"Giving Advice *vs.* Making Decisions: Transparency, Information, and Delegation"

Online Appendix

A Definitions and Proofs

A.1 The Informational Environment

The set of states of nature is denoted by $\Theta = [0,1]$, and the state of nature, $\theta \in \Theta$, is determined according to the Uniform distribution on $[0,1]$.[30] Upon realization of θ, each individual $i \in \{A, P\}$ receives a conditionally independent (and private) signal $s_i \in \{0,1\}$ according to the following probability mass function:

$$\Pr[s_i = x|\theta] = \begin{cases} 1-\theta & \text{if } x = 0, \\ \theta & \text{if } x = 1. \end{cases}$$

After observing a single signal $s_i \in \{0,1\}$, i's posterior beliefs about θ are characterized by the following probability density function:

$$g_i(t|s_i) = \begin{cases} 2(1-t) & \text{if } s_i = 0, \\ 2(t) & \text{if } s_i = 1. \end{cases}$$

More generally, player i's posterior beliefs conditional upon observing m signals $\{s_1, \ldots, s_m\}$ with $k = \sum_{j=1}^{m} s_j$ (*i.e.*, k occurrences of $s = 1$ and $m - k$ occurrences of $s = 0$) is characterized by a Beta$(k+1, m-k+1)$ distribution, so that

$$\begin{aligned} E(\theta|k,m) &= \frac{k+1}{m+2}, \text{ and} \\ V(\theta|k,m) &= \frac{(k+1)(m-k+1)}{(m+2)^2(m+3)}. \end{aligned}$$

[30] This assumption greatly simplifies the calculations and allows us to focus on the key institutional and strategic trade-offs. In addition, note that the uniform distribution is a useful baseline, as it maximizes the *ex ante* informativeness of each agent's signal. In other words, this is the case in which information aggregation is most important to all agents from an *ex ante* perspective. Accordingly, this baseline amplifies the importance of our results when they indicate that information is not aggregated in equilibrium or that optimal institutional design limits information aggregation.

Accordingly, the optimal policy choice for a player $i \in N$, given k and m, is

$$y_i^*(k,m) = \frac{k+1}{m+2} + \beta_i.$$

A.2 Strategies

A (pure) strategy for P is $\sigma_P = (a, z_P)$, where $a \in [0,1]$ is a feasible assignment of authority to A (and P's authority is $1-a$), and $z_P : \mathbf{R} \to \mathbf{R}$ maps observed policy choices by A into policy choices by P. A pure strategy for A is $\sigma_A = z_i : [0,1] \times \{0,1\} \to \mathbf{R}$ specifies a policy choice, given A's authority a and signal $s_A \in \{0,1\}$.

Our notion of equilibrium throughout is perfect Bayesian equilibrium.

A.3 Beliefs

A's beliefs are straightforward in our setting: we require them to be consistent with the informational structure defined above and the strategies of P. P's beliefs are technically more complex, in that P has a continuum of information sets. When considering separating equilibria, we restrict attention to those in which A assigns positive probability only to the sequentially rational policy choices when choosing y_A. We provide further justification for this refinement below, and then consider equilibria ruled out by it.

Taking this refinement as given for the moment, it follows from Equation (2) that in a separating equilibrium, P assigns positive probability only to the following policies:

$$y_A \quad \in \quad \{\,{}^1\!/_3 + \beta_A,\ {}^2\!/_3 + \beta_A\} \tag{14}$$

in the opaque case, and

$$y_A \quad \in \quad \{\,{}^1\!/_4 + \beta_A,\ {}^1\!/_2 + \beta_A,\ {}^3\!/_4 + \beta_A\} \tag{15}$$

in the top-down transparency case.

Now consider P's beliefs about A's signal, as induced by A's policy choice. Denote this belief as $h(y_A) = \Pr[s_A = 1|y_A]$. Suppose without loss of generality that $\beta_A \geq \beta_P$.[31] Given our equilibrium refinement, these beliefs are defined as follows:

$$h(y_A) = \begin{cases} \begin{cases} 0 & \text{for } y_A < 2/3 + \beta_A, \\ 1 & \text{for } y_A \geq 2/3 + \beta_A. \end{cases} & \text{in the opaque case.} \\ \begin{cases} 0 & \text{for } y_A < \frac{s_P+2}{4} + \beta_A, \\ 1 & \text{for } y_A \geq \frac{s_P+2}{4} + \beta_A. \end{cases} & \text{in the top-down transparent case.} \end{cases} \tag{16}$$

A.4 The Agent's Best Response

The following proposition justifies our focus on the choices defined in (14) and (15), above.

Proposition 4 *Suppose that $\beta_A > \beta_P$, $\alpha_A > 0$, and the principal's beliefs satisfy Equation 16. Then the agent's best response y_A^* satisfies the restrictions in* (14) *in the opaque case and* (15) *in the top-down transparent case.*

Proof: Given the principal has beliefs that satisfy (16), A's expected payoff from choosing policy y_A after observing $s_A = 0$ is

$$U_A(y_A|h, s_A = 0) = \begin{cases} -(\beta_P - \beta_A)^2 - \alpha_A(y_A - \frac{1}{3} - \beta_A)^2 & \text{if } y_A < 2/3 + \beta_A, \\ -(\beta_P + \frac{1}{4} - \beta_A)^2 - \alpha_A(y_A - \frac{1}{3} - \beta_A)^2 & \text{if } y_A \geq 2/3 + \beta_A. \end{cases}$$

Note that, except at $y_A = 2/3 + \beta_A$, $U_A(y_A|h, s_A = 0)$ varies with y_A only through the term $\alpha_A(y_A - \frac{1}{3} - \beta_A)^2$. Accordingly, it follows that, if $y_A^* < 2/3 + \beta_A$, then $y_A^* = 1/3 + \beta_A$ since $\alpha_A(y_A - \frac{1}{3} - \beta_A)^2$ is maximized by this choice and, since $\alpha_A(y_A - \frac{1}{3} - \beta_A)^2$ is decreasing in y_A for $y_A > 1/3 + \beta_A$, if $y_A^* \geq 2/3 + \beta_A$, then $y_A^* = 2/3 + \beta_A$. We omit derivation of the case when $s_A = 1$, as it is symmetric.

