

# Supplementary Material to: “Managing Self-Confidence”

August 19, 2014

## S-1 Demand for Information

While some behavioral models study agents who can skew their interpretation of feedback, others focus on selective *acquisition* of feedback as a technique for self-confidence management. Among other things these models predict that agents may strictly prefer not to acquire relevant information about their own types. In this section we test for this mechanism in our data and then relate it to our theoretical framework.

### S-2.1 Evidence

In the final stage of our experiment, after all four noisy signals had been delivered, we elicited subjects’ demand for noiseless feedback on their relative performance. Subjects stated their willingness to pay for receiving \$2 as well as for receiving \$2 and an email containing information on their performance. We bounded responses between \$0.00 and \$4.00. We offered two kinds of information: subjects could learn whether they scored in the top half, or learn their exact quantile in the score distribution. For each subject one of these choices was randomly selected and the subject purchased the corresponding bundle if and only if their reservation price exceeded a randomly generated price. This design is a standard application of the Becker-DeGroot-Marschak mechanism (BDM) except that we measure information values by netting out subjects’ valuations for \$2 alone from their other valuations to address the concern that subjects may under-bid for objective-value prizes.

We use these bids to calculate subjects’ implied value for the various information packages offered to them. For example, a subject’s valuation for learning whether or not she was in the top half is defined as her bid for \$2 and learning this information minus her bid for \$2, all in cents. We take this difference to remove potential bias due to misunderstanding the dominant strategy in the “bid for \$2” decision problem.<sup>1</sup> Subjects also bid on more precise information: learning their exact quantile. Table S-1 summarizes the results. Subjects’ mean value for coarse information is 16.5 (s.d. 47.8), with 9% of subjects reporting a negative value. The mean valuation for precise information is higher at 40.0 (s.d. 78.3), but again 9% of subjects report a negative value.<sup>2</sup>

---

<sup>1</sup>Among our subjects, 89% bid less than \$2, and 80% bid less than \$1.99.

<sup>2</sup>Interestingly, Eliaz and Schotter (2010) find that subjects are willing to pay positive amounts for information (unrelated to ego) even when it cannot improve their decision-making.

Table S-1: Implied Valuations for Information: Summary Statistics

	$N$	Mean	Std. Dev.	$P(v < 0)$
<b>Estimation Sample</b>				
Learning top/bottom half	650	16.5	47.8	0.09
Learning percentile	650	40.0	78.3	0.09
<b>Women</b>				
Learning top/bottom half	338	16.4	49.8	0.11
Learning percentile	338	38.7	82.0	0.11
<b>Men</b>				
Learning top/bottom half	312	16.7	45.5	0.07
Learning percentile	312	41.5	74.1	0.06

Values for information are the differences between subjects bids for \$2 and their bids for the bundle of \$2 and receiving an email containing that information. Values are in cents. The final column reports the fraction of observations with strictly negative valuations. There are fewer than 656 observations because 6 subjects did not provide valuations for information.

**Result 7** (Information Aversion). *A substantial fraction of subjects are willing to pay to avoid learning their type.*

One caveat is that negative valuations could be an artefact of noise in subjects' responses. The strongest piece of evidence that this is not the case is our next result, which shows that confidence has a causal effect on the propensity for aversion. Another clue is the high correlation ( $\rho = 0.77$ ) between having a negative valuation for coarse information and a negative valuation for precise information, which suggests that both measures contain meaningful information. In unreported results we have developed this idea formally and shown that under the structural assumption of i.i.d. normal measurement error the bid data reject the null hypothesis of no aversion (results available on request).

