TOO GOOD TO BE TRUE?

Supplementary Appendix [Not For Publication]

Francisco Espinosa ⓡ Debraj Ray

January 2022

In this not-for-publication Supplementary Appendix we start by studying the extensions described in Section 6.6 of the main text. We then fill some gaps in the proof of our main result, Proposition 1, and provide proofs of Propositions 5, 6 and 7. We give a numerical example of bounded replacement equilibria in the costly noise model (Section 6.2), and provide other missing details, such as the shape of the optimal choice of noise for a generic type, depicted in Figure 3 of the paper.

Numbered references for figures, equations, etc. refer to the main text. References that start with "a" refer to corresponding objects in this Supplementary Appendix.

1. EXTENSIONS SUMMARIZED IN SECTION 6.6

In the main text we covered four variations of the baseline model: the case of costly noise, a dynamic environment with agent term limits, non-normal signals, and the possibility of signal-contingent disclosure. But our setting lends itself easily to other extensions some of which we cover in the following sections. Section 1.1 allows for costly mean-shifting of signals. Section 1.2 considers more than one agent, each with private type. Section 1.3 studies non-binary agent types.

1.1. Mean-Shifting Effort, and Noisy Principals.

We can easily augment the baseline model to include unobserved effort to shift the mean value of one's type. For instance, suppose that each agent k is endowed with some baseline value (or type) $\underline{\theta}_k$ (with $\underline{\theta}_g > \underline{\theta}_b$). He can augment θ using a cost function $d(\theta_k - \underline{\theta}_k)$, common to both types, where d defined on $\mathbb{R}_+$ is increasing, strictly convex and differentiable, with $d(0) = 0$. The signal sent is then given by $x_k = \theta_k + \sigma_k \varepsilon$. Finally, the principal makes a decision to retain or replace.

Parts of this model run fully parallel to our setting. The principal makes her decisions on the basis of conjectured means and variances chosen by each type, leading to the familiar conditions (6)–(8) for the retention edge-points x_- and x_+. Similarly, an agent of type k maximizes the probability of retention net of cost. Whether or not x_- is smaller or larger than x_+ (and even when $x_+ = \infty$ as it will be with monotone retention), the agent always maximizes $\Phi\left([x_+ - \theta_k]/\sigma_k\right) - \Phi\left([x_- - \theta_k]/\sigma_k\right) - d\left(\theta_k - \underline{\theta}_k\right)$, this time by choosing *both* σ_k and θ_k. What this extension adds is a first-order condition for θ_k, given by

$$\text{(a.1)} \qquad \frac{1}{\sigma_k}\phi\left(\frac{x_- - \theta_k}{\sigma_k}\right) - \frac{1}{\sigma_k}\phi\left(\frac{x_+ - \theta_k}{\sigma_k}\right) \leq d'\left(\theta_k - \underline{\theta}_k\right),$$

with equality holding if $\theta_k > \underline{\theta}_k$. This additional condition can be used to show that the extension *fully* mimics the original model: We claim that

$\theta_b < \theta_g$, with choices of noise and principal decisions just as in our baseline setting.

Proof. We begin by eliminating the possibility that $\theta_b > \theta_g$. From the definition in (a.11) is it clear that a bounded retention regime is associated with $\sigma_b > \sigma_g$ and it is of the form $X = [x_-, x_+]$, and a bounded replacement regime is associated with $\sigma_b < \sigma_g$, and the principal replaces inside $X^c = [x_+, x_-]$. Then, under any one of these two regimes, the first-order condition with respect to θ_k is

$$\text{(a.2)} \qquad \frac{1}{\sigma_k}\phi\left(\frac{x_- - \theta_k}{\sigma_k}\right) - \frac{1}{\sigma_k}\phi\left(\frac{x_+ - \theta_k}{\sigma_k}\right) \leq d'\left(\theta_k - \underline{\theta}_k\right),$$

with equality holding if $\theta_k > \underline{\theta}_k$. Under bounded retention, we have

$$x_- < \frac{x_+ + x_-}{2} = \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} < \theta_g < \theta_b,$$

so that

$$\frac{x_- - \theta_k}{\sigma_k} < \frac{x_+ - \theta_k}{\sigma_k} < \frac{\theta_k - x_-}{\sigma_k}.$$

Because $\phi(\cdot)$ is single-peaked and symmetric around 0,

$$\phi\left(\frac{x_+ - \theta_k}{\sigma_k}\right) > \phi\left(\frac{x_- - \theta_k}{\sigma_k}\right),$$

But then (a.2) cannot hold with equality for any k, so $\theta_b > \theta_g$ is impossible if $\underline{\theta}_g > \underline{\theta}_b$. Similarly, under bounded replacement, we have $\sigma_b < \sigma_g$, so that

$$\theta_g < \theta_b < \frac{x_+ + x_-}{2} = \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} < x_-.$$

Then, once again,

$$\frac{\theta_k - x_-}{\sigma_k} < \frac{x_+ - \theta_k}{\sigma_k} < \frac{x_- - \theta_k}{\sigma_k},$$

and the same contradiction follows. Finally, with monotone retention, $\sigma_g = \sigma_b = \sigma$, and the retention rule is: retain iff

$$x \leq x^*(\sigma) := \frac{\theta_g + \theta_b}{2} + \frac{\sigma^2}{\theta_b - \theta_g}\ln(\beta).$$

The first-order derivative with respect to θ_k is then

$$-\frac{1}{\sigma_k}\phi\left(\frac{x^* - \theta_k}{\sigma_k}\right) - d'\left(\theta_k - \underline{\theta}_k\right),$$

which is always negative, so given that $\underline{\theta}_g > \underline{\theta}_b$, $\theta_b > \theta_g$ can never hold.

Moreover, it cannot be that $\theta_g = \theta_b = \theta$. For if so, the induced "second-stage game" with choice of noise must have exactly the same equilibrium payoffs, as well as the same *marginal* payoffs with respect to the common value θ, not counting the effort cost d. But since $\underline{\theta}_g \neq \underline{\theta}_b$, and d' is injective, it is clear that at least one of the agents is not satisfying the optimality conditions in the "first stage", when θ is chosen. Therefore $\theta_g \neq \theta_b$. ∎

This extension is also useful for understanding other aspects of the noisy relationship between principal and agent. For instance, mean-shifting effort for the sake of retention could be directly valuable to the principal, apart from providing information about type.[1] If neither that effort nor the payoff-relevant "output" from it is contractible, then the principal could want to structure her environment to keep agent effort high. Of particular interest is the case in which the background noise $\underline{\sigma}$ is close to zero, so that the agents can communicate their types with very high precision.

In general, this limit model has several equilibria, some pooling and some separating. To see the issue that arises, let's concentrate on a particular parametric configuration in which $\underline{\theta}_g$ and $\underline{\theta}_b$ are sufficiently separated from each other so that

$$d\left(\underline{\theta}_g - \underline{\theta}_b\right) > 1. \tag{a.3}$$

In this case it is easy to see that there can be only separating equilibria in zero-ambient-noise limit. In each such equilibrium, the bad type exerts no effort whatsoever. *The principal cannot incentivize the agent because there is no noise in the signal.* Both types reveal themselves perfectly. There are still many equilibria possible in which the good type is forced to exert effort to raise θ_g beyond $\underline{\theta}_g$, simply because the principal's retention set is some singleton $\{\theta_g\}$ with $\theta_g > \underline{\theta}_g$. But these equilibria are shored up by the "absurd belief" that observations between $\underline{\theta}_g$ and θ_g are attributable to the bad type. These configurations can be eliminated by standard refinements, leaving only the least-cost separating equilibrium in which retention occurs if $x = \underline{\theta}_g$, and no agent exerts any effort at all. Condition (a.3) guarantees that the bad type will not want to mimic the good type in this case.

If mean-shifting effort is separately valuable to the principal, this outcome is undesirable to her. The solution will therefore involve the *principal* adding noise, thereby ensuring that the bad type has some chance of being retained, and so incentivizing him. In any equilibrium of such an extended model in which the principal can move first, the principal will choose $\underline{\sigma} > 0$, endogenously injecting noise into the system.

1.2. **Multiple Agents.** Suppose there are two agents, 1 and 2, who simultaneously signal their types, and the principal must decide which agent to retain. Assume that there is exactly one agent of the good type. The agents know their own types and therefore both types. But they look identical ex ante to the principal, so her prior places equal probability on the two. The communication technology is unchanged:

$$x_i = \theta_{k(i)} + \sigma_{k(i)}\varepsilon_i, \tag{a.4}$$

where $i = 1, 2$, and $k(i)$ denotes i's type. The errors are independent and identically distributed standard normal random variables. A (symmetric) strategy for agent i is a pair (σ_g, σ_b). The principal's strategy is a function $r : \mathbb{R}^2 \to \{1, 2\}$, which indicates for every possible pair of signals (x_1, x_2) the agent she wants to retain. After observing (x_1, x_2) the principal retains agent 1 if (and, modulo indifference, only if)

$$\frac{\frac{1}{\sigma_g}\phi\left(\frac{x_1-\theta_g}{\sigma_g}\right)}{\frac{1}{\sigma_b}\phi\left(\frac{x_1-\theta_b}{\sigma_b}\right)} \geq \frac{\frac{1}{\sigma_g}\phi\left(\frac{x_2-\theta_g}{\sigma_g}\right)}{\frac{1}{\sigma_b}\phi\left(\frac{x_2-\theta_b}{\sigma_b}\right)}. \tag{a.5}$$

[1]For other models of relational contracts in which effort provides both current output and information about match quality, see, Kuvalekar and Lipnowski (2020), Kostadinov and Kuvalekar (2020), and Bhaskar (2017).

In this setting, a *monotone equilibrium* is defined as one where the principal retains the agent with the higher signal value. Once again, monotonicity can only be achieved if both *types* of agent take the same actions, so that $\sigma_g = \sigma_b$, but that won't happen:

Proposition A.1. *In any equilibrium, $\sigma_b > \sigma_g$, and the principal retains agent 1 if and only if $|x_1 - \hat{x}| \leq |x_2 - \hat{x}|$, where $\hat{x} = (\sigma_b^2\theta_g - \sigma_g^2\theta_b)/(\sigma_b^2 - \sigma_g^2)$ maximizes the likelihood ratio $\frac{1}{\sigma_g}\phi\left(\frac{x-\theta_g}{\sigma_g}\right)/\frac{1}{\sigma_b}\phi\left(\frac{x-\theta_b}{\sigma_b}\right)$. In particular, monotone equilibria do not exist.*

The proof of this proposition is long and involved, and we relegate it to the end of this Supplementary Appendix. Intuitively, when both types choose the same level of noise, the principal retains the one with the higher signal realization. But the bad type then wants to inject additional noise, since the good type has a lot of probability mass around his (higher) mean. At the same time, and for the same reason, the good type wants to decrease noise. This proposition bears a broad resemblance to the main result in Hvide (2002), who studies tournaments with moral hazard, when agents can influence both the mean and spread of their output. In equilibrium, there is excessive risk taking. By setting an intermediate value for output and rewarding the agent who gets closer to this threshold, the principal can do better.

1.3. **Multiple Types.** We extend Proposition 2 to many types, in the costly noise model of Section 6.2 of the main text. It is expositionally convenient to assume that there is a prior on types given by some density $q(\theta)$ on $\mathbb{R}$. Let $\mathcal{Q}$ be the space of all such densities and give it any reasonable topology; for concreteness, think of $\mathcal{Q}$ as a subset of the space of all probability measures on $\mathbb{R}$ with the topology of weak convergence. A subset $\mathcal{Q}^0$ of $\mathcal{Q}$ is *degenerate* (relative to $\mathcal{Q}$) if its complement $\mathcal{Q} - \mathcal{Q}^0$ is (relatively) open and dense in $\mathcal{Q}$.

Given $q \in \mathcal{Q}$, each type θ chooses noise $\sigma(\theta)$ as in the model of Section 6.2. Following the choice of noise, a signal is generated. The principal obtains payoff $u(\theta)$ from type θ, where u is some nondecreasing, bounded, continuous function. There is some given continuation payoff — V — from replacing an agent, which reasonably lies somewhere in between the retention utilities: $\lim_{\theta\to-\infty} u(\theta) < V < \lim_{\theta\to\infty} u(\theta)$. We also impose the generic condition that $u(\theta)$ is not locally flat exactly at V.

Proposition A.2. *Fix all the parameters of the model except for the type distribution. Then, under Condition U, an equilibrium with a monotone retention regime can exist only for a degenerate subset of density functions over types.*

See Section 9, at the end of this Supplementary Appendix, for a formal proof.

2. Missing Details in the Proof of Proposition 1

First, we rewrite conditions (11) and (12) in terms of the parameter β. Define

$$\alpha := \frac{\theta_g - \theta_b}{2\underline{\sigma}} > 0, \tag{a.6}$$

and then let

$$\beta_l := \frac{1}{\alpha + \sqrt{1+\alpha^2}} \exp\left[-\frac{\alpha}{\alpha + \sqrt{1+\alpha^2}}\right] < 1, \tag{a.7}$$

and

$$\beta_h := \exp\left[2\alpha^2\right] > 1. \tag{a.8}$$

Lemma A.1. *At $\beta = \beta_l$ (resp. $\beta = \beta_h$), condition (11) (resp. (12)) holds with equality. Furthermore, (12) is equivalent to $\beta < \beta_h$, and (11) is equivalent to $\beta > \beta_l$.*

Proof. The only non-immediate assertion of this lemma is the very last: that (11) is equivalent to $\beta > \beta_l$. To this end, multiplying both sides of (10) by $\alpha(\beta)$ and defining

$$y(\beta) := \frac{\alpha(\beta)}{\alpha(\beta) + \sqrt{1+\alpha(\beta)^2}} \in [0,1), \tag{a.9}$$

we have that (10) is equivalent, for all $\beta \in (0,1)$, to

$$\alpha(\beta)\beta = y(\beta)\exp\left\{-y(\beta)\right\}. \tag{a.10}$$

We claim that $\alpha(\beta)$ is decreasing in β. If false, there is β_1 such that α is locally nondecreasing at β_1. But then $\alpha(\beta)\beta$ is *strictly* locally increasing at β_1. Because (a.10) holds and $y\exp\{-y\}$ is increasing in y when $y \in [0,1)$,[2] $y(\beta)$ is locally strictly increasing at β_1. But from (a.9), it is easy to see that $d\alpha/dy < 0$. These last two observations contradict our presumption that $\alpha(\beta)$ is locally nondecreasing in β. The last assertion of the lemma follows immediately. ■

For convenience, we reproduce here the expressions for the principal's thresholds $x_-(\sigma)$ and $x_+(\sigma)$, when b plays $\sigma_b = \sigma > \underline{\sigma}$ and g plays $\sigma_g = \underline{\sigma}$:

$$x_-(\sigma) := \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b - \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2} \text{ and } x_+(\sigma) := \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b + \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2}, \tag{a.11}$$

where

$$R(\sigma) := +\sqrt{\Delta^2 + (\sigma^2 - \underline{\sigma}^2)\, 2\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right)}, \tag{a.12}$$

with $\Delta := \theta_g - \theta_b$.

Proof of Lemma 5. When $\beta \geq 1$, it is clear that the term within the square root in (a.12) is strictly positive for all $\sigma > \underline{\sigma}$. For $\beta < 1$, consider the set of pairs (σ, β) such that $R = 0$. These pairs satisfy

$$\beta = \frac{\underline{\sigma}}{\sigma}\exp\left[-\frac{\Delta^2}{2(\sigma^2 - \underline{\sigma}^2)}\right] < 1. \tag{a.13}$$

View β in (a.13) as a function of σ, depicted in Figure A.1. Any pair (σ, β) strictly below the $R = 0$ locus (the curve in the diagram) implies that the argument inside the square root in (a.12) is strictly negative, and therefore the functions $x_-(\sigma)$ and $x_+(\sigma)$ are not well-defined for such a

[2] Note that $dy\exp(-y)/dy = \exp(-y)(1-y) > 0$ for $y \in [0,1)$.

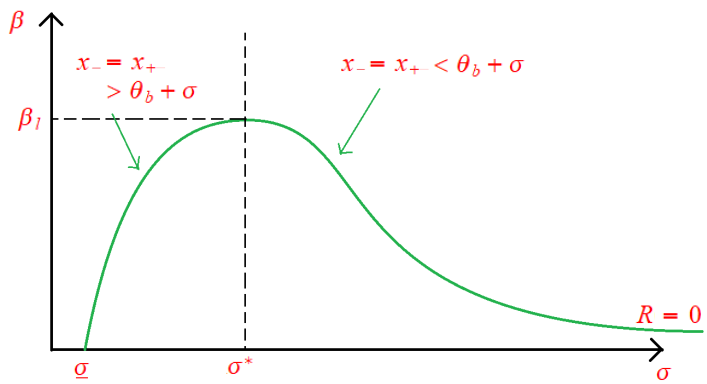


FIGURE A.1. The $R(\sigma,\beta)=0$ locus.

pair: there are no real roots to $\beta\frac{1}{\underline{\sigma}}\phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right)=\frac{1}{\sigma}\phi\left(\frac{x-\theta_b}{\sigma}\right)$. On the other hand, when the pair (σ,β) is strictly above the locus, $R>0$ and, therefore, two distinct real roots exist.

Now consider the $R=0$ locus. We have $\beta\to 0$ as $\sigma\to\underline{\sigma}^-$ and as $\sigma\to\infty$. By computing the derivative with respect to σ, we find that β in (a.13) strictly increases with σ if and only if $-\sigma^2+\sigma\Delta+\underline{\sigma}^2>0$. The two roots to this quadratic polynomial are $\underline{\sigma}\left(\alpha-\sqrt{\alpha^2+1}\right)$ and

$$\sigma^* := \underline{\sigma}\left(\alpha+\sqrt{\alpha^2+1}\right), \tag{a.14}$$

where α is defined in (a.6). The first root is negative, so β is increasing in σ for $\sigma\in[\underline{\sigma},\sigma^*)$, and decreasing for $\sigma>\sigma^*$. At $\sigma=\sigma^*$ the derivative is zero, so a global maximum is attained. Evaluating (a.13) at $\sigma=\sigma^*$, this maximum value equals β_l, as defined in (a.7). So if $\beta>\beta_l$ (i.e. if (11) holds, as per Lemma A.1), then $x_-(\sigma)$ and $x_+(\sigma)$ are well-defined and distinct for all σ. The converse is also true: if the roots are not well-defined for some σ or not distinct for all σ, then (β,σ) is at or below the $R=0$ locus; so $\beta\le\beta_l$, and (11) fails. ■

Proof of Lemma 7. By Lemma 5, if $\beta\ge 1$ or if $\beta<1$ and (11) holds, $x_-(\sigma)$ and $x_+(\sigma)$ are well-defined and distinct for any $\sigma>\underline{\sigma}$.

(i) Inspection of (a.11) immediately reveals that $\lim_{\sigma\to\underline{\sigma}^+}x_+(\sigma)=\infty$. For the corresponding limit of $x_-(\sigma)$, apply L'Hôspital's rule to see that

$$\lim_{\sigma\to\underline{\sigma}^+}x_-(\sigma)=\lim_{\sigma\to\underline{\sigma}^+}\frac{2\sigma\theta_g-\underline{\sigma}R(\sigma)-\sigma\underline{\sigma}\frac{2\sigma\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right)+\left(\sigma^2-\underline{\sigma}^2\right)\frac{1}{\sigma}}{R(\sigma)}}{2\sigma}=\frac{\theta_g+\theta_b}{2}-\frac{\underline{\sigma}^2}{\theta_g-\theta_b}\ln(\beta)=x^*(\underline{\sigma}).$$

(ii) Notice that $\lim_{\sigma\to\infty}\left(\frac{R(\sigma)}{\sigma}\right)^2=\lim_{\sigma\to\infty}\frac{(\theta_g-\theta_b)^2}{\sigma^2}+\left(1-\frac{\underline{\sigma}^2}{\sigma^2}\right)2\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right)=\infty$. So, because $x_-(\sigma)$ and $x_+(\sigma)$ can be respectively written as

$$x_-(\sigma)=\frac{\theta_g-\frac{\underline{\sigma}^2}{\sigma^2}\theta_b-\underline{\sigma}\frac{R(\sigma)}{\sigma}}{1-\frac{\underline{\sigma}^2}{\sigma^2}}\text{ and }x_+(\sigma)=\frac{\theta_g-\frac{\underline{\sigma}^2}{\sigma^2}\theta_b+\underline{\sigma}\frac{R(\sigma)}{\sigma}}{1-\frac{\underline{\sigma}^2}{\sigma^2}},$$

it is clear that $x_-(\sigma)\to-\infty$ and $x_+(\sigma)\to\infty$ as $\sigma\to\infty$.

(iii) Suppose that $\beta \geq 1$ and (12) fails. Using (a.11), we see that

$$x_{-}(\sigma, \beta) - \theta_b = \frac{\sigma^2(\theta_g - \theta_b) - \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2}. \tag{a.15}$$

So the claim is established if the right-hand side in (a.15) is non-positive. But that will be true if $\sigma^4(\theta_g - \theta_b)^2 \leq \sigma^2\underline{\sigma}^2 R(\sigma)^2$, or equivalently, using (a.12), if

$$\sigma^2(\theta_g - \theta_b)^2 \leq \underline{\sigma}^2(\theta_g - \theta_b)^2 + 2\underline{\sigma}^2(\sigma^2 - \underline{\sigma}^2)\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right).$$

Rearranging terms, this is equivalent to $(\theta_g - \theta_b)^2 \leq 2\underline{\sigma}^2 \ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right)$. But this inequality is implied by the failure of (12), because $\sigma \geq \underline{\sigma}$. ■

Proof of Lemma 8. We show that the derivative of $\Psi(\sigma)$ at any fixed point is negative. This, together with the end-point conditions verified in the main text, determines the uniqueness of such a point. The derivative of the left-hand side of (32) with respect to σ is written as

$$\phi\left(\frac{x_{+}(\sigma) - \theta_b}{\Psi(\sigma)}\right)[x_{+}(\sigma) - \theta_b]\left(\frac{x'_{+}(\sigma)}{x_{+}(\sigma) - \theta_b} - \frac{x_{+}(\sigma) - \theta_b}{\Psi(\sigma)}\frac{x'_{+}(\sigma)\Psi(\sigma) - (x_{+}(\sigma) - \theta_b)\Psi'(\sigma)}{\Psi^2(\sigma)}\right),$$

where we use the fact that ϕ is the normal density. For the right-hand side we obtain the same expression but with $x_{-}(\sigma)$ instead of $x_{+}(\sigma)$. Now, the first two terms on each side will cancel each other, because $\Psi(\sigma)$ satisfies (32). Rearranging terms, we obtain

$$\Psi'(\sigma) = \frac{\Psi(\sigma)^3}{2}\frac{\frac{x'_{-}(\sigma)}{x_{-}(\sigma) - \theta_b}\left(1 - \frac{(x_{-}(\sigma) - \theta_b)^2}{\Psi(\sigma)^2}\right) + \frac{x'_{+}(\sigma)}{(x_{+}(\sigma) - \theta_b)}\left(\frac{(x_{+}(\sigma) - \theta_b)^2}{\Psi(\sigma)^2} - 1\right)}{(x_{+}(\sigma) - x_{-}(\sigma))\left(\frac{x_{+}(\sigma) + x_{-}(\sigma)}{2} - \theta_b\right)}. \tag{a.16}$$

By differentiating $\beta\frac{1}{\underline{\sigma}}\phi\left(\frac{x - \theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma}\phi\left(\frac{x - \theta_b}{\sigma}\right)$ with respect to σ we find that

$$x'_{+}(\sigma) = \frac{\underline{\sigma}}{R(\sigma)}\left[1 - \left(\frac{x_{+}(\sigma) - \theta_b}{\sigma}\right)^2\right] \text{ and } x'_{-}(\sigma) = \frac{\underline{\sigma}}{R(\sigma)}\left[\left(\frac{x_{-}(\sigma) - \theta_b}{\sigma}\right)^2 - 1\right], \tag{a.17}$$

where $R(\sigma)$ is defined in (a.12). Substituting (a.17) into (a.16) and evaluating the resulting derivative at $\sigma = \Psi(\sigma)$, we see that

$$\Psi'(\sigma) = -\underline{\sigma}\sigma^3\frac{\frac{1}{(x_{-}(\sigma) - \theta_b)}\left(1 - \frac{(x_{-}(\sigma) - \theta_b)^2}{\sigma^2}\right)^2 + \frac{1}{(x_{+}(\sigma) - \theta_b)}\left(\frac{(x_{+}(\sigma) - \theta_b)^2}{\sigma} - 1\right)^2}{2(x_{+}(\sigma) - x_{-}(\sigma))\left(\frac{x_{+}(\sigma) + x_{-}(\sigma)}{2} - \theta_b\right)R(\sigma)} < 0.$$

■

Lemma A.2. *Let σ^* be defined as in a.14. Along the $R = 0$ locus,*

(i) $x_{-}(\sigma) = x_{+}(\sigma) > \theta_b + \sigma$ *if, and only if,* $\sigma \in (\underline{\sigma}, \sigma^*)$*;*

(ii) $x_{-}(\sigma) = x_{+}(\sigma) < \theta_b + \sigma$ *if, and only if,* $\sigma > \sigma^*$*;*

(iii) $x_{-}(\sigma) = x_{+}(\sigma) = \theta_b + \sigma$ *if, and only if,* $\sigma = \sigma^*$*.*

Proof. Any point (σ, β) along the $R = 0$ locus is characterized by equation (a.13), so $x_-(\sigma) = x_+(\sigma) = \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b}{\sigma^2 - \underline{\sigma}^2}$ (look at (a.11)). So along the locus, $x_-(\sigma) = x_+(\sigma) > \theta_b + \sigma$ if and only if $-\sigma^2 + \sigma(\theta_g - \theta_b) + \underline{\sigma}^2 > 0$. One root of this expression is negative; the other is σ^*, as in (a.14). So $x_-(\sigma) = x_+(\sigma) > \theta_b + \sigma$ if and only if $\sigma \in (\underline{\sigma}, \sigma^*)$. Similarly, $x_-(\sigma) = x_+(\sigma) < \theta_b + \sigma$ if, and only if, $\sigma > \sigma^*$. Finally, at $(\sigma, \beta) = (\sigma^*, \beta_l)$, $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$. ∎

Proof of Lemma 10. We first rule out monotone equilibria: if $\beta < 1$, inspection of the expression of $x^*(\sigma)$ in (5) immediately reveals that $x^*(\sigma) > \frac{\theta_g + \theta_b}{2} > \theta_b$ for any σ, so type b always wants to deviate to inject additional noise. For bounded retention equilibrium, we will claim that if (11) fails when $\beta < 1$, then for any σ such that the roots $x_-(\sigma)$ and $x_+(\sigma)$ of a bounded retention equilibrium are well-defined and distinct, we have either $\theta_b + \sigma < x_-(\sigma) < x_+(\sigma)$ or $\theta_b + \sigma > x_+(\sigma) > x_-(\sigma)$. But this is inconsistent with bounded retention, because by Lemma 3 (iii) type b responds to the retention interval by choosing σ_b that satisfies $x_-(\sigma) < \theta_b + \sigma_b < x_+(\sigma)$, so a fixed-point is impossible.

Now we prove the claim in the previous paragraph. If $\beta \le \beta_l$, Figure A.1 indicates that we have at most two values of σ such that $R = 0$. Denote these by σ_l^β and σ_h^β, with $\sigma_l^\beta \le \sigma^* \le \sigma_h^\beta$. In a bounded retention equilibrium, either $\sigma < \sigma_l^\beta$ or $\sigma > \sigma_h^\beta$.[3]

We treat the cases $\beta < \beta_l$ and $\beta = \beta_l$ separately. In the first case, $x_-(\sigma_l^\beta) > \theta_b + \sigma_l^\beta$ by Lemma A.2(i). So $x_-(\sigma)$ is increasing for $\sigma < \sigma_l^\beta$ but close to σ_l^β; see (a.17). And yet $x_-(\sigma) > \theta_b + \sigma$ for all $\sigma \in (\underline{\sigma}, \sigma_l^\beta)$. For if not, then $x_-(\sigma') \le \theta_b + \sigma'$ for some $\sigma' \in (\underline{\sigma}, \sigma_l^\beta)$, but then, because $x_-(\sigma)$ is well-defined and continuous ($R > 0$), $x_-(\sigma)$ crosses $\theta_b + \sigma$ from below at some σ. That means $x'_-(\sigma) \ge 1$, (a.17) tells is that $x'_-(\sigma) = 0$. Therefore $\theta_b + \sigma < x_-(\sigma) < x_+(\sigma)$ for all $\sigma \in (\underline{\sigma}, \sigma_l^\beta)$. Now consider $\sigma > \sigma_h^\beta$. At σ_h^β we have $x_+(\sigma_h^\beta) < \theta_b + \sigma_h^\beta$ by Lemma A.2(ii). Looking at (a.17), this implies that $x_+(\sigma)$ is increasing for σ close enough to σ_h^β. We claim that $x_+(\sigma) < \theta_b + \sigma$ for all $\sigma > \sigma_h^\beta$. For if this is false for some σ', then $x_+(\sigma') \ge \theta_b + \sigma'$, so because $x_+(\sigma)$ is well-defined and continuous for $\sigma > \sigma_h^\beta$ ($R > 0$), $x_+(\sigma)$ crosses $\theta_b + \sigma$ from below at some σ. That means $x'_+(\sigma) \ge 1$, but (a.17) tells us that $x'_+(\sigma) = 0$ at such a point. So $\theta_b + \sigma > x_+(\sigma) > x_-(\sigma)$ for all $\sigma > \sigma_h^\beta$.

Finally, consider $\beta = \beta_l$. Here, $\sigma_l^\beta = \sigma_h^\beta = \sigma^*$, where σ^* is defined in (a.14). Looking at (a.17) we can see that $x'_-(\sigma^*) = x'_+(\sigma^*) = 0$ and, as established by Lemma A.2(iii), $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$. Then, if we consider $\sigma < \sigma^*$, it is clear that for σ close enough to σ^* we have $x_-(\sigma) > \theta_b + \sigma$, and as we showed before this leads to the conclusion that $\theta_b + \sigma < x_-(\sigma) < x_+(\sigma)$ for all $\sigma \in (\underline{\sigma}, \sigma_l^\beta)$. Similarly, for $\sigma > \sigma^*$ and close enough to σ^*, we have that $x_+(\sigma) < \theta_b + \sigma$, which leads to $\theta_b + \sigma > x_+(\sigma) > x_-(\sigma)$ for all $\sigma > \sigma_h^\beta$. ∎

3. Omitted Proofs for Dynamics with Term Limits, Section 6.3

Lemma A.3. *Assume $\beta \in (\beta_l, \beta_h)$, then*

[3]The equalities are not considered because being on the locus means $x_-(\sigma) = x_+(\sigma)$: It's a trivial equilibrium.

(i) $\frac{\partial x_-^*}{\partial \beta} < 0$ *and* $\frac{\partial x_+^*}{\partial \beta} > 0$*;*

(ii) $\lim_{\beta \to \beta_l^+} x_-^* = \lim_{\beta \to \beta_l^+} x_+^* = \theta_b + \sigma^*$ *and* $\lim_{\beta \to \beta_l^+} \sigma_b^* = \sigma^*$*, where* σ^* *is in* (a.14)*.*

Proof. (i) When $\beta \in (\beta_l, \beta_h)$, both (11) and (12) hold by Lemma A.1, and therefore Proposition 1 tells us that there exists a unique equilibrium, which is a bounded retention equilibrium where $\sigma_b > \sigma_g = \underline{\sigma}$ and the principal retains in a bounded interval $X = [x_-, x_+]$, with $x_- < x_+$. The equilibrium values $\left(\sigma_b^*, x_-^*, x_+^*\right)$ are determined by

$$\beta \frac{1}{\underline{\sigma}} \phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma_b} \phi\left(\frac{x-\theta_b}{\sigma_b}\right), \text{ for } x = x_-^*, x_+^*, \text{ and}$$

$$\phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)(x_- - \theta_b) = \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)(x_+ - \theta_b).$$

Differentiate these equations with respect to β. In the case of the first equation, we obtain

$$\text{(a.18)} \qquad \begin{aligned} \frac{\sigma_b'}{\sigma_b}\left(\left(\frac{x_- - \theta_b}{\sigma_b}\right)^2 - 1\right) &= \frac{R(\sigma_b)}{\sigma_b \underline{\sigma}} x_-' + \frac{1}{\beta}, \\ \frac{\sigma_b'}{\sigma_b}\left(1 - \left(\frac{x_+ - \theta_b}{\sigma_b}\right)^2\right) &= \frac{R(\sigma_b)}{\sigma_b \underline{\sigma}} x_+' - \frac{1}{\beta}. \end{aligned}$$

In the case of the second equilibrium equation we obtain the same expression as in (a.16), where $\Psi(\sigma)$ is now σ_b, and the derivatives are those with respect to β. By combining it with (a.18), and after some heavy algebra, we obtain:

$$\text{(a.19)} \qquad \begin{aligned} x_-' &= -\frac{y_-(y_+ + y_-)}{\beta} \frac{\underline{\sigma}^2}{(\theta_g - \theta_b)} \left(\frac{\frac{\theta_g - \theta_b}{\underline{\sigma}} \frac{\sigma_b}{\underline{\sigma}} (y_+ - y_-) y_+ + (y_- + y_+)\left(y_+^2 - 1\right)}{\frac{\theta_g - \theta_b}{\underline{\sigma}} \frac{\sigma_b}{\underline{\sigma}} y_- y_+ (y_+ - y_-)^2 + \left(1 - y_-^2\right)^2 y_+ + \left(1 - y_+^2\right)^2 y_-} \right) \\ x_+' &= \frac{y_+(y_+ + y_-)}{\beta} \frac{\underline{\sigma}^2}{(\theta_g - \theta_b)} \left(\frac{\frac{\theta_g - \theta_b}{\underline{\sigma}} \frac{\sigma_b}{\underline{\sigma}} (y_+ - y_-) y_- + (y_+ + y_-)\left(1 - y_-^2\right)}{\frac{\theta_g - \theta_b}{\underline{\sigma}} \frac{\sigma_b}{\underline{\sigma}} y_- y_+ (y_+ - y_-)^2 + \left(1 - y_-^2\right)^2 y_+ + \left(1 - y_+^2\right)^2 y_-} \right) \end{aligned}$$

where the notation x_i' means $\frac{\partial x_i}{\partial \beta}$, and $y_i := \frac{x_i - \theta_b}{\sigma_b}$, for $i = -, +$. Now, $\sigma_b > \underline{\sigma}$ implies $x_- > \theta_b$ (Lemma 3(ii)), so $y_- > 0$. Also, from Lemma 3(iii), $y_+ > 1 > y_-$. Then, from (a.19) we see that $x_-' < 0$ and $x_+' > 0$, so the interval shrinks as β decreases.

(ii) By Lemma 3(iii), σ_b satisfies $x_-(\sigma_b) < \theta_b + \sigma_b < x_+(\sigma_b)$. In the limit as $\beta \to \beta_l^+$, condition (11) holds with equality, and $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$ where $\sigma^* = \underline{\sigma}\left(\alpha + \sqrt{\alpha^2 + 1}\right)$ by Lemma A.2 (iii). Then, $\sigma_b^* \to \sigma^*$ as $\beta \to \beta_l^+$, which completes the proof. ∎

Proof of Proposition 5. For some (provisionally given) value of β, use Proposition 1 for the baseline static model to generate retention probabilities Π_g and Π_b. The circle is closed by the additional condition that (β, Π_g, Π_b) must solve (20), reproduced here for convenience:

$$\text{(a.20)} \qquad \beta = \frac{q}{1-q} \frac{1-p}{p} = \frac{1 + \delta \Pi_b}{1 + \delta \Pi_g}.$$

As argued in the main text, it must be that $\Pi_g \geq \Pi_b$, because the principal uses a retention zone that retains the high type at least as often than the low type. So in the dynamic model, $\beta \leq 1$. Then, following Proposition 1 and Lemma A.1, we consider two cases: either (11) fails and $\beta \leq \beta_l < 1$, or (11) holds and $\beta \in (\beta_l, 1]$. In the former case, by Proposition 1, only trivial equilibria exist (see Lemma 10). Then, $\Pi_b = \Pi_g$. But equilibrium condition (a.20) then says that $\beta = 1$, a contradiction. That is, if an equilibrium exists in this dynamic version of the costless model, it must be the case that $\beta \in (\beta_l, 1] \subset (\beta_l, \beta_h)$, so it must have bounded retention. We now prove existence.

For any given $\beta \in (\beta_l, 1]$, by Proposition 1(ii) and Lemma A.1, there is a unique equilibrium in the static model, with bounded retention thresholds $\{x_-(\beta), x_+(\beta)\}$. Given $\{\sigma_b(\beta), \sigma_g(\beta)\}$ in that equilibrium (with $\sigma_b(\beta) > \sigma_g(\beta) = \underline{\sigma}$ as already established), define, for $k = b, g$:

(a.21) $$\Pi_k(\beta) = \int_X \pi_k(x)dx = \frac{1}{\sigma_k(\beta)} \int_{x_-(\beta)}^{x_+(\beta)} \phi\left(\frac{x - \theta_k}{\sigma_k(\beta)}\right) dx.$$

Now, in line with (a.20), define a mapping $\beta' = \psi(\beta)$ by

(a.22) $$\beta' = \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)}$$

Because the equilibrium is unique for every $\beta \in (\beta_l, 1]$, it is easy to see that ψ is a continuous map. Next, when $\beta = 1$, we know from the non-triviality of the corresponding static equilibrium that $\Pi_b(1) < \Pi_g(1)$, so that $\beta' = \psi(1) < 1$. Finally, as $\beta \to {\beta_l}^+$, the boundaries of the static equilibrium retention thresholds x_-^* and x_+^* converge to each other (see Lemma A.3 (ii)), so that $\lim_{\beta \to {\beta_l}^+} \Pi_g = \lim_{\beta \downarrow \beta_l} \Pi_b = 0$, and therefore

$$\beta' = \psi(\beta) = \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)} \to 1$$

as $\beta \to {\beta_l}^+$. This verifies a second end-point condition $\lim_{\beta \to {\beta_l}^+} \psi(\beta) > \beta_l$. By the intermediate value theorem, there is β with $\psi(\beta) = \beta$, and this — along with the corresponding values of σ_b and σ_g — is easily seen to be an equilibrium of the dynamic game.

To complete the proof, we establish uniqueness of equilibrium. Begin by differentiating the expression in (a.21) with respect to β, taking care to use an envelope argument for type b (his first-order condition) and the fact that $\sigma_g(\beta) = \underline{\sigma}$ for type g. We obtain:

(a.23) $$\frac{\partial \Pi_k(\beta)}{\partial \beta} = \frac{1}{\sigma_k(\beta)} \left[\phi\left(\frac{x_+(\beta) - \theta_k}{\sigma_k(\beta)}\right) x_+'(\beta) - \phi\left(\frac{x_-(\beta) - \theta_k}{\sigma_k(\beta)}\right) x_-'(\beta)\right].$$

Next, observe that

(a.24) $$\frac{\partial}{\partial \beta} \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)} = \delta \frac{\frac{\partial \Pi_b(\beta)}{\partial \beta}(1 + \delta\Pi_g(\beta)) - (1 + \delta\Pi_b(\beta))\frac{\partial \Pi_g(\beta)}{\partial \beta}}{(1 + \delta\Pi_g(\beta))^2}.$$

Substitute (a.23) in (a.24) and note that $x_-(\beta)$ and $x_+(\beta)$ solve (6) with equality to obtain:

$$\frac{\partial}{\partial \beta} \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)} = \delta \frac{\frac{1}{\sigma_g(\beta)}\phi\left(\frac{x_+(\beta) - \theta_g}{\sigma_g(\beta)}\right) x_+'(\beta) - \frac{1}{\sigma_g(\beta)}\phi\left(\frac{x_-(\beta) - \theta_g}{\sigma_g(\beta)}\right) x_-'(\beta)}{(1 + \delta\Pi_g(\beta))} \left(\beta - \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)}\right).$$

Because $x'_+(\beta) > 0$ and $x'_-(\beta) < 0$ (Lemma A.3(i)), we must conclude that

(a.25) $$\text{Sign}\left\{\frac{\partial}{\partial\beta}\frac{1+\delta\Pi_b(\beta)}{1+\delta\Pi_g(\beta)}\right\} = \text{Sign}\left\{\beta - \frac{1+\delta\Pi_b(\beta)}{1+\delta\Pi_g(\beta)}\right\}.$$

This, along with $\lim_{\beta\to\beta_l}\phi(\beta) > \beta_l$, eliminates two solutions to:

$$\beta = \frac{1+\delta\Pi_b(\beta)}{1+\delta\Pi_g(\beta)},$$

for that would require the sign equality (a.25) to be violated for some β. ■

4. The Non-Normal Case, Section 6.4

For any σ_b and σ_g, define

(a.26) $$h(x) := \frac{f\left(\frac{x-\theta_b}{\sigma_b}\right)}{f\left(\frac{x-\theta_g}{\sigma_g}\right)},$$

and let $k := \beta\sigma_b/\sigma_g$. Following (3) in the main text, the retention zone is then given by

(a.27) $$X(k) := \{x | h(x) \le k\}.$$

Lemma A.4. (i) *If* $\sigma_b = \sigma_g$, *then* $h(x)$ *is strictly decreasing in* x *with* $\lim_{x\to-\infty} h(x) = \infty$ *and* $\lim_{x\to\infty} h(x) = 0$.

(ii) *If* $\sigma_b > \sigma_g$, *then* $\lim_{x\to\infty} h(x) = \lim_{x\to-\infty} h(x) = \infty$.

(iii) *If* $\sigma_b < \sigma_g$, *then* $\lim_{x\to\infty} h(x) = \lim_{x\to-\infty} h(x) = 0$.

Remark. The symmetric statements of parts (ii) and (iii), despite the fact that $\theta_b < \theta_g$, reflect our observation in the main text that "spreads dominate means."

Proof. (i) Define $z(x) \equiv (x-\theta_b)/\sigma$ and $a \equiv (\theta_g - \theta_b)/\sigma$, where $\sigma = \sigma_b = \sigma_g$. Then

$$h(x) = \frac{f(z(x))}{f(z(x)-a)}.$$

Because $z(x)$ is affine and increasing in x, the result follows directly from strong MLRP.

(ii) There is $\epsilon > 0$ such that for x sufficiently large, $(x-\theta_b)/\sigma_b \le (x - [\theta_g+\epsilon])/\sigma_g$. Because $f'(z) \le 0$ for all $z > 0$ (see main text), it follows that for all x large enough,

$$\frac{f\left(\frac{x-\theta_b}{\sigma_b}\right)}{f\left(\frac{x-\theta_g}{\sigma_g}\right)} \ge \frac{f\left(\frac{x-[\theta_g+\epsilon]}{\sigma_g}\right)}{f\left(\frac{x-\theta_g}{\sigma_g}\right)},$$

and now, using strong MLRP, the right hand side of this inequality goes to infinity as $x\to\infty$. The case $x\to-\infty$ follows parallel lines: switch (θ_b,σ_b) and (θ_g,σ_g) in the argument above, notice that f is increasing for $z<0$, and use part (i) again.

Noticing that the relative magnitudes of θ_b and θ_g played no role in part (ii), the same argument with appropriately switched symbols works for part (iii). ∎

Proof of Proposition 6, Part (i). Suppose the assertion is false. Then, by Lemma A.4 and the definition of the retention zone in (a.27), we must have $\sigma_g > \sigma_b$, and retention for all x large in absolute value. But then, given such a zone, the probability of retention of any type converges to 1 as $\sigma_k \to \infty$. Therefore, for any candidate pair (σ_g, σ_b), any type finds a profitable deviation. ∎

Lemma A.5. *For $z \geq 0$, $f(z)z$ is increasing for $z \in [0, z^*)$, decreasing for $z > z^*$, and maximized at $z^* > 0$, the unique solution to $f'(z)/f(z) = -1/z$.*

Proof. The derivative of $f(z)\,z$ with respect to z is $f(z)\,z\left[\frac{f'(z)}{f(z)} + \frac{1}{z}\right]$. $f'(z)/f(z)$ is non-increasing by MLRP, and $1/z$ is strictly decreasing as well. We also have that $\frac{f'(z)}{f(z)} + \frac{1}{z} \to \infty$ as $z \to 0$ and that $\frac{f'(z)}{f(z)} + \frac{1}{z}$ is negative for z sufficiently large. Then z^*, the unique maximizer of $f(z)\,z$ for $z \geq 0$, satisfies $\frac{f'(z)}{f(z)} + \frac{1}{z} = 0$. ∎

Lemma A.6. (i) *If $X = [x^*, \infty)$ and $\theta_k > x^*$, the agent of type k chooses $\sigma_k = \underline{\sigma}$; if $\theta_k < x^*$, the problem has no solution, in particular, the agent always wants to inject additional noise; if $\theta_k = x^*$, the agent is indifferent across all choices of σ.*

(ii) *Assume a retention zone of the form $[x_-, x_+]$ with $x_- < x_+$. If $x_- \leq \theta_k$ and $x_+ > \theta_k$ then $\sigma_k = \underline{\sigma}$.*

(iii) *Assume a retention zone of the form $[x_-, x_+]$ with $x_- < x_+$. If $x_- > \theta_k$, then for each k define*

$$d_k(\sigma_k) := f\left(\frac{x_- - \theta_k}{\sigma_k}\right)(x_- - \theta_k) - f\left(\frac{x_+ - \theta_k}{\sigma_k}\right)(x_+ - \theta_k) \text{ for all } \sigma_k > 0. \quad \text{(a.28)}$$

Then type k's payoff derivative with respect to her choice of noise σ_k is precisely given by $\sigma_k^2 d_k(\sigma_k)$. The function d_k is continuous, initially positive then negative, with a unique root to $d_k(\sigma_k) = 0$, σ_k^*, satisfying

$$\sigma_k^* \in \left(\frac{x_- - \theta_b}{z^*}, \frac{x_+ - \theta_b}{z^*}\right), \quad \text{(a.29)}$$

where $z^* > 0$ is defined in Lemma A.5, and agent k sets $\sigma_k = \max\{\underline{\sigma}, \sigma_k^*\}$.

Proof. (i) In the case of monotone retention, the first-order derivative with respect to σ_k is

$$f\left(\frac{x^* - \theta_k}{\sigma_k}\right)\frac{x^* - \theta_k}{\sigma_k^2}.$$

It is always negative if $x^* < \theta_k$, so $\sigma_k = \underline{\sigma}$; always positive if $x^* > \theta_k$, so the agent always wants to increase the noise and the problem has no solution; and always equal to 0 if $x^* = \theta_k$, so the agent is indifferent across all choices of σ.

(ii) A type-k agent wishes to maximize the probability of being in the retention zone $[x_-, x_+]$, so he chooses $\sigma_k \geq \underline{\sigma}$, to maximize

$$F\left(\frac{x_+ - \theta_k}{\sigma_k}\right) - F\left(\frac{x_- - \theta_k}{\sigma_k}\right), \tag{a.30}$$

where F is the cdf of f. The first-order derivative of the objective function with respect to σ_k is

$$\frac{d_k(\sigma_k)}{\sigma_k^2} = \frac{1}{\sigma_k^2}\left[f\left(\frac{x_- - \theta_k}{\sigma_k}\right)(x_- - \theta_k) - f\left(\frac{x_+ - \theta_k}{\sigma_k}\right)(x_+ - \theta_k)\right],$$

where d_k is defined in (a.28). If $x_- \leq \theta_k$ and $x_+ > \theta_k$, the sign is negative, and the agent optimally chooses $\sigma_k = \underline{\sigma}$.

(iii). If $x_- > \theta_k$, rewrite the above expression as

$$\frac{d_k(\sigma_k)}{\sigma_k^2} = \frac{1}{\sigma_k^2} f\left(\frac{x_+ - \theta_k}{\sigma_k}\right)(x_- - \theta_k)\left[\frac{f\left(\frac{x_- - \theta_k}{\sigma_k}\right)}{f\left(\frac{x_+ - \theta_k}{\sigma_k}\right)} - \frac{(x_+ - \theta_k)}{(x_- - \theta_k)}\right],$$

so the sign is determined by the sign of the term inside the square brackets. By the strong MLRP, $f\left(\frac{x_- - \theta_k}{\sigma_k}\right)/f\left(\frac{x_+ - \theta_k}{\sigma_k}\right)$ is decreasing in σ_k, with limit ∞ as $\sigma_k \to 0$ and limit 0 as $\sigma_k \to \infty$, so there is a unique $\sigma_k > 0$ with $d_k(\sigma_k) = 0$. Agent k sets $\sigma_k = \max\{\underline{\sigma}, \sigma_k^*\}$.

Since σ_k^* satisfies $d_k(\sigma_k) = 0$, $f\left(\frac{x_- - \theta_k}{\sigma_k^*}\right)\frac{x_- - \theta_k}{\sigma_k^*} = f\left(\frac{x_+ - \theta_k}{\sigma_k^*}\right)\frac{x_+ - \theta_k}{\sigma_k^*}$, and by Lemma A.5 $f(z)z$ is increasing at $z \in [0, z^*)$ and decreasing at $z > z^*$, it must be that $\frac{x_- - \theta_k}{\sigma_k^*} < z^* < \frac{x_+ - \theta_k}{\sigma_k^*}$, so σ_k^* satisfies (a.29). ■

Proof of Proposition 6, Part (ii). By Lemmas A.4 and A.6(i), monotone equilibria are only possible if $\sigma_g = \sigma_b = \underline{\sigma}$.[4] Using the definition of the retention zone in (a.27) and the strong MLRP, the principal retains in such an equilibrium if and only if $x \geq x^*(\underline{\sigma})$, where $x^*(\underline{\sigma})$ is uniquely defined by

$$\beta f\left(\frac{x^*(\underline{\sigma}) - \theta_g}{\underline{\sigma}}\right) \equiv f\left(\frac{x^*(\underline{\sigma}) - \theta_b}{\underline{\sigma}}\right). \tag{a.31}$$

By Lemma A.6(i), $\theta_b \geq x^*(\underline{\sigma})$, and using strong MLRP along with (a.31), we see that

$$\beta f\left(-\frac{\theta_g - \theta_b}{\underline{\sigma}}\right) \geq f(0). \tag{a.32}$$

It follows that $\underline{\sigma} \geq \underline{\sigma}(\beta)$, where recall the definition of $\underline{\sigma}(\beta)$ from (22) in the main text.

Conversely, if $\underline{\sigma} \geq \underline{\sigma}(\beta)$, then allow both types to choose $\sigma_b = \sigma_g = \underline{\sigma}$; then the principal will select the monotone retention threshold $x^*(\underline{\sigma})$, where this threshold solves (a.31). Because $\underline{\sigma} \geq \underline{\sigma}(\beta)$, (a.32) holds, and it follows that $x^*(\underline{\sigma}) \leq \theta_b$. Applying Lemma A.6(i) yet again, we must conclude that $\sigma_b = \sigma_g = \underline{\sigma}$ is an optimal response by each of the types to the retention zone $[x^*(\underline{\sigma}), \infty)$, and the proof is complete. ■

[4]In any monotone equilibrium, $\theta_g > \theta_b \geq x^*(\sigma)$, and so by Lemma A.6(i), $\sigma_g = \underline{\sigma}$. By Lemma A.4, $\sigma_b = \sigma_g$.

Proof of Proposition 7. We proceed in a number of steps. It will be useful to define $s(x) := f'(x)/f(x)$. By MLRP, s is a decreasing function. Moreover, by single-peakedness around 0, $s(x)$ is positive for $x<0$, negative for $x>0$, and zero at $x=0$.

Lemma A.7. *For any $\sigma_b > \sigma_g$:*

(i) *$h(x)$ is decreasing for $x \le \theta_g$ and increasing for $x \ge \frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g} > \theta_g$, In particular, $X(k)$ is an interval for all $k \ge 1$, and because $k = \beta\sigma_b/\sigma_g$, this is a fortiori true for all $\beta \ge 1$.*

(ii) *Under the additional assumption that $\frac{\partial \ln f(x)}{\partial x}$ is convex for all $x>0$. $h(x)$ decreases and then increases on $\left(\theta_g, \frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g}\right)$, with its minimum achieved at the unique solution to*

$$\frac{1}{\sigma_b} s\left(\frac{x-\theta_b}{\sigma_b}\right) = \frac{1}{\sigma_g} s\left(\frac{x-\theta_g}{\sigma_g}\right), \tag{a.33}$$

so that $X(k)$ is an interval for all k higher than the minimum value of h.

(iii) *Combining cases (i) and (ii), a nonempty retention zone is an interval $[x_-, x_+]$, where x_-, x_+ are the two real roots to*

$$\beta\frac{1}{\sigma_g} f\left(\frac{x-\theta_g}{\sigma_g}\right) = \frac{1}{\sigma_b} f\left(\frac{x-\theta_b}{\sigma_b}\right), \tag{a.34}$$

and the upper root always exceeds θ_g.

Proof. For notational convenience, define $z_k(x) := (x-\theta_k)/\sigma_k$ for $k = b, g$. Then differentiate h in (a.26) to see that

$$\text{Sign } h'(x) = \text{Sign } \left[\frac{1}{\sigma_b} s\left(z_b(x)\right) - \frac{1}{\sigma_g} s\left(z_g(x)\right)\right]. \tag{a.35}$$

Figure A.2 can be used to supplement the argument that follows.

