B Online Appendix for "Bend Them but Don't Break Them: Passionate Workers, Skeptical Managers, and Decision-Making in Organizations" by Omar A. Nayeem

B.1 Technical Definitions

Definition 5. Given the appointment of a worker-manager pair $(w, m) \in W \times M$, an *equilibrium* in the resulting subgame is a pure strategy profile, $\left(\tilde{q}, \tilde{\rho}, \tilde{\delta}\right)$, that satisfies the following conditions:

Optimal information acquisition: Given $\tilde{\rho}$ and $\tilde{\delta}$, $\tilde{q}$ maximizes the w-type worker's expected utility:

$$\tilde{q} \in \underset{q \in [0, 1/2)}{\arg \min}\, c(q) + \sum_{\substack{r \in \{\hat{\alpha}, \hat{\beta}\}, \\ s \in \{a, b\}}} f_{\tilde{\rho}}\left(r; s, q\right)$$
$$\cdot \left[g_{\tilde{\delta}}(\alpha; r) \cdot h\left(s; B, q\right) \cdot w + g_{\tilde{\delta}}(\beta; r) \cdot h\left(s; A, q\right) \cdot (1 - w) \right].$$

where

- For any recommendation function ρ, recommendation r, signal realization s, and signal quality q

$$f_{\rho}(r; s, q) \equiv \begin{cases} 1 & \text{if } \rho(s, q) = r, \\ 0 & \text{otherwise.} \end{cases}$$

- For any decision function δ, decision d, and recommendation r,

$$g_{\delta}(d; r) \equiv \begin{cases} 1 & \text{if } \delta(r) = d, \\ 0 & \text{otherwise.} \end{cases}$$

1

- For any signal realization s, state θ, and signal quality q,

$$h(s; \theta, q) \equiv \begin{cases} 1/2 + q & \text{if } (s, \theta) \in \{(a, A), (b, B)\}, \\ 1/2 - q & \text{otherwise.} \end{cases}$$

Optimal recommendation: If the principal's decision is responsive to the recommendation (i.e., if $\tilde{\delta}(\hat{\alpha}) = \alpha$ and $\tilde{\delta}(\hat{\beta}) = \beta$), then, for each signal realization $s \in \{a, b\}$ and signal quality $q \in [0, 1/2)$, $\tilde{\rho}(s, q)$ maximizes the m-type manager's expected utility. That is:

- $\tilde{\rho}(s, q) = \hat{\alpha}$ whenever ($s = a$ and $q > m - 1/2$) or $q < 1/2 - m$; and

- $\tilde{\rho}(s, q) = \hat{\beta}$ whenever ($s = b$ and $q > 1/2 - m$) or $q < m - 1/2$.

Optimal decision making: For each recommendation $r \in \{\hat{\alpha}, \hat{\beta}\}$, given $\tilde{q}$ and $\tilde{\rho}$, $\tilde{\delta}(r)$ maximizes the principal's expected utility. That is, given a recommendation r,

- If $\tilde{\rho}(a, \tilde{q}) \neq \tilde{\rho}(b, \tilde{q})$, then
    - $\tilde{\delta}(r) = \alpha$ whenever $\tilde{q} < 1/2 - p$; and
    - $\tilde{\delta}(r) = \beta$ whenever $r = \hat{\beta}$ and $\tilde{q} > 1/2 - p$.
- If $\tilde{\rho}(a, \tilde{q}) = \tilde{\rho}(b, \tilde{q})$ and $p < 1/2$, then $\tilde{\delta}(\hat{\alpha}) = \tilde{\delta}(\hat{\beta}) = \alpha$.

Note that the above formulation of the solution concept is fairly general; indeed, the only substantive restriction is on the "meanings" of $\hat{\alpha}$ and $\hat{\beta}$. For example, Definition 5 rules out as part of an equilibrium the (redundant) decision function under which the principal implements α if and only if she receives $\hat{\beta}$. Also note that, when the principal expects the manager to send an uninformative recommendation (i.e., one that she believes the manager will send regardless of the signal realization), she makes a decision in accordance with her prior beliefs.[32]

---

[32]In particular, if, given the anticipated signal quality, the principal expects the manager to recommend a specific action regardless of the signal realization but finds that the manager has recommended the other action, the principal treats the recommendation as uninformative. The rationale for this behavior is that, if the principal expects the manager to adopt a recommendation strategy that—based on the anticipated signal quality—does not depend on the signal's realization, the prin-

In defining an "informative equilibrium," it is useful not only to stipulate that the worker acquires a signal of positive quality, but also to apply certain tiebreaking rules that specify the players' actions in cases of indifference.

Definition 6. Given the appointment of a worker-manager pair $(w, m) \in W \times M$, an *informative equilibrium* in the resulting subgame is an equilibrium, $\left(\tilde{q}, \tilde{\rho}, \tilde{\delta}\right)$, in the subgame that satisfies the following additional conditions:

Information transmission: $\tilde{q} > 0$.

Worker acquires signal of highest optimal quality: If, given $\tilde{\rho}$ and $\tilde{\delta}$, the appointed worker's optimal signal quality is not unique, the worker chooses the highest quality level from the set of optima.[33] That is,

$$\tilde{q} = \max \arg \min_{q \in [0,1/2)} c(q) + \sum_{\substack{r \in \left\{\hat{\alpha}, \hat{\beta}\right\}, \\ s \in \{a, b\}}} f_{\tilde{\rho}}(r; s, q)$$
$$\cdot \left[g_{\tilde{\delta}}(\alpha; r) \cdot h(s; B, q) \cdot w + g_{\tilde{\delta}}(\beta; r) \cdot h(s; A, q) \cdot (1 - w)\right].$$

When indifferent, manager follows signal: For each signal realization $s \in \{a, b\}$ and signal quality $q \in [0, 1/2)$, $\tilde{\rho}(s, q) = \hat{\alpha}$ if and only if either ($s = a$ and $q \geq m - 1/2$) or ($s = b$ and $q < 1/2 - m$).

When indifferent, principal follows recommendation: For each recommendation $r \in \left\{\hat{\alpha}, \hat{\beta}\right\}$, given $\tilde{q}$ and $\tilde{\rho}$, $\tilde{\delta}(r)$ maximizes the principal's expected utility. That is, given a recommendation r,

cipal can believe that the worker is rational and that the worker correctly anticipates the manager's recommendation strategy only if the principal expects the worker to acquire a costless and uninformative signal. In treating an unexpected recommendation as uninformative, the principal is guided by a belief that the unexpected recommendation is due merely to an inconsequential deviation by the manager rather than to a costly deviation by the worker.

[33]Because the set of optimal signal quality levels turns out to be nonempty and finite, its maximum is well defined. This tiebreaking assumption provides technical convenience and guarantees the existence of certain types of equilibria, but the paper's main point—that preference misalignment enhances the principal's welfare—would remain valid even if the assumption were relaxed.

- If $\tilde{\rho}(a, \tilde{q}) \neq \tilde{\rho}(b, \tilde{q})$, then

$$\tilde{\delta}(r) = \begin{cases} \alpha & \text{if } r = \hat{\alpha} \text{ or } \tilde{q} < 1/2 - p, \\ \beta & \text{otherwise.} \end{cases}$$

- If $\tilde{\rho}(a, \tilde{q}) = \tilde{\rho}(b, \tilde{q})$, then, for each $r \in \left\{\hat{\alpha}, \hat{\beta}\right\}$,

$$\tilde{\delta}(r) = \begin{cases} \alpha & \text{if } p < 1/2 \text{ or } r = \hat{\alpha}, \\ \beta & \text{otherwise.} \end{cases}$$

As noted above, Definition 6 specifies tiebreaking rules for all players in cases of indifference. For example, a manager that is indifferent between α and β (based upon the observed play of the game and her beliefs about others' actions) acts in accordance with the signal's realization. Similarly, if the principal is indifferent between the two actions given her posterior belief, she acts in accordance with her inference about the signal's realization.[34] Furthermore, given the messages' meanings, the manager's recommendation function is sincere; that is, the manager recommends the action that she prefers under the posterior belief that the worker's signal induces, even if the manager does not expect the principal to heed her recommendation.[35]

Definition 7. Let $d \in \{\alpha, \beta\}$ be given. Define

$$Q_w^d \equiv \left\{q \in [0, 1/2) : \varphi_w^d(q) = q\right\},$$

and

$$q^d(w) \equiv \begin{cases} 0 & \text{if } Q_w^d = \varnothing, \\ \max Q_w^d & \text{otherwise.} \end{cases}$$

---

[34]These tiebreaking assumptions, like the one for workers, provide technical convenience and guarantee the existence of certain types of equilibria, but the paper's main point—that preference misalignment enhances the principal's welfare—would remain valid even if these assumptions were relaxed.

[35]Note that it is a weakly dominant strategy for the manager to sincerely recommend her preferred action to the principal.

In words, if φ_w^d has any fixed points, $q^d(w)$ is its largest fixed point. Otherwise, $q^d(w)$ is set to 0.[36]

B.2 Technical Results

Lemma 2. *Suppose that the worker's type is w, that the manager's type is m, and that the principal's strategy is to rubberstamp the manager's recommendation. The quality, $\hat{q}(w, m)$, of the signal that the worker acquires depends upon the manager's type as follows:*

Case I: $m \leq 1/2$. *In this case, the manager is weakly biased toward α.*

>**Subcase 1: $w < 1/2 - (q^* - 2c(q^*))$.** *In this subcase, the worker holds an appreciable bias toward α,[37] and $q^\alpha(w) = 0$. Furthermore,*
>
>$$\hat{q}(w, m) = \begin{cases} q^* & \text{if } m = 1/2, \\ 0 & \text{if } 0 < m < 1/2. \end{cases}$$

>**Subcase 2: $w \geq 1/2 - (q^* - 2c(q^*))$.** *In this subcase, the worker is, at most, slightly biased toward α, and $q^\alpha(w) \geq q^*$. (Note that the worker may be neutral or even biased toward β.) Then*
>
>$$\hat{q}(w, m) = \begin{cases} 0 & \text{if } 0 < m < 1/2 - q^\alpha(w), \\ 1/2 - m & \text{if } 1/2 - q^\alpha(w) \leq m < 1/2 - q^*, \\ q^* & \text{if } 1/2 - q^* \leq m \leq 1/2. \end{cases}$$

Case II: $m > 1/2$. *In this case, the manager is biased toward β.*

>**Subcase 1: $w > 1/2 + (q^* - 2c(q^*))$.** *In this subcase, the worker holds an appreciable bias toward β; $q^\beta(w) = 0$, and $\hat{q}(w, m) = 0$.*

---

[36]Lemma 5 in Appendix B.2 shows that the set of fixed points is finite, so that $q^d(w)$ is well defined.

[37]The definition of q^* and the assumptions on $c(\cdot)$ imply that $q^* - 2c(q^*) > 0$. Thus $w < 1/2 - (q^* - 2c(q^*)) < 1/2$.

Subcase 2: $w \leq 1/2 + (q^* - 2c(q^*))$. *In this subcase, the worker is, at most, slightly biased toward β, and $q^\beta(w) \geq q^*$. (Note that the worker may be neutral or even biased toward α.) Furthermore,*

$$
\hat{q}(w, m) = \begin{cases} q^* & \text{if } 1/2 < m \leq 1/2 + q^*, \\ m - 1/2 & \text{if } 1/2 + q^* < m \leq 1/2 + q^\beta(w), \\ 0 & \text{if } 1/2 + q^\beta(w) < m < 1. \end{cases}
$$

Proof of Lemma 2. I prove Case I, in which $m \leq 1/2$. (The proof for Case II, in which $m > 1/2$, is analogous.) Given that $m \leq 1/2$, the set of optimal signal quality levels for the worker is given by

$$
\underset{q \in [0, 1/2)}{\arg \min} \, \kappa(q; w, m),
$$

which can be expressed as

$$
\underset{q \in Q_m}{\arg \min} \, \kappa(q; w, m),
$$

where

$$
Q_m \equiv \underbrace{\left[\underset{q \in [0, 1/2 - m)}{\arg \min} \, c(q) + w/2 \right]}_{\substack{= \{0\} \text{ if } m < 1/2; \\ \varnothing \text{ otherwise}}} \cup \underbrace{\left[\underset{q \in [1/2 - m, 1/2)}{\arg \min} \, c(q) - q/2 + 1/4 \right]}_{= \max\{1/2 - m, q^*\}}
$$

$$
= \begin{cases} \{0, 1/2 - m\} & \text{if } 0 < m < 1/2 - q^*, \\ \{0, q^*\} & \text{if } 1/2 - q^* \leq m < 1/2, \\ \{q^*\} & \text{if } m = 1/2. \end{cases}
$$

Thus $\hat{q}(w, 1/2) = q^*$ for all $w \in W$. From this point, suppose that $m < 1/2$. Consider the two subcases (under Case I) from the statement of the result.

Subcase 1: $w < 1/2 - (q^* - 2c(q^*))$. The fact that $q^\alpha(w) = 0$ follows from

Lemma 6. Furthermore, for any $m < 1/2$,

$$\underbrace{c(q^*) - q^*/2 + 1/4}_{=\min_{q \in [0,1/2)} c(q) - q/2 + 1/4} > \underbrace{w/2}_{=\kappa(0;w,m)} ,$$

so

$$\arg\min_{q \in Q_m} \kappa(q; w, m) = \{0\}.$$

Thus $\hat{q}(w, m) = 0$.

Subcase 2: $w \geq 1/2 - (q^* - 2c(q^*))$. Lemma 6 implies that $q^\alpha(w) \geq q^*$. First suppose that $0 < m < 1/2 - q^*$. In this case, $Q_m = \{0, 1/2 - m\}$. Note that $\kappa(0; w, m) = w/2$ and that

$$\kappa(1/2 - m; w, m) = c(1/2 - m) - (1/2 - m)/2 + 1/4.$$

Lemma 8 implies that

$$\arg\min_{q \in Q_m} \kappa(q; w, m) = \begin{cases} \{0\} & \text{if } 0 < m < 1/2 - q^\alpha(w), \\ \{0, 1/2 - m\} & \text{if } m = 1/2 - q^\alpha(w), \\ \{1/2 - m\} & \text{if } 1/2 - q^\alpha(w) < m < 1/2 - q^*. \end{cases}$$

Now suppose that $1/2 - q^* \leq m < 1/2$. In this case, $Q_m = \{0, q^*\}$. Note that

$$\underbrace{c(q^*) - q^*/2 + 1/4}_{=\kappa(q^*;w,m)} \leq \underbrace{w/2}_{=\kappa(0;w,m)} ,$$

where equality holds if and only if $w = 2c(q^*) - q^* + 1/2$. Thus,

$$\arg\min_{q \in Q_m} \kappa(q; w, m) =$$

$$\begin{cases} \{q^*\} & \text{if } w > 2c(q^*) - q^* + 1/2 \text{ and } 1/2 - q^* \leq m, \\ \{0, q^*\} & \text{if } w = 2c(q^*) - q^* + 1/2 \text{ and } 1/2 - q^* \leq m. \end{cases}$$

Now the assumption that the worker chooses the highest of optimal signal

7

qualities implies that

$$\hat{q}(w,m) = \begin{cases} 0 & \text{if } 0 < m < 1/2 - q^{\alpha}(w), \\ 1/2 - m & \text{if } 1/2 - q^{\alpha}(w) \leq m < 1/2 - q^*, \\ q^* & \text{if } 1/2 - q^* \leq m < 1/2. \end{cases}$$

$\square$

Lemma 3. *Let $w^* \in W$ and $m^* \in M$ be given. If the subgame that follows the appointment of (w^*, m^*) has no informative equilibrium, $V(w^*, m^*) = 0$. Otherwise,*

$$V(w^*, m^*) = \frac{p + \hat{q}(w^*, m^*)}{2} - \frac{1}{4}.$$

Proof. In a babbling equilibrium, the principal chooses her ex-ante-preferred option, so her expected payoff is $-p/2$. In an informative equilibrium, the principal rubberstamps the appointed manager's recommendation. By Lemma 2, the appointed worker acquires a signal according to $\hat{q}$. A direct computation shows that, by following such a signal, the principal achieves an expected payoff of $-1/2 \cdot (1/2 - \hat{q}(w^*, m^*))$. By definition, then,

$$\begin{aligned} V(w^*, m^*) &\equiv -\frac{1}{2} \cdot \left(\frac{1}{2} - \hat{q}(w^*, m^*) \right) - \frac{-p}{2} \\ &= \frac{p + \hat{q}(w^*, m^*)}{2} - \frac{1}{4}. \end{aligned}$$

$\square$

Lemma 4. *There exist two real numbers, $w^{\alpha} \in (0, 1/2)$ and $w^{\beta} \in (1/2, 1)$, for which the following conditions hold:*

(i) $q^{\alpha}(\cdot)$ *is strictly increasing on $W \cap [w^{\alpha}, 1)$.*

(ii) $q^{\beta}(\cdot)$ *is strictly decreasing on $W \cap (0, w^{\beta}]$.*

Proof. This result follows immediately from Lemmas 10 and 11. $\square$

Lemma 5. *Let $w \in W$ be given, and consider the functions*

$$\varphi_w^\alpha : [0, 1/2) \to \mathbb{R},$$
$$q \mapsto 2c(q) - (w - 1/2)$$

and

$$\varphi_w^\beta : [0, 1/2) \to \mathbb{R},$$
$$q \mapsto 2c(q) - (1/2 - w).$$

Each of these two functions has at most two fixed points.

Proof. Fix $w \in W$ and $d \in \{\alpha, \beta\}$. Consider the following function:

$$\psi_w^d : [0, 1/2) \to \mathbb{R},$$
$$q \mapsto \varphi_w^d(q) - q.$$

Observe that $q \in [0, 1/2)$ is a fixed point of φ_w^d if and only if $\psi_w^d(q) = 0$. Since $\left(\psi_w^d\right)'(q) = 0$ if and only if $q = q^*$, Rolle's Theorem implies that $\psi_w^d(q) = 0$ for no more than two values of q. The result follows. $\square$

Because the set of fixed points of φ_w^d is finite, it is either empty or has a well-defined maximum.

Lemma 6. *Let $w \in W$ be given.*

   (i) $w < 2c(q^*) - q^* + 1/2$ *if and only if* $q^\alpha(w) = 0$.

  (ii) $w = 2c(q^*) - q^* + 1/2$ *if and only if* $q^\alpha(w) = q^*$.

 (iii) $w > 2c(q^*) - q^* + 1/2$ *if and only if* $q^\alpha(w) > q^*$.

Proof. Given that the left hand side statements are mutually exclusive and exhaustive and that the right hand side statements are mutually exclusive, it suffices to prove only the three "only if" statements. Fix $w \in W$, and observe that

$$\arg\min_{q \in [0,1/2)} \psi_w^\alpha(q) = \{q^*\}, \tag{12}$$

9

where ψ_w^α is as defined in the proof of Lemma 5. Also, observe that

$$q^\alpha(w) = \begin{cases} 0 & \text{if } \{q \in [0, 1/2) : \psi_w^\alpha(q) = 0\} = \varnothing, \\ \max\{q \in [0, 1/2) : \psi_w^\alpha(q) = 0\} & \text{otherwise.} \end{cases}$$

(i) Suppose that $w < 2c(q^*) - q^* + 1/2$. Then, by (12),

$$\min_{q \in [0,1/2)} \psi_w^\alpha(q) = \psi_w^\alpha(q^*) > 0,$$

and $\psi_w^\alpha(q) = 0$ has no solution. Hence, φ_w^α has no fixed point: $q^\alpha(w) = 0$.

(ii) Suppose that $w = 2c(q^*) - q^* + 1/2$. In this case,

$$\min_{q \in [0,1/2)} \psi_w^\alpha(q) = \psi_w^\alpha(q^*) = 0.$$

Since q^* is the unique minimizer, it is the unique fixed point: $q^\alpha(w) = q^*$.

(iii) Suppose that $w > 2c(q^*) - q^* + 1/2$. In this case,

$$\min_{q \in [0,1/2)} \psi_w^\alpha(q) = \psi_w^\alpha(q^*) < 0.$$

Since $\lim_{q \uparrow 1/2} \psi_w^\alpha(q) = \infty$, the Intermediate Value Theorem implies that there exists $q' \in (q^*, 1/2)$ such that $\psi_w^\alpha(q') = 0$. Thus $q^\alpha(w) \geq q' > q^*$.[38]

□

Lemma 7. *Let $w \in W$ be given.*

(i) $w > q^* - 2c(q^*) + 1/2$ *if and only if* $q^\beta(w) = 0$.

(ii) $w = q^* - 2c(q^*) + 1/2$ *if and only if* $q^\beta(w) = q^*$.

(iii) $w < q^* - 2c(q^*) + 1/2$ *if and only if* $q^\beta(w) > q^*$.

Proof. The proof is analogous to that of Lemma 7. □

---

[38]In fact, since $\psi_w'(q) > 0$ for $q \in (q^*, 1/2)$, Rolle's Theorem implies that $\psi_w^\alpha(q) \neq 0$ for all $q \in (q^*, q') \cup (q', 1/2)$. Thus $q^\alpha(w) = q'$.

Lemma 8. *Suppose that $q^\alpha(w) > q^*$.*

(i) *For all $q \in (q^*, q^\alpha(w))$, $c(q) - q/2 + 1/4 < w/2$.*

(ii) *For all $q \in (q^\alpha(w), 1/2)$, $c(q) - q/2 + 1/4 > w/2$.*

Proof. Suppose that $q^\alpha(w) > q^*$. Lemma 6 implies that $\psi_w^\alpha(q^*) < 0$, where ψ_w^α is defined in the proof of Lemma 5.

(i) Suppose that there exists $q' \in (q^*, q^\alpha(w))$ such that $c(q') - q'/2 + 1/4 \geq w/2$. Then $\psi_w^\alpha(q') \geq 0$. By the Intermediate Value Theorem, there exists $q'' \in (q^*, q']$ such that $\psi_w^\alpha(q'') = 0$. Recall that $\psi_w^\alpha(q^\alpha(w)) = 0$ as well. Rolle's Theorem implies that there exists $q''' \in (q'', q^\alpha(w))$ such that $(\psi_w^\alpha)'(q''') = 0$, which contradicts the fact that q^* is the unique solution to $(\psi_w^\alpha)'(q) = 0$.

(ii) Suppose that there exists $q' \in (q^\alpha(w), 1/2)$ such that $c(q') - q'/2 + 1/4 \leq w/2$. Then $\psi_w^\alpha(q') \leq 0$. Since $\lim_{q \uparrow 1/2} \psi_w^\alpha(q) = \infty$, the Intermediate Value Theorem implies that there exists $q'' \in [q', 1/2)$ such that $\psi_w^\alpha(q'') = 0$. That is, $q'' \in Q_w^\alpha$, which contradicts the defining characteristic $q^\alpha(w) \equiv \max Q_w^\alpha$.

$\square$

Lemma 9. *Suppose that $q^\beta(w) > q^*$.*

(i) *For all $q \in \left(q^*, q^\beta(w)\right)$, $c(q) - q/2 + 1/4 < (1-w)/2$.*

(ii) *For all $q \in \left(q^\beta(w), 1/2\right)$, $c(q) - q/2 + 1/4 > (1-w)/2$.*

Proof. The proof is analogous to that of Lemma 8. $\square$

Lemma 10. *Consider any $w \in W$.*

(i) *If $q^\alpha(w) = 0$, then $q^\alpha(w') = 0$ for every $w' \in W$ that satisfies $w' < w$.*

(ii) *If $q^\alpha(w) > 0$, then $q^\alpha(w') < q^\alpha(w)$ for every $w' \in W$ that satisfies $w' < w$.*

(iii) *If $q^\alpha(w) > 0$, then $q^\alpha(w') > q^\alpha(w)$ for every $w' \in W$ that satisfies $w' > w$.*

Proof. Let $w \in W$ be given.

(i) Take any $w' \in W$ that satisfies $w' < w$. Since $q^\alpha(w) = 0$, Lemma 6 implies that

$$w' < w < 2c(q^*) - q^* + 1/2,$$

and that $q^\alpha(w') = 0$.

(ii) Take any $w' \in W$ that satisfies $w' < w$. Suppose that $q^\alpha(w) > 0$. If $q^\alpha(w') = 0$, the result follows immediately. Suppose that $q^\alpha(w') > 0$. Recall the function ψ_w^α, defined in the proof of Lemma 5. Note that

$$\psi_w^\alpha(q^\alpha(w')) = \psi_w^\alpha(q^\alpha(w')) - \underbrace{\psi_{w'}^\alpha(q^\alpha(w'))}_{=0}$$
$$= \varphi_w^\alpha(q^\alpha(w')) - \varphi_{w'}^\alpha(q^\alpha(w'))$$
$$= w' - w$$
$$< 0.$$

Since $\lim_{q \uparrow 1/2} \psi_w^\alpha(q) = \infty$, the Intermediate Value Theorem implies that there exists $q' \in (q^\alpha(w'), 1/2)$ such that $\psi_w^\alpha(q') = 0$, meaning that $q' \in Q_w^\alpha$. Since $q^\alpha(w) \equiv \max Q_w^\alpha$, we conclude that $q^\alpha(w) \geq q' > q^\alpha(w')$.

(iii) Take any $w' \in W$ that satisfies $w' > w$. Suppose that $q^\alpha(w) > 0$. In this case,

$$\psi_{w'}^\alpha(q^\alpha(w)) = \psi_{w'}(q^\alpha(w)) - \underbrace{\psi_w^\alpha(q^\alpha(w))}_{=0}$$
$$= \varphi_{w'}(q^\alpha(w)) - \varphi_w(q^\alpha(w))$$
$$= w - w'$$
$$< 0.$$

Since $\lim_{q \uparrow 1/2} \psi_{w'}^\alpha(q) = \infty$, the Intermediate Value Theorem implies that there exists $q' \in (q^\alpha(w), 1/2)$ such that $\psi_{w'}^\alpha(q') = 0$, meaning that $q' \in Q_{w'}^\alpha$. Because $q^\alpha(w') \equiv \max Q_{w'}^\alpha$, we have $q^\alpha(w') \geq q' > q^\alpha(w)$.

$\square$

Lemma 11. *Consider any $w \in W$.*

(i) If $q^\beta(w) = 0$, then $q^\beta(w') = 0$ for every $w' \in W$ that satisfies $w' > w$.

(ii) If $q^\beta(w) > 0$, then $q^\beta(w') < q^\beta(w)$ for every $w' \in W$ that satisfies $w' > w$.

(iii) If $q^\beta(w) > 0$, then $q^\beta(w') > q^\beta(w)$ for every $w' \in W$ that satisfies $w' < w$.

Proof. The proof is analogous to that of Lemma 10. □

B.3 Numerical Examples

This section presents several numerical examples that illustrate the effects of preference misalignment on the principal's welfare.

Example 1. Suppose that $p = 1/2$, $W = \{1/2\}$, $M = (0, 1/2]$, and $c(q) = q^2/(1 - 2q)$. In this case, the principal is unbiased, as is the only candidate worker. However, the principal can choose from a broad array of candidate managers, each of whom holds at least a weak bias toward α. It is straightforward to verify that $q^* = (2 - \sqrt{2})/4 \approx 0.146$, and $q^\alpha(1/2) = 1/4$. Because the principal is unbiased, she is willing to follow any recommendation that she believes to be consistent with the worker's signal. By Lemma 2,

$$\hat{q}(1/2, m) \equiv \begin{cases} 0 & \text{if } 0 < m < 1/4, \\ 1/2 - m & \text{if } 1/4 \leq m < \sqrt{2}/4, \\ (2 - \sqrt{2})/4 & \text{otherwise;} \end{cases}$$

Furthermore, in any informative equilibrium, the manager's recommendation strategy is

$$\tilde{\rho}(s, q) \equiv \begin{cases} \hat{\alpha} & \text{if } s = a \text{ or } q < 1/2 - m, \\ \hat{\beta} & \text{otherwise;} \end{cases}$$

and the principal's decision strategy is

$$\tilde{\delta}(r) \equiv \begin{cases} \alpha & \text{if } r = \hat{\alpha}, \\ \beta & \text{otherwise.} \end{cases}$$

13

The principal maximizes her expected utility by appointing the most biased manager (of type $m^* \equiv 1/4$) that the worker finds it worthwhile to convince; the worker acquires a signal of quality $\tilde{q} = 1/4$.

In Example 1, the principal and worker have aligned interests over the decision, but the principal deliberately appoints a manager whose ex-ante preferences differ from hers.[39] Such a manager's bias serves to strengthen the worker's incentive to acquire information. Given the worker's signal, the manager (at least weakly) prefers the same action as the principal, so the principal realizes an unambiguous welfare increase by appointing the biased manager. Example 2 shows how the principal can exploit workers' biases to further improve the quality of the decision.

Example 2. Suppose that $p = 1/2$, $W = \{1/4, 1/2, 3/4\}$, $M = (0, 1/2]$, and $c(q) = q^2/(1 - 2q)$. Just as in Example 1, the principal is unbiased, every candidate manager is at least weakly biased toward α, and we have $q^* = (2 - \sqrt{2})/4$ and $q^\alpha(1/2) = 1/4$. Here, however, there are two additional candidate workers of opposite biases. Furthermore, $1/2 - (q^* - 2c(q^*)) = \sqrt{2} - 1 \approx 0.414$, $q^\alpha(1/4) = 0$, and $q^\alpha(3/4) = (1 + \sqrt{17})/16 \approx 0.320$. In this case,

$$\hat{q}(1/4, m) \equiv \begin{cases} 0 & \text{if } m < 1/2, \\ (2 - \sqrt{2})/4 & \text{otherwise,} \end{cases}$$

$$\hat{q}(1/2, m) \equiv \begin{cases} 0 & \text{if } 0 < m < 1/4, \\ 1/2 - m & \text{if } 1/4 \leq m < \sqrt{2}/4, \\ (2 - \sqrt{2})/4 & \text{otherwise,} \end{cases}$$

$$\hat{q}(3/4, m) \equiv \begin{cases} 0 & \text{if } 0 < m < (7 - \sqrt{17})/16, \\ 1/2 - m & \text{if } (7 - \sqrt{17})/16 \leq m < \sqrt{2}/4, \\ (2 - \sqrt{2})/4 & \text{otherwise.} \end{cases}$$

---

[39]This example highlights an important distinction between this paper and that of Dessein (2002). In Dessein's model, information is exogenous, and the expert (the analogue of the worker from this paper) knows the state. Therefore, in Dessein's environment, the manager—who functions as a delegate to whom the principal cedes control of the decision—serves no useful purpose when the principal and worker have identical preferences.

The optimal appointment for the principal is to pair $w^* \equiv 3/4$ with $m^* \equiv 1/2 - q^\alpha(3/4) = (7 - \sqrt{17})/16 \approx 0.180$.

The structure of this example is similar to that of Example 1, but, in this case, the principal can exploit the stronger information acquisition incentive of a worker that is biased toward β. In particular, such a worker will be willing to acquire a more precise (though costlier) signal to persuade an α-biased manager. Thus, the principal achieves a higher expected payoff in this case. Note, also, that the worker of type $1/4$ has no incentive to exert any effort (except in the knife-edge case in which the appointed manager is neutral), since, with an uninformative signal, he can ensure that his preferred choice of α is implemented. By Lemma 2, the worker of type $1/4$ prefers this outcome to any one in which his signal determines the decision (since $1/4 < 1/2 - (q^* - 2c(q^*))$).

Example 3. Suppose that $p = 1/4$, $W = \{1/8, 1/2, 3/4\}$, $M = (0, 1)$, and $c(q) = q^2/(1 - 2q)$. Just as in the previous two examples, we have $q^* = (2 - \sqrt{2})/4$. As in Example 2, there is an unbiased candidate worker and two additional candidate workers of opposite biases (though, unlike in Example 2, the magnitudes of the two biased candidate workers' biases are different here). Now, however, the principal is biased toward α, and candidate managers of all biases are available. Thus, the candidate workers' values for both q^α and q^β are relevant. These values are summarized in Table 1.

w	$q^\alpha(w)$	$q^\beta(w)$
$1/8$	0	$\left(1 + \sqrt{97}\right)/32 \approx 0.339$
$1/2$	$1/4$	$1/4$
$3/4$	$\left(1 + \sqrt{17}\right)/16 \approx 0.320$	0

Table 1: Values of q^α and q^β for each of the available candidate workers in Example 3

Given $m \in M$,

$$\hat{q}(1/8, m) \equiv \begin{cases} (2 - \sqrt{2})/4 & \text{if } 1/2 \leq m \leq (4 - \sqrt{2})/4, \\ m - 1/2 & \text{if } (4 - \sqrt{2})/4 < m \leq (17 + \sqrt{97})/32, \\ 0 & \text{otherwise}, \end{cases}$$

$$\hat{q}(1/2, m) \equiv \begin{cases} 0 & \text{if } |m - 1/2| > 1/4, \\ |m - 1/2| & \text{if } (2 - \sqrt{2})/4 \leq |m - 1/2| \leq 1/4, \\ (2 - \sqrt{2})/4 & \text{otherwise}, \end{cases}$$

$$\hat{q}(3/4, m) \equiv \begin{cases} (2 - \sqrt{2})/4 & \text{if } \sqrt{2}/4 \leq m \leq 1/2, \\ 1/2 - m & \text{if } (7 - \sqrt{17})/16 \leq m < \sqrt{2}/4, \\ 0 & \text{otherwise}. \end{cases}$$

In an informative equilibrium, a manager of type m follows the recommendation strategy given by

$$\tilde{\rho}(s, q) \equiv \begin{cases} \hat{\alpha} & \text{if } [s = a \text{ and } q \geq m - 1/2] \text{ or } [s = b \text{ and } q < 1/2 - m], \\ \hat{\beta} & \text{otherwise}. \end{cases}$$

The optimal appointment here pairs the worker of type $w^* \equiv 1/8$ with the manager of type $m^* \equiv (17 + \sqrt{97})/32$. Note that the principal appoints the most biased worker that is available: the one of type $1/8$ (who holds a strong bias toward α). Given that candidate managers of all types are available, this type of worker will have the strongest incentive to acquire a signal that is informative enough to influence a manager that is fairly strongly biased toward β. To be more precise, given that $q^\beta(1/8) \approx 0.339$, this worker would be willing to acquire a signal that is strong enough to influence a β-biased manager of type as high as $1/2 + q^\beta(1/8) \approx 0.839$ (provided that the principal would implement this manager's recommendation). Neither of the two other workers, whose biases are less severe than this worker's, would be willing to acquire a signal of quality as high as $q^\beta(1/8)$. The appointed manager, of type $1/2 + q^\beta(1/8)$, is the one that exploits the worker's information acquisition incentive to the greatest possible extent.

16

Note, also, that there is no informative equilibrium in any subgame in which an unbiased manager is appointed. In particular, since $p = 1/4$, the principal will implement α unless she has sufficient reason to believe that the evidence favors state B (i.e., that $s = b$ and $q \geq |p - 1/2| = 1/4$). Of course, given her inability to observe the signal directly, the principal draws her inference regarding the signal's quality based on her belief—which, in equilibrium, is correct—regarding the worker's choice of signal based on the appointed manager's type. Since no worker can commit to acquiring a signal of a given quality, and every worker prefers to acquire a signal of quality $(2 - \sqrt{2})/4 < 1/4$ when facing an unbiased manager, the principal (who cannot observe the signal's quality) always discards the unbiased manager's recommendation. More generally, no subgame following an appointment of a manager of type $m \in (1/4, 3/4)$ has an informative equilibrium.

Example 4. Let $p \equiv 3/8$, $W \equiv \{1/2, 3/4\}$, $M \equiv \{3/5\}$, and $c(q) \equiv q^2/(1 - 2q)$. As before, $q^* = (2 - \sqrt{2})/4$. Since there is a single candidate manager, the principal's only nontrivial appointment decision relates to her choice of worker. As noted in Example 3, $q^\beta(3/4) = 0$. Given that the only available candidate manager has a bias toward β, if the candidate worker of type $3/4$ is appointed, he will not acquire information. Thus, there is no informative equilibrium in a subgame in which the appointed worker has type $3/4$. On the other hand, if the principal appoints the candidate worker of type $1/2$, then, since a signal of quality q^* is sufficiently informative to persuade the manager (i.e., $q^* \geq 3/5 - 1/2$) as well as the principal (i.e., $q^* \geq 1/2 - 3/8$), and the worker prefers to influence the manager than to shirk and allow the implementation of β (i.e., $q^\beta(1/2) > |3/5 - 1/2|$), there is an informative equilibrium in the subgame following the appointment of $(w^*, m^*) = 1/2, 3/5$. The worker acquires a signal of quality q^* in this equilibrium.

Note that, in this case, the manager's bias serves no purpose: the worker acquires the same signal that he would have if the manager had been unbiased. Furthermore, unlike in previous examples, an unbiased worker here is preferred to a biased one. The reason is that a worker's bias is helpful only if it opposes the manager's bias. In this example, the biased worker and the manager both prefer β ex ante.

Example 5. Suppose that $p \equiv 1/4$, $W \equiv (0,1)$, $M \equiv (1/4, 3/4)$, and $c(q) \equiv q^2/(1-2q)$. In this case, there is no informative equilibrium for any appointment, because, for every candidate worker $w \in W$ and candidate manager $m \in M$,

$$1/2 - p > \max\left\{q^*, |m - 1/2|\right\} \geq \hat{q}(w, m).^{40}$$

That is, the principal's bias toward α is strong enough that, if she believes that the manager's recommendation is based on a signal of quality q^* (which would be sufficient to influence any manager whose type lies in the interval $(1/2 - q^*, 1/2 + q^*)$), or even of quality $|m - 1/2|$ for any $m \in M$, the principal would ignore the recommendation and implement α. Since no worker (no matter how biased) can credibly acquire a signal of quality greater than $\hat{q}(w, m) \leq \max\left\{q^*, |m - 1/2|\right\}$ while facing a manager of type m, no worker can credibly acquire a signal that will induce the principal to rubberstamp the manager's recommendation.[41]

B.4   Different Cost Functions

To the extent that q^*, $q^\alpha(\cdot)$, and $q^\beta(\cdot)$ implicitly depend on the cost function $c(\cdot)$, a change in the information-acquisition technology typically will alter the set of appointments whose associated subgames have informative equilibria. Recall that Footnote 13 describes a family of cost functions, $\left\{c_{h,k,n} : (h, k, n) \in \mathbb{R}^3_{++}\right\}$. The panels in Figure 7 depict how the choice of cost function affects q^*, $q^\alpha(\cdot)$, $q^\beta(\cdot)$, and the set of appointments with informative equilibria.[42] As Figures 7(c) and 7(d) show, the comparative statics are not always clear; for example, for some (but not all) values of w, $q^\alpha(w)$ is greater under $c_{1,1,2}$ than under $c_{2,1,1}$, and the same is true for $q^\beta(w)$. However, given the stated assumptions on the cost function, if the second derivative of one cost function, $\hat{c}$, is uniformly greater than that of another, $\tilde{c}$, then it also will be the case that $\hat{c}(\cdot) > \tilde{c}(\cdot)$, and hence that q^*, $q^\alpha(\cdot)$, and $q^\beta(\cdot)$ all will be greater for $\tilde{c}$ than for $\hat{c}$. The intuition is clear: because information is more expensive under $\hat{c}$, workers of all biases will acquire no more information under

---

[40]See Proposition 2.

[41]See the discussion of Example 3 for a detailed discussion of this point.

[42]As shown in Lemmas 6 and 7, $q^* = \min\left\{q^d(w) : w \in (0,1), q^d(w) > 0\right\}$ for each $d \in \{\alpha, \beta\}$.

(a) Appointments with informative equilibria under $c_{1,1,1}$

(b) Appointments with informative equilibria under $c_{2,1,1}$

(c) Comparison of q^α across cost functions

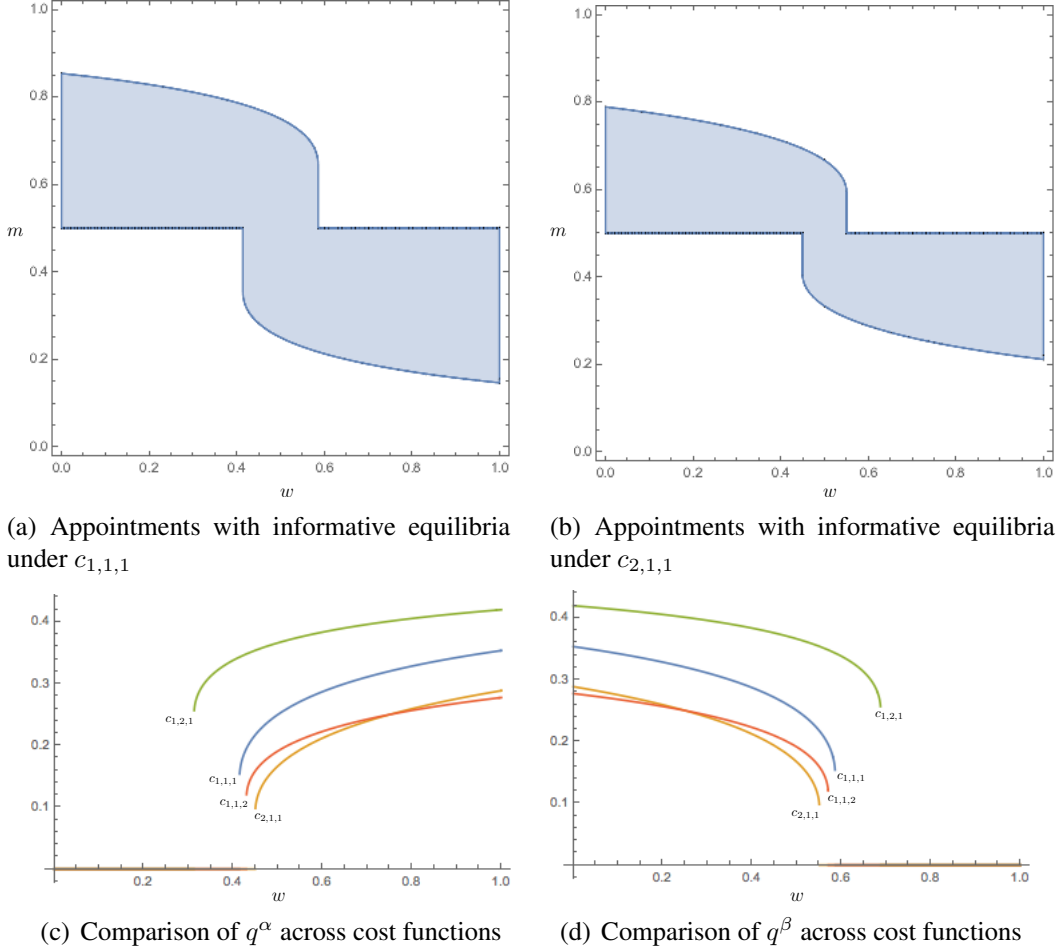(d) Comparison of q^β across cost functions

Figure 7: Let $W = M = (0, 1)$, and consider the family of cost functions defined, in Footnote 13, by $\{c_{h,k,n} : (h, k, n) \in \mathbb{R}^3_{++}\}$. The set of appointments whose subgames have informative equilibria changes with the cost function parameters as shown in the panels above, as shown in Panels 7(a) and 7(b). Panels 7(c) and 7(d) show how $q^\alpha(\cdot)$ and $q^\beta(\cdot)$ change with the cost function.

$\hat{c}$ than under $\tilde{c}$. Hence the set of appointments with informative equilibria will be larger under $\tilde{c}$ than under $\hat{c}$.

B.5 Discussion of Mixed-Strategy Equilibria

The analysis has focused on pure-strategy equilibria. This section considers mixed strategies and equilibria in mixed strategies. The key finding—to which the main

text alludes—is that the principal does no better under mixed-strategy equilibria than under pure-strategy equilibria in this environment. In particular, given an appointment and associated subgame with an informative equilibrium in pure strategies, the subgame admits no mixed-strategy equilibrium under which the principal's expected welfare is higher than under the pure-strategy informative equilibrium.

In demonstrating this fact, I will not employ a constructive method of proof. That is, I will not explicitly characterize the set of mixed-strategy equilibria or even construct such an equilibrium. Rather, I will define the set of mixed-strategy equilibria (as well as a subclass of that set that is analogous to the class of informative equilibria from the main text) and apply key properties to argue that the principal's welfare, for a fixed appointment, is no greater under a mixed-strategy equilibrium than under a pure-strategy equilibrium.

To begin, define a mixed strategy profile as follows:

Definition 8. Given an appointment and associated subgame, a *mixed strategy profile* is an ordered triple, (ι, ρ, δ), that consists of:

- An *information acquisition strategy*, $\iota \in \Delta([0, 1/2))$, which is a probability distribution on $[0, 1/2)$;

- A *recommendation strategy*, $\rho : \{a, b\} \times [0, 1/2) \to [0, 1]$, that maps each signal-realization-and-quality pair to a probability distribution on $\left\{ \hat{\alpha}, \hat{\beta} \right\}$ (so that $\rho(s, q) \in [0, 1]$ represents the probability that the manager sends recommendation $\hat{\alpha}$ when the signal is of realization s and quality q);

- A *decision strategy*, $\delta : \left\{ \hat{\alpha}, \hat{\beta} \right\} \to [0, 1]$, that maps a recommendation to a probability distribution on $\{\alpha, \beta\}$ (so that $\delta(r) \in [0, 1]$ represents the probability that the principal chooses α when the manager sends recommendation r).

An equilibrium in mixed strategies is a mixed strategy profile in which, for each player, each action that is in the support of the player's strategy is a best response given the other players' strategies and the player's observations about the current history of the game:

Definition 9. Given an appointment, (w, m), and associated subgame, an *equiib-rium in mixed strategies* is a mixed strategy profile, $(\iota^*, \rho^*, \delta^*)$, that satisfies the following conditions:

Optimal information acquisition: $\text{supp}(\iota^*) \subseteq \arg\min_{q \in [0,1/2)} c(q) + \ell_w(q; m, \rho^*, \delta^*)$, where $\ell_w(q; m, \rho^*, \delta^*)$ denotes the w-type worker's expected loss (not counting the information acquisition cost) when he acquires a signal of quality q, assuming that the m-type manager and principal follow the strategies described by ρ^* and δ^*, respectively:

$$
\begin{aligned}
\ell_w(q; m, \rho^*, \delta^*) \equiv w \cdot \frac{1}{2} \cdot & \left[\left(\frac{1}{2} - q \right) \cdot \left(\rho^*(a, q) \cdot \delta^*(\hat{\alpha}) + (1 - \rho^*(a, q)) \cdot \delta^*(\hat{\beta}) \right) \right. \\
& \left. + \left(\frac{1}{2} + q \right) \cdot \left(\rho^*(b, q) \cdot \delta^*(\hat{\alpha}) + (1 - \rho^*(b, q)) \cdot \delta^*(\hat{\beta}) \right) \right] \\
+ (1 - w) \cdot \frac{1}{2} \cdot & \left[\left(\frac{1}{2} - q \right) \cdot \left(\rho^*(b, q) \cdot (1 - \delta^*(\hat{\alpha})) + (1 - \rho^*(b, q)) \cdot \left(1 - \delta^*(\hat{\beta}) \right) \right) \right. \\
& \left. + \left(\frac{1}{2} + q \right) \cdot \left(\rho^*(a, q) \cdot (1 - \delta^*(\hat{\alpha})) + (1 - \rho^*(a, q)) \cdot \left(1 - \delta^*(\hat{\beta}) \right) \right) \right].
\end{aligned}
$$

Optimal recommendation: For every $q \in [0, 1/2)$, given the principal's decision strategy, δ^*, the manager's recommendation strategy, ρ^*, maximizes the manager's expected utility. That is, the following conditions hold:

- If $\rho^*(a, q) > 0$, then

$$
\left(m - q - \frac{1}{2} \right) \cdot \delta^*(\hat{\alpha}) \leq \left(m - q - \frac{1}{2} \right) \cdot \delta^*(\hat{\beta}).
$$

- If $\rho^*(a, q) < 1$, then

$$
\left(m - q - \frac{1}{2} \right) \cdot \delta^*(\hat{\alpha}) \geq \left(m - q - \frac{1}{2} \right) \cdot \delta^*(\hat{\beta}).
$$

- If $\rho^*(b, q) > 0$, then

$$
\left(\frac{1}{2} - m - q \right) \cdot (1 - \delta^*(\hat{\alpha})) \leq \left(\frac{1}{2} - m - q \right) \cdot \left(1 - \delta^*(\hat{\beta}) \right).
$$

- If $\rho^*(b, q) < 1$, then

$$\left(\frac{1}{2} - m - q\right) \cdot (1 - \delta^*(\hat{\alpha})) \geq \left(\frac{1}{2} - m - q\right) \cdot \left(1 - \delta^*(\hat{\beta})\right).$$

Optimal decision making: For every recommendation, $r \in \left\{\hat{\alpha}, \hat{\beta}\right\}$, the randomized decision rule, $\delta^*(r)$, maximizes the principal's expected utility under a particular belief (to be defined shortly) over histories that is induced by r. In particular, whenever the manager sends r with positive probability given ι^* and ρ^*, the principal applies Bayes' Rule to obtain the relevant belief over histories. When the manager sends r with probability zero given ι^* and ρ^*, the principal's belief, upon observing r, is that the worker acquired an uninformative signal, and that the recommendation is also uninformative. Thus, the principal's posterior belief over the state space remains unchanged from her prior belief. More concretely, let $\pi_{(\iota^*, \rho^*)}(r)$ denote the probability that the principal's posterior belief assigns to state A given recommendation r, and given that the principal believes that the worker and manager play according to ι^* and ρ^*, respectively. That is, assuming that $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] > 0$ (i.e., given ι^* and ρ^*, the manager recommends α with positive probability),

$$\pi_{(\iota^*, \rho^*)}(\hat{\alpha}) \equiv \frac{1}{2} + \frac{\mathbb{E}_{\iota^*}[q \cdot (\rho^*(a, q) - \rho^*(b, q))]}{\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)]}.$$

Similarly, assuming that $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] < 2$ (i.e., given ι^* and ρ^*, the manager recommends β with positive probability),

$$\pi_{(\iota^*, \rho^*)}(\hat{\beta}) \equiv \frac{1}{2} - \frac{\mathbb{E}_{\iota^*}[q \cdot (\rho^*(a, q) - \rho^*(b, q))]}{2 - \mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)]}.$$

The following conditions hold:

- $\delta^*(\hat{\alpha}) \geq \delta^*(\hat{\beta})$ (i.e., the meanings of $\hat{\alpha}$ and $\hat{\beta}$ are restricted to eliminate redundancy).

- If $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] = 0$ (so that, if the worker and manager follow ι^* and ρ^*, respectively, the probability that the manager recommends α

is zero), then $\delta^*(\hat{\alpha}) = 1$. That is, the principal selects α unconditionally when she receives an unexpected recommendation to implement α.

- If $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] = 2$ (so that, if the worker and manager follow ι^* and ρ^*, respectively, the probability that the manager recommends α is one), then $\delta^*(\hat{\beta}) = 1$. That is, the principal selects α unconditionally when she receives an unexpected recommendation to implement β.

- If $\delta^*(\hat{\alpha}) > 0$ and $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] > 0$,

$$p \cdot \left(1 - \pi_{(\iota^*, \rho^*)}(\hat{\alpha})\right) \leq (1 - p) \cdot \pi_{(\iota^*, \rho^*)}(\hat{\alpha}).$$

- If $\delta^*(\hat{\alpha}) < 1$ and $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] > 0$,

$$p \cdot \left(1 - \pi_{(\iota^*, \rho^*)}(\hat{\alpha})\right) \geq (1 - p) \cdot \pi_{(\iota^*, \rho^*)}(\hat{\alpha}).$$

- If $\delta^*(\hat{\beta}) > 0$ and $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] < 2$,

$$p \cdot \left(1 - \pi_{(\iota^*, \rho^*)}(\hat{\beta})\right) \leq (1 - p) \cdot \pi_{(\iota^*, \rho^*)}(\hat{\beta}).$$

- If $\delta^*(\hat{\beta}) < 1$ and $\mathbb{E}_{\iota^*}[\rho^*(a, q) + \rho^*(b, q)] < 2$,

$$p \cdot \left(1 - \pi_{(\iota^*, \rho^*)}(\hat{\beta})\right) \geq (1 - p) \cdot \pi_{(\iota^*, \rho^*)}(\hat{\beta}).$$

Just like in the case of pure-strategy equilibria, we can refine the set of mixed-strategy equilibria:

Definition 10. An *informative equilibrium in mixed strategies* is an equilibrium in mixed strategies, $(\iota^*, \rho^*, \delta^*)$, in which the worker acquires an informative signal with positive probability: $\mathbb{E}_{\iota^*}[q] > 0$.

With the requisite terminology in hand, it is helpful to begin by observing that, in an informative equilibrium, the manager will not randomize except at possibly one signal quality, which depends on the manager's type.

Lemma 12. *Let (w, m) be an appointment and $(\iota^*, \rho^*, \delta^*)$ be an informative equilibrium in mixed strategies in the associated subgame. There is at most one signal quality, $\hat{q} \in [0, 1/2)$, for which either $\rho^*(a, \hat{q}) \in (0, 1)$ or $\rho^*(b, \hat{q}) \in (0, 1)$, and that quality satisfies the condition $\hat{q} = |m - 1/2|$.*

Proof. Definition 9 indicates that $\rho^*(a, q) \in (0, 1)$ only if $q = m - 1/2$ or $\delta^*(\hat{\alpha}) = \delta^*(\hat{\beta})$. In the latter case, the principal effectively disregards the manager's recommendation. However, it must then follow (again from Definition 9) that ι^* puts probability one on $q = 0$, which contradicts the assumption that $(\iota^*, \rho^*, \delta^*)$ is an informative equilibrium. Thus, it must be the case that $q = m - 1/2$. A symmetric argument shows that $\rho^*(b, q) \in (0, 1)$ only if $q = 1/2 - m$. Note that $\rho^*(a, q_1) \in (0, 1)$ and $\rho^*(b, q_2) \in (0, 1)$ cannot both hold for any $q_1, q_2 \in [0, 1/2)$ unless $m = 1/2$ and $q_1 = q_2 = 0$. $\qquad\square$

Based on Lemma 12, it is possible to simplify the form of the worker's objective function, which is described in Definition 9. For example, if $m \leq 1/2$,[43] then, for $q < 1/2 - m$, the (α-biased) manager recommends α unconditionally (i.e., $\rho^*(a, q) = \rho^*(b, q) = 1$), and, for $q > 1/2 - m$, the manager's recommendation matches the signal's realization (i.e., $\rho^*(a, q) = 1$ and $\rho^*(b, q) = 0$). When $q = 1/2 - m$, a signal of realization b makes the manager indifferent between α and β. Thus, the manager may randomize her recommendation given a signal of quality $1/2 - m$ and realization b: that is, $0 < \rho^*(b, 1/2 - m) < 1$ is possible. In this case, the worker's expected loss when he acquires a signal of quality $1/2 - m$ is a convex combination of his expected loss when the manager discards the signal and his expected loss when the manager follows the signal. In particular, given a manager of type $m \leq 1/2$ that follows the recommendation strategy ρ^* as part of an informative equilibrium in mixed strategies,

$$
\ell_w(q; m, \rho^*, \delta^*) = \begin{cases} L(w, \delta^*) & \text{if } q < 1/2 - m, \\ \rho^*(b, q) \cdot L(w, \delta^*) + (1 - \rho^*(b, q)) \cdot R(q; w, \delta^*) & \text{if } q = 1/2 - m, \\ R(q; w, \delta^*) & \text{if } 1/2 - m < q, \end{cases}
$$

---

[43]For much of the remainder of this section, I will focus on appointments in which $0 < m \leq 1/2$; naturally, cases in which $1/2 < m < 1$ are symmetric.

where

$$L(w, \delta^*) \equiv \frac{w \cdot \delta^*(\hat{\alpha}) + (1 - w) \cdot (1 - \delta^*(\hat{\alpha}))}{2},$$

$$R(q; w, \delta^*) \equiv -\frac{q \cdot \left[\delta^*(\hat{\alpha}) - \delta^*(\hat{\beta})\right]}{2} + \frac{2 - \delta^*(\hat{\alpha}) - \delta^*(\hat{\beta}) + 2w \cdot \left(\delta^*(\hat{\alpha}) + \delta^*(\hat{\beta}) - 1\right)}{4}.$$

Note the similarity between the worker's objective function, $q \mapsto c(q) + \ell_w(q; m, \rho^*, \delta^*)$ (which maps from individual signal qualities to expected loss levels under ρ^* and δ^*), from this setting and the worker's objective function under pure strategies, given by (4). Indeed, (4) is a special case; the two objective functions are identical when ρ^* and δ^* are deterministic. Essentially the same argument as the one provided in the proof of Lemma 2 can be applied here to characterize the solution to the worker's problem. As in the main text, the argument is illustrated graphically below. Figure 8—which is analogous to Figure 3—shows, for an unbiased worker (of type $w = 1/2$), how randomization by α-biased managers of different types (and possibly by the principal as well) affects the worker's objective function and optimal signal quality. As illustrated by Figure 8, it is only in the case of a mildly biased manager (in which the solution to the worker's objective function is denoted by $\hat{q}(\delta^*)$) that an informative equilibrium exists in which the manager randomizes her recommendation when indifferent between α and β; however, that randomization occurs off the equilibrium path. In Figure 8(b) (which shows a moderately biased manager), the discontinuity in the worker's objective function occurs at the signal quality that, in the absence of randomization, would have been optimal for the worker. Under randomization by the manager, though, the worker's objective function does not have a well-defined minimum, so there is no informative equilibrium in which the manager randomizes in this subgame. In Figure 8(c) (which shows a severely biased manager), the worker finds an uninformative signal uniquely optimal; no informative equilibrium in which the manager randomizes exists in this subgame either.

Based on the analysis thus far, randomization by the appointed manager is inconsequential: in an informative equilibrium, only a mildly biased manager (i.e., one with $|m - 1/2| < \hat{q}(\delta^*)$) may possibly randomize her recommendation, and
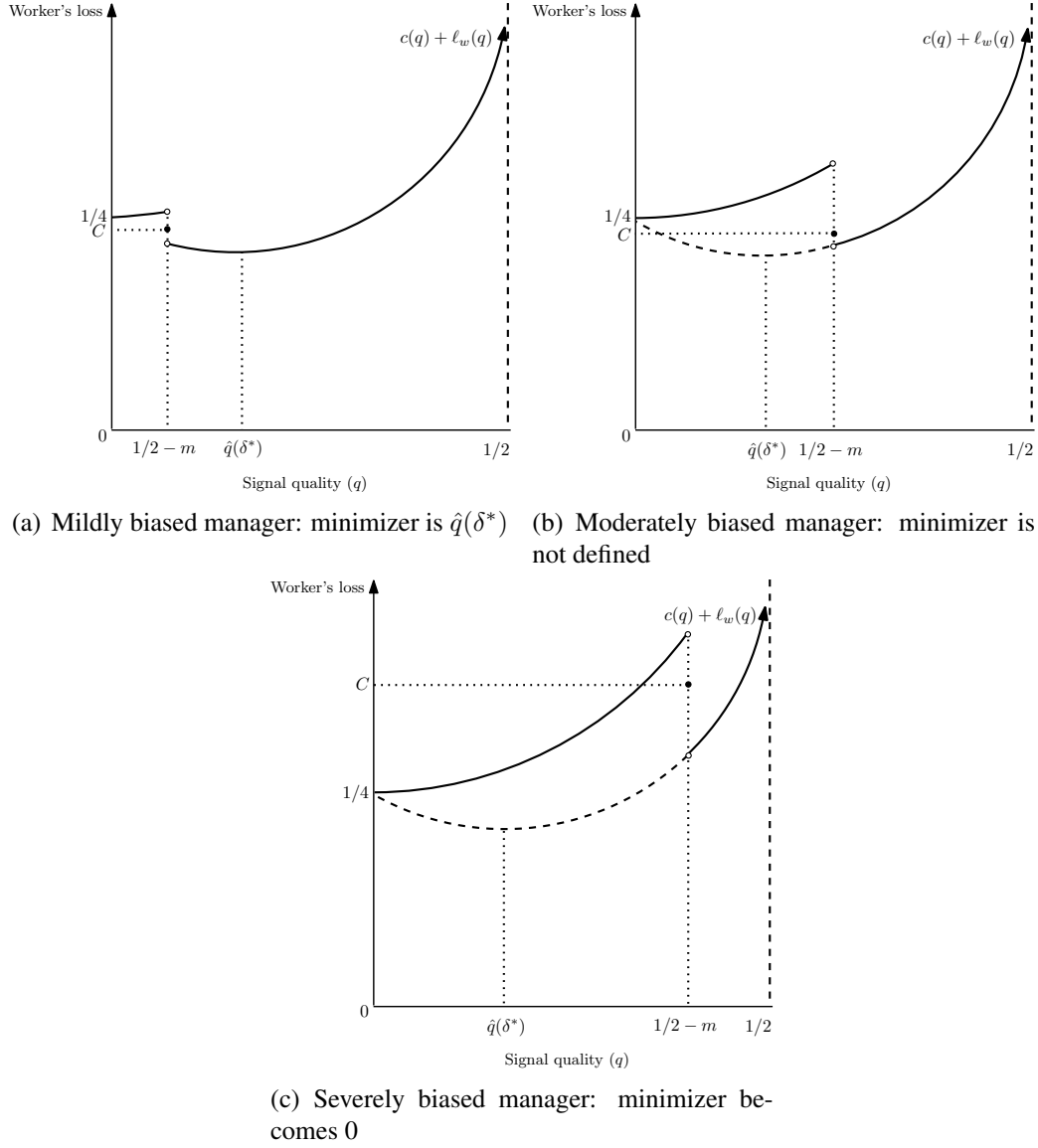
(a) Mildly biased manager: minimizer is $\hat{q}(\delta^*)$

(b) Moderately biased manager: minimizer is not defined

(c) Severely biased manager: minimizer becomes 0

Figure 8: The diagrams above are analogous to the panels in Figure 3 and show how randomization by the manager affects the relationship between the manager's bias and the solution to the unbiased worker's problem. In each of the diagrams, $C \equiv c(1/2 - m) + \rho^*(b, 1/2 - m) \cdot L(w, \delta^*) + [1 - \rho^*(b, 1/2 - m)] \cdot R(1/2 - m; w, \delta^*)$ represents the worker's expected loss from acquiring a signal of quality $1/2 - m$ given mixing by the (indifferent) manager in response to a signal of realization b and quality $1/2 - m$, and $\hat{q}(\delta^*)$ (analogous to q^*) is the global minimum of $c(\cdot) + R(\cdot; w, \delta^*)$.

it is only at a signal quality that is suboptimal for the worker and is thus off the equilibrium path. On the other hand, if managers are assumed to behave as in the baseline model (i.e., to follow the signal with probability one under indifference), the candidate informative equilibria that have been identified for mildly biased managers remain candidate informative equilibria (with no welfare loss to the principal), and, for moderately biased managers, the worker's optimization problem has a well-defined solution, so the set of candidate informative equilibria is potentially larger. From this point, therefore, I assume that any appointed manager will base her recommendation on the signal when she is indifferent. With this assumption in place, I restrict attention to informative equilibria in which only the appointed worker and the principal may randomize.

I now make two observations about the effects of randomization by the principal. First, note that the solution, $\hat{q}(\delta^*)$, to the worker's problem in Figure 8(a) is characterized by the following first-order condition:

$$c'(\hat{q}(\delta^*)) = \frac{\delta^*(\hat{\alpha}) - \delta^*(\hat{\beta})}{2}. \tag{13}$$

Observe that (13) is a more general version of the first-order condition that characterizes q^* (which is analogous to $\hat{q}(\delta^*)$) in the baseline model:

$$c'(q^*) = 1/2. \tag{14}$$

Given the strict convexity of $c(\cdot)$, (13) and (14) imply that $\hat{q}(\delta^*) \leq q^*$, where equality holds precisely when the principal's decision strategy satisfies $\delta^*(\hat{\alpha}) = 1$ and $\delta^*(\hat{\beta}) = 0$; that is, when the principal does not randomize.

The second observation relates to the analogue—which I denote by $\hat{q}^\alpha(w, \delta^*)$—of $q^\alpha(w)$ in this setting. Let $Q_{w,\delta^*} \equiv \{q \in [0, 1/2) : L(w, \delta^*) = c(q) + R(q; w, \delta^*)\}$. Then $\hat{q}^\alpha(w, \delta^*)$ can be characterized as

$$\hat{q}^\alpha(w, \delta^*) \equiv \begin{cases} \max Q_{w,\delta^*} & \text{if } Q_{w,\delta^*} \neq \varnothing, \\ 0 & \text{otherwise.} \end{cases}$$

A straightforward application of the implicit function theorem shows that, when-

ever $\hat{q}^{\alpha}(w, \delta^*)$ is positive, it is increasing in $\delta^*(\hat{\alpha})$ and decreasing in $\delta^*(\hat{\beta})$.[44] As a consequence, $\hat{q}^{\alpha}(w, \delta^*) \leq q^{\alpha}(w)$, and equality holds exactly when $\delta^*(\hat{\alpha}) = 1$ and $\delta^*(\hat{\beta}) = 0$; that is, when the principal follows the recommendation with probability one.

The above two observations can be combined to demonstrate that the principal can do no better (and often will do worse) in an informative equilibrium in which she randomizes than in a pure-strategy informative equilibrium for the same appointment. In particular, fix a randomized decision strategy, δ^*, for the principal, and suppose that a manager of type $m > 1/2 - \hat{q}(\delta^*)$ has been appointed. Then $1/2 - m < \hat{q}(\delta^*) < q^*$. In an informative equilibrium in pure strategies for this subgame, the worker would acquire a signal of quality q^*. However, in a mixed-strategy informative equilibrium in which the principal employs δ^*, the worker would acquire a signal of lower quality: $\hat{q}(\delta^*)$. Similarly, if the manager has type $m \in [1/2 - q^*, 1/2 - \hat{q}(\delta^*)]$, then, in a pure-strategy informative equilibrium, the worker would acquire a signal of quality q^*, while, in a mixed-strategy informative equilibrium, the worker would acquire a signal of quality at most $1/2 - m$, which is no greater than q^*.

Now suppose that $m \in [1/2 - \hat{q}^{\alpha}(w, \delta^*), 1/2 - q^*)$ has been appointed.[45] In this case, regardless of whether the principal follows the manager's recommendation with probability one or follows δ^*, the worker's optimal signal quality is $1/2 - m$, so randomization does not have an adverse effect, but it also does not improve the principal's welfare. However, if $m \in [1/2 - q^{\alpha}(w), 1/2 - \hat{q}^{\alpha}(w, \delta^*))$, then the worker's optimal signal quality is $1/2 - m$ when the principal follows the manager's recommendation with probability one, but her optimal signal quality is zero when the principal employs δ^*. That is, there is no informative equilibrium in which

---

[44]To be precise, the implicit function theorem shows that $\partial \hat{q}^{\alpha}(w, \delta^*)/\partial \delta^*(\hat{\alpha}) > 0$ and $\partial \hat{q}^{\alpha}(w, \delta^*)/\partial \delta^*(\hat{\beta}) < 0$ whenever $\hat{q}^{\alpha}(w, \delta^*) > \hat{q}(\delta^*)$. The desired conclusion follows by observing that $\hat{q}(\delta^*) = \min \{\hat{q}^{\alpha}(w, \delta^*) : 0 < w < 1, \hat{q}^{\alpha}(w, \delta^*) > 0\}$. (Note that the partial derivatives are not defined at $\hat{q}^{\alpha}(w, \delta^*) = \hat{q}(\delta^*)$.)

[45]Of course, this condition assumes that $q^* < \hat{q}^{\alpha}(w, \delta^*)$. If the strict inequality is reversed, and, if $m \in [1/2 - q^*, 1/2 - \hat{q}^{\alpha}(w, \delta^*))$, then the worker will acquire a signal of quality q^* under the pure-strategy informative equilibrium and an uninformative signal if the principal employes δ^*. If $q^* = \hat{q}^{\alpha}(w, \delta^*)$, then a manager of type $m = 1/2 - q^*$ will acquire a signal of quality q^* in both the pure-strategy informative equilibrium and as a best response to the principal's use of δ^*.

the manager of type $m \in [1/2 - q^\alpha(w), 1/2 - \hat{q}^\alpha(w, \delta^*))$ is appointed and the principal follows δ^*. In general, by randomizing, the principal reduces the worker's capacity to persuade an α-biased manager, thus limiting the potential payoff that the principal could earn by appointing a highly skeptical manager. Finally, if a manager of type $m < 1/2 - q^\alpha(w)$ is appointed, there is no informative equilibrium, regardless of the principal's strategy.

To summarize, the worker acquires more information (and the principal does better) in an informative equilibrium in which the principal does not randomize. The reason is intuitively clear: if there is a chance that the signal will be discarded, the worker's marginal benefit from information acquisition is lower than it would be if the signal were heeded.

To this point, I have shown that, for a fixed appointment, randomization by the manager is inconsequential, and that it cannot improve the principal's welfare relative to a pure-strategy informative equilibrium (regardless of the other players' strategies). I have also shown that, for any fixed appointment, randomization by the principal often induces the worker to acquire less—and never to acquire more—information. The only scenario that remains is one in which the worker plays a mixed strategy, while the manager and principal play pure strategies. In this case, the worker's objective function is the same as in the baseline model. Note that the worker is willing to play a mixed strategy only when, given the appointed manager's type, the worker is indifferent between acquiring an uninformative signal and persuading the manager. That is, either $1/2 - m = q^\alpha(w)$ or $m - 1/2 = q^\beta(w)$ is necessary for the worker to randomize. It is clear, however, that, if the worker randomizes in either case, the expected quality of the decision suffers (since, with positive probability, the decision is made based on an uninformative signal), as does the principal's expected welfare.

B.6 Extensions

This appendix explores three extensions to the model: a relaxation of the assumption of common priors, the introduction of richer information structures, and the use of transfer payments by the principal. The results are generally robust to these

modifications.

B.6.1 Private Prior Beliefs

In the baseline model of this paper, each player believes that each of the two states is equally likely. The analysis of this section demonstrates that the results are robust to allowing nonuniform prior beliefs and open disagreement among players, since the possibility of private prior beliefs does not introduce any noteworthy effects into the model that are not already captured by differences in biases. In particular, for the principal and managers, differences in prior beliefs are observationally equivalent to differences in biases.[46] For workers, differences in prior beliefs can affect information acquisition incentives, but not in a manner that substantively alters the main results.

Maintain all features of the baseline model, but suppose that the principal's prior beliefs assign probability $\pi \in (0, 1)$ to state A. Note that the principal's ex-ante preference between α and β is determined completely by the relative magnitudes of p and π; for example, the principal weakly prefers α to β ex ante if and only if $p \leq \pi$. Candidate workers and candidate managers also may vary in their prior beliefs. In particular, the pool, W, of candidate workers is a subset of $(0, 1)^2$, and a candidate worker may be identified by an ordered pair of the form (w, ω), where ω is the probability that the candidate worker's prior beliefs assign to A. Similarly, a candidate manager can be identified by an ordered pair of the form (m, μ), where μ is the probability that the candidate manager's prior beliefs assign to A.

The first point to note is that, for managers, the modified type space can be reduced, without loss of generality, to the one of the original model. To make this point precise, a notion of preference alignment (introduced in Nayeem (2014)) is useful:

Definition 11. Two candidate managers have *aligned preferences* if, given any signal quality and realization, the managers' preferences between α and β coincide.

---

[46]This point is demonstrated for managers by Nayeem (2014), who allows both the principal and workers to vary in both their prior beliefs and their biases. The arguments of this section follow the analysis of Nayeem (2014).

Thus, given a common belief about the principal's decision making rule, two candidate managers with aligned preferences will send the same message in response to any signal that a worker may acquire: their types are observationally indistinguishable in this game. With this point in mind, Lemma 13, also due to Nayeem (2014), shows that the augmented type space of candidate managers exhibits some redundancy.

Lemma 13. *Let $(m, \mu) \in M$. For any $\mu' \in (0, 1)$, there exists a unique $m' \in (0, 1)$ such that a manager of type (m, μ) and a manager of type (m', μ') have aligned preferences.*

Proof. Let m, μ, and μ' be given, and set

$$m' \equiv \left[1 + \frac{\mu \cdot (1 - \mu') \cdot (1 - m)}{m\mu' \cdot (1 - \mu)}\right]^{-1}.$$

It is straightforward to verify that the following condition characterizes the above choice of m':

$$\frac{m \cdot (1 - \mu)}{\mu \cdot (1 - m)} = \frac{m' \cdot (1 - \mu')}{\mu' \cdot (1 - m')}. \tag{15}$$

The following condition is equivalent to (15): for every $q \in [0, 1/2)$,

$$- m \cdot \left(1 + \frac{\mu}{1 - \mu} \cdot \frac{1/2 + q}{1/2 - q}\right)^{-1} \lesseqqgtr -(1 - m) \cdot \left(1 + \frac{1 - \mu}{\mu} \cdot \frac{1/2 - q}{1/2 + q}\right)^{-1} \text{ iff}$$

$$- m' \cdot \left(1 + \frac{\mu'}{1 - \mu'} \cdot \frac{1/2 + q}{1/2 - q}\right)^{-1} \lesseqqgtr -(1 - m') \cdot \left(1 + \frac{1 - \mu'}{\mu'} \cdot \frac{1/2 - q}{1/2 + q}\right)^{-1} \tag{16}$$

and

$$- m \cdot \left(1 + \frac{\mu}{1 - \mu} \cdot \frac{1/2 - q}{1/2 + q}\right)^{-1} \lesseqqgtr -(1 - m) \cdot \left(1 + \frac{1 - \mu}{\mu} \cdot \frac{1/2 + q}{1/2 - q}\right)^{-1} \text{ iff}$$

$$- m' \cdot \left(1 + \frac{\mu'}{1 - \mu'} \cdot \frac{1/2 - q}{1/2 + q}\right)^{-1} \lesseqqgtr -(1 - m') \cdot \left(1 + \frac{1 - \mu'}{\mu'} \cdot \frac{1/2 + q}{1/2 - q}\right)^{-1}. \tag{17}$$

Note that the above condition formally states that the candidate managers of types (m, μ) and (m', μ') have aligned preferences. Given its equivalence with (15), which characterizes the choice of m', the preference alignment condition determines m' uniquely as a function of m, μ, and μ'. $\square$

Since two candidate managers with aligned preferences are indistinguishable, Lemma 13 indicates that, in the context of the model, differences in prior beliefs cannot be separated from differences in biases. Therefore it is sufficient to allow candidate managers to vary only in their biases; in particular, setting $\mu' \equiv 1/2$ in the statement of the result shows that any candidate manager type in the augmented space has an analogue in the original space (in which all types have uniform prior beliefs). As a result, allowing candidate managers to hold nonuniform prior beliefs, or even to vary in their prior beliefs, does not affect the results.[47] A similar argument shows that the principal also can be assumed to hold uniform prior beliefs.

For candidate workers, there is a caveat: differences in prior beliefs can affect workers' information acquisition incentives. To see how, suppose that the principal has appointed a worker of type (w, ω) and a manager of type $(m, 1/2)$, where $0 < m < 1$, and that it is common knowledge that the principal will rubberstamp the manager's recommendation. Then the worker's problem can be written as

$$\min_{q \in [0, 1/2)} \kappa(q; w, \omega, m),$$

where, if $m \leq 1/2$,

$$\kappa(q; w, \omega, m) \equiv \begin{cases} c(q) + w \cdot (1 - \omega) & \text{if } q \in [0, 1/2 - m), \\ c(q) + (1/2 - q) \cdot (w + \omega - 2w\omega) & \text{if } q \in [1/2 - m, 1/2), \end{cases}$$

and, if $m > 1/2$,

$$\kappa(q; w, \omega, m) \equiv \begin{cases} c(q) + \omega \cdot (1 - w) & \text{if } q \in [0, m - 1/2), \\ c(q) + (1/2 - q) \cdot (w + \omega - 2w\omega) & \text{if } q \in [m - 1/2, 1/2). \end{cases}$$

---

[47]In fact, given the identification problem between prior beliefs and biases, one may argue that allowing both to vary is not even meaningful.

The result of Lemma 2, and the analysis that follows it, generalizes in a straightforward manner. In particular, the worker-optimal signal quality under rubberstamping, $q^*(w, \omega)$, is now characterized by $c'(q^*(w, \omega)) = w + \omega - 2w\omega$. The maximal signal quality, denoted by $q^\alpha(w, \omega)$, that a worker of type (w, ω) would be willing to acquire when facing an α-biased manager is now defined to be the largest fixed point of the function

$$\varphi^\alpha_{(w,\omega)} : [0, 1/2) \to \mathbb{R}$$
$$q \mapsto \frac{2c(q) + \omega - w}{2w + 2\omega - 4w\omega},$$

if it has any fixed points, and 0 otherwise. The maximal signal quality, $q^\beta(w, \omega)$, that the worker would be willing to acquire when facing a β-biased manager is defined analogously with respect to the function

$$\varphi^\beta_{(w,\omega)} : [0, 1/2) \to \mathbb{R}$$
$$q \mapsto \frac{2c(q) + w - \omega}{2w + 2\omega - 4w\omega}.$$

Two key factors affect the information acquisition decision for a worker of type (w, ω). These factors are evident in the following decompositions of $\varphi^\alpha_{(w,\omega)}(q)$ and $\varphi^\beta_{(w,\omega)}(q)$:

$$\varphi^\alpha_{(w,\omega)}(q) = \frac{c(q)}{w + \omega - 2w\omega} - \frac{w - \omega}{2 \cdot (w + \omega - 2w\omega)},$$
$$\varphi^\beta_{(w,\omega)}(q) = \frac{c(q)}{w + \omega - 2w\omega} + \frac{w - \omega}{2 \cdot (w + \omega - 2w\omega)}.$$

Consider the quantities $w - \omega$ and $w + \omega - 2w\omega$, which are depicted as functions of w and ω in Figure 9. The former is an obvious analogue of $w - 1/2$ from the baseline model. Recall that, in the baseline model, a high value of w (and hence of $w - 1/2$) is associated with a high value of $q^\alpha(w)$ and a low value of $q^\beta(w)$. Similarly, in this setting, a high value of $w - \omega$ is associated with a high value of $q^\alpha(w, \omega)$ and a low value of $q^\beta(w, \omega)$. Thus, given the appointment of an α-biased manager, and holding $w + \omega - 2w\omega$ fixed, candidate worker types with high values of $w - \omega$
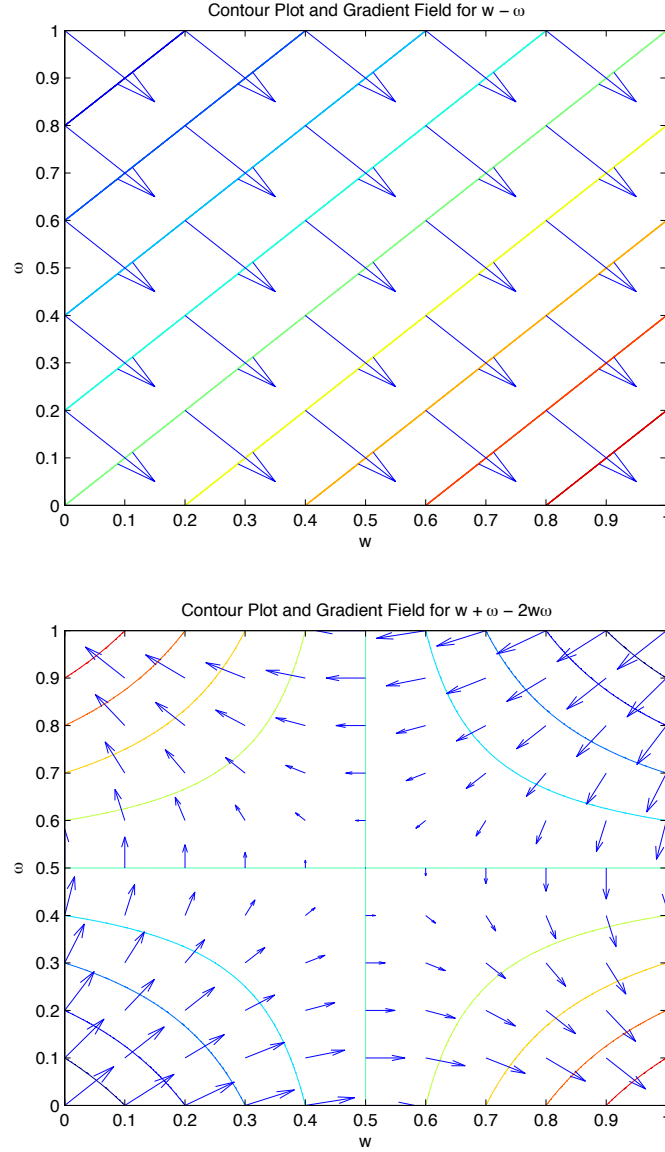
33

Figure 9: The curves in each plot are isoquants (in (w, ω) space) of the associated function. Each arrow points in the direction of the function's steepest increase from the arrow's origin. [Source: Nayeem (2014)]

can be induced to acquire better information. Similarly, given the appointment of a β-biased manager, and holding $w + \omega - 2w\omega$ fixed, candidate worker types with high values of $\omega - w$ can be induced to acquire better information.

Now consider the second quantity, $w + \omega - 2w\omega$.[48] As the expressions for $\kappa(q; w, \omega, m)$ and the above first-order condition indicate, this quantity represents the marginal benefit of an increase in information quality, provided that the signal influences the principal's decision. Thus, holding $w - \omega$ fixed, the larger this quantity is, the stronger is the worker's incentive to acquire information. Figure 9 illustrates that the candidate worker types with the highest values of $w + \omega - 2w\omega$ are those for which $|w - 1/2|$ is close to $1/2$ and ω is close to $1 - w$; incidentally, $|w - \omega|$ is high for such types as well. Hence, such types (provided that the manager is appointed appropriately) can be induced to acquire the best information. A worker of such a type believes that the state of the world in which his utility is highly dependent on the principal's decision is very probable. For example, a candidate worker with $w \approx 1$ and $\omega \approx 0$ believes with near certainty that B is the correct state (and hence that β is optimal) and also suffers a relatively severe loss when the principal errs by choosing α. Note that, just as in the baseline model, the workers that exhibit the greatest potential for information acquisition are fierce supporters of their ex-ante-preferred actions.

It is clear that allowing candidate workers to differ in their prior beliefs as well as biases adds complexity to the candidate worker type space. In the baseline model, candidate workers can be "ranked" (as in Proposition 4, for example) by their unidimensional biases, which completely characterize their information acquisition incentives. In this more general environment, workers vary in $q^*(w, \omega)$, $q^\alpha(w, \omega)$, and $q^\beta(w, \omega)$. Each of these quantities is a factor in a worker's information acquisition decision. Thus, ranking the candidate worker types is not straightforward in this setting. Nevertheless, the analysis above indicates that the paper's main message—that the principal benefits by deliberately introducing dissonance within the organization—remains valid whether dissonance takes the form of preference misalignment or open disagreement (or both).

---

[48]Note that, in the baseline model (where $\omega = 1/2$), $w + \omega - 2w\omega = 1/2$ for all values of w.

B.6.2 Richer Information Structures

The above results might appear to be driven by the discontinuity of the worker's objective function, which, in turn, is implied by the assumed signal structure. More generally, the signal structure of the baseline model—with only two realizations and symmetric informativeness levels—might appear overly stylized and restrictive. In this section, I show that the binary signal structure is not an essential feature that drives the results. In particular, preference misalignment can benefit the principal under alternate signal structures, including ones that produce continuous objective functions.

Just as in the original game, the principal must choose between α, which is optimal in state A, and β, which is optimal in state B. The focus here will be on informative equilibria, so it will be common knowledge that the principal acts according to the manager's recommendation. All actors hold the prior belief that each of the two states is equally likely,[49] and an appointed worker can acquire information about the state by observing the realization of a signal that he chooses from an exogenously given set of signals. Each signal is a random variable with a commonly known joint distribution with the state. In principle, each signal may take values in an arbitrary space, which might differ from the spaces in which other signals take values. However, for any signal, the posterior belief that is induced by a given realization captures the realization's informational content, and thus the realization can be identified with a value from $[0, 1]$. Therefore, without loss of generality, each signal is a random variable (with commonly known joint distribution with the state) with support in that interval.[50] (In fact, a result due to Kamenica and Gentzkow (2011) implies that a random variable, σ, that has support in $[0, 1]$ is a signal of the type just described if and only if $\mathbb{E}[\sigma] = 1/2$, which is equivalent to the condition $\mathbb{E}[\sigma|A] + \mathbb{E}[\sigma|B] = 1$.) Let Σ denote the set of available signals. To reveal the realization of any $\sigma \in \Sigma$, a worker must incur a nonnegative cost, denoted by $c(\sigma)$.

---

[49]The assumption of common and uniform prior beliefs can be relaxed as shown in Appendix B.6.1.

[50]Note that, while it is possible for a given posterior belief to be associated with more than one realization of a given signal, an appointed manager will treat any two realizations that correspond to the same posterior belief equivalently. Thus, each realization from $[0, 1]$ corresponds to an equivalence class of realizations from the original set.

Σ contains exactly one uninformative signal, denoted by σ_0, which takes the value $1/2$ with probability one and has zero cost. All other signals are informative (i.e., have supports that are different from the singleton $\{1/2\}$) and have positive costs.

Under this setup, a manager of type $m \leq 1/2$,[51] upon observing a realization $s \in [0,1]$, recommends α if $m \leq s$ and recommends β otherwise.[52] The worker's problem, then, can be expressed as

$$\min_{\sigma \in \Sigma} c(\sigma) + w \cdot \Pr(\{\sigma \geq m\} | B)/2 + (1-w) \cdot \Pr(\{\sigma < m\} | A)/2. \quad (18)$$

Suppose that the principal believes that, for each worker type $w \in W$ and each manager type $m \in M$, $\hat{\sigma}(w, m) \in \Sigma$ is the signal that the worker of type w will acquire given that the manager of type m is appointed.[53] Then, given these beliefs, the principal's problem becomes

$$\min_{w \in W, m \in M} p \cdot \Pr(\{\hat{\sigma}(w,m) \geq m\} | B)/2 + (1-p) \cdot \Pr(\{\hat{\sigma}(w,m) < m\} | A)/2. \quad (19)$$

Note that, in this more general setting, the informativeness of a signal's realization (i.e., the realization's distance from $1/2$) may depend stochastically—rather than deterministically—on the signal's cost. Roughly speaking, a signal, σ, may be viewed as having high quality if the error probabilities that it induces (i.e., $\Pr(\{\sigma < m\} | A)$ and $\Pr(\{\sigma \geq m\} | B)$) are relatively low across different manager types, $m \in M$. However, it is not generally possible to order signals according to these quantities. For example, given two signals, σ and σ', and a manager type, m, it is not necessarily the case that either of the following inequalities implies the other:

$$\Pr(\{\sigma < m\} | A) < \Pr(\{\sigma' < m\} | A),$$

---

[51]For simplicity, I consider only α-biased managers—so that $M \subseteq (0, 1/2]$—in this section, but it is straightforward to extend the analysis to cover β-biased managers as well.

[52]I assume that the manager will act upon her ex-ante preference for α if she is indifferent under her posterior belief. This assumption is inconsequential if the worker chooses a signal σ for which $\Pr\{\sigma = m\} = 0$ (which occurs, for example, if the distribution of σ is atomless).

[53]Since the solution to (18) need not be not unique, the structure of an informative equilibrium might depend on these beliefs.

$$\Pr\left(\{\sigma \geq m\} | B\right) < \Pr\left(\{\sigma' \geq m\} | B\right).$$

Thus, the notion of quality may be less clear-cut than in the baseline model. As it turns out, however, the main insights of the previous sections do not rely on strong assumptions related to signal ordering. The following example illustrates this fact.

Example 6. Let $\{\psi_q\}_{q \geq 0}$ be a collection of random variables that are independent conditional on the state. Suppose that ψ_q has the following conditional distributions:

$$\Pr\left(\{\psi_q \leq t\} | A\right) = \begin{cases} 0 & \text{if } t < 0, \\ t^{q+1} & \text{if } 0 \leq t \leq 1, \\ 1 & \text{otherwise,} \end{cases}$$

$$\Pr\left(\{\psi_q \leq t\} | B\right) = \begin{cases} 0 & \text{if } t < 0, \\ t^{1/(q+1)} & \text{if } 0 \leq t \leq 1, \\ 1 & \text{otherwise.} \end{cases}$$

The conditional distributions of ψ_q are depicted in Figure 10. Note that, for each $q \geq 0$, $\text{supp}\left(\psi_q\right) = [0, 1]$, and that the following two functions are conditional densities for ψ_q:

$$f_{\psi_q}(t|A) \equiv \begin{cases} (q+1) \cdot t^q & \text{if } 0 < t < 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$f_{\psi_q}(t|B) \equiv \begin{cases} \left[(q+1) \cdot t^{q/(q+1)}\right]^{-1} & \text{if } 0 < t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that, for any $q \geq 0$, the ratio $f_{\psi_q}(\cdot|A)/f_{\psi_q}(\cdot|B)$ is nondecreasing—and strictly increasing for positive values of q—on $(0, 1)$. Thus, a higher realization of ψ_q leads to a stronger posterior belief that the state is A. Furthermore, for $q_1 > q_2 \geq 0$, the conditional distribution of ψ_{q_1} under state A first-order stochastically dominates that of ψ_{q_2} under state A, and the conditional distribution of ψ_{q_2} under state B first-order stochastically dominates that of ψ_{q_1} under state B. These relationships suggest that the index q can be viewed as an informativeness parame-
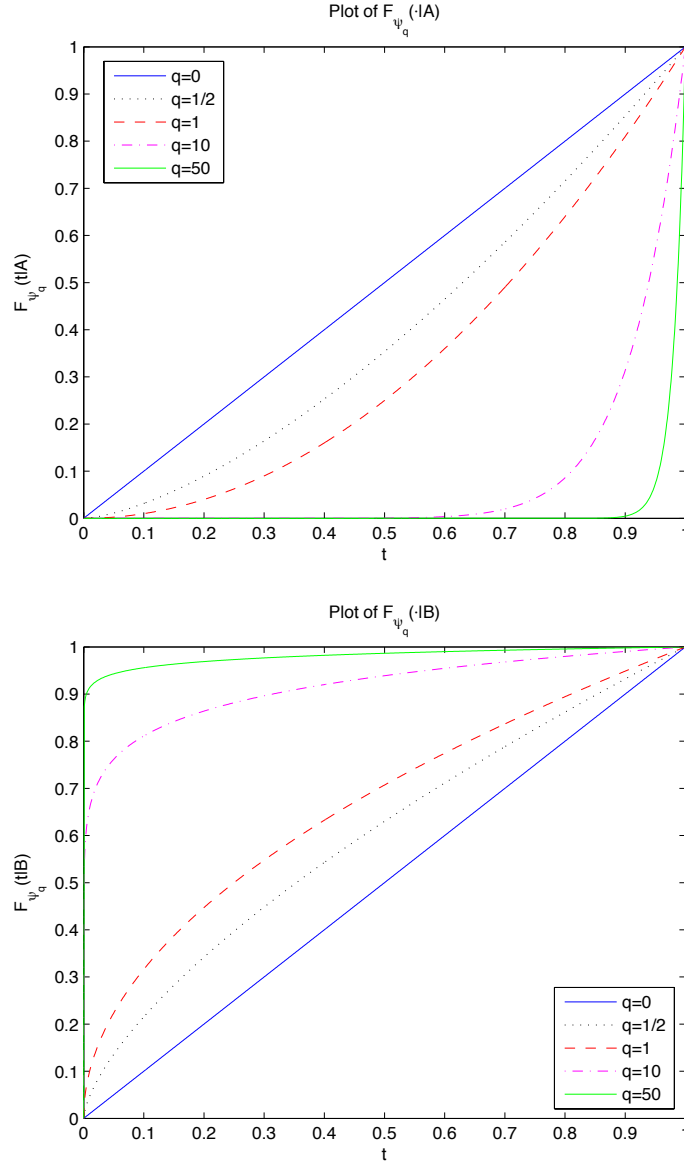
Figure 10: The conditional cumulative distribution functions of ψ_q for various values of q

ter.

For each $q \geq 0$, let σ_q be a signal that represents the information structure defined by ψ_q. In particular, the realization of σ_q is the posterior belief (assuming

that this belief is well defined) that the state is A given the realization of ψ_q:

$$\sigma_q \equiv \begin{cases} \left[1 + f_{\psi_q}(\psi_q|B)/f_{\psi_q}(\psi_q|A)\right]^{-1} & \text{if } f_{\psi_q}(\psi_q|A) \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$= \begin{cases} \left[1 + (q+1)^{-2}\psi_q^{-(q^2+2q)/(q+1)}\right]^{-1} & \text{if } \psi_q \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

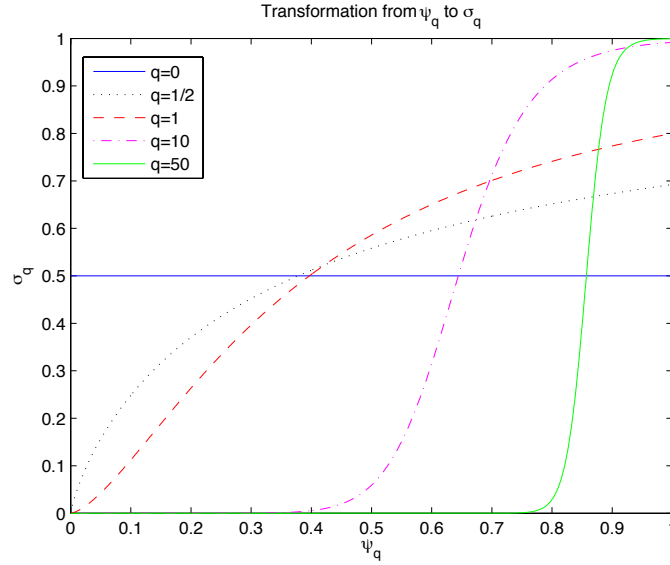The relationship between ψ_q and σ_q is illustrated in Figure 11. Under this setup, a

Figure 11: For each realization of ψ_q, the corresponding value of σ_q represents the induced posterior belief that the state is A. Note that ψ_0 is uninformative, and that extreme values of ψ_q become more informative as q increases.

worker of type w facing a manager of type m solves

$$\min_{q \geq 0} c(\sigma_q) + w \cdot \left[1 - F_{\sigma_q}(m|B)\right]/2 + (1-w) \cdot F_{\sigma_q}(m|A)/2,$$

40

where, for $s \in [0, 1]$,

$$
F_{\sigma_q}(s|A) = \begin{cases} 0 & \text{if } s = 0 \\ & \text{or } (q = 0 \text{ and } s < 1/2), \\ [(q+1)^2 \cdot (1-s)/s]^{-(q^2+2q+1)/(q^2+2q)} & \text{if } q > 0 \\ & \text{and } 0 < s \le \frac{q^2+2q+1}{q^2+2q+2}, \\ 1 & \text{otherwise}, \end{cases}
$$

$$
F_{\sigma_q}(m|B) = \begin{cases} 0 & \text{if } s = 0 \\ & \text{or } (q = 0 \text{ and } s < 1/2), \\ [(q+1)^2 \cdot (1-s)/s]^{-1/(q^2+2q)} & \text{if } q > 0 \\ & \text{and } 0 < s \le \frac{q^2+2q+1}{q^2+2q+2}, \\ 1 & \text{otherwise}, \end{cases}
$$

are the conditional cumulative distribution functions for σ_q. Note that the signal structure is asymmetric; it exhibits a bias toward β. In particular, for any $q > 0$ and $\varepsilon > 0$, $\Pr\{\sigma_q < \varepsilon\} > 0$; that is, for any positive q, the posterior belief that the state is B can be arbitrarily close to one. On the other hand, $\Pr\left\{\sigma_q \le \frac{q^2+2q+1}{q^2+2q+2}\right\} = 1$ for any q; the belief that the state is A is bounded away from one for any fixed q. Furthermore, as Figure 12 illustrates, although the "error probabilities" (i.e., $\Pr(\{\sigma_q < m\}|A)$ and $\Pr(\{\sigma_q \ge m\}|B)$) are eventually decreasing in q for various values of m, $\Pr(\{\sigma_q < m\}|A)$ (the probability of the (β, A) error) is actually increasing in q when q is small. This analysis underscores the point that the concept of signal ordering can take many forms in a general setting. Indeed, the notion of signal quality in this example is rather weak.

Let $c(q) \equiv q^2/16$, $p \equiv 1/2$, and $M \equiv (0, 1/2]$. If W is a singleton, the principal's problem reduces to the choice of a manager. The basic tradeoff is the same as in the original model: although the principal can use a manager's bias to improve the worker's information acquisition incentives, this bias may lead to ex-post disagreement between the principal and manager about the decision given the signal. For example, if $W = \{1/2\}$, then it turns out that the principal does best by
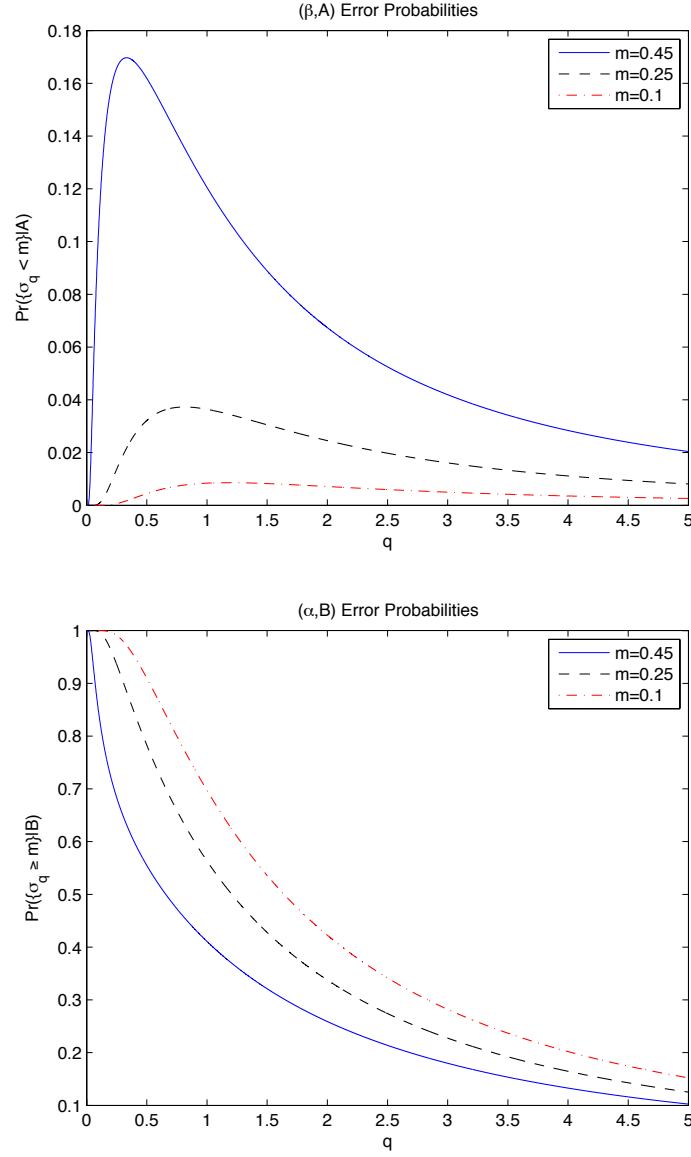
41

Figure 12: The error probabilities generally decline in q, but not monotonically in the case of the (β, A) error.

choosing a like-minded manager, with $m = 1/2$.

When biased workers are available, though, the principal can do better. For example, if $W = \{1/2, 3/4\}$, then the principal's optimal appointment is the worker of type $3/4$ and a manager of type approximately 0.388. If $W = \{1/2, 3/4, 7/8\}$,

then the optimal appointment is the worker of type $7/8$ and a manager of type approximately 0.369. As in the baseline model, a more biased worker has stronger incentives to counteract a manager's opposing bias and thus can be induced to acquire better information. Thus, when appointed in tandem with a worker of opposing bias, a biased manager can be helpful.

Figure 13 illustrates the solution to the principal's problem in this example. Note the following features and their similarity to the result of Proposition 4:

- Given any appointed manager, a worker with a stronger bias toward β will find it optimal to acquire a (weakly) better signal than one with a weaker bias toward β.

- For any appointed manager, the principal will (weakly) decrease her expected loss by appointing a worker with a stronger bias toward β rather than one with a weaker bias toward β.

- The principal's optimal choice of manager, given a worker, becomes more biased toward α as the appointed worker type becomes more biased toward β.

Example 6 clearly demonstrates that the main insights of the baseline model extend to other informational environments. However, the example contains more structure (e.g., signals that are parametrized in a clean and easily interpretable fashion) than is needed for such extensions to be possible. To see this point, let $\Sigma \equiv \{\sigma, \sigma'\}$, where $c(\sigma') > c(\sigma)$. For any $m \in M \equiv \{m_1, m_2\}$ (where $1/2 \geq p = m_1 > m_2$), let

$$\Delta P_A(m) \equiv [\Pr(\{\sigma' < m\} | A) - \Pr(\{\sigma < m\} | A)] / 2,$$
$$\Delta P_B(m) \equiv [\Pr(\{\sigma' \geq m\} | B) - \Pr(\{\sigma \geq m\} | B)] / 2.$$
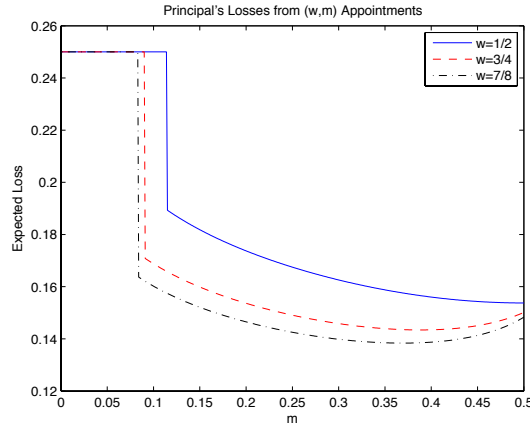
Let $W = \{w_1, w_2\}$ (where $w_1 > w_2 = p$), and suppose that, for each $m \in M$,

$$w_2 \cdot \Delta P_B(m) + (1 - w_2) \cdot \Delta P_A(m) < c(\sigma') - c(\sigma),$$

meaning that the worker of type w_2 prefers to acquire σ rather than σ', regardless

43
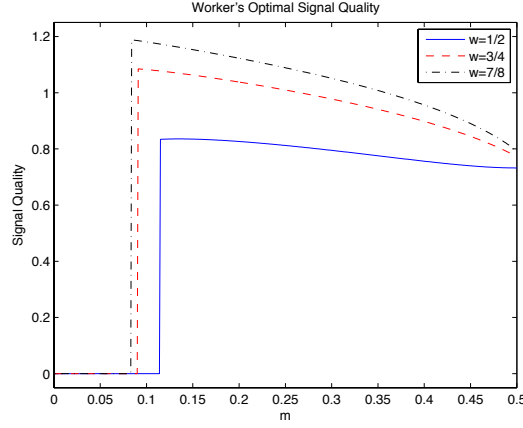
(a) For any fixed manager type m, a more biased worker from W (i.e., one with a higher bias) will acquire a signal of weakly higher quality when facing m.

(b) The principal does best by appointing the most biased worker from W (of type $7/8$) while appointing a moderately biased manager (of type approximately 0.369).

Figure 13: In Example 6, the principal can exploit a biased worker's desire to persuade a manager that opposes her bias. Unlike in the baseline model, though, the principal's optimal choice of manager for a given worker is not the one that maximizes the worker's information acquisition incentives.

of which manager is appointed. Suppose also that

$$p \cdot \left[\Pr\left(\{m_2 \le \sigma' < m_1\} \,|\, B \right)/2 + \Delta P_B(m_1) \right]$$

44

$$< (1 - p) \cdot \left[\Pr\left(\{ m_2 \leq \sigma < m_1 \} \,|\, A \right) / 2 - \Delta P_A(m_2) \right],$$

which implies that the principal's expected welfare is higher when she follows the recommendation of the manager of type m_2 based on the realization of σ' as compared to the case in which she follows the recommendation of the manager of type m_1 based on the realization of σ. Then, if

$$w_1 \cdot \Delta P_B(m_1) + (1 - w_1) \cdot \Delta P_A(m_1) < c(\sigma') - c(\sigma)$$

and

$$w_1 \cdot \Delta P_B(m_2) + (1 - w_1) \cdot \Delta P_A(m_2) > c(\sigma') - c(\sigma)$$

hold, the worker of type w_1 will prefer to acquire σ' if and only if the (more strongly biased) manager of type m_2 is appointed.[54] Hence, the principal does best by appointing w_1 and m_2—that is, by deliberately introducing preference misalignment. Intuitively, the signal σ' is costlier than σ, and its main advantage is in reducing the probability of an error made under state B, particularly when m_2 is the manager. The worker of type w_1, who suffers more under such errors than the worker of type w_2, will have stronger incentives to acquire σ'. The principal also benefits from the acquisition of the higher quality signal and is even willing to appoint the manager of type m_2 (despite this manager's greater tendency to err under state A) so that the worker of type w_1 will acquire it. Note that, despite the principal's preference for appointing w_1 and m_2 (thereby ensuring that the signal σ' is acquired), there need not be a general statistical sense in which σ' dominates σ.

The above heuristic analysis can be formalized using the terminology of monotone comparative statics, as developed by Milgrom and Shannon (1994). For a fixed principal type, p, and manager type, m, define the following preference relation on Σ:

$$\sigma \succsim_m^p \sigma' \overset{\Delta}{\Longleftrightarrow}$$

---

[54] A necessary condition for the inequalities $w_2 \cdot \Delta P_B(m_2) + (1 - w_2) \cdot \Delta P_A(m_2) < c(\sigma') - c(\sigma)$ and $w_1 \cdot \Delta P_B(m_2) + (1 - w_1) \cdot \Delta P_A(m_2) > c(\sigma') - c(\sigma)$ to be logically consistent is $\Delta P_B(m_2) > \Delta P_A(m_2)$.

$$p \cdot \Pr\left(\{\sigma \geq m\} \,|B\right) + (1-p) \cdot \Pr\left(\{\sigma < m\} \,|A\right)$$
$$\leq p \cdot \Pr\left(\{\sigma' \geq m\} \,|B\right) + (1-p) \cdot \Pr\left(\{\sigma' < m\} \,|A\right).$$

That is, $\succsim_m^p$ orders the available signals by the expected loss to the principal of type p, assuming that a manager of type m has been appointed. The following assumption will be important for this section's results.

Assumption 1. $\succsim_m^p$ is antisymmetric: if $\sigma \succsim_m^p \sigma'$ and $\sigma' \succsim_m^p \sigma$, then $\sigma = \sigma'$.

Assumption 1 rules out indifference (on the principal's part, given any appointed manager) between signals. It holds if, for example, for any $m \in M$, the signals can be ordered such that they are simultaneously decreasing in both $\Pr\left(\{\cdot \geq m\} \,|B\right)$ and $\Pr\left(\{\cdot < m\} \,|A\right)$.[55]

Under Assumption 1, the pair $(\Sigma, \succsim_m^p)$ constitutes a chain, and hence a lattice. In particular, given any $\sigma, \sigma' \in \Sigma$, $\sigma \wedge \sigma'$ and $\sigma \vee \sigma'$ are well defined:

$$\sigma' \wedge \sigma'' \equiv \begin{cases} \sigma' & \text{if } \sigma'' \succsim_m^p \sigma', \\ \sigma'' & \text{if } \sigma' \succsim_m^p \sigma'', \end{cases}$$

$$\sigma' \vee \sigma'' \equiv \begin{cases} \sigma'' & \text{if } \sigma'' \succsim_m^p \sigma', \\ \sigma' & \text{if } \sigma' \succsim_m^p \sigma''. \end{cases}$$

The main insight of this section is conveyed through two results, which have a common conclusion but rely on different combinations of assumptions. Before stating these assumptions, it is convenient to define some notation: for any $w \in W$ and $\sigma \in \Sigma$, let

$$\ell_w(\sigma; m) \equiv c(\sigma) + w \cdot \Pr\left(\{\sigma \geq m\} \,|B\right)/2 + (1-w) \cdot \Pr\left(\{\sigma < m\} \,|A\right)/2$$

denote the loss that a worker of type w expects to incur from acquiring the signal σ, given that a manager of type m has been appointed.

Assumption 2. Let $m \in M$ be given. Suppose that $w, w' \in W$ and $\sigma, \sigma' \in \Sigma$ satisfy $w > w'$ and $\sigma \succsim_m^p \sigma'$. Then:

---

[55]Note that Assumption 1 does not hold for the signals either in the baseline model or in Example 6. This analysis therefore should be viewed as an extension rather than as a generalization.

46

(i) $\ell_{w'}(\sigma; m) \leq \ell_{w'}(\sigma'; m)$ implies $\ell_w(\sigma; m) \leq \ell_w(\sigma'; m)$;

(ii) $\ell_{w'}(\sigma; m) < \ell_{w'}(\sigma'; m)$ implies $\ell_w(\sigma; m) < \ell_w(\sigma'; m)$.

Assumption 2 is a version of the single-crossing property. In this context, for any fixed manager type and any pair of signals, if a worker that is relatively less biased toward β prefers to acquire the signal that the principal prefers, then so does any worker that is more biased toward β. The condition might reflect, for example, that signals tend to differ to a relatively high degree in the error probability $\Pr(\{\cdot \geq m\}|B)$ as compared to the error probability $\Pr(\{\cdot < m\}|A)$ and to the cost $c(\cdot)$.

Assumption 3. An appointed worker that is indifferent among optimal signals selects the principal's most preferred signal (from the set of optima).[56] That is, given any $w \in W$, $m \in M$, and $\sigma \in \Sigma$, $\ell_w(\hat{\sigma}(w, m); m) \leq \ell_w(\sigma; m)$. Furthermore, if equality holds in the previous expression, then $\hat{\sigma}(w, m) \succsim^p_m \sigma$.

Proposition 5. *Under Assumptions 1, 2 and 3, for any fixed manager type, the principal's welfare is nondecreasing in the type of worker that is appointed.*

Proof. Let $w', w'' \in W$ and $m \in M$ be given. Suppose that $w' > w''$. Let $\sigma' \equiv \hat{\sigma}(w', m)$, and let $\sigma'' \equiv \hat{\sigma}(w'', m)$. Note that $\ell_{w''}(\sigma' \vee \sigma''; m) \leq \ell_{w''}(\sigma'; m)$. By Assumption 2, $\ell_{w'}(\sigma' \vee \sigma''; m) \leq \ell_{w'}(\sigma'; m)$. Since $\sigma' \in \arg\min_{\sigma \in \Sigma} \ell_{w'}(\sigma; m)$ by hypothesis, $\ell_{w'}(\sigma' \vee \sigma''; m) = \ell_{w'}(\sigma'; m)$. By Assumption 3, $\sigma' \succsim^p_m \sigma' \vee \sigma''$. Hence $\sigma' \succsim^p_m \sigma''$ (i.e., $\hat{\sigma}(w', m) \succsim^p_m \hat{\sigma}(w'', m)$). $\square$

A variant of Proposition 5 holds in the absence of Assumption 3 but requires a stronger version of Assumption 2: a *strict* single-crossing property.[57]

Assumption 4. Let $m \in M$ be given. Suppose that $w, w' \in W$ and $\sigma, \sigma' \in \Sigma$ satisfy $w > w'$ and $\sigma \succ^p_m \sigma'$ (i.e., $\sigma \succsim^p_m \sigma'$ and $\sigma' \not\succsim^p_m \sigma$). Then $\ell_{w'}(\sigma; m) \leq \ell_{w'}(\sigma'; m)$ implies $\ell_w(\sigma; m) < \ell_w(\sigma'; m)$.

---

[56]For this tiebreaking rule to produce a well-defined result, the sublattice $\arg\min_{\sigma \in \Sigma} \ell_w(\sigma; m) \subseteq \Sigma$ needs to contain its supremum (for each $w \in W$ and $m \in M$). Note that, because Σ is a chain, $\ell_w(\cdot; m)$ is trivially quasisupermodular, and that Σ is trivially a sublattice of itself. These two conditions are sufficient to ensure that $\arg\min_{\sigma \in \Sigma} \ell_w(\sigma; m)$ is a complete sublattice (Milgrom and Shannon, 1994), and hence that it contains its supremum.

[57]Propositions 5 and 6 are analogous to Propositions 4 and $4'$ (respectively) of Milgrom and Shannon (1994).

Proposition 6. *For each $w \in W$ and $m \in M$, let $\hat{\sigma}(w, m) \in \arg\min_{\sigma \in \Sigma} \ell_w(\sigma; m)$.*[58] *Under Assumptions 1 and 4, for any $m \in M$, $\hat{\sigma}(\cdot, m)$ is nondecreasing (i.e., for any $w', w'' \in W$ with $w'' > w'$, $\hat{\sigma}(w'', m) \succsim_m^p \hat{\sigma}(w', m)$).*

Proof of Proposition 6. Fix $m \in M$, and let $w', w'' \in W$ with $w'' > w'$ be given. Take any $\sigma' \in \arg\min_{\sigma \in \Sigma} \ell_{w'}(\sigma; m)$ and $\sigma'' \in \arg\min_{\sigma \in \Sigma} \ell_{w''}(\sigma; m)$. Suppose that $\sigma' \succ_m^p \sigma''$. Note that $\ell_{w'}(\sigma'; m) \leq \ell_{w'}(\sigma''; m)$. By Assumption 4, $\ell_{w''}(\sigma'; m) < \ell_{w''}(\sigma''; m)$, which contradicts the choice of σ'' as a minimizer of $\ell_{w''}(\cdot; m)$. Thus $\sigma'' \succsim_m^p \sigma'$. □

Proposition 6 illustrates that, even in informative equilibria in which workers choose among optimal signals without regard to the principal's preference, the principal's welfare does not suffer from the appointment of a more biased worker. In other words, the main insight of Proposition 5 does not hinge on a favorable tiebreaking assumption, though it might require a stronger condition on the ordering of workers' preferences than Assumption 2 imposes.

B.6.3   Transfer Payments

As noted earlier, the model of the employment relationship that is developed here ignores the role that transfer payments, particularly outcome-contingent bonuses to workers, can play in providing incentives for information acquisition. This section augments the baseline model and illustrates that the results do not change substantively if such payments are allowed. In particular, while bonuses can certainly help the principal in strengthening a worker's incentive to acquire information, preference misalignment can further strengthen that incentive. Thus, even if the principal is able to implement a transfer scheme that rewards workers directly for information acquisition, preference misalignment can still benefit the principal.

First, it is helpful to simplify matters by observing that the baseline model is flexible enough to accommodate a profit-sharing scheme, $\tilde{t} : \{\alpha, \beta\} \times \{A, B\} \to \mathbb{R}$, that rewards a worker based on the ex-post quality of the decision. To this end, suppose that the worker's payoffs, based both on his intrinsic preferences regarding the decision and on those that are induced by $\tilde{t}(\cdot)$, are as shown in Table 2. Observe

---

[58]Note that Assumption 3 is relaxed here.

State

	A	B
α	$\overline{v}_\alpha$	$\underline{v}_\alpha$
β	$\underline{v}_\beta$	$\overline{v}_\beta$

Action is the label for the rows (α, β).

Table 2: The payoff structure shown in this matrix reflects both the worker's intrinsic preferences regarding the decision as well as the effects of the profit-sharing scheme on his welfare. Throughout the analysis, maintain the assumption that $\overline{v}_\alpha > \underline{v}_\beta$ and $\overline{v}_\beta > \underline{v}_\alpha$ (i.e., that the worker's ex-post preferences over the decision are aligned with those of the principal and all candidate managers, and that the decision is nontrivial to the worker).

that, when facing a manager of type $m \leq 1/2$, this worker's objective function (which he seeks to minimize) is

$$
\tilde{\kappa}(q; m) \equiv
\begin{cases}
c(q) - \left(\overline{v}_\alpha + \underline{v}_\alpha\right)/2 & \text{if } 0 \leq q < 1/2 - m, \\
c(q) - q \cdot \left(\overline{v}_\alpha - \underline{v}_\alpha + \overline{v}_\beta - \underline{v}_\beta\right)/2 & \\
\quad - \left(\overline{v}_\alpha + \underline{v}_\alpha + \overline{v}_\beta + \underline{v}_\beta\right)/4 & \text{if } 1/2 - m \leq q < 1/2.
\end{cases}
\tag{20}
$$

Now, in the environment of the baseline model, consider a worker of type $\hat{w} \equiv \left(\overline{v}_\beta - \underline{v}_\alpha\right)/\left(\overline{v}_\alpha - \underline{v}_\alpha + \overline{v}_\beta - \underline{v}_\beta\right)$, and suppose that the worker's cost function is given by $\hat{c}(\cdot) \equiv c(\cdot)/\left(\overline{v}_\alpha - \underline{v}_\alpha + \overline{v}_\beta - \underline{v}_\beta\right)$.[59] Observe that each of these two workers weakly prefers α if and only his posterior belief assigns probability of at least $\hat{w}$ to state A. Furthermore, the latter worker's objective function, when facing a manager of type $m \leq 1/2$, is given by

$$
\hat{\kappa}(q; \hat{w}, m) \equiv
\begin{cases}
\hat{c}(q) + \hat{w}/2 & \text{if } 0 \leq q < 1/2 - m, \\
\hat{c}(q) - q/2 + 1/4 & \text{if } 1/2 - m \leq q < 1/2.
\end{cases}
$$

Observe that

$$
\tilde{\kappa}(q, m) = \hat{\kappa}(q; \hat{w}, m) \cdot \left(\overline{v}_\alpha - \underline{v}_\alpha + \overline{v}_\beta - \underline{v}_\beta\right) - \left(\overline{v}_\alpha + \overline{v}_\beta\right)/2
$$

---

[59]Note that $\hat{c}(\cdot)$ satisfies all of the assumed properties of $c(\cdot)$, so it is an admissible cost function in the model.

It follows that, for any $m \in (0, 1/2]$,

$$\arg\min_{q \in [0,1/2)} \tilde{\kappa}(q, m) = \arg\min_{q \in [0,1/2)} \hat{\kappa}(q; \hat{w}, m). \tag{21}$$

It is straightforward to show that (21) holds for any $m \in (1/2, 1)$ as well. Thus, the introduction of a profit-sharing scheme of the form of $\tilde{t}(\cdot)$ does not affect a worker's behavior in a manner that cannot be replicated by simply adjusting the parameters of the baseline model. As a result, such schemes can be ignored in the baseline model without loss of generality.

Return now to the baseline model, and suppose that, for an appointed worker and a signal quality $q \in [0, 1/2)$, $\hat{c}(q)$ represents the *net cost*—relative to the worker's outside option—of accepting the principal's appointment and acquiring a signal of quality q.[60] Suppose that $\hat{c}(\cdot)$ is convex and twice differentiable and satisfies the following conditions:

- $\hat{c}(0) \leq -1/2$;

- $\hat{c}'(0) < 0$;

- There exists $\hat{q} \in (0, 1/2)$ such that $\hat{c}'(q) > 1/2$ for $q \in (\hat{q}, 1/2)$;

- $\lim_{q \uparrow 1/2} \hat{c}(q) > 0$.

The first assumption implies that any candidate worker would prefer to accept an appointment and shirk than to settle for his outside option. The second assumption implies that, from a worker's perspective, some positive level of information acquisition—even if it is insufficient to influence the manager's recommendation—is preferable to shirking. The third assumption implies that a worker eventually

---

[60]For example, suppose that the worker is paid a fixed wage, $T \in \mathbb{R}$ (net of the expected wage that he would earn if he were to reject the principal's offer of appointment), and receives additional remuneration through a bonus scheme, $t : [0, 1/2) \to \mathbb{R}$, that specifies a payment as a function of the acquired signal's quality. (Of course, given that the signal's quality is not observed by the principal, it seems implausible that such a scheme would be implementable, but the point here is to demonstrate that, even if such a scheme were admissible, the results would be substantively unchanged.) As before, let $c : [0, 1/2) \to \mathbb{R}$ represent the agent's private cost of information acquisition. Then, for any $q \in [0, 1/2)$, $\hat{c}(q) \equiv c(q) - t(q) - T$.

experiences marginal declines in his expected welfare from information acquisition, even after accounting for the benefits of marginal improvements in decision quality. Finally, the fourth assumption implies that the acquisition of information above a certain threshold is prohibitively costly; in particular, any candidate worker prefers to settle for his outside option than to accept an appointment and influence the decision by acquiring perfect information.

Under this setup, a worker of type w, when facing a manager of type $m \leq 1/2$, has objective function

$$\hat{\kappa}(q; w, m) \equiv \begin{cases} \hat{c}(q) + w/2 & \text{if } 0 \leq q < 1/2 - m, \\ \hat{c}(q) - q/2 + 1/4 & \text{if } 1/2 - m \leq q < 1/2. \end{cases} \tag{22}$$

If the appointed manager has type $m > 1/2$, the worker's objective function is

$$\hat{\kappa}(q; w, m) \equiv \begin{cases} \hat{c}(q) + (1-w)/2 & \text{if } 0 \leq q < m - 1/2, \\ \hat{c}(q) - q/2 + 1/4 & \text{if } m - 1/2 \leq q < 1/2. \end{cases} \tag{23}$$

Note the similarity between (22) and (4), and also between (23) and (5). These comparisons demonstrate that, in this new model, $\hat{c}(\cdot)$ takes the place of $c(\cdot)$ from the baseline model. An analysis that is similar to the one of the baseline model can be conducted here to obtain essentially the same results.[61] This point can be illustrated through a comparison of Figure 14—which depicts the two branches of the objective function in (22)—with Figure 5(c), its analogue from the baseline model. Note that, in this version of the model, q^* is the minimum of the function $q \mapsto \hat{c}(q) - q/2 + 1/4$, and $q^\alpha(w)$ is the largest value of $q \in [0, 1/2)$ for which $\hat{c}(q) - q/2 + 1/4 = \hat{c}(0) + w/2$ (or is 0 if no such value exists).

---

[61]There are two minor caveats here. First, the transfer payments also affect the principal's welfare. Second, different candidate workers may have different outside options and thus may require different compensation schemes for participation. In either case, the principal's expected welfare, in principle, may not be increasing in the equilibrium quality of information. Assuming, however, that the cost to the principal of compensating a worker is small compared to the principal's stake in the decision, these details can be ignored.
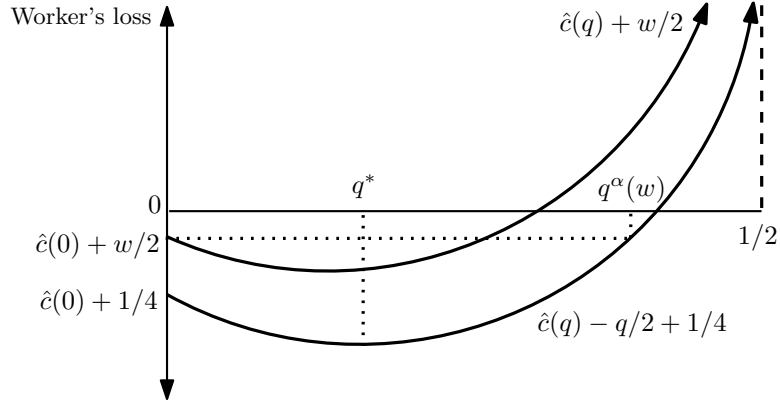
Figure 14: Suppose that $0 < m \le 1/2$ and $1/2 < w < 1$. In this case, the worker has a mild ex-ante preference for β. If the worker believes that the principal will implement the manager's recommendation, then, provided that the manager is not overly biased in favor of α (i.e., as long as $1/2 - m \le q^\alpha(w)$), the worker will find it worthwhile to investigate the alternative by acquiring information. (Note the similarity to Figure 5(c).)

References

Aghion, Philippe and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 1997, *105* (1), 1–29.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek, "When Does Coordination Require Centralization?," *American Economic Review*, 2008, *98* (1), 145–179.

Ambrus, Attila, Eduardo Azevedo, and Yuichiro Kamada, "Hierarchical Cheap Talk," *Theoretical Economics*, 2013, *8* (1), 233–261.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani, "Strategic Information Acquisition and Transmission," *American Economic Journal: Microeconomics*, 2016, *8* (3), 119–155.

Che, Yeon-Koo and Navin Kartik, "Opinions as Incentives," *Journal of Political Economy*, 2009, *117* (5), 815–860.

Crawford, Vincent P. and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, *50* (6), 1431–1451.

Dessein, Wouter, "Authority and Communication in Organizations," *Review of Economic Studies*, 2002, *69* (4), 811–838.

Eső, Péter and Ádám Galambos, "Disagreement and Evidence Production in Strategic Information Transmission," *International Journal of Game Theory*, 2013, *42* (1), 263–282.

﹘ **and Deszö Szalay**, "Incomplete Language As an Incentive Device," 2010. Mimeo, Oxford University and University of Bonn.

Gerardi, Dino and Leeat Yariv, "Costly Expertise," *American Economic Review: Papers & Proceedings*, 2008, *98* (2), 187–193.

Gershkov, Alex and Balázs Szentes, "Optimal Voting Schemes with Costly Information Acquisition," *Journal of Economic Theory*, 2009, *144* (1), 36–68.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani, "Mediation, Arbitration and Negotiation," *Journal of Economic Theory*, 2009, *144* (4), 1397–1420.

Hiltzik, Michael, *Dealers of Lightning: Xerox PARC and the Dawn of the Computer Age*, HarperBusiness, 1999.

Ivanov, Maxim, "Communication via a Strategic Mediator," *Journal of Economic Theory*, 2010, *145* (2), 869–884.

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

Milgrom, Paul and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, 1994, *62* (1), 57–80.

Myerson, Roger B., "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 1979, *47* (1), 61–74.

Nayeem, Omar A., "The Value of 'Useless' Bosses," 2014. `http://ssrn.com/abstract=2424029`.

Rotemberg, Julio J. and Garth Saloner, "Overt Interfunctional Conflict (and Its Reduction through Business Strategy)," *RAND Journal of Economics*, 1995, *26* (4), 630–653.

Smith, Douglas K. and Robert C. Alexander, *Fumbling the Future: How Xerox Invented, Then Ignored, the First Personal Computer*, William Morrow and Company, Inc., 1988.

Van den Steen, Eric J., "Culture Clash: The Costs and Benefits of Homogeneity," *Management Science*, 2010, *56* (10), 1718–1738.