**Result 8.** *More confident subjects are causally less information-averse.*

To examine whether information aversion is more pronounced among more or less confident subjects we regress an indicator  $I(v_i \geq 0)$  on subjects' logit posterior belief after all four rounds of updating, which is when they bid for information. Columns I–III of Table S-2 show that subjects with higher posterior beliefs are indeed significantly more likely to have (weakly) positive information values. The point estimate is slightly larger and remains strongly significant when we control for ability (Column II) and gender and age (Column III). There could, however, be some other unobserved factor orthogonal to these controls that explains the positive correlation. To address this issue Columns IV and V report instrumental variables estimates. We use two instruments. First, the average score of other subjects randomly assigned to the same quiz type remains a valid instrument for beliefs, as in Section 4 above. In addition, once we control for whether or not the subject scored in the top half the number of positive signals she received during the updating stage is a valid instrument since signals were random conditional on ability. Estimates using these instruments are similar to the OLS estimates, slightly larger, and though less precise, still significant at the 10% level.

Table S-2: Confidence and Positive Information Value

Regressor	OLS			IV	
	I	II	III	IV	V
logit( $\mu$ )	0.017 (0.007)**	0.023 (0.009)***	0.023 (0.009)**	0.027 (0.016)*	0.027 (0.017)*
Top Half		-0.033 (0.028)	-0.035 (0.028)	-0.038 (0.034)	-0.042 (0.034)
Male			0.029 (0.023)		0.027 (0.023)
YOG			0.018 (0.012)		0.018 (0.012)
First-Stage $F$ -Statistic	-	-	-	118.48	113.19
N	609	609	609	609	609
$R^2$	0.007	0.010	0.016	-	-

Notes: Each column is a separate regression. Estimation is via OLS in Columns I–III and by IV in Columns IV–V using the instruments described in the text. The outcome variable in all regressions is an indicator equal to 1 if the subject’s valuation for information was positive; the mean of this variable is 0.91. “Top Half” is an indicator equal to one if the subject scored above the median on his/her quiz type; “YOG” is the subject’s year of graduation. Heteroskedasticity-robust standard errors in parenthesis. Statistical significance is denoted as: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

## S-2.2 Theory

The result that low-confidence agents are (causally) likely to be information-averse is broadly consistent with a number of behavioral models which generate information aversion. In this section we examine more specifically how our data compare to the predictions of our own model. Towards this aim, extend the model and suppose that with probability  $\epsilon > 0$  the agent is presented with the opportunity to purchase a perfectly informative signal at time  $\tilde{T}$  just before learning the cost  $c$  for making costly investment. It is easy to calculate the unbiased Bayesian’s willingness to pay for information,  $WTP^{PB}(\mu_{\tilde{T}})$ :

$$WTP^{PB}(\mu_{\tilde{T}}) = \mu_{\tilde{T}} \left( 1 - \int_0^1 cdG(c) \right) - \int_0^{\mu_{\tilde{T}}} (\mu_{\tilde{T}} - c)dG(c) \quad (19)$$

Importantly, an unbiased Bayesian’s value of information is always positive and single-peaked: the value of information is zero when the agent is very sure about her type and largest when she is the least sure. This valuation is generally sub-optimal for an agent with belief utility, however, who wishes to balance this motive against the needs of decision-making. If a low type were to learn the truth at time  $\tilde{T}$  her carefully calibrated self-belief management would break down and she would enjoy no belief utility between periods  $\tilde{T}$  and  $T$ .

We therefore calculate the optimal willingness to pay  $WTP^{OB}(\hat{\mu}_\tau, \tau)$  at relative time  $\tau$  which the agent would commit to at time  $t = 0$ . To simplify our analysis and build on the results from the previous section, we assume that the decision-maker does not take the

possibility of buying information into account when choosing her bias. This assumption seems appropriate when the probability of purchasing information,  $\epsilon$ , is small.

**Proposition 4.** *Assume that an agent with positive belief utility chooses an optimal biased Bayesian updating process. Let the subjective belief at relative time  $\tau$  be  $0 < \hat{\mu}_\tau < 1$ . The agent's willingness to pay evaluated at period 0,  $WTP^{OB}(\hat{\mu}_\tau, \tau)$ , satisfies*

$$\lim_{T \rightarrow \infty} WTP^{OB}(\hat{\mu}_\tau, \tau) = -\tilde{L}(\hat{\mu}_\tau) \quad (20)$$

where  $\tilde{L}(\hat{\mu}) = (1 - \tau)b(\hat{\mu}) - \int_0^{\hat{\mu}} cdG(c)$  is the per-period utility of a low type with belief utility  $(1 - \tau)b(\hat{\mu})$ .