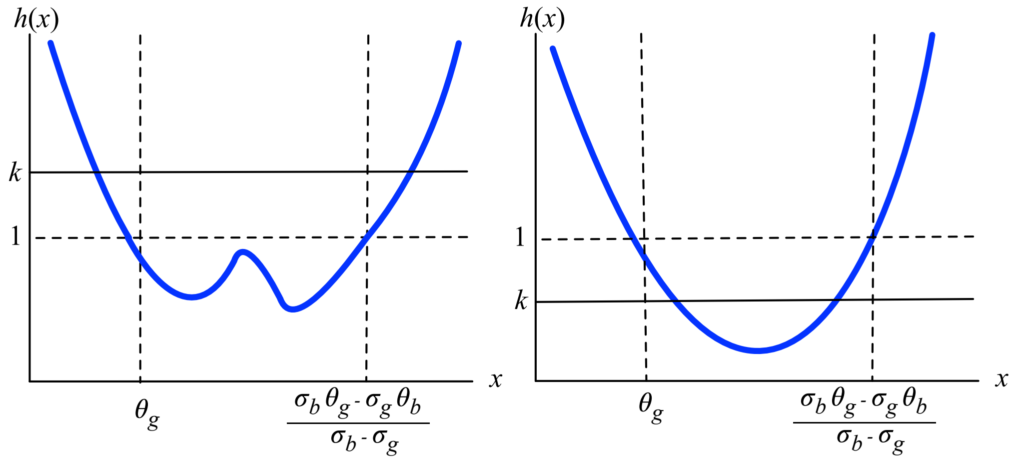
Part (i). Break $x \le \theta_g$ into two regions. If $x \le \theta_b < \theta_g$, then $0 \ge z_b(x) > z_g(x)$, so $0 \le s(z_b(x)) < s(z_g(x))$. Therefore, $\frac{1}{\sigma_b} s(z_b(x)) < \frac{1}{\sigma_g} s(z_g(x))$, and so by (a.35), $h'(x) < 0$. If $x \in (\theta_b, \theta_g)$, then $z_b(x) > 0$ but $z_g(x) \le 0$, so (a.35) implies right away that $h'(x) < 0$.

At the other extreme, if $x \ge \frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g} > \theta_b$, it is easy to verify that $0 < z_b(x) \le z_g(x)$. It follows that $0 > s(z_b(x)) \ge s(z_g(x))$, so that $\frac{1}{\sigma_b} s(z_b(x)) > \frac{1}{\sigma_g} s(z_g(x))$ and therefore (a.35) implies that $h'(x) > 0$.

By Lemma A.4(ii), $\lim_{|x|\to\infty} h(x) = \infty$. Also, $h(\theta_g) < 1$, and $h\left(\frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g}\right) = 1$. Finally, if $x \in \left(\theta_g, \frac{\sigma_g\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g}\right)$, $z_b(x) > z_g(x) > 0$, so by the single-peakedness of f around 0, $f(z_b(x)) < f(z_g(x))$ and therefore $h(x) < 1$. So $X(k)$ must be an interval for all $k \ge 1$.

(ii) Suppose that the assertion is false. Then there exist $y, w \in \left(\theta_g, \frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g}\right)$, with $y > w$ such that $h(y) = h(w)$, $h'(y) \le 0$ and $h'(w) \ge 0$. These inequalities together imply

$$\frac{\sigma_g}{\sigma_b} s\left(\frac{y-\theta_b}{\sigma_b}\right) \le s\left(\frac{y-\theta_g}{\sigma_g}\right) < s\left(\frac{w-\theta_g}{\sigma_g}\right) \le \frac{\sigma_g}{\sigma_b} s\left(\frac{w-\theta_b}{\sigma_b}\right). \tag{a.36}$$



(A) Lemma A.7, Part (i)

(B) Lemma A.7, Part (ii)

FIGURE A.2. Diagram to Accompany Lemma A.7.

Since $y, w \in \left(\theta_g, \frac{\sigma_b\theta_g - \sigma_g\theta_b}{\sigma_b - \sigma_g}\right)$, we have $\frac{y-\theta_b}{\sigma_b} > \frac{y-\theta_g}{\sigma_g}$ and $\frac{w-\theta_b}{\sigma_b} > \frac{w-\theta_g}{\sigma_g}$. Then, $s(x)$ convex for all $x > 0$ implies

$$\frac{s\left(\frac{y-\theta_b}{\sigma_b}\right) - s\left(\frac{w-\theta_b}{\sigma_b}\right)}{\frac{y-\theta_b}{\sigma_b} - \frac{w-\theta_b}{\sigma_b}} \geq \frac{s\left(\frac{y-\theta_g}{\sigma_g}\right) - s\left(\frac{w-\theta_g}{\sigma_g}\right)}{\frac{y-\theta_g}{\sigma_g} - \frac{w-\theta_g}{\sigma_g}},$$

or, equivalently,

$$\frac{\sigma_b}{\sigma_g}\left(s\left(\frac{w-\theta_b}{\sigma_b}\right) - s\left(\frac{y-\theta_b}{\sigma_b}\right)\right) \leq s\left(\frac{w-\theta_g}{\sigma_g}\right) - s\left(\frac{y-\theta_g}{\sigma_g}\right).$$

Since $s\left(\frac{w-\theta_b}{\sigma_b}\right) - s\left(\frac{y-\theta_b}{\sigma_b}\right) > 0$ and $\sigma_b > \sigma_g$ we also have

$$\frac{\sigma_g}{\sigma_b}\left(s\left(\frac{w-\theta_b}{\sigma_b}\right) - s\left(\frac{y-\theta_b}{\sigma_b}\right)\right) < s\left(\frac{w-\theta_g}{\sigma_g}\right) - s\left(\frac{y-\theta_g}{\sigma_g}\right),$$

which contradicts (a.36).

(iii) The assertion that the retention zone is an interval follows from the arguments in parts (i) and (ii). The equation (a.34) is equivalent to $X(k) = k$ and therefore must define the edges of the retention zone. Because h is decreasing all the way up to $x = \theta_g$, the upper root x_+ that defines the retention zone must exceed θ_g. See Figure A.2 for an illustration. ■

Lemma A.8. *Under the conditions of Lemma A.7, consider any situation in which $\sigma_b > \sigma_g \geq \underline{\sigma}$, in which the principal retains if and only if $x \in [x_-, x_+]$ with $x_+ > x_-$, where these roots solve (a.34), and in which type b is playing a best response to the principal's choice of retention zone. Then, the derivative of the payoff of type g evaluated at σ_g is strictly negative.*

Proof. Because $\sigma_b > \sigma_g \geq \underline{\sigma}$, σ_b is an interior solution to type b's optimization problem. Therefore, the first order condition for type b's optimization holds with equality, and

(a.37) $$f\left(\frac{x_+ - \theta_b}{\sigma_b}\right)(x_+ - \theta_b) = f\left(\frac{x_- - \theta_b}{\sigma_b}\right)(x_- - \theta_b),$$

which also shows in passing that $x_- > \theta_b$. (For $x_+ > \theta_g > \theta_b$ by Lemma A.7(iii), so that every term in (a.37) must be strictly positive.)

Now, let's study the derivative of type-g's retention probability evaluated at σ_g. This is:

$$\frac{1}{\sigma_g^2} f\left(\frac{x_- - \theta_g}{\sigma_g}\right)(x_- - \theta_g) - \frac{1}{\sigma_g^2} f\left(\frac{x_+ - \theta_g}{\sigma_g}\right)(x_+ - \theta_g)$$
$$= \frac{1}{\beta\sigma_b\sigma_g}\left[f\left(\frac{x_- - \theta_b}{\sigma_b}\right)(x_- - \theta_g) - f\left(\frac{x_+ - \theta_b}{\sigma_b}\right)(x_+ - \theta_g)\right]$$
$$= \frac{1}{\beta\sigma_g} f\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\left[\frac{(x_+ - \theta_b)}{(x_- - \theta_b)}(x_- - \theta_g) - (x_+ - \theta_g)\right],$$

where the first equality invokes (a.34) for the roots, and the second equality uses the first order condition (a.37) for the low type. Because $x_- > \theta_b$, $x_+ > x_-$ and $\theta_g > \theta_b$, this derivative is negative. It follows that σ_g must be at the corner $\underline{\sigma}$, and the proof is complete. ∎

Lemma A.8 is suggestive of the fact that in any bounded retention equilibrium, $\sigma_g = \underline{\sigma}$. So in our hunt for such equilibria, we will provisionally fix σ_g at $\underline{\sigma}$, and to save on notation we denote σ_b by simply σ.

Lemma A.12 below will guarantee that the principal will employ nontrivial bounded retention intervals for any $\sigma > \underline{\sigma}$, under some conditions. For this, we first prove some technical results (Lemmas A.9–A.11 below). Let $x^{**}(\sigma)$ be the unique minimizer of $h(x)$, defined by (a.33) in Lemma A.7, when type b employs $\sigma_b = \sigma > \underline{\sigma}$ and type g plays $\sigma_g = \underline{\sigma}$. That is,

$$\text{(a.38)} \qquad \frac{1}{\sigma} s\left(\frac{x^{**}(\sigma) - \theta_b}{\sigma}\right) \equiv \frac{1}{\underline{\sigma}} s\left(\frac{x^{**}(\sigma) - \theta_g}{\underline{\sigma}}\right).$$

Lemma A.9. *$\frac{x^{**}(\sigma)-\theta_b}{\sigma}$ is strictly decreasing in σ, and there is a unique σ^{**} that solves*

$$\text{(a.39)} \qquad x^{**}(\sigma) = \theta_b + z^*\sigma.$$

Proof. Differentiate (a.38) with respect to σ to obtain

$$\text{(a.40)} \qquad x^{**\prime}(\sigma) = \frac{\frac{1}{\sigma^2} s'\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right)\frac{x^{**}(\sigma)-\theta_b}{\sigma} + \frac{1}{\sigma^2} s\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right)}{\frac{1}{\sigma^2} s'\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right) - \frac{1}{\underline{\sigma}^2} s'\left(\frac{x^{**}(\sigma)-\theta_g}{\underline{\sigma}}\right)}.$$

By Lemma A.7(ii), $x^{**}(\sigma) > \theta_g$, and $s(x)$ is decreasing and negative for $x > 0$, so the numerator is negative. By Lemma A.7(ii), $h(x)$ is decreasing for $x < x^{**}(\sigma)$ and increasing for $x > x^{**}(\sigma)$. That means that

$$\frac{1}{\underline{\sigma}} s\left(\frac{x - \theta_g}{\underline{\sigma}}\right) > \frac{1}{\sigma} s\left(\frac{x - \theta_b}{\sigma}\right) \text{ for all } x < x^{**}(\sigma), \text{ and}$$
$$\frac{1}{\underline{\sigma}} s\left(\frac{x - \theta_g}{\underline{\sigma}}\right) < \frac{1}{\sigma} s\left(\frac{x - \theta_b}{\sigma}\right) \text{ for all } x > x^{**}(\sigma).$$

Therefore, at $x = x^{**}(\sigma)$ we have

$$\frac{1}{\sigma^2} s'\left(\frac{x^{**}(\sigma) - \theta_b}{\sigma}\right) > \frac{1}{\underline{\sigma}^2} s'\left(\frac{x^{**}(\sigma) - \theta_g}{\underline{\sigma}}\right)$$

so the denominator in (a.40) is positive, and $x^{**\prime}(\sigma) < 0$: $x^{**}(\sigma)$ is decreasing. Then so is $\frac{x^{**}(\sigma)-\theta_b}{\sigma}$.

Finally, as $\sigma \to \underline{\sigma}$, $x^{**}(\sigma)$ cannot converge to a finite value, because $x^{**}(\sigma)$ solves (a.38) and $s(x)$ is strictly decreasing. Since $x^{**}(\sigma) > \theta_g$ it must be that $x^{**}(\sigma) \to \infty$ as $\sigma \to \underline{\sigma}$. So $\frac{x^{**}(\sigma)-\theta_b}{\sigma} \to \infty$ as $\sigma \to \underline{\sigma}$. And $\frac{x^{**}(\sigma)-\theta_b}{\sigma} \to 0$ as $\sigma \to \infty$. Then, since $z^* > 0$, there exists a unique $\sigma = \sigma^{**}$ that satisfies (a.39). ∎

To proceed further, define

(a.41) $$\beta_l := \frac{\frac{1}{\sigma^{**}} f\left(\frac{x^{**}(\sigma^{**})-\theta_b}{\sigma^{**}}\right)}{\frac{1}{\underline{\sigma}} f\left(\frac{x^{**}(\sigma^{**})-\theta_g}{\underline{\sigma}}\right)}.$$

Lemma A.10. $\beta_l < 1$.

Proof. By Lemma A.7, the minimizer of $h(x)$, $x^{**}(\sigma)$, is in the interval $\left(\theta_g, \frac{\sigma\theta_g - \underline{\sigma}\theta_b}{\sigma - \underline{\sigma}}\right)$. That means that $\frac{x^{**}(\sigma)-\theta_b}{\sigma} > \frac{x^{**}(\sigma)-\theta_g}{\underline{\sigma}} > 0$ and therefore $f\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right) < f\left(\frac{x^{**}(\sigma)-\theta_g}{\underline{\sigma}}\right)$. Then, for $\sigma > \underline{\sigma}$ we also have $\frac{1}{\sigma} f\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right) < \frac{1}{\underline{\sigma}} f\left(\frac{x^{**}(\sigma)-\theta_g}{\underline{\sigma}}\right)$. $\beta_l < 1$ results from taking $\sigma = \sigma^{**}$, defined as the solution to (a.39). ∎

Lemma A.11. *Let $\beta(\sigma)$ be defined as*

(a.42) $$\beta(\sigma)\frac{1}{\underline{\sigma}} f\left(\frac{x^{**}(\sigma)-\theta_g}{\underline{\sigma}}\right) \equiv \frac{1}{\sigma} f\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right).$$

(i) *If $\beta = \beta(\sigma)$, $X = \{x^{**}(\sigma)\}$;*

If $\beta > \beta(\sigma)$, $X = [x_-(\sigma), x_+(\sigma)]$ with $x_+(\sigma) > x_-(\sigma)$, which are the two roots to (a.34)*;*

If $\beta < \beta(\sigma)$, $X = \emptyset$;

(ii) *$\beta(\sigma)$ is increasing at all $\sigma \in (\underline{\sigma}, \sigma^{**})$; it is decreasing at all $\sigma > \sigma^{**}$; it attains a maximum at $\sigma = \sigma^{**}$, and its maximum value is β_l, defined in* (a.41).

Proof. (i) By Lemma A.7(ii), $x^{**}(\sigma)$ is the unique minimizer of $h(x)$. Recall the retention zone is $X = \{x : h(x) \le k\}$.

If $\beta = \beta(\sigma)$ or, equivalently, if $h(x^{**}(\sigma)) = k$, $X = \{x^{**}(\sigma)\}$.

If $\beta > \beta(\sigma)$, $h(x^{**}(\sigma)) < k$. By Lemma A.7(ii), $X = [x_-(\sigma), x_+(\sigma)]$ with $x_+(\sigma) > x_-(\sigma)$, which are the two roots to (a.34).

If $\beta < \beta(\sigma)$, $h(x^{**}(\sigma)) > k$ and therefore $h(x) > k$ for all x, so $X = \emptyset$.

(ii) Take (a.42) and differentiate with respect to σ. After some algebra, and using the fact that $\beta(\sigma)$ satisfies (a.42) and $x^{**}(\sigma)$ satisfies (a.38), we obtain

$$\text{(a.43)} \qquad \beta'(\sigma) = -\frac{\beta(\sigma)}{\sigma} \cdot \frac{1}{f\left(\frac{x^{**}(\sigma)-\theta_b}{\sigma}\right)} \cdot \left[\frac{\partial f(z)\, z}{\partial z}\right]_{|_{z=\frac{x^{**}(\sigma)-\theta_b}{\sigma}}}.$$

By Lemma A.7(ii), $x^{**}(\sigma) > \theta_g > \theta_b$. By Lemma A.5, for $z \geq 0$, $\partial f(z)z/\partial z$ is first positive, then negative, and zero at z^*. By Lemma A.9, $\frac{x^{**}(\sigma)-\theta_b}{\sigma}$ is decreasing in σ and there exists a unique σ^{**}, defined in (a.39), such that $\frac{x^{**}(\sigma)-\theta_b}{\sigma} = z^*$. Then, for $\sigma \in (\underline{\sigma}, \sigma^{**})$, $\beta'(\sigma) > 0$, and for $\sigma > \sigma^{**}$, $\beta'(\sigma) < 0$. Finally, this means that $\beta(\sigma)$ attains a *maximum* at σ^{**}. This maximum value of β is β_l, defined in (a.41). ∎

We now establish a sufficient condition for the existence of bounded retention intervals for any possible pair (σ_g, σ_b) with $\sigma_b > \sigma_g = \underline{\sigma}$.

Lemma A.12. *For any (σ_g, σ_b) with $\sigma_b = \sigma > \underline{\sigma}$ and $\sigma_g = \underline{\sigma}$, the principal employs a bounded retention interval if $\beta > \beta_l$, or equivalently, if $\underline{\sigma} < \widehat{\underline{\sigma}}(\beta)$, where $\widehat{\underline{\sigma}}(\beta)$ is defined as the value of $\underline{\sigma}$ that solves (a.41) when we replace β_l with $\beta < 1$ and $\hat{\underline{\sigma}}(\beta) = \infty$ for $\beta \geq 1$.*

Proof. By Lemma A.11(ii), if $\beta > \beta_l$, $\beta > \beta(\sigma)$ for all $\sigma > \underline{\sigma}$. Then, by Lemma A.11(i), the principal retains if and only if $x \in [x_-(\sigma), x_+(\sigma)]$, with $x_-(\sigma) < x_+(\sigma)$, for all $\sigma > \underline{\sigma}$.