Turning to the top-down transparency case, given the principal has beliefs that satisfy (16), A's

[31] The arguments that follow are symmetric in case $\beta_A < \beta_P$.

expected payoff from choosing policy y_A after observing $s_A = 0$ (and leaving s_P arbitrary) is

$$U_A(y_A|h, s_A = 0, s_P) = \begin{cases} -(\beta_P - \beta_A)^2 - \alpha_A(y_A - \frac{1+s_P}{4} - \beta_A)^2 & \text{if } y_A < \frac{s_P+2}{4} + \beta_A, \\ -(\beta_P + \frac{1}{5} - \beta_A)^2 - \alpha_A(y_A - \frac{1+s_P}{4} - \beta_A)^2 & \text{if } y_A \geq \frac{s_P+2}{4} + \beta_A. \end{cases}$$

The reasoning explained for the opaque case can be applied immediately to this case as well, and leads to the analogous conclusion that if $y_A^* < \frac{s_P+2}{4} + \beta_A$, then $y_A^* = \frac{1+s_P}{4} + \beta_A$ and if $y_A^* \geq \frac{s_P+2}{4} + \beta_A$, then $y_A^* = \frac{s_P+2}{4} + \beta_A$.

Accordingly, the beliefs defined in (16) induce the agent's best response (regardless of whether the incentive compatibility conditions hold or not) to be restricted to the choices defined in (14) and (15), as was to be shown. ∎

A.5 The Equilibrium Set

There are three classes of equilibria in this model:

1. *Pooling equilibria.* In a pooling equilibrium, A's decision is uninformative in the sense that P's beliefs are independent of y_A. Such a profile can be an equilibrium if and only if $\alpha_A = 0$, as otherwise such beliefs give the agent a strict incentive to set y_A "truthfully" based on his or her signal.

2. *Separating equilibria.* In a separating equilibrium, P's beliefs (correctly) assign probability one to the signal that was observed by the agent based on y_A. Such a profile can be an equilibrium only if there are exactly two choices of y_A that are observed with positive probability on the equilibrium path of play. Equation (16) refines this set of equilibria, a point to which we return below (Section A.6).

3. *Semi-separating equilibria.* A "semi-separating" profile is one in which A uses a strategy of

the following form:

$$y^*(s_A) = \begin{cases} y_0 & \text{with probability } p \text{ if } s_A = 0 \text{ and with probability } 1-q \text{ if } s_A = 1, \\ y_1 & \text{with probability } 1-p \text{ if } s_A = 0 \text{ and with probability } q \text{ if } s_A = 1. \end{cases} \tag{17}$$

Such a strategy can be part of an equilibrium profile only if either $p = 0$ or $q = 0$.[32] Furthermore, given our normalization that $\beta_A > \beta_P$, such a profile can be part of an equilibrium profile only if $q = 0$.

Any strategy-belief profile not falling into one of those three categories cannot be an equilibrium in this setting: unless $\alpha_A = 0$, the agent must be assign positive probability to exactly two choices of y_A in equilibrium.

A.6 What Equilibria Are Eliminated by the Refinement in Equation (16)?

As stated above, we consider only those separating equilibria in which the principal holds beliefs consistent with Equation (16). There are two qualitative cases to consider when asking what other equilibria are ruled out by our refinement, corresponding to $\alpha_A = 0$ and $\alpha_A > 0$, respectively. We treat each in turn.

- **$\alpha_A = 0$.** In this case, the game is equivalent to a cheap talk game. Accordingly, when the incentive compatibility conditions are satisfied, there are uncountably infinite separating equilibria (including ones in which the agent might use more than two "messages"). However, these equilibria are all payoff equivalent to both players, and accordingly, eliminating them is irrelevant. Semi-separating and pooling equilibria—which are not payoff-equivalent to the separating equilibrium considered in our analysis—can also exist.

  However, when the incentive compatibility conditions for the separating equilibrium are satisfied, it Pareto dominates (in terms of ex ante expected payoffs) any other, non-separating equilibrium.[33] The principal's ex ante preferences are clear because, conditional on the agent's

[32] Of course, if $p = q = 0$, then this is a separating strategy.

[33] This is obvious from Crawford and Sobel (1982), but for completeness we include another argument demonstrating this here.

signal, s_A, the principal prefers a separating equilibrium regardless of the realization of the signal. To see that the agent also prefers a separating profile from an ex ante perspective, Consider any strategy profile where there is some policy z such that $\Pr[y_A = z|s_A = 0] = p_0 > 0$ and $\Pr[y_A = z|s_A = 1] = p_1 > 0$. When the principal observes z, his or sequentially rational response is

$$y_P^*(z, s_P) = \begin{cases} \beta_P + \frac{p_0+p_1}{4p_0+2p_1} & \text{if } s_P = 0, \\ \beta_P + \frac{p_0+3p_1}{2p_0+4p_1} & \text{if } s_P = 1. \end{cases}$$

Substituting $\alpha_A = 0$ and $\alpha_P = 1$, the agent's state-unconditional expected payoff from playing z in equilibrium (*i.e.*, prior to observing s_A) is

$$-(\beta_A - \beta_P)^2 - \frac{1}{2}\int_0^1 \left[\left(\beta_P + \frac{p_0 + p_1}{4p_0 + 2p_1} - \theta - \beta_A\right)^2 (1 - \theta) + \left(\beta_P + \frac{p_0 + 3p_1}{2p_0 + 4p_1} - \theta - \beta_A\right)^2 \theta\right] d\theta.$$

which reduces to

$$-(\beta_A - \beta_P)^2 - \frac{1}{24}\left(\frac{3p_0^2 + 2p_1p_0 + p_1^2}{(2p_0 + p_1)^2} + \frac{p_0^2 + 2p_1p_0 + 3p_1^2}{(p_0 + 2p_1)^2}\right) \tag{18}$$

This is maximized by $p_0 = p_1$ and achieves a value of $-(\beta_A - \beta_P)^2 - \frac{1}{18}$. This is, of course, equivalent to the pooling equilibrium. The agent's ex ante payoff from the separating equilibrium is

$$-(\beta_A - \beta_P)^2 - \frac{1}{24}.$$

Note, also, that (18) must equal the agent's unconditional, ex ante payoff if the strategy is part of an equilibrium because the agent must be indifferent between z and any other policy he or she chooses with positive probability. Thus, in ex ante terms, the agent strictly prefers a separating equilibrium.

- $\alpha_A > 0$**.** In this case, there is no equilibrium in which the principal does not condition his or her policy on the agent's policy. If the principal ignored the agent's decision, then $\alpha_A > 0$ implies that the agent would have a strict best response of setting $y_A = \frac{s_A+1}{3} + \beta_A$ in the opaque case and $y_A = \frac{s_A+s_P}{4} + \beta_A$ in the transparent case. Of course, if the agent does that, then the principal should not ignore the agent's decision when making his or her own policy choice.