*Proof.* We know that high-type beliefs converge to 1 while low type beliefs stay close to  $\mu_L^*$ . We also know that  $\hat{\sigma}_T^L \rightarrow 0$  and  $\hat{\sigma}_T^H \rightarrow 0$  and that there are constants  $m_1, m_2 > 0$  such that  $m_1 < \hat{\sigma}_T^L / \hat{\sigma}_T^H < m_2$ . Hence, the probability at relative time  $\tau$  that the agent is a low type provided that  $\hat{\mu}_{\lfloor \tau T^j \rfloor} < 1$  converges to 1. Therefore, learning one's type decreases the agent's total utility to 0 with probability approaching 1 as  $T \rightarrow \infty$  and destroys belief utility  $(1 - \tau)b(\hat{\mu}_\tau)$  (since low type logit-beliefs follow a driftless random walk with vanishing variance).  $\square$

Intuitively, an agent with subjective belief below 1 is asymptotically likely to be a low type, as otherwise her beliefs would have converged rapidly to 1. Proposition 2 implies that her beliefs in the low state follow a driftless random walk with vanishing variance and hence stay around  $\hat{\mu}_\tau$ . This implies that her belief utility over the remaining relative time  $1 - \tau$  is approximately  $(1 - \tau)b(\hat{\mu}_\tau)$ . Buying information, on the other hand, would reveal her to be a low type immediately and yield a payoff of 0.

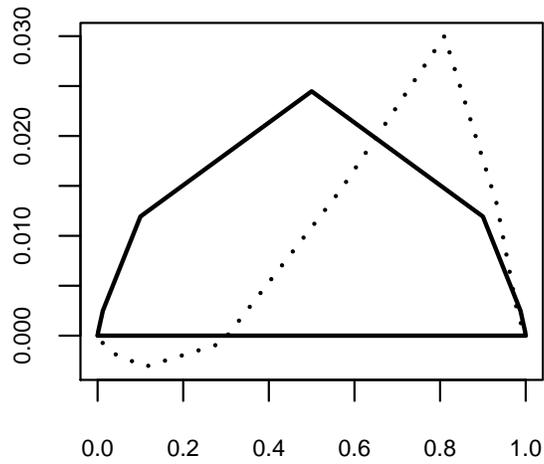
The economic significance of this result is that for low subjective beliefs  $\hat{\mu}$  (and  $\tau$  not too large) the optimal willingness to pay is negative, since the benefits of sustaining belief utility exceed the costs of mistaken choices, while for high subjective beliefs the optimal WTP is positive, since this relationship is reversed.<sup>3</sup> Thus Proposition 4 implies that, consistent with our empirical findings, the optimally biased agent will have a negative value of information when her belief is low and a positive value of information when her belief is high. This effect is mitigated for larger  $\tau$  when belief utility is aggregated over fewer periods and hence becomes relatively less important; in this case information demands begin to resemble traditional, unbiased demands.

Figure S-1 plots an example of the finite- $T$  numerical demands generated by our model for both an unbiased and an optimally biased Bayesian. The unbiased Bayesian always values information positively, and values it most at intermediate beliefs where uncertainty is highest. The optimally biased agent, on the other hand, places a negative value on information for low levels of confidence and only assigns a positive value above a threshold level of confidence.

---

<sup>3</sup>Note, that  $WTP^{OB}(\hat{\mu}_\tau, \tau)$  equals  $-L(\hat{\mu}_\tau)$  for  $\tau = 0$ . Therefore, the biased Bayesian's willingness to pay for information is negative for low beliefs because  $L(\mu_L^*) > 0$ .