Now, we show that $\beta > \beta_l$ is equivalent to $\underline{\sigma} < \widehat{\underline{\sigma}}(\beta)$. First, recall that for any $\underline{\sigma} > 0$, σ^{**} and $x^{**}(\sigma)$ in (a.39) and (a.38) respectively are well-defined. Then, so is $\beta_l(\underline{\sigma})$, defined in (a.41). We show that $\beta_l(\underline{\sigma})$ is strictly increasing in $\underline{\sigma}$. Define $\widetilde{x}(\underline{\sigma})$ by $\widetilde{x}(\underline{\sigma}) := x^{**}(\sigma^{**}(\underline{\sigma}), \underline{\sigma})$; then $\widetilde{x}(\underline{\sigma})$ is just $x^{**}(\sigma, \underline{\sigma})$, defined in (a.38), evaluated at $\sigma^{**}(\underline{\sigma})$, defined in (a.39). Because

$$\frac{\widetilde{x}(\underline{\sigma}) - \theta_b}{\sigma^{**}(\underline{\sigma})} = z^*,$$

the equation (a.38) evaluated at $\sigma^{**}(\underline{\sigma})$ becomes

$$\text{(a.44)} \qquad \frac{\widetilde{x}(\underline{\sigma}) - \theta_b}{\underline{\sigma}} s\left(\frac{\widetilde{x}(\underline{\sigma}) - \theta_g}{\underline{\sigma}}\right) \equiv -1,$$

and (a.41), also evaluated at $\sigma^{**}(\underline{\sigma})$, becomes

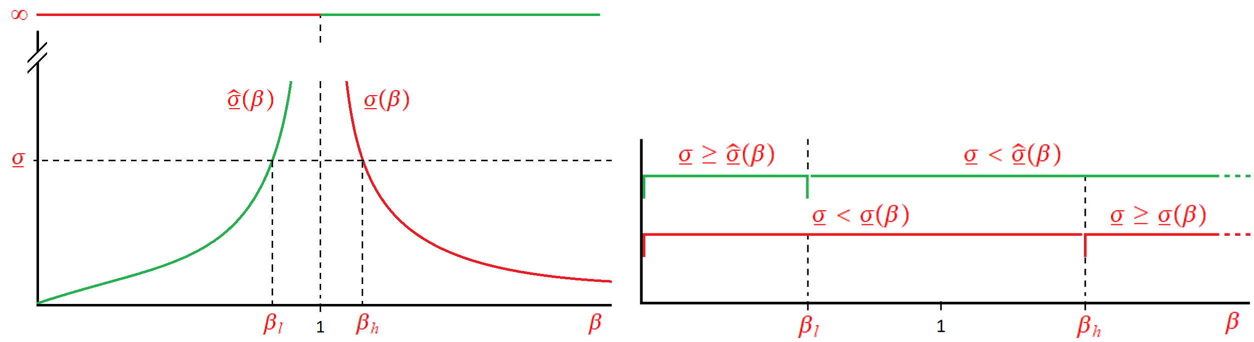
$$\text{(a.45)} \qquad \beta_l(\underline{\sigma}) = \frac{\frac{z^* f(z^*)}{\widetilde{x}(\underline{\sigma})-\theta_b}}{\frac{1}{\underline{\sigma}} f\left(\frac{\widetilde{x}(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right)}.$$

Differentiate (a.45) with respect to $\underline{\sigma}$ to get

$$\beta_l'(\underline{\sigma}) = \beta_l(\underline{\sigma}) \cdot \left[\frac{1}{\underline{\sigma}}\left(1 + \frac{1}{\underline{\sigma}} s\left(\frac{\widetilde{x}(\underline{\sigma}) - \theta_g}{\underline{\sigma}}\right)(\widetilde{x}(\underline{\sigma}) - \theta_g)\right) - \left(\frac{1}{\widetilde{x}(\underline{\sigma}) - \theta_b} + \frac{1}{\underline{\sigma}} s\left(\frac{\widetilde{x}(\underline{\sigma}) - \theta_g}{\underline{\sigma}}\right)\right) \cdot \widetilde{x}'(\underline{\sigma})\right].$$

Using (a.44) we can replace $\frac{1}{\underline{\sigma}} s\left(\frac{\widetilde{x}(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right)$ by $-1/(\widetilde{x}(\underline{\sigma}) - \theta_b)$, and we obtain

$$\beta_l'(\underline{\sigma}) = \beta_l(\underline{\sigma}) \cdot \frac{1}{\underline{\sigma}}\left(1 - \frac{\widetilde{x}(\underline{\sigma}) - \theta_g}{\widetilde{x}(\underline{\sigma}) - \theta_b}\right) = \beta_l(\underline{\sigma}) \cdot \frac{1}{\underline{\sigma}} \frac{\theta_g - \theta_b}{\widetilde{x}(\underline{\sigma}) - \theta_b} > 0.$$



(A) The two upper bounds as functions of β.

(B) Relationship between the two conditions.

FIGURE A.3. The two upper bounds for $\underline{\sigma}$.

We describe the limit of $\beta_l(\sigma)$ as $\underline{\sigma} \to 0$. If $\underline{\sigma} \to 0$, the value that maximizes the likelihood of the good type, $x^{**}(\sigma)$, converges to θ_g for any σ. That means that $\frac{1}{\underline{\sigma}} f\left(\frac{\widetilde{x}(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right) \to \infty$ as $\underline{\sigma} \to 0$, and therefore

$$\beta_l\left(\underline{\sigma}\right) = \frac{\frac{z^* f(z^*)}{\widetilde{x}(\underline{\sigma})-\theta_b}}{\frac{1}{\underline{\sigma}} f\left(\frac{\widetilde{x}(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right)} \to 0.$$

So we can conclude that $\beta > \beta_l$ is equivalent to $\underline{\sigma} < \widehat{\underline{\sigma}}\left(\beta\right)$, where $\widehat{\underline{\sigma}}\left(\beta\right)$ is defined as the value of $\underline{\sigma}$ that solves (a.41) with $\beta = \beta_l$. ■

Now we determine a second upper bound on $\underline{\sigma}$ (recall that a parallel bound (12) was used to negate the existence of a monotone retention equilibrium):

$$\underline{\sigma} < \hat{\underline{\sigma}}(\beta) \text{ if } \beta \in (0,1). \tag{a.46}$$

By Lemma A.10, $\beta_l < 1$, and therefore if $\beta \geq 1$, (a.46) is trivially satisfied.

For the arguments to follow, it will be useful to indicate clearly the way in which the two upper bounds on $\underline{\sigma}$ relate to each other. To do so, let β_h be the value of $\beta > 1$ such that $\underline{\sigma} = \underline{\sigma}(\beta)$. Now look at Figure A.3. Notice that (a) $\underline{\sigma} \geq \hat{\underline{\sigma}}(\beta)$ implies $\underline{\sigma} < \underline{\sigma}(\beta)$; (b) $\underline{\sigma} \geq \underline{\sigma}(\beta)$ implies $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$; whereas (c) $\underline{\sigma} < \underline{\sigma}(\beta)$ and $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$ can occur simultaneously.

The condition $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$ is a sufficient condition under which the principal, when conjecturing that the agent will play $\sigma_g = \underline{\sigma}$ and $\sigma_b = \sigma > \underline{\sigma}$, will employ a nontrivial, bounded retention interval, for *any* such σ. The next step is to show that there exists a fixed point between the noise σ_b conjectured by the principal and the one optimally chosen by type b. For this, we need to analyze the way in which the retention interval $[x_-(\sigma), x_+(\sigma)]$ behaves, in particular as $\sigma \to \underline{\sigma}$ and as $\sigma \to \infty$. This is what we do next.

Lemma A.13. *Let $x_-\left(\sigma\right)$ and $x_+\left(\sigma\right)$ be the roots to*

$$\beta \frac{1}{\underline{\sigma}} f\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma} f\left(\frac{x-\theta_b}{\sigma}\right). \tag{a.47}$$

for $\sigma > \underline{\sigma}$. *Then,*

(i) $\lim_{\sigma\to\underline{\sigma}} x_-(\sigma) = x^*(\underline{\sigma})$ *and* $\lim_{\sigma\to\underline{\sigma}} x_+(\sigma) = \infty$

(ii) $\lim_{\sigma\to\infty} x_-(\sigma) < \theta_b$.

(iii) *For* $i = -, +$, *the derivatives are*

(a.48) $$x_i'(\sigma) = -\frac{1}{\sigma^2}\frac{\left[\frac{\partial f(z)z}{\partial z}\right]\big|_{z=\frac{x_i(\sigma)-\theta_b}{\sigma}}}{\beta\frac{1}{\underline{\sigma}^2}f'\left(\frac{x_i(\sigma)-\theta_g}{\underline{\sigma}}\right)-\frac{1}{\sigma^2}f'\left(\frac{x_i(\sigma)-\theta_b}{\sigma}\right)}.$$

(iv) *If* $\underline{\sigma} \geq \underline{\sigma}(\beta)$, *then* $x_-(\sigma) < \theta_b$ *for all* $\sigma > \underline{\sigma}$.

Proof. (i) $x_-(\sigma)$ must satisfy $h'(x_-) < 0$, whereas $x_+(\sigma)$ must satisfy $h'(x_+) > 0$. Recalling (a.35), this means that:

(a.49) $$\frac{1}{\underline{\sigma}}s\left(\frac{x_-(\sigma)-\theta_g}{\underline{\sigma}}\right) > \frac{1}{\sigma}s\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right),$$

and

(a.50) $$\frac{1}{\underline{\sigma}}s\left(\frac{x_+(\sigma)-\theta_g}{\underline{\sigma}}\right) < \frac{1}{\sigma}s\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right).$$

Similarly, the monotone threshold $x^*(\underline{\sigma})$ satisfies

(a.51) $$\beta f\left(\frac{x^*(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right) = f\left(\frac{x^*(\underline{\sigma})-\theta_b}{\underline{\sigma}}\right)$$

and

(a.52) $$s\left(\frac{x^*(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right) > s\left(\frac{x^*(\underline{\sigma})-\theta_b}{\underline{\sigma}}\right).$$

In the limit as $\sigma \to \underline{\sigma}$, the condition (a.47) is the same as (a.51), and the condition (a.49) is the same as (a.52) (equality of (a.49) at $\sigma = \underline{\sigma}$ cannot hold because $s(x)$ is decreasing). Because the MLRP implies that $x^*(\underline{\sigma})$ is uniquely defined, we have $\lim_{\sigma\to\underline{\sigma}} x_-(\sigma) = x^*(\underline{\sigma})$.

Notice that, as $\sigma \to \underline{\sigma}$, (a.50) becomes

$$s\left(\frac{x_+(\underline{\sigma})-\theta_b}{\underline{\sigma}}\right) \geq s\left(\frac{x_+(\underline{\sigma})-\theta_g}{\underline{\sigma}}\right).$$

But since $s(x)$ is decreasing and $\theta_g > \theta_b$, it must be either true that $\lim_{\sigma\to\underline{\sigma}} x_+(\sigma) = \infty$, or $\lim_{\sigma\to\underline{\sigma}} x_+(\sigma) = -\infty$. But the latter cannot hold because $x_+(\sigma) \geq x_-(\sigma)$ for all $\sigma > \underline{\sigma}$, so it must be that $\lim_{\sigma\to\underline{\sigma}} x_+(\sigma) = \infty$.

(ii) Notice that, for σ large enough, at $x = \theta_b$ we have

$$\beta\frac{1}{\underline{\sigma}}f\left(\frac{\theta_b-\theta_g}{\underline{\sigma}}\right) > \frac{1}{\sigma}f(0).$$

So for σ large enough, the principal retains the agent at $x = \theta_b$; i.e., $\lim_{\sigma\to\infty} x_-(\sigma) < \theta_b$.

(iii) Differentiate (a.55) with respect to σ to obtain (a.48).

(iv) By part (i) of this Lemma, $x_-(\sigma) \to x^*(\underline{\sigma})$ as $\sigma \to \underline{\sigma}$. If $\underline{\sigma} \geq \underline{\sigma}(\beta)$, a monotone equilibrium exists by Proposition 6(ii), and therefore $x^*(\underline{\sigma}) \leq \theta_b$, by Lemma A.6(i). $\underline{\sigma} \geq \underline{\sigma}(\beta)$ also implies that $\beta > 1 > \beta_l$ and therefore $x_-(\sigma)$ is well-defined for any $\sigma > \underline{\sigma}$ by Lemma A.11(i), and it is clearly continuous in σ.

We note again that $x_-(\sigma)$ must satisfy $h'(x_-) < 0$, which implies (a.49). Then, the denominator in (a.48) for $x_-(\sigma)$ is positive, so that $x'_-(\sigma)$ is also well-defined and continuous in $\sigma > \underline{\sigma}$, and in addition:

$$\text{Sign}\left(x'_-(\sigma)\right) = \text{Sign}\left(-\left[\frac{\partial f(z)z}{\partial z}\right]_{\big|_{z=\frac{x_i(\sigma)-\theta_b}{\sigma}}}\right). \tag{a.53}$$

If $x^*(\underline{\sigma}) < \theta_b$, the continuity of $x_-(\sigma)$ implies that $x_-(\sigma) < \theta_b$ for $\sigma > \underline{\sigma}$ close enough to $\underline{\sigma}$.

If $x^*(\underline{\sigma}) = \theta_b$, (a.53) and Lemma A.5 together say that $x'_-(\sigma) < 0$ for $\sigma > \underline{\sigma}$ close enough to $\underline{\sigma}$, so once again $x_-(\sigma) < \theta_b$ for $\sigma > \underline{\sigma}$ close enough to $\underline{\sigma}$.

Then, for the assertion to be false, it is required that $x'_-(\sigma) \geq 0$ for some $\sigma > \underline{\sigma}$ at which $x'_-(\sigma) = \theta_b$, but this contradicts (a.53) and Lemma A.5. ∎

Guided by Lemma A.6, let us now consider the following mapping, defined for all $\sigma > \underline{\sigma}$:

$$\Psi(\sigma) = \max\left\{\underline{\sigma}, \sigma^*(\sigma)\right\},$$

where $\sigma^*(\sigma)$ is the unrestricted maximizer of type-b's retention probability when the principal retains if and only if $x \in [x_-(\sigma), x_+(\sigma)]$. So by Lemma A.6(iii), $\sigma^*(\sigma)$ solves:

$$f\left(\frac{x_-(\sigma)-\theta_b}{\sigma^*(\sigma)}\right)[x_-(\sigma)-\theta_b] = f\left(\frac{x_+(\sigma)-\theta_b}{\sigma^*(\sigma)}\right)[x_+(\sigma]-\theta_b),\tag{a.54}$$

and $x_-(\sigma)$ and $x_+(\sigma)$ are the roots to

$$\beta\frac{1}{\underline{\sigma}}f\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma}f\left(\frac{x-\theta_b}{\sigma}\right).\tag{a.55}$$

The following Lemma determines an important feature of $\Psi(\sigma)$.

Lemma A.14. *At any fixed point $\sigma = \Psi(\sigma)$, $\Psi'(\sigma) < 0$.*

Proof. At any fixed point σ we have $\underline{\sigma} < \sigma = \Psi(\sigma)$, so $\Psi(\sigma)$ must solve (a.54). Differentiate (a.54) with respect to σ, replacing $\sigma^*(\sigma)$ by $\Psi(\sigma)$, to obtain

$$\Psi'(\sigma) = \frac{\left(f'\left(\frac{x_+(\sigma)-\theta_b}{\Psi(\sigma)}\right)\frac{x_+(\sigma)-\theta_b}{\Psi(\sigma)} + f\left(\frac{x_+(\sigma)-\theta_b}{\Psi(\sigma)}\right)\right)x'_+(\sigma) - \left(f'\left(\frac{x_-(\sigma)-\theta_b}{\Psi(\sigma)}\right)\frac{x_-(\sigma)-\theta_b}{\Psi(\sigma)} + f\left(\frac{x_-(\sigma)-\theta_b}{\Psi(\sigma)}\right)\right)x'_-(\sigma)}{f'\left(\frac{x_+(\sigma)-\theta_b}{\Psi(\sigma)}\right)\left(\frac{x_+(\sigma)-\theta_b}{\Psi(\sigma)}\right)^2 - f'\left(\frac{x_-(\sigma)-\theta_b}{\Psi(\sigma)}\right)\left(\frac{x_-(\sigma)-\theta_b}{\Psi(\sigma)}\right)^2}.$$

Evaluating at $\Psi(\sigma) = \sigma$ and plugging the expressions for $x'_-(\sigma)$ and $x'_+(\sigma)$ in (a.48) yields

$$\Psi'(\sigma) = \frac{1}{\sigma^2}\frac{\frac{\left(f'\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)\frac{x_-(\sigma)-\theta_b}{\sigma}+f\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)\right)^2}{\beta\frac{1}{\underline{\sigma}^2}f'\left(\frac{x_-(\sigma)-\theta_g}{\underline{\sigma}}\right)-\frac{1}{\sigma^2}f'\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)} - \frac{\left(f'\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)\frac{x_+(\sigma)-\theta_b}{\sigma}+f\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)\right)^2}{\beta\frac{1}{\underline{\sigma}^2}f'\left(\frac{x_+(\sigma)-\theta_g}{\underline{\sigma}}\right)-\frac{1}{\sigma^2}f'\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)}}{f'\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)^2 - f'\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)^2}\tag{a.56}$$

Since $h'(x_-(\sigma)) < 0$ and $h'(x_+(\sigma)) > 0$,

$$\beta \frac{1}{\underline{\sigma}^2} f'\left(\frac{x_+(\sigma) - \theta_g}{\underline{\sigma}}\right) < \frac{1}{\sigma^2} f'\left(\frac{x_+(\sigma) - \theta_b}{\sigma}\right), \text{ and}$$
$$\beta \frac{1}{\underline{\sigma}^2} f'\left(\frac{x_-(\sigma) - \theta_g}{\underline{\sigma}}\right) > \frac{1}{\sigma^2} f'\left(\frac{x_-(\sigma) - \theta_b}{\sigma}\right),$$

so the numerator in (a.56) is positive. The second-order condition of σ is

$$f'\left(\frac{x_+ - \theta_b}{\sigma}\right) \frac{(x_+ - \theta_b)^2}{\sigma^2} - f'\left(\frac{x_- - \theta_b}{\sigma}\right) \frac{(x_- - \theta_b)^2}{\sigma^2} < 0,$$

which is the denominator in (a.56). So $\Psi'(\sigma) < 0$ at any fixed point. ■

Lemma A.15. *If both the conditions $\underline{\sigma} < \underline{\sigma}(\beta)$ and $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$ hold, there is a unique nontrivial equilibrium. It has bounded retention.*

Proof. Under the condition $\underline{\sigma} < \underline{\sigma}(\beta)$, both bounded replacement equilibria and monotone equilibria are ruled out by Proposition 6. Importantly, the nonexistence of monotone equilibrium is equivalent to $x^*(\underline{\sigma}) > \theta_b$.

By Lemma A.13(i), $[x_-(\sigma), x_+(\sigma)] \to [x^*(\underline{\sigma}), \infty)$ as $\sigma \to \underline{\sigma}$. Since $x^*(\sigma) \geq \theta_b$, by Lemma A.6(i), we have that $\Psi(\sigma) \to \infty$ as $\sigma \to \underline{\sigma}$.

By Lemma A.13(ii), for σ large enough, $x_-(\sigma) < \theta_b$, and by Lemma A.7, $x_+(\sigma) > \theta_b$, so $\theta_b \in [x_-(\sigma), x_+(\sigma)]$ for σ large enough. Then, $\Psi(\sigma) = \underline{\sigma}$ for σ large enough, by Lemma A.6(ii).

By Lemma A.12, (a.46) guarantees that the principal always employs a bounded retention interval for any pair (σ_g, σ_b) with $\sigma_b = \sigma > \underline{\sigma} = \sigma_g$. Then, it is clear that $x_-(\sigma)$ and $x_+(\sigma)$ are continuous in σ, and therefore so is $\Psi(\sigma)$. Then, the above end-point verifications and continuity guarantee that Ψ has at least one fixed point. Lemma A.14 guarantees that such a fixed point is unique. At this fixed point, both the principal and the bad type are playing best responses.

It remains to show that the good type is also playing a best response at $\underline{\sigma}$. Notice that $x_(\sigma) > \theta_g$, so that Lemma A.6(iii) applies, and the derivative of the payoff of the good type with respect to σ_g is given by $\sigma_g^2 d_g(\sigma_g)$, where d_g is defined in equation (a.28). By Lemma A.8, this derivative is negative at $\sigma_g = \underline{\sigma}$, and by part (iii) of Lemma A.6, it must continue to be negative for all $\sigma_g > \underline{\sigma}$. Therefore the best response of the good type is indeed to play $\underline{\sigma}$, as claimed. ■

Lemma A.16. *If $\underline{\sigma} \geq \underline{\sigma}(\beta)$ (in which case $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$ automatically holds), then a bounded retention equilibrium cannot exist.*

Proof. Every bounded retention equilibrium involves $\sigma_g = \underline{\sigma}$. This follows from Lemma A.8 and the decreasing payoff derivative as noted in part (iii) of Lemma A.6. It follows that every bounded retention equilibrium can be expressed as a fixed point of the mapping $\Psi(\sigma)$, where at the fixed point, $\sigma > \underline{\sigma}$. However, if $\underline{\sigma} \geq \underline{\sigma}(\beta)$, $x_-(\sigma) < \theta_b$ for all $\sigma > \underline{\sigma}$, by Lemma A.13(iv). Moreover, $x_+(\sigma) > \theta_b$ for all $\sigma > \underline{\sigma}$, by Lemma A.7(iii), so $\theta_b \in [x_-(\sigma), x_+(\sigma)]$ for all $\sigma > \underline{\sigma}$. Then, by Lemma A.6(ii), $\Psi(\sigma) = \underline{\sigma}$ for all $\sigma > \underline{\sigma}$, and a fixed point of Ψ cannot exist. ■

Lemma A.17. *If $\underline{\sigma} \geq \hat{\underline{\sigma}}(\beta)$, then $\underline{\sigma} < \underline{\sigma}(\beta)$ and a nontrivial equilibrium does not exist.*

Proof. By Proposition 6, a bounded replacement equilibrium does not exist. Moreover, $\underline{\sigma} \geq \hat{\underline{\sigma}}(\beta)$ implies $\beta \leq 1$. But then $\underline{\sigma} < \underline{\sigma}(\beta) = \infty$, and again by Proposition 6, a monotone retention equilibrium cannot exist. It only remains to show that a bounded retention equilibrium cannot exist either.

Consider $\beta < \beta_l$. By Lemma A.11 a nontrivial bounded retention equilibrium is possible only if $\sigma < \sigma_l^\beta$ or if $\sigma > \sigma_h^\beta$, where $\beta\left(\sigma_l^\beta\right) = \beta\left(\sigma_h^\beta\right) = \beta$, and $\sigma_l^\beta < \sigma^{**} < \sigma_h^\beta$. What we now have to show is that, for any σ such that a nontrivial retention regime is employed by the principal, type-b's best response will never coincide with such σ.

Suppose σ_l^β exists, and consider $\sigma < \sigma_l^\beta$. By Lemma A.11(i), $x_-\left(\sigma_l^\beta\right) = x^{**}\left(\sigma_l^\beta\right)$, and by Lemma A.11(ii) $\sigma_l^\beta < \sigma^{**}$ implies $\beta'\left(\sigma_l^\beta\right) > 0$ or, equivalently (see (a.43)),

$$\frac{\partial f(z)\, z}{\partial z}\Big|_{z=\frac{x_-(\sigma)-\theta_b}{\sigma}} < 0.$$

By Lemma A.5, $\frac{x_-\left(\sigma_l^\beta\right)-\theta_b}{\sigma_l^\beta} > z^*$, but type-$b$'s optimal σ satisfies $\frac{x_-(\sigma)-\theta_b}{\sigma} < z^*$ by Lemma A.6(iii), so σ_l^β cannot be a fixed point. The next step is to show that $\frac{x_-(\sigma)-\theta_b}{\sigma} > z^*$ for all $\sigma < \sigma_l^\beta$. So suppose not: there exists σ such that $\frac{x_-(\sigma)-\theta_b}{\sigma} \leq z^*$. That means that $x_-(\sigma)$ crosses $\sigma z^* + \theta_b$ from below at some point, which requires $x_-'(\sigma) > z^* > 0$ at such intersection point. But $x_-(\sigma) = \sigma z^* + \theta_b$ implies $\frac{x_-(\sigma)-\theta_b}{\sigma} = z^*$, and therefore by Lemma A.5 and inspection of (a.48), we have $x_-'(\sigma) = 0$, a contradiction.

Now suppose σ_h^β exists, and consider $\sigma > \sigma_h^\beta$. By Lemma A.11(i), $x_+\left(\sigma_h^\beta\right) = x^{**}\left(\sigma_h^\beta\right)$, and by Lemma A.11(ii), $\sigma_h^\beta > \sigma^{**}$ implies $\beta'\left(\sigma_h^\beta\right) < 0$ or, equivalently (see (a.43) again),

$$\frac{\partial f(z)\, z}{\partial z}\Big|_{z=\frac{x_+(\sigma)-\theta_b}{\sigma}} > 0.$$

By Lemma A.5, $\frac{x_+\left(\sigma_h^\beta\right)-\theta_b}{\sigma_h^\beta} < z^*$, but type-$b$'s optimal σ satisfies $\frac{x_+(\sigma)-\theta_b}{\sigma} > z^*$ by Lemma A.6(iii), so σ_h^β cannot be a fixed point. The next step is to show that $\frac{x_+(\sigma)-\theta_b}{\sigma} < z^*$ for all $\sigma > \sigma_h^\beta$. Suppose not: then there is σ such that $\frac{x_+(\sigma)-\theta_b}{\sigma} \geq z^*$. That means that $x_+(\sigma)$ crosses $\sigma z^* + \theta_b$ from below at some point, which requires $x_+'(\sigma) \geq z^* > 0$ at such intersection point. This implies that $\frac{x_+(\sigma)-\theta_b}{\sigma} = z^*$, but by Lemma A.5 and inspection of (a.48), that means $x_+'(\sigma) = 0$, a contradiction.

Finally, consider $\beta = \beta_l$. In this case $\sigma_l^\beta = \sigma^{**} = \sigma_h^\beta$, so at $\sigma = \sigma^{**}$, $x_-(\sigma) = x_+(\sigma) = \sigma z^* + \theta_b$ and $x_-'(\sigma) = x_+'(\sigma) = 0$. Then, for $\sigma < \sigma^{**}$, it is clear that for σ close enough to σ^{**} we have $x_-(\sigma) > \theta_b + z^*\sigma$, and as we showed before this leads to the conclusion that $\theta_b + z^*\sigma < x_-(\sigma)$ for all $\sigma \in (\underline{\sigma}, \sigma_l^\beta)$. Similarly, for $\sigma > \sigma^{**}$ and close enough to σ^{**}, we have that $x_+(\sigma) < \theta_b + z^*\sigma$, which leads to $\theta_b + z^*\sigma > x_+(\sigma)$ for all $\sigma > \sigma_h^\beta$. ∎

We now complete the proof of Proposition 7. For parts (i) and (ii), consider first the case $\underline{\sigma} \geq \hat{\underline{\sigma}}(\beta)$. Lemma A.17 says a nontrivial equilibrium does not exist. Now let $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$. Bounded replacement equilibria are ruled out by Proposition 6(i). If $\underline{\sigma} \geq \underline{\sigma}(\beta)$, by Lemma A.16 a bounded retention equilibrium cannot exist, but by Proposition 6(ii) a unique monotone retention equilibrium does exist. If $\underline{\sigma} < \underline{\sigma}(\beta)$, by Proposition 6(ii) and Lemma A.15 there is a unique nontrivial equilibrium, and it has bounded retention. The equilibrium strategies are given by Lemmas A.7, A.8, and A.6(iii).

To prove Part (iii): when $\beta \leq 1$, we have $\underline{\sigma} < \hat{\underline{\sigma}}(\beta)$, and therefore if a nontrivial equilibrium exists, it must involve bounded retention by Proposition 6(ii). ■

5. Signal-Contingent Disclosure, Section 6.5

In Section 6.5 we observed that there is a potential multiplicity of equilibria with a pessimistic future ($q > p$) and low non-disclosure costs ($\kappa < 1$). To see this, consider an arbitrary pair of variances (σ_g, σ_b). Consider the principal's best responses to these in the benchmark model; call that retention set X. Likewise, consider the principal's best response to (σ_g, σ_b) in our extended scenario, which produces a possibly distinct retention set; call it R. Because all disclosed signal realizations are transmitted by both types, if the principal retains the agent after observing $\tilde{x}$, it must be that $\tilde{x} \in X$. So $R \subset X$. The multiplicity arises because R can be any possible subset of X: after all, the principal can threaten to replace the agent for some specific realization $\hat{x} \in X$ is observed, which induces both agent types to hide the signal after that realization. Consequently, $\hat{x} \in X$ is not observed in equilibrium and the principal's threat can be sustained in a sequential equilibrium. To close the equilibrium circle, we observe that σ_k is a best response for type-k, given R.

This multiplicity is sustained by the principal's threat to replace the agent if she sees some $\hat{x} \in X \setminus R$. A simple refinement eliminates such threats, and yields an equality between R and X. With that observation in hand, we will be able to pin down a unique equilibrium, which coincides with that in the benchmark setting.

Consider a sequential equilibrium and a signal value $\hat{x}$ that is not transmitted in that equilibrium. Suppose the principal asks herself which agent type could have conceivably benefited from transmitting that particular realization. We will say that a subset of types is a *beneficiary set* if — whenever the principal evaluates her beliefs using equilibrium play but restricted to that beneficiary set — she has a best response that benefits every type, and *only* those types, within the beneficiary set. We will say that an equilibrium is *robust* if for every unsent signal realization, there is no beneficiary set.

Fix an equilibrium with agent strategies (σ_g, σ_b), associated non-disclosure rules, and principal retention set R. Then we know that $R \subseteq X$. Now suppose that an unsent signal $\hat{x}$ is received. Then, if $\hat{x} \in X \setminus R$, $\{b, g\}$ is a beneficiary set. For under that presumption, and using the equilibrium strategies of the agent, it is optimal for the principal to retain, because $\hat{x} \in X$. Therefore both types must benefit from the deviation to sending $\hat{x}$. It follows that for this equilibrium to be robust, $X \setminus R = \emptyset$, so $R = X$.

If $\hat{x} \notin X$, we observe that there is no beneficiary set. For if the set consists of g alone, the principal will retain, but then b will also want to transmit the same signal. If the set consists of b alone, the principal will surely replace, but then b is worse off. Finally, if the set is $\{b, g\}$, the principal will *also* replace, given that $\hat{x} \notin X$. Therefore all the robustness conditions are satisfied for any equilibrium with $R = X$.

Finally, let's analyze the agent's best response under a robust equilibrium. The agent knows that if a signal realization falls within R, then it will be disclosed and the agent will be retained, receiving a payoff of 1. If the signal realization is not in R, the agent will hide it and will still be retained, but this time at a cost of κ, thus collecting a total of $1-\kappa$ in such case. Type-k's best response to R, σ_k, therefore solves

$$\max_{\sigma} 1-\kappa+\kappa \int_R \frac{1}{\sigma}\phi\left(\frac{x-\theta_k}{\sigma}\right)dx.$$

If $R = X$, we conclude that the robust equilibrium with noise choices (σ_g, σ_b, R) coincides with that of the benchmark model.

6. Examples

6.1. **Bounded Replacement Equilibria in the Costly Noise Model.**

We first describe necessary conditions for bounded replacement equilibria in the model of Section 6.2.

Proposition A.3. *Suppose a bounded replacement equilibrium exists. Then, either*

(i) $\underline{\sigma} > \sigma_g > \sigma_b$ *and* $\beta > 1$ *and large enough so that* $x_+ < x_- < \theta_b < \theta_g$*; or*

(ii) $\sigma_g > \sigma_b > \underline{\sigma}$ *and* $\beta < 1$ *and small enough so that* $x_+ < \theta_b < \theta_g < x_-$*.*

The proof of this Proposition is available on request from the authors. Guided by it, we now construct two examples of bounded replacement, one for $\beta < 1$ and another for $\beta > 1$. For $\beta < 1$, Proposition A.3 says that both σ_b and σ_g must exceed $\underline{\sigma}$, and that both types are inside the replacement zone, with the bad type deeper it. Then, the good type will pay a bigger cost of escaping the zone, whereas there is not much the bad type can do. Provided we construct the "right" marginal cost function, this yields $\sigma_g > \sigma_b$.

Take $\theta_b = 1$, $\theta_g = 2$ and $x_- = 2.3$. Impose $\sigma_g = 0.42$ and $\sigma_b = 0.250001$. Then, because

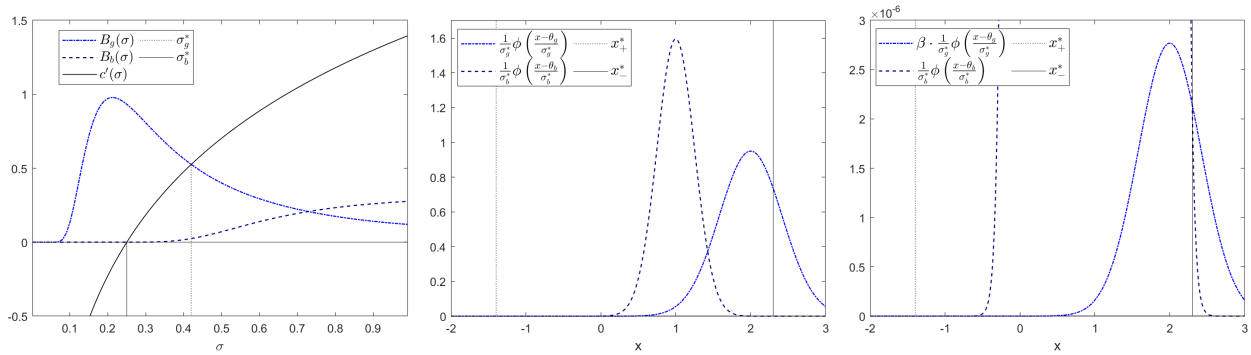
$$\frac{x_+ + x_-}{2} = \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2},$$

we have $x_+ = 2\frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} - x_- \approx -1.38$. The value of $\beta < 1$ is now also determined:

$$\beta = \frac{\frac{1}{\sigma_b}\phi\left(\frac{x_+-\theta_b}{\sigma_b}\right)}{\frac{1}{\sigma_g}\phi\left(\frac{x_+-\theta_g}{\sigma_g}\right)} = \frac{\frac{1}{\sigma_b}\phi\left(\frac{x_--\theta_b}{\sigma_b}\right)}{\frac{1}{\sigma_g}\phi\left(\frac{x_--\theta_g}{\sigma_g}\right)} \approx 2.92 \cdot 10^{-6}.$$

Finally, the two first-order conditions need to be satisfied. To this end, we choose

$$c'(\sigma) = A\ln(\sigma) + B.$$



(A) Marginal Benefits and Costs

(B) Equilibrium Likelihoods

(C) Adjusted Eq. Likelihoods

FIGURE A.4. A Bounded Replacement Equilibrium for β Small.

The cost function that yields this expression for the marginal cost is

$$c(\sigma) = A\left(\sigma \ln(\sigma) - \underline{\sigma} \ln(\underline{\sigma})\right) + (B - A)(\sigma - \underline{\sigma}).$$

We have two free parameters, for the two first-order conditions:

$$\begin{aligned} \phi\left(\frac{x_- - \theta_g}{\sigma_g}\right)\frac{x_- - \theta_g}{\sigma_g^2} - \phi\left(\frac{x_+ - \theta_g}{\sigma_g}\right)\frac{x_+ - \theta_g}{\sigma_g^2} &= A\ln(\sigma_g) + B, \\ \phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)\frac{x_- - \theta_b}{\sigma_b^2} - \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\frac{x_+ - \theta_b}{\sigma_b^2} &= A\ln(\sigma_b) + B. \end{aligned}$$

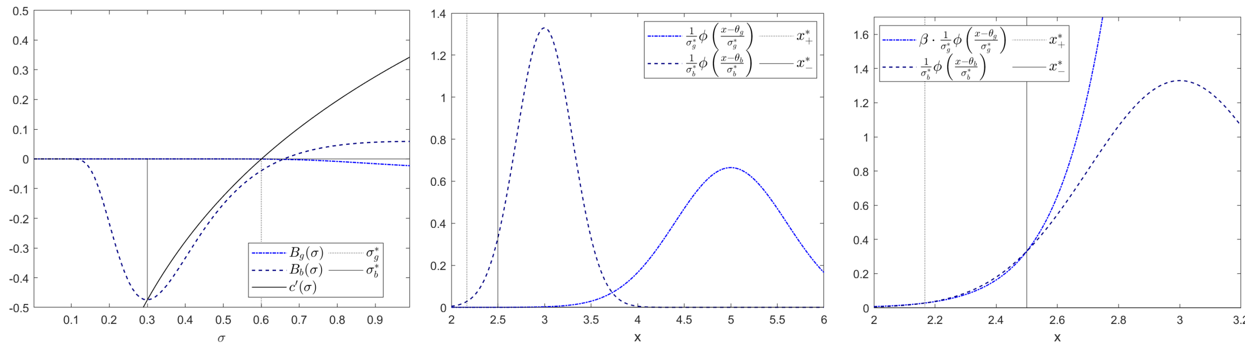
Therefore,

$$\begin{aligned} A &= \frac{\left(\phi\left(\frac{x_- - \theta_g}{\sigma_g}\right)\frac{x_- - \theta_g}{\sigma_g^2} - \phi\left(\frac{x_+ - \theta_g}{\sigma_g}\right)\frac{x_+ - \theta_g}{\sigma_g^2}\right) - \left(\phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)\frac{x_- - \theta_b}{\sigma_b^2} - \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\frac{x_+ - \theta_b}{\sigma_b^2}\right)}{\ln(\sigma_g) - \ln(\sigma_b)} \\ &\approx 1 \\ B &= \phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)\frac{x_- - \theta_b}{\sigma_b^2} - \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\frac{x_+ - \theta_b}{\sigma_b^2} - A\ln(\sigma_b) \approx 1.39 \end{aligned}$$

The resulting value of ambient noise is $\underline{\sigma} \approx \frac{1}{4}$. Figure A.4 depicts the equilibrium.

Now we find an example for the case $\beta > 1$. By Proposition A.3 it must be that $x_+ < x_- < \theta_b < \theta_g$. Both agents are in the retention zone, so $\sigma_b, \sigma_g < \underline{\sigma}$, but the bad type is closer to replacement, and so will make a bigger effort than the good type: $\sigma_b < \sigma_g < \underline{\sigma}$. Let us then choose $\theta_b = 3$, $\theta_g = 5$ and $x_- = 2.5$. For the choices of noise, let's take $\sigma_b = 0.3$ and $\sigma_g = 0.6$. All this again pins down x_+ at $x_+ = 2\frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} - x_- \approx 2.17$, while

$$\beta = \frac{\frac{1}{\sigma_b}\phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)}{\frac{1}{\sigma_g}\phi\left(\frac{x_+ - \theta_g}{\sigma_g}\right)} = \frac{\frac{1}{\sigma_b}\phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)}{\frac{1}{\sigma_g}\phi\left(\frac{x_- - \theta_g}{\sigma_g}\right)} \approx 2936.$$



(A) Marginal Benefits and Costs

(B) Equilibrium Likelihoods

(C) Adjusted Eq. Likelihoods

FIGURE A.5. A Bounded Replacement Equilibrium for β Large.

For the cost function, take: $c'(\sigma) = A \ln(\sigma) + B$ to implement the first-order conditions, so:

$$\begin{aligned} A &= \frac{\left(\phi\left(\frac{x_- - \theta_g}{\sigma_g}\right)\frac{x_- - \theta_g}{\sigma_g^2} - \phi\left(\frac{x_+ - \theta_g}{\sigma_g}\right)\frac{x_+ - \theta_g}{\sigma_g^2}\right) - \left(\phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)\frac{x_- - \theta_b}{\sigma_b^2} - \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\frac{x_+ - \theta_b}{\sigma_b^2}\right)}{\ln(\sigma_g) - \ln(\sigma_b)} \\ &\approx 0.68 \\ B &= \phi\left(\frac{x_- - \theta_b}{\sigma_b}\right)\frac{x_- - \theta_b}{\sigma_b^2} - \phi\left(\frac{x_+ - \theta_b}{\sigma_b}\right)\frac{x_+ - \theta_b}{\sigma_b^2} - A\ln(\sigma_b) \approx 0.35. \end{aligned}$$

Ambient noise is now $\underline{\sigma} \approx 0.6004$. Figure A.5 depicts the equilibrium.

7. MISCELLANEOUS DETAILS

7.1. **The Behavior of $\sigma(\theta)$ in the Costly Noise Model.** We describe the behavior of $\sigma(\theta)$ in the costly noise model of Section 6.2, when the agents face a retention rule $X = [x_-, x_+]$, including $x_+ = \infty$ (the monotone regime). To begin, the first-order condition of type θ is

$$\phi\left(\frac{x_- - \theta}{\sigma}\right)\frac{x_- - \theta}{\sigma^2} - \phi\left(\frac{x_+ - \theta}{\sigma}\right)\frac{x_+ - \theta}{\sigma^2} = c'(\sigma). \tag{a.57}$$

Differentiation of (a.57) shows that

$$\sigma'(\theta) = -\frac{1}{\sigma(\theta)^2}\frac{h(\theta)}{\frac{\partial}{\partial\sigma}\left[\phi\left(\frac{x_- - \theta}{\sigma}\right)\frac{x_- - \theta}{\sigma^2} - \phi\left(\frac{x_+ - \theta}{\sigma}\right)\frac{x_+ - \theta}{\sigma^2} - c'(\sigma)\right]|_{\sigma=\sigma(\theta)}}, \tag{a.58}$$

where

$$h(\theta) := \phi\left(\frac{x_- - \theta}{\sigma(\theta)}\right)\left(\left(\frac{x_- - \theta}{\sigma(\theta)}\right)^2 - 1\right) - \phi\left(\frac{x_+ - \theta}{\sigma(\theta)}\right)\left(\left(\frac{x_+ - \theta}{\sigma(\theta)}\right)^2 - 1\right). \tag{a.59}$$

The denominator is the second-order derivative, which is negative at the optimum. Therefore:

$$\text{Sign}\{\sigma'(\theta)\} = \text{Sign}\{h(\theta)\}.$$

For monotone retention (Panel A of Figure 3), the term with x_+ in $h(\theta)$ disappears, and so:

$$\text{Sign}\{\sigma'(\theta)\} = \text{Sign}\{|x_- - \theta| - \sigma(\theta)\}.$$

Remember that $\sigma(x_-) = \underline{\sigma} > 0$, so $\sigma(\theta)$ is decreasing as we enter the retention zone, and it will be so until $\sigma(\theta) = \theta - x_-$, at which stage the derivative is 0. From this point onwards $\sigma(\theta)$ is always increasing.[5] However, $\sigma(\theta)$ cannot grow unboundedly, since $\sigma(\theta) \in (\sigma_*, \sigma^*)$. Then, $\frac{x_- - \theta}{\sigma(\theta)} \to -\infty$ as $\theta \to \infty$. Given that $\phi(z)z \to 0$ as $|z| \to \infty$, $c'(\sigma(\theta)) \to 0$ by (a.57), which means that $\sigma(\theta)$ approaches $\underline{\sigma}$ from below as $\theta \to \infty$.

If we decrease θ from x_-, $\sigma(\theta)$ cannot always stay above $x_- - \theta$ because this would mean $\sigma(\theta) \to \infty$ as $\theta \to -\infty$, but this is inconsistent with optimality: the first-order condition would imply that $c'(\sigma(\theta)) \to 0$, which means $\sigma(\theta) \to \underline{\sigma}$. This means there is an intersection point at which $\sigma(\theta) = x_- - \theta$ and $\sigma(\theta)$ reaches its maximum. Then, $\sigma(\theta) \downarrow \underline{\sigma}$ as $\theta \to -\infty$.

Now consider bounded retention, depicted in Panel B of Figure 3. The symmetry of $\sigma(\theta)$ around the midpoint of the retention interval, $\bar{x} := \frac{x_- + x_+}{2}$, is evident from the first-order condition, so we study $\sigma(\theta)$ only for $\theta \geq \bar{x}$. Start with $\theta = \bar{x}$: evaluating h at $\theta = \bar{x}$ we have $h(\bar{x}) = \sigma'(\bar{x}) = 0$. Next, let $\theta \in [\bar{x}, x_+]$. The LHS of the first-order condition in (a.57) is negative, so $\sigma < \underline{\sigma}$ for every θ in this interval. We cannot pin down the sign of $\sigma'(\theta)$ for every θ in this range, but we can claim that $\sigma(x_+) > \sigma(\bar{x})$ and $\sigma'(x_+) > 0$. To prove this claim, first note that for any σ:

$$-\phi\left(\frac{x_+ - x_-}{\sigma}\right)\frac{x_+ - x_-}{\sigma^2} > -\phi\left(\frac{x_+ - x_-}{2\sigma}\right)\frac{x_+ - x_-}{\sigma^2}.$$

Since c' is increasing, we may use (a.57) to see that $\sigma(x_+) > \sigma(\bar{x})$. Next, by (a.59):

$$h(x_+) = \left(\phi\left(\frac{x_+ - x_-}{\sigma(x_+)}\right)\left(\frac{x_+ - x_-}{\sigma(x_+)}\right)^2 + \phi(0) - \phi\left(\frac{x_+ - x_-}{\sigma(x_+)}\right)\right) > 0,$$

so that by (a.58), we have $\sigma'(x_+) > 0$, as claimed. Together, $\sigma'(\bar{x}) = 0$, $\sigma'(x_+) > 0$, and $\sigma(x_+) > \sigma(\bar{x})$ suggest that $\sigma(\theta)$ is nondecreasing on $[\bar{x}, x_+]$. On the other hand, if the retention zone is large, the midpoint type could choose greater noise than a type closer to the edge, who needs more precision for his signal to stay in the zone. We claim that at $\theta = \bar{x}$, $\sigma(\theta)$ is concave (convex) for $x_+ - x_-$ sufficiently large (small). From (a.58) we have that

$$\sigma''(\bar{x}) = -\frac{1}{\sigma(\bar{x})^2}\frac{h'(\bar{x})}{\frac{\partial}{\partial\sigma}\left[\phi\left(\frac{x_- - \theta}{\sigma}\right)\frac{x_- - \theta}{\sigma^2} - \phi\left(\frac{x_+ - \theta}{\sigma}\right)\frac{x_+ - \theta}{\sigma^2} - c'(\sigma)\right]|_{\sigma=\sigma(\bar{x})}},$$

so, at $\theta = \bar{x}$, the sign of the second-order derivative of $\sigma(\theta)$ coincides with the sign of $h'(\theta)$. We compute h' from (a.59), and evaluating it at $\theta = \bar{x}$ we conclude that

$$\text{Sign}\{\sigma''(\bar{x})\} = \text{Sign}\left\{\sqrt{3} - \frac{x_+ - x_-}{2\sigma(\bar{x})}\right\}.$$

Because $\sigma(\theta) \in (\sigma_*, \sigma^*)$, so $\frac{x_+ - x_-}{2\sigma^*} < \frac{x_+ - x_-}{2\sigma(\bar{x})} < \frac{x_+ - x_-}{2\sigma_*}$. So if $x_+ - x_-$ is sufficiently large, $\sigma(\theta)$ is concave at $\theta = \bar{x}$, which eliminates the monotonicity of $\sigma(\theta)$ on all of $[\bar{x}, x_+]$.

[5] $\sigma(\theta)$ cannot cross the $\theta - x_-$ function again since this would require $\sigma'(\theta) \geq 1$ at the intersection point, but crossing $\theta - x_-$ means $\sigma'(\theta) = 0$.

Now consider $\theta > x_+$. We will show that there is only one type, $\underline{\theta}$, such that $\sigma(\underline{\theta}) = \underline{\sigma}$, all types in $(x_+, \underline{\theta})$ choose $\sigma(\theta) < \underline{\sigma}$, and all types above $\underline{\theta}$ choose $\sigma(\theta) > \underline{\sigma}$. Additionally, in this range above x_+, $\sigma(\theta)$ is first increasing and then decreasing, and it converges to $\underline{\sigma}$ as $\theta \to \infty$. First, to see how the optimal choice $\sigma(\theta)$ compares to the ambient noise $\underline{\sigma}$, consider type-θ's first-order payoff derivative with respect to noise at $\sigma = \underline{\sigma}$:

$$\phi\left(\frac{\theta - x_+}{\underline{\sigma}}\right)\frac{\theta - x_+}{\underline{\sigma}^2} - \phi\left(\frac{\theta - x_-}{\underline{\sigma}}\right)\frac{\theta - x_-}{\underline{\sigma}^2}. \tag{a.60}$$

A type chooses $\underline{\sigma}$ has this expression equal to zero. So, we study $\phi(z)z$ for $z > 0$, looking for z_1 $(= \frac{\theta - x_+}{\underline{\sigma}})$ and z_2 $(= \frac{\theta - x_-}{\underline{\sigma}})$ such that $z_1 < z_2$, $\phi(z_1) z_1 = \phi(z_2) z_2$, and $z_2 - z_1 = \frac{x_+ - x_-}{\underline{\sigma}}$. It's easy to see that there exists only one such pair of values (look at Figure A.6), and type $\underline{\theta} = z_2 \cdot \underline{\sigma} + x_-$ chooses $\sigma = \underline{\sigma}$. For types $\theta \in (x_+, \underline{\theta})$, their corresponding values of z_1 and z_2 are both smaller than the ones of type $\underline{\theta}$ so $\phi(z_1)z_1 < \phi(z_2)z_2$. That is, the sign of (a.60) is negative for these types, and they will choose $\sigma(\theta) < \underline{\sigma}$. Similarly, for those types above $\underline{\theta}$, the sign of (a.60) is positive, so they will choose $\sigma(\theta) > \underline{\sigma}$.

Notice that, since $\sigma(\theta) \in (\sigma_*, \sigma^*)$, $\frac{\theta - x_+}{\sigma(\theta)}$ and $\frac{\theta - x_-}{\sigma(\theta)}$ go to ∞ as $\theta \to \infty$. It must then be that $c'(\sigma(\theta)) \to 0$ as $\theta \to \infty$, by the first-order condition and the fact that $\phi(z)z \to 0$ as $z \to \infty$. This implies that $\sigma(\theta) \to \underline{\sigma}$ as $\theta \to \infty$.

Now consider any type $\theta^* > x_+$ such that $\sigma'(\theta^*) = 0$.[6] $\sigma'(\theta^*) = 0$ means $h(\theta^*) = 0$, so take a look at Figure A.6, which plots the function $\phi(z)(z^2 - 1)$, which relates to $h(\theta)$ (see (a.59)): $h(\theta^*) = 0$ means that there are two points on the x axis, z_1 and z_2, that reach the same height: $\phi(z_1)(z_1^2 - 1) = \phi(z_2)(z_2^2 - 1)$. Recall the smaller point, z_1, corresponds to $\frac{\theta^* - x_+}{\sigma(\theta^*)}$, and the larger point, z_2, to $\frac{\theta^* - x_-}{\sigma(\theta^*)}$. Since $\sigma'(\theta^*) = 0$, we have that both z_1 and z_2 are increasing in θ at θ^*. It is easy to see that this implies that, for $\theta > \theta^*$ and close to θ^*, $h(\theta) < 0$ and therefore $\sigma'(\theta) < 0$. Then, $\frac{\theta - x_+}{\sigma(\theta)}$ and $\frac{\theta - x_-}{\sigma(\theta)}$ will always be increasing forever after, so $\sigma'(\theta) < 0$ for all $\theta > \theta^*$. We can the conclude that, in the range $[x_+, \infty)$, $\sigma(\theta)$ is first increasing and then decreasing.

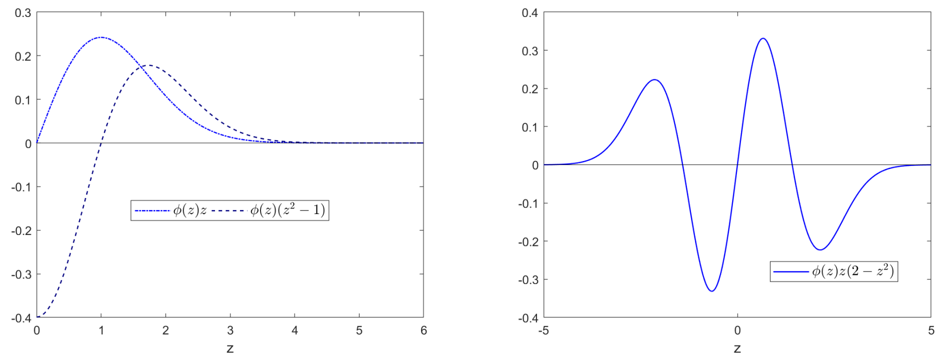
7.2. **Sufficient Conditions for Uniqueness.** We describe a condition on $c(\sigma)$ that implies Condition U in Section 6.2 . Consider the case $\sigma_b \geq \sigma_g$, so $X = [x_-, x_+]$ with $x_+ < \infty$ iff $\sigma_b > \sigma_g$. Recall the necessary first-order condition:

$$-\phi\left(\frac{x_+ - \theta}{\sigma}\right)\frac{x_+ - \theta}{\sigma^2} + \phi\left(\frac{x_- - \theta}{\sigma}\right)\frac{x_- - \theta}{\sigma^2} = c'(\sigma).$$

We want to impose conditions such that the objective function is always strictly concave, this generating an unique optimal choice for each parameter. For this, we will ask $c''(\sigma)$ to be always bigger than the second derivative of the marginal benefit, which is the derivative of the left-hand side with respect to σ :

$$\frac{1}{\sigma^2}\left[\phi\left(\frac{x_+ - \theta}{\sigma}\right)\frac{x_+ - \theta}{\sigma}\left(2 - \left(\frac{x_+ - \theta}{\sigma}\right)^2\right) - \phi\left(\frac{x_- - \theta}{\sigma}\right)\frac{x_- - \theta}{\sigma}\left(2 - \left(\frac{x_- - \theta}{\sigma}\right)^2\right)\right].$$

[6] By the previous results, at least one such type exists: $\sigma(\theta) < \underline{\sigma}$ for all $\theta \in [x_+, \underline{\theta})$; $\sigma(\theta) > \underline{\sigma}$ for all $\theta > \underline{\theta}$; and $\sigma(\theta) \to \underline{\sigma}$ as $\theta \to \infty$.



(A) $\phi(z)\,z$ and $\phi(z)\,(z^2-1)$ (B) $\phi(z)\,z\,(2-z^2)$

FIGURE A.6. Variations of the Function ϕ.