Moving to non-babbling equilibria, matters become more complicated. As with the cheap-talk case of $\alpha_A = 0$, the existence of a semi-separating equilibrium for a given α_A and β implies that an ex ante Pareto superior separating equilibrium also exists, so we set semi-separating equilibria aside. However, the real complication arises because there are other separating equilibria.

Specifically, with the proper specification of beliefs, it is always possible to construct a separating equilibrium, *regardless of the players' biases*. The next proposition, drawn from Patty and Penn (2014), demonstrates that beliefs can always be properly constructed so that it is incentive compatible for the agent to truthfully reveal his or her signal through his or her policy choice. The proposition applies to the opaque information case, but the logic and calculations can be straightforwardly applied to the top-down transparency case.

Proposition (Patty and Penn (2014)) *For any $\alpha_A > 0$ and $(\beta_A, \beta_P) \in \mathbf{R}^2$, there exists a belief $h(\cdot|\alpha_A, \beta_A, \beta_P)$ such that a separating equilibrium exists.*

Proof: Suppose that $\beta_A > \beta_P$ and that there exists some $\tau \geq 0$ such that P's sequentially rational strategy as a function of his or her signal, s_P, and the policy chosen by A, y_A, given beliefs $h(y_A|\alpha_A, \beta_A, \beta_P)$, is

$$y_P(y_A;\tau) = \begin{cases} \frac{s_P+1}{4} & \text{if } y_A < \frac{2}{3} + \tau + \beta_A, \\ \frac{s_P+2}{4} & \text{if } y_A \geq \frac{2}{3} + \tau + \beta_A. \end{cases} \tag{19}$$

Given the behavior described by (19), there are only 2 policy choices that are potentially optimal for A *and informative to P*: $y_A \in \{\frac{1}{3} + \beta_A, \frac{2}{3} + \tau + \beta_A\}$. Let $y_A^0 = \frac{1}{3} + \beta_A$ and $y_A^1 = \frac{2}{3} + \tau + \beta_A$ denote these two policies. Note that $y_A^1 - y_A^0 = \frac{1}{3} + \tau$. The expected payoff difference between these two choices is

$$\begin{aligned} U_A(y_A^0|s_A) - U_A(y_A^1|s_A) &= \begin{cases} \alpha_P\left(\beta_A - \beta_P + \frac{1}{4}\right)^2 + \alpha_A\left(\frac{1}{3} + \tau\right)^2 - \alpha_P\left(\beta_A - \beta_P\right)^2 & \text{if } s_A = 0, \\ \alpha_P\left(\beta_A - \beta_P\right)^2 + \alpha_A\tau^2 - \alpha_A\left(\frac{1}{3}\right)^2 - \alpha_P\left(\beta_A - \beta_P - \frac{1}{4}\right)^2 & \text{if } s_A = 1, \end{cases} \\ &= \begin{cases} \alpha_P\left(\frac{1}{16} - \frac{1}{2}(\beta_A - \beta_P)\right) + \alpha_A\left(\frac{1}{3} + \tau\right)^2 & \text{if } s_A = 0, \\ \alpha_A\left(\tau^2 - \frac{1}{9}\right) - \frac{\alpha_P}{16} - \frac{1}{2}\alpha_P(\beta_A - \beta_P) & \text{if } s_A = 1. \end{cases} \end{aligned}$$

Because $\beta_A > \beta_P$, consider first the incentive compatibility condition in the case of $s_A = 0$ (this is the only one that can bind for sufficiently small τ). In this case, truthfulness is incentive compatible with these beliefs if and only if

$$
\begin{aligned}
\frac{\alpha_P}{16} + \alpha_A\left(\frac{1}{3}+\tau\right)^2 &\geq \frac{\alpha_P}{2}(\beta_A - \beta_P), \\
\frac{2}{\alpha_P}\left[\frac{\alpha_P}{16} + \alpha_A\left(\frac{1}{3}+\tau\right)^2\right] &\geq \beta_A - \beta_P, \\
\frac{1}{8} + \frac{2\alpha_A}{\alpha_P}\left(\frac{1}{3}+\tau\right)^2 &\geq \beta_A - \beta_P, \qquad (20)
\end{aligned}
$$

so that, as intuition suggests, for any fixed β_A and β_P, truthfulness is incentive compatible—under the supposition that player 1 must choose between y_A^0 and y_A^1—for "sufficiently demanding" beliefs (*i.e.*, for sufficiently large values of τ).

However, as alluded to above, setting τ too large can lead to a violation of incentive compatibility when $s_A = 1$. In particular, player 1 could choose his or her sequentially rational policy following $s_A = 1$ even if that leads to player 2 inferring that $s_A = 0$. Doing so optimally (given $\tau \geq 0$) involves choosing $y_A = \frac{s_A+2}{3} + \beta_A$. The resulting incentive compatibility condition in this situation is based on the following expected payoff difference calculation:

$$
\begin{aligned}
U_A\left(\frac{s_A+2}{3} + \beta_A | s_A = 1\right) - U_A(y_A^1 | s_A = 1) &= \alpha_P\left(\beta_A - \beta_P\right)^2 + \alpha_A\tau^2 - \alpha_P\left(\beta_A - \beta_P - \frac{1}{4}\right)^2, \\
&= \alpha_A\tau^2 - \frac{\alpha_P}{16} - \frac{\alpha_P}{2}(\beta_A - \beta_P),
\end{aligned}
$$

so that incentive compatibility is satisfied only if

$$
\tau \leq \sqrt{\frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}|\beta_A - \beta_P|\right]}, \qquad (21)
$$

So, suppose that $\tau = \sqrt{\frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}(\beta_A - \beta_P)\right]}$. Then substituting this into inequality 20,

$$
\frac{1}{8} + \frac{2\alpha_A}{\alpha_P}\left(\frac{1}{3} + \sqrt{\frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}(\beta_A - \beta_P)\right]}\right)^2 \geq |\beta_A - \beta_P|,
$$

and letting $M \equiv \frac{1}{8} + \frac{2\alpha_A}{\alpha_P}\left(\frac{1}{3} + \sqrt{\frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}(\beta_A - \beta_P)\right]}\right)^2$,

$$\begin{aligned} M > \frac{1}{8} + \frac{2\alpha_A}{\alpha_P}\left(\frac{1}{9} + \frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}(\beta_A - \beta_P)\right]\right) &\geq |\beta_A - \beta_P|, \\ \frac{1}{4} + \frac{2\alpha_A}{9\alpha_P} + \beta_A - \beta_P &\geq \beta_A - \beta_P. \end{aligned}$$

Accordingly, setting $\tau = \sqrt{\frac{\alpha_P}{\alpha_A}\left[\frac{1}{16} + \frac{1}{2}(\beta_A - \beta_P)\right]}$ implies that inequalities 20 and 21 are simultaneously satisfied. ∎

We now turn to the question of the principal's preferences over the equilibria for $\alpha > 0$. Specifically, we show that *the principal prefers the equilibrium we consider (defined by α_A^{*O} in Proposition 6) to all of the other equilibria covered by Proposition A.6.*

A.7 The Principal's Preferences over Equilibria

We establish the optimality (from the principal's perspective) of the equilibria considered in the paper in two steps. First, we consider the separating equilibria in which the principal's beliefs do not satisfy Equation (16). Then, we demonstrate that any semi-separating equilibrium is dominated by a separating equilibrium from the perspective of the principal's ex ante equilibrium expected payoff.

A.7.1 Equilibria that do not satisfy Equation (16)

While incentive compatibility can always be obtained, it is not necessarily the case that obtaining incentive compatibility is in P's interests. That is, it is not clear that the value of the agent's information for the principal in setting his or her own policy is sufficient to outweigh the distortion that the agent must impose on his or her policy choice to sustain incentive compatibility of truthfulness. Based on the analysis above, we can establish an upper bound on the preference divergence between the principal and agent, above which the principal receives a strictly higher payoff from setting $\alpha_A = 0$ and the two players playing a babbling equilibrium of the resulting cheap talk signaling game.

Proposition 5 *For all β, the Principal's equilibrium expected payoff from choosing $\alpha_A^{*O}(\beta)$, given that the equilibrium strategies and beliefs satisfy Equation* (16) *is strictly higher than any equilibrium following any other choice α_A.*

Proof: To establish this bound, note the "cheapest" pure strategy separating equilibrium in terms of α, β is constructed by setting τ so as to satisfy inequality (20) with equality. This value is

$$\tau^*(\alpha,\beta) = \begin{cases} 0 & \text{if } \beta_A - \beta_P \leq \frac{1}{8} + \frac{2\alpha_A}{9\alpha_P}, \\ \sqrt{\frac{\alpha_P}{\alpha_A}\left(\frac{1}{2}(\beta_A-\beta_P) - \frac{1}{16}\right)} & \text{otherwise.} \end{cases} \tag{22}$$

Note that player 2's expected payoff from the pure strategy separating equilibrium supported by the beliefs based on equation (19), with $\tau = \tau^*(\alpha,\beta)$ as defined in (22), is

$$\begin{aligned} U_P(\tau^*(\alpha,\beta)) &= -\left[\frac{\alpha_P}{24} + \alpha_A\left(\frac{1}{18} + \frac{(\tau^*(\alpha,\beta)+\beta_A-\beta_P)^2 + (\beta_A-\beta_P)^2}{2}\right)\right], \\ &= -\left[\frac{\alpha_P}{24} + \alpha_A\left(\frac{1}{18} + (\beta_A-\beta_P)^2 + \frac{\tau^*(\alpha,\beta)^2}{2} + \tau^*(\alpha,\beta)(\beta_A-\beta_P)\right)\right] \end{aligned} \tag{23}$$

Presuming without loss of generality that $\beta_A - \beta_P > 1/8$, substituting

$$\tau^*(\alpha,\beta) = \sqrt{\frac{\alpha_P}{\alpha_A}\left(\frac{1}{2}(\beta_A-\beta_P) - \frac{1}{16}\right)},$$

$\alpha_P \equiv 1-\alpha_A$, and $B \equiv \beta_A - \beta_P$ into (23) yields

$$\begin{aligned} U_P(\tau^*(\alpha,\beta)) &= -\frac{1-\alpha_A}{24} - \frac{\alpha_A}{18} - \alpha_A B^2 - (1-\alpha_A)\left(\frac{1}{4}B - \frac{1}{32}\right) - B\sqrt{(1-\alpha_A)\alpha_A\left(\frac{1}{2}B - \frac{1}{16}\right)}, \\ &= \alpha_A\left(\frac{B}{4} - B^2 - \frac{13}{288}\right) - \frac{1}{4}B - \frac{1}{96} - B\sqrt{(1-\alpha_A)\alpha_A\left(\frac{1}{2}B - \frac{1}{16}\right)}, \end{aligned}$$

which is decreasing in α_A for all $\alpha_A \in [0, 1/2]$.

The principal's equilibrium payoff from the equilibrium constructed in our analysis for the opaque case (Proposition 6) is

$$U_P(\alpha_A^{*O}(\beta);\beta) = -\frac{1}{24} - \left(\frac{8(\beta_A - \beta_P) - 1}{7/9 + 8(\beta_A - \beta_P)}\right)\left(\frac{1}{72} + (\beta_A - \beta_P)^2\right),$$

and comparing this with (24) evaluated at $\alpha_A = 0$ yields the following:[34]

$$\begin{aligned} U_P(\alpha_A^{*O}(\beta);\beta) &\geq U_P(\tau^*(\alpha,\beta)), \\ -\frac{1}{24} - \left(\frac{8(\beta_A - \beta_P) - 1}{7/9 + 8(\beta_A - \beta_P)}\right)\left(\frac{1}{72} + (\beta_A - \beta_P)^2\right) &\geq -\frac{|\beta_A - \beta_P|}{4} - \frac{1}{96}, \\ &\Rightarrow \frac{1}{8} \leq |\beta_A - \beta_P| \leq \frac{3 + \sqrt{15}}{24} \approx 0.286 > \rho^* \approx 0.201. \end{aligned}$$

These inequalities imply (as shown in Proposition 6) that, in the opaque case, the preference divergence must be greater than the maximal amount of divergence for which the principal prefers to delegate any authority to the agent. To verify this further, note that the principal's expected equilibrium payoff in the babbling equilibrium of the cheap talk case ($\alpha_A = 0$) is greater than (24) evaluated at $\alpha_A = 0$ whenever

$$\begin{aligned} -\frac{1}{18} &\geq -\frac{|\beta_A - \beta_P|}{4} - \frac{1}{96}, \\ &\Rightarrow |\beta_A - \beta_P| \geq \frac{13}{72} \approx 0.18. \end{aligned}$$

Thus, the principal strictly prefers delegating $\alpha_A^{*O}(\beta)$ to the agent and playing the equilibrium described in the article to delegating any other value α_A and playing the equilibrium with beliefs described by supported by the beliefs based on equation (19) with $\tau = \tau^*(\alpha_A, \beta)$, as defined in (22).