Figure S-1: Numerical optimum information demand functions for finite  $T$  and binary signals



Plots information values for realizable values of  $\hat{\mu}_{[\tau T]}$  for  $T = 31$ , and  $[\tau T] = 10$  for the unbiased Bayesian (solid lines) and agent with optimal simple updating bias (dotted lines) cases. The remaining parameters are fixed in both cases at  $\mu_0 = 0.5$ ,  $c \sim U[0, 1]$ ,  $b(\hat{\mu}) = \frac{1}{4}\hat{\mu}$ ,  $p = 0.75$ ,  $q = 0.25$

## S-2 Additional Tables

Table S-3: Quiz Performance: Summary Statistics

	<i>N</i>	<b>Correct</b>		<b>Incorrect</b>		<b>Score</b>	
		Mean	SD	Mean	SD	Mean	SD
<b>Overall</b>							
Restricted Sample	656	10.2	4.3	2.7	2.1	7.4	4.8
Full Sample	1058	9.7	4.3	3.0	2.4	6.8	4.9
<b>By Quiz Type</b>							
1	79	8.1	3.1	1.7	1.2	6.4	3.3
2	85	13.0	2.9	2.7	2.1	10.3	3.4
3	69	8.9	3.3	3.0	2.1	5.9	3.8
4	74	12.2	3.8	3.1	2.3	9.2	4.6
5	75	6.5	1.6	4.0	2.3	2.5	2.8
6	63	14.5	4.5	2.3	1.7	12.3	4.7
7	73	7.6	2.6	2.2	1.7	5.4	3.1
8	69	13.6	2.8	3.2	1.8	10.4	3.3
9	69	7.3	3.5	2.7	2.8	4.7	4.5
<b>By Gender</b>							
Male	314	10.6	4.2	2.7	2.3	7.9	4.8
Female	342	9.7	4.4	2.8	2.0	6.9	4.8

Table S-4: Conservative and Asymmetric Belief Updating

Regressor	Round 1	Round 2	Round 3	Round 4	All Rounds	Unrestricted
<b>Panel A: OLS</b>						
$\delta$	0.777 (0.042)***	0.946 (0.020)***	0.943 (0.030)***	1.009 (0.027)***	0.937 (0.016)***	0.888 (0.014)***
$\beta_H$	0.448 (0.021)***	0.400 (0.020)***	0.456 (0.024)***	0.568 (0.035)***	0.487 (0.016)***	0.264 (0.013)***
$\beta_L$	0.477 (0.033)***	0.422 (0.025)***	0.457 (0.027)***	0.471 (0.027)***	0.454 (0.016)***	0.211 (0.011)***
$\mathbb{P}(\beta_H = 1)$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathbb{P}(\beta_L = 1)$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathbb{P}(\beta_H = \beta_L)$	0.471	0.492	0.989	0.030	0.083	0.000
N	420	413	422	458	1713	3996
$R^2$	0.754	0.882	0.874	0.864	0.846	0.798
<b>Panel B: IV</b>						
$\delta$	1.262 (0.325)***	0.953 (0.098)***	1.058 (0.136)***	0.943 (0.157)***	1.032 (0.078)***	0.977 (0.060)***
$\beta_H$	0.617 (0.129)***	0.401 (0.024)***	0.456 (0.025)***	0.578 (0.041)***	0.496 (0.016)***	0.273 (0.013)***
$\beta_L$	0.414 (0.052)***	0.421 (0.025)***	0.450 (0.028)***	0.477 (0.033)***	0.446 (0.015)***	0.174 (0.027)***
$\mathbb{P}(\beta_H = 1)$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathbb{P}(\beta_L = 1)$	0.000	0.000	0.000	0.000	0.000	0.000
$\mathbb{P}(\beta_H = \beta_L)$	0.231	0.567	0.864	0.031	0.044	0.004
First Stage $F$ -statistic	4.85	14.47	11.24	8.40	14.86	20.61
N	420	413	422	458	1713	3996
$R^2$	-	-	-	-	-	-

Notes:

