Agreeing to be fooled: Optimal ignorance about information sources Online Appendix

Takuma Habu

18th February 2024

B Extensions

B.1 Investigation strategies that can depend on the realisation of the Sender's experiment

Proposition 2. *The Receiver is no better off with the ability to choose an investigation strategy that is a function of both the Sender's choice of an experiment as well as the realisation of the Sender's message*

Proof. Recall from Theorem 2 that, for sufficiently low prior belief about the Sender's reliability (i.e., $\rho_0 \leq \rho_{0,1}$), the Receiver can obtain his ideal payoff using investigations that do not depend on the realisation of experiments—hence, the Receiver would not benefit from the additional flexibility in this case.

Let us first show that the Sender's maxmin payoff is the same against a Receiver who can (additionally) condition his investigation strategy on the realisation of the Sender's experiment. To see this, first observe that, given the simplification in Lemma 1, the Receiver knows with probability one that the Sender is reliable after observing the message m_0. Therefore, the only investigation that can affect the Receiver's belief is the one after the Receiver observes the message m_1. Hence, so long as the simplification holds, the assumption that the investigation strategies do not depend on realisations is without loss. Second, notice that the Receiver can always choose not to condition investigations on realisations even if he could. Together, these observations mean that the Sender must be weakly worse off against a Receiver who minimises the Sender's payoff if she chose more complex experiments than the experiments identified by Lemma 1.

Because the Sender's maxmin payoff is unaffected by this extra flexibility, the additional flexibility could (only) induce the Sender to choose an experiment outside of the set of experiments identified in Lemma 1 on the equilibrium path of the commitment equilibrium. Moreover, it must also be that the case that both the Sender's and the Receiver's payoffs must be strictly greater than the case in which the Receiver's investigation can only depend on the experiment. For sufficiently high prior belief about the Sender's reliability (specifically, $\rho_0 \geq \rho_{0,2}$), the total surplus between the Sender and the Receiver is constant (equals $\frac{\mu_0}{\mu^*}$), meaning that it is not possible to improve both players' equilibrium payoffs simultaneously. For intermediate prior beliefs, $\rho_0 \in (\rho_{0,1}, \rho_{0,2})$, recall that the Receiver maximises his payoff subject to the Sender getting at least her maxmin payoff, and that the Sender is choosing the fully informative experiment. Thus, for the Receiver to obtain a higher payoff, it must be that he conducts a more informative investigation; however, doing so would necessarily result in lower payoffs for the Sender. Thus, it follows that the Receiver can do no better even for intermediate prior beliefs about reliability. ■

Remark 1. The additional flexibility in the Receiver's investigation would not change the payoffs that the Receiver can achieve by delegating investigations to a λ-balanced Third Party either. This is because a λ-balanced Third Party either conducts a punishing or a full investigation, and the Sender's optimal experiments given such investigations are contained in the canonical set of experiments identified in Lemma 1.

B.2 Character witnesses

Recall that a purely Sender-aligned Third Party is a $(\lambda_{\mathrm{S}}, \lambda_{\mathrm{R}})$-balanced Third Party with $\lambda_{\mathrm{S}} > 0$ and $\lambda_{\mathrm{R}} = 0$.

Proposition 3. *The Sender's delegation equilibrium payoff with a purely Sender-aligned Third Party is weakly higher than her equilibrium payoff with no investigations—and strictly so if $\rho_0 \leq \underline{\rho}$. The Receiver's delegation equilibrium payoff with a purely Sender-aligned Third Party is zero.*

Proof. Standard arguments (Aumann and Maschler, 1968; Kamenica and Gentzkow, 2011) mean that the maximal payoff for the Sender from any experiment $\widehat{\rho} \in [\underline{\rho}, 1]$ that can be induced by some investigation strategy is given by the concave envelope of the function $\max V_S(\cdot|\widehat{\rho})$ (evaluated at the prior ρ_0):

$$\operatorname{cav}\max V_S(\cdot|\widehat{\rho})(\rho_0) = \begin{cases} \frac{\mu_0}{\mu^*}\frac{\rho_0}{\widehat{\rho}} & \text{if } \rho_0 \in (0, \widehat{\rho}] \\ \left(1 - \frac{\mu^*-\mu_0}{\mu^*}\frac{1}{\widehat{\rho}}\right)\frac{\rho_0-\widehat{\rho}}{1-\widehat{\rho}} & \text{if } \rho_0 \in (\widehat{\rho}, 1) \end{cases}.$$

If $\rho_0 < \underline{\rho}$, the Sender maximises her payoff by choosing $\widehat{\rho} = \underline{\rho}$ induced by an investigation with support $\{0, \widehat{\rho}\}$. Since the Receiver's payoffs are zero under these two possible posterior beliefs. Note that with $\rho_0 < \underline{\rho}$, the Sender's payoff without an investigation is zero because the Sender is unable to induce the Receiver to take action even with the fully informative experiment. If, instead, $\rho_0 \geq \underline{\rho}$, the Sender maximises her payoff by choosing $\widehat{\rho} = \rho_0$ (under no investigation) and since this makes the Receiver indifferent between taking action or not, his payoff is zero. ■

B.3 Limited commitment

Define a *limited commitment equilibrium* as a Receiver-preferred i-commitment equilibrium in which the Receiver can only choose constant investigations; i.e., $i(\cdot) = \iota \in \mathscr{I}$. Recall that, for the Receiver to attain a strictly positive payoff, the Receiver must induce the Sender to choose $\widehat{\rho} < \max \operatorname{supp}(\iota)$. Thus, Receiver would choose $\iota \in \mathscr{I}$ such that $\operatorname{supp}(\iota) \cap [\underline{\rho}, 1]$ contains at least two elements, say $\widehat{\rho}_1$ and $\widehat{\rho}_2$ with $\widehat{\rho}_1 < \widehat{\rho}_2$, while ensuring that $V_S(\widehat{\rho}_1, \iota) \geq V_S(\widehat{\rho}_2, \iota)$. Moreover, since $\widehat{\rho}_1 < 1$, the Receiver's payoff is strictly convex but linear in $\rho \in [\widehat{\rho}_1, 1]$. Hence, the Receiver prefers an investigation that spreads mass to the extremes while providing the Sender with sufficient incentive to choose an informative experiment, $\widehat{\rho}_1$. Therefore, in any limited commitment equilibrium, the Receiver's investigation induces posterior beliefs at $\{0, \widehat{\rho}^\circ, 1\}$ for some $\widehat{\rho}^\circ \in [\underline{\rho}, 1)$ while ensuring that $V_S(\widehat{\rho}^\circ, \iota) \geq V_S(1, \iota)$. The following proposition formalises the intuition described above. Let $\mathscr{I}^\circ$ be a set of investigations with support that is a subset of $\{0, \widehat{\rho}, 1\}$

for some $\widehat{\rho} \in [\underline{\rho}, 1]$; i.e.,

$$\mathscr{I}^{\circ} := \left\{ \iota \in \mathscr{I} : \operatorname{supp}(\iota) \subseteq \left\{ \{0, \widehat{\rho}, 1\} : \widehat{\rho} \in \left[\underline{\rho}, 1 \right] \right\} \right\}.$$

Lemma 4. *In any limited commitment equilibrium, the Sender chooses*

$$\widehat{\rho} \in \min \left\{ \operatorname{supp}(\iota) \cap \left[\underline{\rho}, 1 \right] \right\}.$$

Proof. Fix a finite support investigation $\iota \in \mathscr{I}$ and let $\{\widehat{\rho}_k\}_{k=1}^{K}$ be such that $\operatorname{supp}(\iota) \cap [\underline{\rho}, 1] = \{\widehat{\rho}_1, \widehat{\rho}_2 \ldots, \widehat{\rho}_K\}$ and $\widehat{\rho}_1 < \widehat{\rho}_2 < \cdots < \widehat{\rho}_K$. Suppose Sender chooses $\widehat{\rho} \in (\widehat{\rho}_{\ell-1}, \widehat{\rho}_{\ell}]$ for some $\ell \in \{2, \ldots K\}$. Then,

$$V_{\mathrm{S}}(\widehat{\rho}, \tau) = \sum_{k=\ell}^{K} \left(1 - \frac{\mu^* - \mu_0}{\mu^*} \frac{\widehat{\rho}_k}{\widehat{\rho}} \right) \iota(\widehat{\rho}_k)$$

and so the Sender can do strictly better by choosing $\widehat{\rho} = \widehat{\rho}_{\ell}$. Hence, in any limited commitment equilibrium Sender chooses $\widehat{\rho} \in \operatorname{supp}(\iota) \cap [\underline{\rho}, 1]$. Now let $k^* \in \{1, \ldots, K\}$ be such that

$$V_{\mathrm{S}}(\widehat{\rho}_{k^*}, \iota) \geq V_{\mathrm{S}}(\widehat{\rho}_k, \iota) \;\; \forall k \in \{1, \ldots, K\}.$$

Suppose first that $k^* > 1$. We will show that there exists a mean preserving spread of ι, say $\iota' \in \mathscr{I}$, obtained by spreading mass at $\widehat{\rho}_1$ that does not affect the Sender's IC and the Receiver's payoff is strictly increased. Observe that for any $\ell \in \{2, \ldots, K-1\}$,

$$\begin{aligned}
&V_{\mathrm{S}}(\widehat{\rho}_{\ell+1}, \iota) - V_{\mathrm{S}}(\widehat{\rho}_{\ell}, \iota) \\
&= \sum_{k=\ell+1}^{K} \left(1 - \frac{\mu^* - \mu_0}{\mu^*} \frac{\widehat{\rho}_k}{\widehat{\rho}_{\ell+1}} \right) \iota(\widehat{\rho}_k) - \sum_{k=\ell}^{K} \left(1 - \frac{\mu^* - \mu_0}{\mu^*} \frac{\widehat{\rho}_k}{\widehat{\rho}_{\ell}} \right) \iota(\widehat{\rho}_k) \\
&= \frac{\mu^* - \mu_0}{\mu^*} \left(\frac{1}{\widehat{\rho}_{\ell}} - \frac{1}{\widehat{\rho}_{\ell+1}} \right) \sum_{k=\ell+1}^{K} \widehat{\rho}_k \iota(\widehat{\rho}_k) - \frac{\mu_0}{\mu^*} \iota(\widehat{\rho}_{\ell}). \qquad (1)
\end{aligned}$$

Suppose we obtain $\iota'(\widehat{\rho}_K)$ by adding $\varepsilon_K > 0$ to $\iota(\widehat{\rho}_K)$; i.e., $\iota'(\widehat{\rho}_K) = \iota(\widehat{\rho}_K) + \varepsilon_K$. To ensure that Sender's IC remains unchanged, we can set $\iota'(\widehat{\rho}_{K-1}) = \iota(\widehat{\rho}_{K-1}) + \varepsilon_{K-1}$

such that the expression (1) equals zero with $\ell = K-1$; i.e., set $\varepsilon_{K-1} > 0$ to

$$\varepsilon_{K-1} = \frac{\mu^* - \mu_0}{\mu_0}\left(\frac{1}{\widehat{\rho}_{K-1}} - \frac{1}{\widehat{\rho}_K}\right)\widehat{\rho}_K\left[\iota\left(\widehat{\rho}_K\right) + \varepsilon_K\right] - \iota\left(\widehat{\rho}_{K-1}\right).$$

We can use (1) in a similar manner to obtain $\{\varepsilon_k\}_{k=2}^K$ such that Sender's IC is unaffected. To ensure that ι' is a mean-preserving spread of ι, we set $\iota'(0) = \varepsilon_0$ to satisfy

$$\varepsilon_0 + \sum_{k=2}^{K} \varepsilon_k \leq \iota\left(\widehat{\rho}_1\right) \text{ and } 0 \cdot \varepsilon_0 + \sum_{k=2}^{K} \widehat{\rho}_k \varepsilon_k = \widehat{\rho}_1 \iota\left(\widehat{\rho}_1\right).$$

Since $\{\varepsilon_k\}_{k=2}^K$ is pinned down via (1) when setting $\varepsilon_K > 0$, we can choose $\varepsilon_K, \varepsilon_0 > 0$ to ensure $\{\varepsilon_0, \varepsilon_2, \ldots, \varepsilon_K\}$ satisfies the expressions above. Since the Sender's IC remains unchanged, the Sender is willing to choose $\widehat{\rho}_{k^*}$ under ι' and, by construction, ι' is a mean-preserving spread of ι. Thus, the Receiver (with convex preferences) prefers ι' over ι and strictly so if $\widehat{\rho}_{k^*} < 1$. ■

Lemma 5. *Any investigation is dominated by an investigation in $\mathscr{I}^\circ$ for the Receiver.*

Proof. By the previous lemma, we can focus on $\iota \in \mathscr{I}$ such that $\operatorname{supp}(\iota) \cap [\underline{\rho}, 1] = \{\widehat{\rho}_1, \widehat{\rho}_2 \ldots, \widehat{\rho}_K\}$ with $\widehat{\rho}_1 < \widehat{\rho}_2 < \cdots < \widehat{\rho}_K$ and

$$V_{\mathrm{S}}\left(\widehat{\rho}_1, \iota\right) \geq V_{\mathrm{S}}\left(\widehat{\rho}_k, \iota\right) \ \forall k \in \{1, \ldots, K\}.$$

Consider a mean-preserving spread of ι, denoted ι', obtained by moving mass from $\widehat{\rho}_2$ to $\{\widehat{\rho}_1, \widehat{\rho}_3, \ldots, \widehat{\rho}_K\}$. Letting $\iota'(\widehat{\rho}_K) = \iota(\widehat{\rho}_K) + \varepsilon_K$ with $\varepsilon_K > 0$, we can again obtain $\{\varepsilon_1, \varepsilon_3, \varepsilon_4, \ldots, \varepsilon_K\}$ that ensures that Sender still finds it incentive compatible to choose $\widehat{\rho}_1$ and we choose $\varepsilon_0 > 0$ be such that

$$\varepsilon_0 + \sum_{k \in \{1,3,4,\ldots,K\}} \varepsilon_k \leq \iota\left(\widehat{\rho}_1\right) \text{ and } 0 \cdot \varepsilon_0 + \sum_{k \in \{1,3,4,\ldots,K\}} \widehat{\rho}_k \varepsilon_k = \widehat{\rho}_2 \iota\left(\widehat{\rho}_2\right).$$

Convexity of the Receiver's preferences means that the Receiver strictly prefers ι' over ι. Iterating the process for $k = \{3, \ldots, K-1\}$ means that the Receiver prefers ι with $K = 2$. So fix $K = 2$ and suppose now that $\widehat{\rho}_2 < 1$. Then, consider $\iota' \in \mathscr{I}$

obtained by spreading the mass at $\widehat{\rho}_2$ to $\widehat{\rho} = 1$ such that

$$\mathbb{E}_\iota \left[\rho | \rho \geq \widehat{\rho}_1\right] = \mathbb{E}_\iota \left[\rho | \rho \geq \widehat{\rho}_1\right],$$

which can be done because $\widehat{\rho}_2 \in (\widehat{\rho}_1, 1)$. Then, because $V_{\mathrm{S}}(\widehat{\rho}_1, \delta_\rho)$ is linear on $\rho \in [\widehat{\rho}_1, 1]$,

$$V_{\mathrm{S}}\left(\widehat{\rho}_1, \iota\right) = V_{\mathrm{S}}\left(\widehat{\rho}_1, \iota'\right)$$

and

$$V_{\mathrm{S}}\left(\widehat{\rho}_2, \iota\right) - V_{\mathrm{S}}\left(1, \iota'\right) = \frac{\mu_0}{\mu^*}\left[\iota\left(\widehat{\rho}_2\right) - \iota'(1)\right] > 0$$

since

$$\begin{aligned} \frac{\iota'(1)}{\iota'\left(\rho \geq \widehat{\rho}_1\right)} = \frac{\iota'(1)}{\iota\left(\rho \geq \widehat{\rho}_1\right)} &= \frac{\mathbb{E}_\iota\left[\rho | \rho \geq \widehat{\rho}_1\right] - \widehat{\rho}_1}{1 - \mathbb{E}_\iota\left[\rho | \rho \geq \widehat{\rho}_1\right]} \\ &< \frac{\mathbb{E}_\iota\left[\rho | \rho \geq \widehat{\rho}_1\right] - \widehat{\rho}_1}{\widehat{\rho}_2 - \mathbb{E}_\iota\left[\rho | \rho \geq \widehat{\rho}_1\right]} = \frac{\iota\left(\widehat{\rho}_2\right)}{\iota\left(\rho \geq \widehat{\rho}_1\right)}. \end{aligned}$$

It follows that the Sender is willing to choose $\widehat{\rho}_1$ over $\widehat{\rho} = 1$ under ι'. Together with the fact that ι' is a mean-preserving spread of ι, convexity of the Receiver's payoff implies that the Receiver's payoff is higher with ι' than under ι. Finally, if $\mathrm{supp}(\iota) \cap (0, \underline{\rho})$ is nonempty, then we can find an improvement without affecting IC by moving the mass from any $(0, \underline{\rho})$ to 0 in a mean-preserving manner while preserving Sender's IC in the manner described above. ■

Proposition 4. *The Receiver's limited commitment payoff strictly positive* for any $\rho_0 \in (0, 1)$.

Proof. By the previous lemmata, we can focus on $\iota \in \mathscr{I}$ such that $\mathrm{supp}(\iota) \subseteq$

$\{0, \widehat{\rho}, 1\}$, then

$$\begin{aligned}
V_{\mathrm{S}}(\widehat{\rho}, \iota) &\geq V_{\mathrm{S}}(1, \iota) \\
\Leftrightarrow 1 - \frac{\mu^* - \mu_0}{\mu^*} \frac{\mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}]}{\widehat{\rho}} &\geq \frac{\mu_0}{\mu^*} \frac{\iota(1)}{\iota(\rho \geq \widehat{\rho})} = \frac{\mu_0}{\mu^*} \frac{\mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}] - \widehat{\rho}}{1 - \widehat{\rho}} \\
\Leftrightarrow \mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}] &\leq \frac{1 + \frac{\mu_0}{\mu^*} \frac{\widehat{\rho}}{1-\widehat{\rho}}}{\frac{\mu_0}{\mu^*} \frac{1}{1-\widehat{\rho}} + \frac{\mu^* - \mu_0}{\mu^*} \frac{1}{\widehat{\rho}}} \\
&= \frac{\frac{\mu_0}{\mu^*}\widehat{\rho} + (1 - \widehat{\rho})}{\frac{\mu_0}{\mu^*}\widehat{\rho} + \left(1 - \frac{\mu_0}{\mu^*}\right)(1 - \widehat{\rho})} \widehat{\rho}. \qquad (2)
\end{aligned}$$

The right-hand side is strictly increasing in $\widehat{\rho}$. Note that

$$U_{\mathrm{R}}(\widehat{\rho}, \iota) = \frac{\mu^* - \mu_0}{\mu^*} \left(\frac{1}{\widehat{\rho}} - 1\right) \iota(1) = \frac{\mu^* - \mu_0}{\mu^*} \rho_0 \left(\frac{1}{\widehat{\rho}} - \frac{1}{\mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}]}\right),$$

where we used the fact that $\rho_0 = \widehat{\rho}\iota(\widehat{\rho}) + 1 \cdot \iota(1) = \mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}]\iota(\rho \geq \widehat{\rho})$. Suppose that (2) does not bind. Then, the Receiver can choose a lower $\widehat{\rho}$ while keeping $\mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}]$ constant by changing ι appropriately which would strictly increase his payoff. Thus, (2) must bind in any limited commitment equilibrium. Then,

$$U_{\mathrm{R}}(\widehat{\rho}, \iota) = \frac{\mu^* - \mu_0}{\mu^*} \rho_0 \frac{1 - \widehat{\rho}}{\widehat{\rho}} \frac{1}{1 - \frac{\mu^* - \mu_0}{\mu^*} \widehat{\rho}},$$

which is strictly decreasing in $\widehat{\rho}$. Since $\mathbb{E}_\iota[\rho|\rho \geq \widehat{\rho}] \geq \rho_0$, using (2) gives

$$\frac{\frac{\mu_0}{\mu^*}\widehat{\rho} + (1 - \widehat{\rho})}{\frac{\mu_0}{\mu^*}\widehat{\rho} + \left(1 - \frac{\mu_0}{\mu^*}\right)(1 - \widehat{\rho})} \widehat{\rho} \geq \rho_0 \Leftrightarrow 0 \geq \widehat{\rho}^2 - \frac{\mu^* + (\mu^* - 2\mu_0)\rho_0}{\mu^* - \mu_0} \widehat{\rho} + \rho_0, \qquad (3)$$

which is equivalent to

$$\widehat{\rho} \in \left[\frac{\mu^* + (\mu^* - 2\mu_0)\rho_0 + \sqrt{(1 - \rho_0)\left((\mu^*)^2 - (\mu^* - 2\mu_0)^2 \rho_0\right)}}{2(\mu^* - \mu_0)}, 1\right].$$

We also need that $\widehat{\rho} \geq \underline{\rho}$ and we obtain $\widehat{\rho}^{\circ}$ by noting that

$$\frac{\mu^* + (\mu^* - 2\mu_0)\,\rho_0 + \sqrt{(1-\rho_0)\left((\mu^*)^2 - (\mu^* - 2\mu_0)^2\,\rho_0\right)}}{2\,(\mu^* - \mu_0)} \geq \underline{\rho} = \frac{\mu^* - \mu_0}{\mu^*\,(1-\mu_0)}$$
$$\Leftrightarrow \frac{(2-\mu^*)\,\mu^* - \mu_0}{(2-\mu^*)\,\mu^*\,(1-\mu_0)} \leq \rho_0.$$

That the Receiver's payoff is strictly positive follows from the fact that Sender always chooses $\widehat{\rho}^{\circ} < 1$. ■

References

Aumann, Robert J., and Michael Maschler. 1968. "Repeated games of incomplete information: the zero-sum extensive case." *Mathematica*.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review*, 101(6): 2590–2615.