∎

[34]Evaluating (24) at α_A is actually inconsistent with equilibrium for $|\beta_A - \beta_P| > 1/8$, but this values equals the supremum of the principal's best possible equilibrium payoff using the τ^* beliefs based on equation (19), which is sufficient for our purposes here.

A.7.2 Semi-separating Equilibria

When $\alpha_A > 0$, any equilibrium must involve the agent assigning positive probability to exactly two distinct policy choices. Therefore, suppose that the agent plays a strategy of the following form for some value $p \in [0,1]$:

$$y^*(s_A) = \begin{cases} \beta_A + 1/3 & \text{with probability } 1-p \text{ if and only if } s_A = 0, \\ \beta_A + 2/3 & \text{with probability } p \text{ if } s_A = 0 \text{ or with probability 1 if } s_A = 1. \end{cases} \tag{24}$$

Because

$$\begin{aligned} \Pr[s_A = 1 | s_P = 1] &= \frac{\int_0^1 \theta^2 d\theta}{\int_0^1 \theta^2 d\theta + \int_0^1 \theta(1-\theta) d\theta}, \\ &= 2/3, \end{aligned}$$

the principal's "on-the-path" beliefs regarding s_A are

$$h(s_A|y_A, s_P) \equiv \Pr[s_A = 1|y_A, s_P] = \begin{cases} 0 & \text{if } y_A = \beta_A + 1/3, \\ \frac{2}{2+p} & \text{if } y_A = \beta_A + 2/3 \text{ and } s_P = 1, \text{ and} \\ \frac{1}{1+2p} & \text{if } y_A = \beta_A + 2/3 \text{ and } s_P = 0. \end{cases}$$

$$U_P(z|y_A = \beta_A + 2/3, s_P) = \begin{cases} -\left[\frac{2}{2+p}\left((z-\frac{3}{4})^2 + \frac{3}{80}\right) + \frac{p}{2+p}\left((z-\frac{1}{2})^2 + \frac{1}{20}\right)\right] & \text{if } s_P = 1, \text{ and} \\ -\left[\frac{1}{1+2p}\left((z-\frac{1}{2})^2 + \frac{1}{20}\right) + \frac{2p}{1+2p}\left((z-\frac{1}{4})^2 + \frac{3}{80}\right)\right] & \text{if } s_P = 0. \end{cases}$$

$$\frac{\partial}{\partial z} U_P(z|y_A = \beta_A + 2/3, s_P) = \begin{cases} -\left[\frac{1}{1+2p} 2(z-\frac{1}{2}) + \frac{2p}{1+2p} 2(z-\frac{1}{4})\right] & \text{if } s_P = 0, \text{ and} \\ -\left[\frac{2}{2+p} 2(z-\frac{3}{4}) + \frac{p}{2+p} 2(z-\frac{1}{2})\right] & \text{if } s_P = 1. \end{cases}$$

Solving the FOC for $s_P = 0$:

$$\begin{aligned} \frac{1}{1+2p} 2\left(z - \frac{1}{2}\right) + \frac{2p}{1+2p} 2\left(z - \frac{1}{4}\right) &= 0, \\ z &= \frac{1+p}{2(1+2p)}, \end{aligned}$$

and solving the FOC for $s_P = 1$:

$$\begin{aligned} \frac{2}{2+p} 2\left(z - \frac{3}{4}\right) + \frac{p}{2+p} 2\left(z - \frac{1}{2}\right) &= 0, \\ z &= \frac{3+p}{2(2+p)}. \end{aligned}$$

Setting $\frac{\partial}{\partial z} U_P(z | y_A = \beta_A + 2/3, s_P) = 0$ implies that $y_P^*(y_A = \beta_A + 2/3, s_P)$ is as follows:

$$y_P^*(y_A = \beta_A + 2/3, s_P) = \begin{cases} \frac{1+p}{2(1+2p)} + \beta_P & \text{if } s_P = 0, \text{ and} \\ \frac{3+p}{2(2+p)} + \beta_P & \text{if } s_P = 1. \end{cases}$$

Solving with an uninformed principal. If the principal receives no signal, the principal's beliefs after observing (on the path values of) y_A are

$$h(s_A | y_A) \equiv \Pr[s_A = 1 | y_A] = \begin{cases} 0 & \text{if } y_A = \beta_A + 1/3, \\ \frac{1}{1+p} & \text{if } y_A = \beta_A + 2/3. \end{cases}$$

Given these beliefs, the principal's sequentially rational response to observing $y_A = \beta_A + 2/3$ is derived as follows:

$$U_P(z | y_A = \beta_A + 2/3) = -\left[\frac{1}{1+p} \left(\left(z - \frac{2}{3} - \beta_P\right)^2 + \frac{1}{18} \right) + \frac{p}{1+p} \left(\left(z - \frac{1}{3} - \beta_P\right)^2 + \frac{1}{18} \right) \right]$$

$$\frac{\partial}{\partial z}U_P(z|y_A = \beta_A + {}^2\!/_3) \quad = \quad -\left[\frac{2}{1+p}\left(z - \frac{2}{3} - \beta_P\right) + \frac{2p}{1+p}\left(z - \frac{1}{3} - \beta_P\right)\right]$$

Solving the FOC:

$$\begin{aligned}\frac{2}{1+p}\left(z - \frac{2}{3} - \beta_P\right) + \frac{2p}{1+p}\left(z - \frac{1}{3} - \beta_P\right) &= 0, \\ z &= \beta_P + \frac{2+p}{3(1+p)},\end{aligned}$$

Thus,

$$y_P^*(y_A = \beta_A + {}^2\!/_3) \quad = \quad \beta_P + \frac{2+p}{3(1+p)}.$$

Solving for the Agent's IC Conditions. Omitting the variance terms, the agent's expected payoff from choosing $y_A = \beta_A + {}^2\!/_3$ when $s_A = 0$ is

$$U_A(\beta_A + {}^2\!/_3|s_A = 0) = -\left[\alpha_P\left(\beta_P + \frac{2+p}{3(1+p)} - {}^1\!/_3 - \beta_A\right)^2 + \frac{\alpha_A}{9}\right]$$

and the agent's expected payoff from choosing $y_A = \beta_A + {}^1\!/_3$ when $s_A = 0$ is

$$U_A(\beta_A + {}^1\!/_3|s_A = 0) = -\alpha_P\left(\beta_P - \beta_A\right)^2.$$