1. Each column in each panel is a regression. The outcome in all regressions is the log posterior odds ratio.  $\delta$  is the coefficient on the log prior odds ratio;  $\beta_H$  and  $\beta_L$  are the estimated effects of the log likelihood ratio for positive and negative signals, respectively. Bayesian updating (for both biased and unbiased Bayesians) corresponds to  $\delta = \beta_H = \beta_L = 1$ .
2. Estimation samples are restricted to subjects whose beliefs were always within  $(0, 1)$ . Columns 1-5 further restrict to subjects who updated their beliefs in every round and never in the wrong direction; Column 6 includes subjects violating this condition. Columns 1-4 examine updating in each round separately, while Columns 5-6 pool the 4 rounds of updating.
3. Estimation is via OLS in Panel A and via IV in Panel B, using the average score of other subjects who took the same (randomly assigned) quiz variety as an instrument for the log prior odds ratio.
4. Heteroskedasticity-robust standard errors in parenthesis; those in the last two columns are clustered by individual. Statistical significance is denoted as: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table S-5: Updating is not Differential by Prior

Regressor	Round 1	Round 2	Round 3	Round 4	All Rounds	Unrestricted
<b>Panel A: OLS</b>						
$\delta$	0.908 (0.029)***	0.944 (0.019)***	0.952 (0.027)***	0.958 (0.022)***	0.944 (0.012)***	0.885 (0.014)***
$\delta_H$	-0.150 (0.053)***	-0.040 (0.030)	-0.020 (0.046)	0.058 (0.046)	-0.037 (0.023)	0.006 (0.026)
$\beta_H$	0.361 (0.018)***	0.295 (0.017)***	0.334 (0.021)***	0.434 (0.030)***	0.369 (0.013)***	0.264 (0.013)***
$\beta_L$	0.268 (0.026)***	0.270 (0.020)***	0.302 (0.022)***	0.354 (0.024)***	0.298 (0.012)***	0.212 (0.011)***
N	612	612	612	612	2448	3996
$R^2$	0.808	0.891	0.875	0.860	0.854	0.798
<b>Panel B: IV</b>						
$\delta$	0.876 (0.513)*	1.070 (0.187)***	1.398 (0.266)***	0.830 (0.164)***	1.071 (0.109)***	0.976 (0.097)***
$\delta_H$	0.092 (0.530)	-0.287 (0.215)	-0.544 (0.311)*	0.166 (0.238)	-0.167 (0.131)	0.002 (0.124)
$\beta_H$	0.409 (0.045)***	0.292 (0.018)***	0.335 (0.021)***	0.437 (0.037)***	0.369 (0.012)***	0.273 (0.014)***
$\beta_L$	0.277 (0.149)*	0.243 (0.044)***	0.216 (0.063)***	0.385 (0.050)***	0.268 (0.027)***	0.175 (0.041)***
N	612	612	612	612	2448	3996
$R^2$	-	-	-	-	-	-

Notes:

- Each column in each panel is a regression. The outcome in all regressions is the log posterior odds ratio.  $\delta$  is the coefficient on the log prior odds ratio;  $\delta_H$  is the coefficient on an interaction between the log prior odds ratio and an indicator for a positive signal;  $\beta_H$  and  $\beta_L$  are the estimated effects of the log likelihood ratio for positive and negative signals, respectively. Bayesian updating corresponds to  $\delta = \beta_H = \beta_L = 1$  and  $\delta_H = 0$ .
- Estimation samples are restricted to subjects whose beliefs were always within  $(0, 1)$ . Columns 1-5 further restrict to subjects who updated their beliefs at least once and never in the wrong direction; Column 6 includes subjects violating this condition. Columns 1-4 examine updating in each round separately, while Columns 5-6 pool the 4 rounds of updating.
- Estimation is via OLS in Panel A and via IV in Panel B, using the average score of other subjects who took the same (randomly assigned) quiz variety as an instrument for the log prior odds ratio.
- Heteroskedasticity-robust standard errors in parenthesis; those in the last two columns are clustered by individual. Statistical significance is denoted as: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .