum_{j=1}^{10} P(s_1|j)q_0(j)}
\end{aligned}
$$

This shows that the Bayesian posterior is fully determined by the beliefs on rank i in round 0. The distribution over different priors does not matter. It is possible that subjects report only one element out of the support of their distribution over priors in round 0, as well as only one element of the support of higher order beliefs in later rounds. While any number of stories about how this element is selected could result in the data we observe while maintaining Bayesian updating of each individual distribution, it is impossible for us to investigate this question given our data.

Tables

Appendix Table 1: Total Change in Mean Beliefs Relative to Bayesian
($(\mu_t - \mu_0)$-$(\mu_t^B - \mu_0^B)$) by Gender

	Round 1		Round 2		Round 3		All
Signal History							
Signal History	F	M	F	M	F	M	Pooled t test
Beauty Condition							
0 $s=1$	-0.13	0.14	-0.27	0.20	-0.47	0.02	-1.47
1 $s=1$	0.04	-0.12	-0.15	-0.07	-0.16	0.01	0.05
2 $s=1$			0.11	-0.19	-0.35	-0.31	0.90
3 $s=1$					0.18	-0.17	0.86
IQ Condition							
0 $s=1$	0.04	-0.28	0.04	-0.19	0.25	-0.44	1.50
1 $s=1$	0.44	0.22	0.11	-0.28	-0.41	-0.06	0.48
2 $s=1$			0.13	0.27	-0.09	0.01	-0.63
3 $s=1$					0.45	0.13	0.68

Experimental Instructions

Instructions along with other experimental materials are provided in the online documentation.