This expression is related to the function $\phi(z)\,z\,(2-z^2)$, where the value of z could be anywhere in the real line: $x_+ - \theta$ is always positive, but $x_- - \theta$ can take either sign. Forget about the term $\frac{1}{\sigma^2}$ on the left: we will find the biggest possible value of the term inside the square brackets, which will be a number, say κ. Then we ask for $c''(\sigma) \geq \frac{\kappa}{\sigma^2}$ for all σ. We plot $\phi(z)\,z\,(2-z^2)$ function in Panel B of Figure A.6 In order to find the critical values of this function, compute the first-order derivative and set it equal to zero:

$$\frac{\partial}{\partial z}\phi(z)\,z\left(2-z^2\right) = \phi(z)\left(z^4 - 5z^2 + 2\right) = 0.$$

We have 4 values of z that satisfy the condition:

$$z = \pm\sqrt{\frac{5}{2} \pm \sqrt{\frac{17}{4}}} \Rightarrow z = \{-2.14, -0.66, 0.66, 2.14\}.$$

Finally, to find the maximum value of $\phi(z_2)\,z_2\,(2-z_2^2) - \phi(z_1)\,z_1\,(2-z_1^2)$ with $z_2 > z_1$ and $z_2 > 0$, it is clear that we have to consider $z_2 = \sqrt{\frac{5}{2} - \sqrt{\frac{17}{4}}}$ and $z_1 = -\sqrt{\frac{5}{2} - \sqrt{\frac{17}{4}}}$. So we ask for

$$c''(\sigma) \geq \frac{\kappa}{\sigma^2}\forall\sigma$$

where

$$\begin{aligned}\kappa &= \phi(z_2)\,z_2\left(2-z_2^2\right) - \phi(z_1)\,z_1\left(2-z_1^2\right)\\ &\approx 0.662594.\end{aligned}$$

8. PROOF OF PROPOSITION A.1 IN SECTION 1.2

The inequality (a.5) implies that

$$\left(\sigma_b^2 - \sigma_g^2\right)x_1^2 - 2\left(\sigma_b^2\theta_g - \sigma_g^2\theta_b\right)x_1 \leq \left(\sigma_b^2 - \sigma_g^2\right)x_2^2 - 2\left(\sigma_b^2\theta_g - \sigma_g^2\theta_b\right)x_2. \quad \text{(a.61)}$$

To prove Proposition A.1, we fist eliminate $\sigma_b = \sigma_g = \sigma$ in equilibrium. If that were the case, then (a.61) reduces to $x_1 \geq x_2$: the principal retains the agent with the higher signal. In this case, it is easy to compute the retention probability for agent j for both realizations of types:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma_g}\phi\left(\frac{x_j-\theta_g}{\sigma_g}\right)\left(1-\Phi\left(\frac{x_j-\theta_b}{\sigma_b}\right)\right)dx_j = 1-\Phi\left(\frac{\theta_g-\theta_b}{\sqrt{\sigma_g^2+\sigma_b^2}}\right) \text{ if } k(i)=b,$$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma_b}\phi\left(\frac{x_j-\theta_b}{\sigma_b}\right)\left(1-\Phi\left(\frac{x_j-\theta_g}{\sigma_g}\right)\right)dx_j = \Phi\left(\frac{\theta_g-\theta_b}{\sqrt{\sigma_b^2+\sigma_g^2}}\right) \text{ if } k(i)=g,$$

where we have used the property that $\int_{-\infty}^{\infty}\phi(w)\,\Phi\left(\frac{w-a}{b}\right)dw = \Phi\left(\frac{-a}{\sqrt{1+b^2}}\right)$. But it is clear from these expressions that b will want to increase σ_b, whereas g will seek to lower σ_g — there will always be an agent who would deviate, and therefore there is no equilibrium in which both types choose the same noise. Also, as we will see (but it's already quite clear) there can be no monotonic equilibrium either, since the only way the principal will keep the agent with the higher signal is when both agents communicate with the same level of noise.

Next, we eliminate $\sigma_b < \sigma_g$. In this case, let $\hat{x}$ be the value of x that minimizes the likelihood ratio $\left[\frac{1}{\sigma_g}\phi\left(\frac{x-\theta_g}{\sigma_g}\right)\right] / \left[\frac{1}{\sigma_b}\phi\left(\frac{x-\theta_b}{\sigma_b}\right)\right]$. It is easy enough to verify that

$$\hat{x} = \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2-\sigma_g^2} < \theta_b,$$

and that (a.61) becomes $|x_1-\hat{x}| \geq |x_2-\hat{x}|$; that is, the principal retains the agent whose signal is further away from $\hat{x}$. So player i's retention probability, when his type is θ_i, is:

$$\Pi_i = \int_{-\infty}^{\hat{x}} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(1-\Phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)+\Phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)\right)dx_j$$
$$+\int_{\hat{x}}^{\infty} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(1-\Phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)+\Phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)\right)dx_j$$

We analyze the derivative of Π_i with respect to σ_i at $\hat{x} = \frac{\sigma_b^2\theta_g-\sigma_g^2\theta_b}{\sigma_b^2-\sigma_g^2}$, which is given by:

$$\sigma_i\frac{\partial \Pi_i}{\partial \sigma_i} = \int_{-\infty}^{\hat{x}} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(\frac{1}{\sigma_i}\phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)(2\hat{x}-x_j-\theta_i)-\frac{1}{\sigma_i}\phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)(x_j-\theta_i)\right)dx_j$$
$$+\int_{\hat{x}}^{\infty} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(\frac{1}{\sigma_i}\phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)(x_j-\theta_i)-\frac{1}{\sigma_i}\phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)(2\hat{x}-x_j-\theta_i)\right)dx_j$$
$$=\int_{-\infty}^{\hat{x}} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\frac{1}{\sigma_i}\phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)(2\hat{x}-x_j-\theta_i)\,dx_j$$
$$-\int_{-\infty}^{\hat{x}} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\frac{1}{\sigma_i}\phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)(x_j-\theta_i)\,dx_j$$
$$+\int_{\hat{x}}^{\infty} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\frac{1}{\sigma_i}\phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)(x_j-\theta_i)\,dx_j$$
$$-\int_{\hat{x}}^{\infty} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\frac{1}{\sigma_i}\phi\left(\frac{2\chi-x_j-\theta_i}{\sigma_i}\right)(2\hat{x}-x_j-\theta_i)\,dx_j.$$

Lengthy and tedious manipulation of this equation(details available on request) shows that

$$\begin{aligned}
\sigma_i \frac{\partial \Pi_i}{\partial \sigma_i} &= \frac{\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}} \phi\left(\frac{\theta_j+\theta_i-2\hat{x}}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)-1\right)\left(\frac{2\hat{x}-\theta_i-\theta_j}{\sigma_i^2+\sigma_j^2}\right) \\
&+ \frac{\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}} \phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)-1\right)\left(\frac{\theta_i-\theta_j}{\sigma_i^2+\sigma_j^2}\right) \\
&+ 2\frac{\sigma_j\sigma_i}{\sigma_i^2+\sigma_j^2}\phi\left(\frac{\theta_j+\theta_i-2\hat{x}}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right) \\
&+ 2\frac{\sigma_j\sigma_i}{\sigma_i^2+\sigma_j^2}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right). \qquad \text{(a.62)}
\end{aligned}$$

Now notice that

$$\phi\left(\frac{\theta_j+\theta_i-2\hat{x}}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right) = \phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right),$$

so that

$$\begin{aligned}
\sigma_i \frac{\partial \Pi_i}{\partial \sigma_i} &= \frac{\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}} \phi\left(\frac{\theta_j+\theta_i-2\hat{x}}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)-1\right)\left(\frac{2\hat{x}-\theta_i-\theta_j}{\sigma_i^2+\sigma_j^2}\right) \\
&+ \frac{\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}} \phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)-1\right)\left(\frac{\theta_i-\theta_j}{\sigma_i^2+\sigma_j^2}\right) \\
&+ 4\frac{\sigma_j\sigma_i}{\sigma_i^2+\sigma_j^2}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)
\end{aligned}$$

Evaluated at $\hat{x} = \frac{\sigma_b^2\theta_g-\sigma_g^2\theta_b}{\sigma_b^2-\sigma_g^2}$, we obtain

$$\begin{aligned}
\sigma_i \frac{\partial \Pi_i}{\partial \sigma_i} &= \frac{\sigma_i^2}{\sqrt{\sigma_i^2+\sigma_j^2}} \phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)-1\right)\left(\frac{\theta_i-\theta_j}{\sigma_i^2+\sigma_j^2}\right) \\
&+ 4\frac{\sigma_j\sigma_i}{\sigma_i^2+\sigma_j^2}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right)
\end{aligned}$$

Because $\hat{x} < \theta_b < \frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}$, we have that $\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right) < \frac{1}{2}$. Therefore $\partial\Pi_b/\partial\sigma_b > 0$, whereas the sign of $\partial\Pi_g/\partial\sigma_g$ is ambiguous. The fact that $\partial\Pi_b/\partial\sigma_b > 0$ indicates that the bad type wants to deviate, and the equilibrium falls apart.

We are therefore left with only one possibility, $\sigma_b > \sigma_g$, where the principal retains 1 if and only if

$$|x_1 - \hat{x}| \leq |x_2 - \hat{x}|\,;$$

that is, for the signal closer to $\hat{x} := \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2-\sigma_g^2} > \theta_g$. ■

9. Proof of Proposition A.2 in Section 1.3

Suppose each type θ chooses some noise $\sigma(\theta)$. Then signal emitted by type θ has density $\pi_\theta(x) = \frac{1}{\sigma(\theta)}\phi\left(\frac{x-\theta}{\sigma(\theta)}\right)$. Let $U(x)$ be the expected payoff to the principal when the signal x is received. This is just the expected value of $u(\theta)$ weighted by the posterior distribution of θ using Bayes' Rule and the strategies, as described above. So

$$U(x) \equiv \frac{1}{\int \pi_\theta(x)q(\theta)d\theta}\int_{-\infty}^{\infty} u(\theta)\frac{1}{\sigma(\theta)}\phi\left(\frac{x-\theta}{\sigma(\theta)}\right)q(\theta)d\theta. \tag{a.63}$$

Lemma A.18. *Suppose that $\sigma(\theta)$ is continuous in θ and has a unique maximum at θ^*. Then $U(x)$ converges to $u(\theta^*)$ as $|x| \to \infty$.*

Proof. Pick any sequence x_n such that $x_n \to \infty$ (the argument for $x_n \to -\infty$ will be identical). Define a corresponding sequence of density functions on $\mathbb{R}$, h_n, by

$$h_n(\theta) = \frac{1}{\int \pi_t(x_n)q(t)dt}\frac{q(\theta)}{\sigma(\theta)}\phi\left(\frac{x_n-\theta}{\sigma(\theta)}\right),$$

and let $H_n(\theta) = \int_{-\infty}^{\theta} h_n(s)ds$ be the corresponding cdfs. We claim that this sequence converges weakly to the degenerate probability measure placing probability 1 on θ^*.

To prove the claim, first pick any $\theta < \theta^*$. Let σ_1 be the maximum value of $\sigma(s)$ for $s \leq \theta$. Because $\sigma(\theta)$ is uniquely maximized at θ^* and $\theta^* > \theta$, there exists an interval of length ϵ around θ^* such that $\min \sigma(s)$ for s in that interval — call it σ_2 — strictly exceeds σ_1. Denote by $Q(\theta)$ the prior mass of types up to θ, and by Δ_Q the prior mass in the ϵ-interval around θ^*. With these values fixed, observe that for n large enough so that $x_n > \theta$,

$$H_n(\theta) = \frac{\int_{-\infty}^{\theta} \frac{q(s)}{\sigma(s)} \exp\left\{-\frac{1}{2}\left[\frac{x_n - s}{\sigma(s)}\right]^2\right\} ds}{\int_{-\infty}^{\infty} \frac{q(t)}{\sigma(t)} \exp\left\{-\frac{1}{2}\left[\frac{x_n - t}{\sigma(t)}\right]^2\right\} dt}$$
$$\leq \frac{\frac{Q(\theta)}{\sigma_*} \exp\left\{-\frac{1}{2}\left[\frac{x_n - \theta}{\sigma_1}\right]^2\right\}}{\frac{\Delta_Q}{\sigma^*} \exp\left\{-\frac{1}{2}\left[\frac{x_n - (\theta^* - \epsilon)}{\sigma_2}\right]^2\right\}} \to 0$$

as $n \to \infty$,[7] where the very last conclusion uses $\sigma_1 < \sigma_2$. By applying the same logic to the "other side" of θ^*, we also conclude that $1 - H_n(\theta) \to 0$ for each $\theta > \theta^*$. It follows that $H_n(\theta) \to 1$ for each $\theta > \theta^*$, so H_n converges to the degenerate cdf placing all weight on θ^*. Because $u(\theta)$ is a bounded, continuous function, it follows that

$$U(x_n) = \int_{-\infty}^{\infty} u(\theta) h_n(\theta) d\theta \to u(\theta^*).$$

■

Lemma A.19. *Assume Condition* U. *Consider any monotone retention threshold x^*. Then any optimal choice function by an agent of type θ only depends on the difference $t \equiv x^* - \theta$ and on that agent's payoffs; in particular, it does not depend on the type distribution $q(\theta)$. Call this function $s(t)$. It is continuous. If the retention zone is $[x^*, \infty)$, then $s(t)$ attains a unique maximum at some $t_1 > 0$. If the retention zone is $(-\infty, x^*]$, then $s(t)$ attains a unique maximum at some $t_2 < 0$.*

Proof. An agent of type θ chooses σ to maximize

$$1 - \Phi\left(\frac{x^* - \theta}{\sigma}\right) - c(\sigma)$$

if the retention zone is $[x^*, \infty)$, and

$$\Phi\left(\frac{x^* - \theta}{\sigma}\right) - c(\sigma)$$

if the retention zone is $(-\infty, x^*]$. Just these expressions make it clear that the solution σ can only depend on $t = x^* - \theta$. By Condition U, the solution is unique and therefore easily seen to be continuous. The first order condition with retention zone $[x^*, \infty)$ is given by

$$\phi\left(\frac{x^* - \theta}{\sigma}\right)\frac{x^* - \theta}{\sigma^2} - c'(\sigma) = 0. \tag{a.64}$$

By Condition U, (a.64) is necessary and sufficient for a maximum. When $x^* > \theta$, the corresponding value of σ exceeds $\underline{\sigma}$, and using the fact that $\sigma c'(\sigma)$ is increasing when $\sigma \geq \underline{\sigma}$, we see that the

[7]The values σ_* and σ^* are the lowest and highest values that noise could optimally have; see main text.

maximum possible value of σ satisfying (a.64) is achieved when

$$\sigma c'(\sigma) = \phi\left(\frac{x^* - \theta}{\sigma}\right)\frac{x^* - \theta}{\sigma} = \phi(z^*)z^*,$$

where z^* is the value that maximizes $\phi(z)z$. That is, define σ_1 by the first and last terms in the equality above and then set $x^* - \theta = t_1 = \sigma_1 z^*$ to define t_1. When the retention zone is $(-\infty, x^*]$, the first order condition is given by

$$-\phi\left(\frac{x^* - \theta}{\sigma}\right)\frac{x^* - \theta}{\sigma^2} - c'(\sigma) = 0. \tag{a.65}$$

Now the corresponding value of σ exceeds $\underline{\sigma}$ when $x^* < \theta$. By a parallel argument to the one just made, the maximum possible value of σ satisfying (a.64) is achieved when

$$\sigma c'(\sigma) = -\phi\left(\frac{x^* - \theta}{\sigma}\right)\frac{x^* - \theta}{\sigma} = -\phi(z_*)z_*,$$

where z_* is the value that minimizes $\phi(z)z$ (z_* will be negative). Define σ_2 by the first and last terms in the equality above and then set $x^* - \theta = t_2 = \sigma_2 z_*$ to define t^*. ■

Lemma A.20. *Let t^* stand for t_1 or t_2 as defined in Lemma A.19. Then $u(x^* - t^*) = V$.*

Proof. We consider the retention zone $[x^*, \infty)$ where $t^* = t_1$; the other case is dealt with in identical fashion. By Lemmas A.18 and A.19, $U(x)$ converges to $u(x^* - t_1)$ as $|x| \to \infty$. Suppose that $u(x^* - t_1) > V$. Then for x negative and large in absolute value — in particular for some $x < x^*$ — we would have $U(x) > V$, so that the principal must retain for such values. That contradicts monotone retention. Similarly, if $u(x^* - t_1) < V$, then for x large — in particular for some $x > x^*$ — we would have $U(x) < V$, so that the principal must replace for such values. Once again, that contradicts monotone retention. We are therefore left with just one possibility: $u(x^* - t_1) = V$. ■

Lemma A.21. $U(x^*) = V$.

Proof. By monotone retention, $U(x^* - \epsilon) \leq V \leq U(x^* + \epsilon)$ (or $U(x^* - \epsilon) \geq V \geq U(x^* + \epsilon)$). U is obviously continuous, so the result follows. ■

Lemma A.21 combined with (a.63) tells us that

$$\frac{1}{\int \pi_\theta(x^*)q(\theta)d\theta}\int_{-\infty}^{\infty} u(\theta)\frac{1}{s(x^* - \theta)}\phi\left(\frac{x^* - \theta}{s(x^* - \theta)}\right)q(\theta)d\theta = V,$$

where $s(t)$ is the optimal noise choice function as defined in Lemma A.19. Using the formula for $\pi_\theta(x)$ and transposing terms, we have

$$\int_{-\infty}^{\infty} \frac{u(\theta) - V}{s(x^* - \theta)}\phi\left(\frac{x^* - \theta}{s(x^* - \theta)}\right)q(\theta)d\theta = 0.$$

Lemma A.20 pins down x^* uniquely:

$$x^* = u^{-1}(V) + t^*,$$

so that combining these two inequalities, we conclude that

$$\int_{-\infty}^{\infty} h(\theta)q(\theta)d\theta = 0, \tag{a.66}$$

where

$$h(\theta) = \frac{u(\theta) - V}{s(u^{-1}(V) + t^* - \theta)}\phi\left(\frac{u^{-1}(V) + t^* - \theta}{s(u^{-1}(V) + t^* - \theta)}\right)$$

is a function that depends on model parameters but is entirely independent of the particular density $\{q(\theta)\}$; see Lemma A.19. Let $\mathcal{Q}$ be the set of all densities on $\mathbb{R}$ equipped with the topology induced by the sup norm, and let $\mathcal{Q}^0$ be the subset of densities in $\mathcal{Q}$ that satisfy (a.66). It is obvious that $\mathcal{Q} - \mathcal{Q}^0$ is open and dense in $\mathcal{Q}$. ■

REFERENCES

BHASKAR, D. (2017): "The Value of Monitoring in Dynamic Screening," mimeo, Department of Economics, New York University.

HVIDE, H. (2002): "Tournament Rewards and Risk Taking," *Journal of Labor Economics*, 20, 877–898.

KOSTADINOV, R., AND A. KUVALEKAR (2020): "Learning in Relational Contracts," *American Economic Journal: Microeconomics*, Forthcoming.

KUVALEKAR, A., AND E. LIPNOWSKI (2020): "Job Insecurity," *American Economic Journal: Microeconomics*, 12, 188–229.