The agent must be indifferent between these two policies after $s_A = 0$ for the semi-separating strategy to be a best response. Thus, it must be the case that the equilibrium mixing probability, p^*, satisfy the following:

$$\alpha_P\left(\beta_P + \frac{2+p^*}{3(1+p^*)} - {}^1\!/_3 - \beta_A\right)^2 + \frac{\alpha_A}{9} \quad = \quad \alpha_P\left(\beta_P - \beta_A\right)^2.$$

Substituting $\alpha_P = 1 - \alpha_A$ to solve for p^* only as a function of α_A, β_A, and β_P yields

$$p^*(\alpha_A, \beta) = \frac{3(1-\alpha_A)(\beta_A - \beta_P) - \alpha_A \pm \sqrt{(1-\alpha_A)\left((1-\alpha_A)\left(9(\beta_A-\beta_P)^2\right) - \alpha_A\right)}}{\alpha_A}.$$

Equivalently, α_A is uniquely identified by a choice of p^* (given β_P and β_A) as follows:

$$\alpha_A(p^*) = \frac{6\beta_A(p+1) - 6\beta_P(p+1) - 1}{6\beta_A + p^2 + 6\beta_A p - 6\beta_P(p+1) + 2p}.$$

The Principal's Expected Welfare. By choosing α_A, the principal can effectively determine the equilibrium mixing probability for the agent. The principal's equilibrium ex ante expected payoff as a function of p is

$$\begin{aligned}
EU_P(p) &= -(1-\alpha_A(p))\int_0^1 \left(\frac{2+p}{3(1+p)} - \omega\right)^2 (\omega + p(1-\omega))\,\mathrm{d}\omega - (1-\alpha_A(p))\int_0^1 \left(\left(\frac{1}{3}-\omega\right)^2 (1-p)(1-\omega)\right)\mathrm{d}\omega \\
&\quad -\alpha_A(p)\int_0^1 \left(\beta_A + \frac{2}{3} - \beta_P - \omega\right)^2 (\omega + p(1-\omega))\,\mathrm{d}\omega - \alpha_A(p)\int_0^1 \left(\left(\beta_A + \frac{1}{3} - \beta_P - \omega\right)^2 (1-p)(1-\omega)\right)\mathrm{d}\omega, \\
&= -\beta_A^2 - \beta_P^2 + \frac{1}{36}\left(6\beta_A - \frac{\left(p^2-2\right)p^2}{6\beta_A + p^2 + 6\beta_A p - 6\beta_P(p+1) + 2p} + p^2 + 6\beta_P(12\beta_A + p - 1) - 6\beta_A p - 2p - 2\right)
\end{aligned}$$

Substituting $\beta_P = 0$ without loss of generality, this reduces to

$$EU_P(p) = -\beta_A^2 + \frac{1}{36}\left(6\beta_A(1-p) - \frac{\left(p^2-2\right)p^2}{6\beta_A + p^2 + 6\beta_A p + 2p} + p^2 - 2p - 2\right).$$

Notice that

$$\frac{\partial EU_P(p)}{\partial p} = -\left[\frac{(3\beta_A+1)\left(18\beta_A^2 + p^2\left(6\left(3\beta_A^2+\beta_A\right)+1\right) + 6p\beta_A(6\beta_A+1)\right)}{9\left(6\beta_A + p^2 + 6\beta_A p + 2p\right)^2}\right],$$

where $\beta_A > \beta_P \equiv 0$ and $p \geq 0$ imply that the numerator is positive and the denominator is positive, so that

$$\frac{\partial EU_P(p)}{\partial p} < 0.$$

This implies that the principal wants to choose α_A so that the agent uses the minimum feasible equilibrium value of p, given β_A. Of course, the principal can achieve $p^* = 0$: this is simply the separating equilibrium examined in the article. Accordingly, the principal always strictly prefers the separating equilibrium and therefore prefers the babbling equilibrium to any mixed strategy equilibrium whenever he or she prefers the babbling equilibrium to the separating equilibrium. On a technical note, this is because $EU_P(1)$ does not equal the expected payoff of the babbling equilibrium: when $p^* = 1$, $\alpha^*(p^*) > 0$ for all $|\beta_A - \beta_P| > \frac{1}{12}$. But recall that for all $|\beta_A - \beta_P| < \frac{1}{8}$, $\alpha_A = 0$ yields a separating equilibrium.

Solving with an Informed Principal: The Opaque Case. Duplicating the analysis above with the principal having a private signal yields the following expression for $\alpha_A(p^*)$:

$$\alpha_A(p^*) = 1 - \frac{8(p^*+2)^2}{2(72\beta_A - 72\beta_P + 7) + p^*(120\beta_A + 24\beta_A p^* - 24\beta_P(p^*+5) + 5p^* + 20)}$$

and, setting $\beta_P = 0$, one obtains the following expression for the principal's expected payoff in a p-semi-separating equilibrium:

$$EU_P(p) = -\left[\frac{720\beta_A^3(p+2)(p+3) + 30\beta_A^2(8p^3 + 37p^2 + 36p - 18) + 10\beta_A(p^3 + 12p^2 + 26p + 24) + (p+2)^2(p+9)}{30\,(5(p)^2 + 24\beta_A(p+2)(p+3) + 20p + 14)}\right],$$

such that

$$\frac{\partial EU_P(p)}{\partial p} = -\frac{40(7\beta_A - 4) + 5760\beta_A^3(p+2)^3(p+4) + 480\beta_A^2(p+2)^2(3p^2 + 13p + 18) + 2\beta_A(37p^4 + 320p^3 + 1276p^2 + 1488p) + 5p^4 + 40p^3 + 102p^2 + 4p}{30\left(5p^2 + 24\beta_A(p+2)(p+3) + 20p + 14\right)^2}.$$

It can be verified that, for $\beta_A > 0.0566026$,

$$\frac{\partial EU_P(p)}{\partial p} < 0.$$

For $\beta_A \leq \frac{1}{8}$, the principal prefers the separating equilibrium that can be achieved with α_A, so that if the principal can not achieve a separating equilibrium in the cheap-talk case, the principal's payoff is strictly decreasing in p, implying that he or she strictly prefers $p = 0$ to any $p > 0$.

B Assorted Proofs

This appendix contains proofs of several claims made throughout the body of the paper.

B.1 Optimal Delegation

Proposition 6 *Letting* $\rho^* \approx 0.201$ *denote the first root of* $f(x) = 648x^3 - 81x^2 - 2$*, in the opaque case in which agent* A *does not observe the principal's information,* s_P*, prior to choosing* y_A*, the principal will delegate discretionary authority to the agent as follows:*

$$\alpha_A(\beta) = \begin{cases} 0 & \textit{if } (\beta_A - \beta_P) < \frac{1}{8}, \\ \frac{8(\beta_A-\beta_P)-1}{7/9+8(\beta_A-\beta_P)} & \textit{if } (\beta_A - \beta_P) \in \left[\frac{1}{8}, \rho^*\right], \\ 0 & \textit{if } (\beta_A - \beta_P) > \rho^*. \end{cases}$$

Proof: Note that if $|\beta_A - \beta_P| \leq \frac{1}{8}$, then agent A will be truthful regardless of α_A, so that the optimal choice for principal P is clear: $\alpha_A = 0$. Thus, presuming that $|\beta_A - \beta_P| > \frac{1}{8}$ (so that eliciting truthful communication requires delegating positive discretionary authority to the agent) and recalling that

$$\alpha_A^{*O}(\beta) \equiv \max\left[\frac{8(\beta_A - \beta_P) - 1}{7/9 + 8(\beta_A - \beta_P)}, 0\right]$$

denotes the minimal level of discretionary authority required to elicit truthful revelation from the message sender when the principal is known to be informed but his or her information is opaque to the message sender, the principal's expected payoff from inducing truthful revelation as cheaply as possible in this setting is

$$\begin{aligned} U_P(\alpha_A^{*O}(\beta);\beta) &= -\frac{1-\alpha_A^{*O}(\beta)}{24} - \frac{\alpha_A^{*O}(\beta)}{18} - \alpha_A^{*O}(\beta)(\beta_A-\beta_P)^2, \\ &= -\frac{1}{24} - \alpha_A^{*O}(\beta)\left(\frac{1}{72} + (\beta_A - \beta_P)^2\right), \\ &= -\frac{1}{24} - \left(\frac{8(\beta_A-\beta_P)-1}{7/9+8(\beta_A-\beta_P)}\right)\left(\frac{1}{72} + (\beta_A-\beta_P)^2\right), \end{aligned}$$

whereas the expected payoff from delegating zero discretionary authority is

$$U_P(\alpha_A = 0; \beta) \quad = \quad -\frac{1}{18},$$

so that the delegation of positive discretionary authority is in the principal's interest only if

$$\begin{aligned} \Delta U_P &\equiv U_P(\alpha_A^{*O}(\beta); \beta) - U_P(\alpha_A = 0; \beta), \\ &= \frac{1}{18} - \frac{1}{24} - \left(\frac{8(\beta_A - \beta_P) - 1}{7/9 + 8(\beta_A - \beta_P)}\right)\left(\frac{1}{72} + (\beta_A - \beta_P)^2\right) \\ &= \frac{1}{72} - \left(\frac{8(\beta_A - \beta_P) - 1}{7/9 + 8(\beta_A - \beta_P)}\right)\left(\frac{1}{72} + (\beta_A - \beta_P)^2\right) \\ &\geq 0 \end{aligned}$$

which holds if and only if

$$|\beta_A - \beta_P| \leq \rho^* \approx 0.201,$$

as was to be shown. ∎

Proposition 7 *In the top-down transparency case in which agent A observes principal P's information, s_P, prior to choosing y_A, principal P will delegate discretionary authority to agent A as follows:*

$$\alpha_A(\beta) = \begin{cases} 0 & \text{if } (\beta_A - \beta_P) < \frac{1}{8}, \\ 1 - \frac{1}{8(\beta_A - \beta_P)} & \text{if } (\beta_A - \beta_P) \in \left[\frac{1}{8}, \frac{3+\sqrt{41}}{48}\right], \\ 0 & \text{if } (\beta_A - \beta_P) > \frac{3+\sqrt{41}}{48}. \end{cases}$$

Proof: Recalling

$$\alpha_A^{*T}(\beta) = \max\left[0, 1 - \frac{1}{8(\beta_A - \beta_P)}\right]$$

denote the minimal level of discretionary authority required to elicit truthful revelation from the message sender, note first that principal P's optimal delegation is $\alpha_A = 0$ when $|\beta_A - \beta_P| \leq \frac{1}{8}$, because agent j's optimal choice is to be truthful even with no discretionary authority in those cases. Thus, presuming that $|\beta_A - \beta_P| > \frac{1}{8}$, the principal's expected payoff from inducing truthful revelation as cheaply as possible (*i.e.*, setting j's discretionary authority equal to $\alpha_A^{*T}(\beta)$) is

$$U_P(\alpha_A^{*T}(\beta);\beta) \quad = \quad -\frac{1}{24} - \left(1 - \frac{1}{8(\beta_A - \beta_P)}\right)(\beta_A - \beta_P)^2,$$

whereas the expected payoff from delegating zero discretionary authority is

$$U_P(\alpha_A = 0;\beta) \quad = \quad -\frac{1}{18},$$

so that the delegation of positive discretionary authority is in the principal's interest only if

$$U_P(\alpha_A^{*T}(\beta);\beta) - U_P(\alpha_A = 0;\beta) \quad = \quad \frac{1}{72} - \left(1 - \frac{1}{8(\beta_A - \beta_P)}\right)(\beta_A - \beta_P)^2 \geq 0, \qquad (25)$$

which holds only only if

$$|\beta_A - \beta_P| \leq \frac{1}{48}\left(3 + \sqrt{41}\right) \approx 0.196,$$

so that the principal will delegate positive discretionary authority equal to $\alpha^{**}(\beta)$ if and only if

$$|\beta_A - \beta_P| \in \left(\frac{1}{8}, \frac{3 + \sqrt{41}}{48}\right),$$

as was to be shown. ∎

B.2 Derivation of Equations 5 & 6 and Inequality 7

The Principal's Decision Problem. Given his or her information, (y_A, s_P), the principal's expected payoff from a policy choice y_P is

$$\begin{aligned}
U_P(y_P) &= -\int_0^1 \left((1-\alpha_A)(y_P-\theta-\beta_P)^2 + \alpha_A(y_A-\theta-\beta_P)^2\right) g(\theta|y_A,s_P)\mathrm{d}\theta, \\
&= -\int_0^1 (1-\alpha_A)\left(y_P^2+\beta_P^2+\theta^2-2(\theta(y_P-\beta_P)+y_P\beta_P)\right) g(\theta|y_A,s_P)\mathrm{d}\theta \\
&\quad -\int_0^1 \alpha_A\left(y_A^2+\beta_P^2+\theta^2-2(\theta(y_A-\beta_P)+y_A\beta_P)\right) g(\theta|y_A,s_P)\mathrm{d}\theta, \\
&= -(1-\alpha_A)\left(y_P^2+\beta_P^2-2y_P\beta_P\right) - \int_0^1 (1-\alpha_A)\left(\theta^2-2\theta(y_P-\beta_P)\right) g(\theta|y_A,s_P)\mathrm{d}\theta \\
&\quad -\alpha_A\left(y_A^2+\beta_P^2-2y_A\beta_P\right) - \int_0^1 \alpha_A\left(\theta^2-2\theta(y_A-\beta_P)\right) g(\theta|y_A,s_P)\mathrm{d}\theta, \\
&= -(1-\alpha_A)\left(y_P-\beta_P\right)^2 - (1-\alpha_A)\left(\mathrm{Var}[\theta|y_A,s_P]+E[\theta|y_A,s_P]^2-2E[\theta|y_A,s_P](y_P-\beta_P)\right) \\
&\quad -\alpha_A\left(y_A-\beta_P\right)^2 - \alpha_A\left(\mathrm{Var}[\theta|y_A,s_P]+E[\theta|y_A,s_P]^2-2E[\theta|y_A,s_P](y_A-\beta_P)\right), \\
&= -(1-\alpha_A)\left(\left(y_P-\beta_P\right)^2-2E[\theta|y_A,s_P](y_P-\beta_P)\right) \\
&\quad -\alpha_A\left(\left(y_A-\beta_P\right)^2-2E[\theta|y_A,s_P](y_A-\beta_P)\right) - \left(\mathrm{Var}[\theta|y_A,s_P]+E[\theta|y_A,s_P]^2\right).
\end{aligned}$$

The first order necessary condition for maximization of $U_P(y_P;\beta,y_A,s_P)$ with respect to y_P is

$$\frac{dU_P(y_P;\beta,y_A,s_P)}{dy_P} = -(1-\alpha_A)2\left(y_P-\beta_P-E[\theta|y_A,s_P]\right) = 0,$$

which yields

$$y_P^* = \beta_P + E[\theta|y_A,s_P]. \tag{26}$$

In the separating equilibrium considered here, the principal's beliefs, $E[\theta|y_A,s_P]$, satisfy the following:

$$E[\theta|y_A,s_P] = \begin{cases} 1/4 & \text{if } s_P=0 \text{ and } y_A=\beta_A+1/3, \\ 1/2 & \text{if } s_P=0 \text{ and } y_A=\beta_A+2/3, \\ 1/2 & \text{if } s_P=1 \text{ and } y_A=\beta_A+1/3, \\ 3/4 & \text{if } s_P=1 \text{ and } y_A=\beta_A+2/3. \end{cases} \tag{27}$$

Substituting (27) into (26) yields equation (4):

$$y_P^*(\tilde{s}_A, s_P) = \frac{1 + \tilde{s}_A + s_P}{4} + \beta_P.$$

We now turn to the agent's decision problem to obtain the agent's best response to the principal's beliefs (and sequentially rational behavior based on those beliefs) when choosing y_A based on s_A.

The Agent's Decision Problem. Let $h(s_P|\theta)$ denote the conditional probability of s_P, given θ. Then, letting $y_A^0 \equiv \beta_A + 1/3$ and $y^1 \equiv \beta_A + 2/3$, then the agent's expected payoffs from y_A^0 and y_A^1, respectively, given the agent's information, s_A, are[35]

$$\begin{aligned} U_A(y_A^0; s_A) &= -\int_0^1 \left[\sum_{s_P=0}^{1} \left((1-\alpha_A)\left(\frac{s_P+1}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A (y_A^0 - \theta - \beta_A)^2 \right) h(s_P|\theta) \right] g(\theta|s_A) \mathrm{d}\theta, \text{ and} \\ U_A(y_A^1; s_A) &= -\int_0^1 \left[\sum_{s_P=0}^{1} \left((1-\alpha_A)\left(\frac{s_P+2}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A (y_A^1 - \theta - \beta_A)^2 \right) h(s_P|\theta) \right] g(\theta|s_A) \mathrm{d}\theta. \end{aligned}$$

Focusing on the case of $s_A = 0$ (this is the "tempting" case for the agent when $\beta_A > \beta_P$) and noting that $h(s_P = 1|\theta) = \theta$, $h(s_P = 0|\theta) = 1 - \theta$, and $g(\theta|s_A) = 2(1-\theta)$,

$$\begin{aligned} U_A(y_A^0; 0) &= -\int_0^1 \left[\left((1-\alpha_A)\left(\frac{1}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A \left(\frac{1}{3} - \theta\right)^2 \right) (1-\theta) \right] 2(1-\theta) \mathrm{d}\theta \\ &\quad -\int_0^1 \left[\left((1-\alpha_A)\left(\frac{2}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A \left(\frac{1}{3} - \theta\right)^2 \right) \theta \right] 2(1-\theta) \mathrm{d}\theta, \\ &= -(1-\alpha_A)(\beta_A - \beta_P)^2 - \frac{\alpha_A}{72} - \frac{1}{24}, \\ U_A(y_A^1; 0) &= -\int_0^1 \left[\left((1-\alpha_A)\left(\frac{2}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A \left(\frac{2}{3} - \theta\right)^2 \right) (1-\theta) \right] 2(1-\theta) \mathrm{d}\theta \\ &\quad -\int_0^1 \left[\left((1-\alpha_A)\left(\frac{3}{4} + \beta_P - \theta - \beta_A\right)^2 + \alpha_A \left(\frac{2}{3} - \theta\right)^2 \right) \theta \right] 2(1-\theta) \mathrm{d}\theta, \\ &= -(1-\alpha_A)(\beta_A - \beta_P)^2 + (1-\alpha_A)\frac{\beta_A - \beta_P}{2} - \frac{\alpha_A}{16} - \frac{5}{48}, \end{aligned}$$

as stated in equations (5) and (6). Following this, a few steps of algebra yields the following

[35] The analysis of the principal's incentives can be applied directly to show that, given these beliefs by the principal, either (and only) y_A^0 or y_A^1 are strictly best responses for the agent.

incentive compatibility condition:

$$\begin{aligned} U_A(y_A^0;0) &\geq U_A(y_A^1;0), \\ \beta_A - \beta_P &\leq \frac{2}{9}\frac{\alpha_A}{(1-\alpha_A)} + \frac{1}{8}, \end{aligned}$$

as stated in inequality (7).