Noisy Agents: Online Appendix

FRANCISCO ESPINOSA AND DEBRAJ RAY

ABSTRACT. In this not-for-publication Online Appendix we provide proofs of the results in Section 7 of the main text. We also provide numerical examples of bounded replacement equilibria and monotone equilibria, together with some missing details, such as the shape of the optimal choice of noise for a generic type $\theta \in \mathbb{R}$, depicted in Figure 5 of the paper.

1. PROOFS OF PROPOSITIONS IN SECTION 7

Proof of Proposition 7. Suppose that $\sigma_b > \sigma_g$. Define $h(x) \equiv f\left(\frac{x-\theta_b}{\sigma_b}\right)/f\left(\frac{x-\theta_g}{\sigma_g}\right)$ for all x. The retention zone is given by

$$X = \{x | h(x) \leq k \equiv \sigma_b \beta / \sigma_g\}.$$

Our claim is that X is a bounded interval. If this is not true, then — using $\sigma_b > \sigma_g$ and Lemma 6(ii) — there exist (a) $y \in \mathbb{R}$ such that $h(y) = k$ and $h'(y) < 0$, and (b) $w < y$ such that $h(w) = k$ and $h'(w) > 0$. We record here that:

$$h'(x) = \frac{\frac{1}{\sigma_b} f\left(\frac{x-\theta_g}{\sigma_g}\right) f'\left(\frac{x-\theta_b}{\sigma_b}\right) - \frac{1}{\sigma_g} f\left(\frac{x-\theta_b}{\sigma_b}\right) f'\left(\frac{x-\theta_g}{\sigma_g}\right)}{f\left(\frac{x-\theta_g}{\sigma_g}\right)^2}$$

$$= \frac{\frac{1}{\sigma_b}\left[f'\left(\frac{x-\theta_b}{\sigma_b}\right)\right] / \left[f\left(\frac{x-\theta_b}{\sigma_b}\right)\right] - \frac{1}{\sigma_g}\left[f'\left(\frac{x-\theta_g}{\sigma_g}\right)\right] / \left[f\left(\frac{x-\theta_g}{\sigma_g}\right)\right]}{f\left(\frac{x-\theta_g}{\sigma_g}\right) / f\left(\frac{x-\theta_b}{\sigma_b}\right)} \quad \text{(a.1)}$$

We divide the rest of the proof into two cases.

Case 1. $f'\left(\frac{y-\theta_g}{\sigma_g}\right) \leq 0$. Because $h'(y) < 0$, it follows from (a.1) that $f'\left(\frac{y-\theta_b}{\sigma_b}\right) < 0$. Because $h(y) = k > 1$, it follows that $(y-\theta_b)/\sigma_b < (y-\theta_g)/\sigma_g$. So, by MLRP, we must conclude that

$$\frac{f'\left(\frac{y-\theta_b}{\sigma_b}\right)}{f\left(\frac{y-\theta_b}{\sigma_b}\right)} > \frac{f'\left(\frac{y-\theta_g}{\sigma_g}\right)}{f\left(\frac{y-\theta_g}{\sigma_g}\right)}.$$

April 2018. Espinosa: New York University. Ray: New York University and University of Warwick. We thank Dilip Abreu, Dhruva Bhaskar and Gaute Torvik for useful comments. Ray's research was funded by National Science Foundation grant SES-1261560.

Because these objects are negative, and $\sigma_b > \sigma_g$, we must conclude that

$$\frac{\sigma_g}{\sigma_b}\frac{f'\left(\frac{y-\theta_b}{\sigma_b}\right)}{f\left(\frac{y-\theta_b}{\sigma_b}\right)} > \frac{f'\left(\frac{y-\theta_g}{\sigma_g}\right)}{f\left(\frac{y-\theta_g}{\sigma_g}\right)},$$

but invoking (a.1), this contradicts our presumption that $h'(y) < 0$.

Case 2. $f'\left(\frac{y-\theta_g}{\sigma_g}\right) > 0$. Then, because $w < y$, it must also be the case that $f'\left(\frac{w-\theta_g}{\sigma_g}\right) > 0$. Because $h'(w) > 0$, it follows from (a.1) that $f'\left(\frac{w-\theta_b}{\sigma_b}\right) > 0$. Combining this information with the fact that $h(w) = k > 1$, we must conclude that $(w-\theta_b)/\sigma_b > (w-\theta_g)/\sigma_g$. But then, by MLRP, we have

$$\frac{f'\left(\frac{w-\theta_b}{\sigma_b}\right)}{f\left(\frac{w-\theta_b}{\sigma_b}\right)} < \frac{f'\left(\frac{yw\theta_g}{\sigma_g}\right)}{f\left(\frac{w-\theta_g}{\sigma_g}\right)},$$

and because both these terms are positive, it follows that

$$\frac{\sigma_g}{\sigma_b}\frac{f'\left(\frac{w-\theta_b}{\sigma_b}\right)}{f\left(\frac{w-\theta_b}{\sigma_b}\right)} < \frac{f'\left(\frac{yw\theta_g}{\sigma_g}\right)}{f\left(\frac{w-\theta_g}{\sigma_g}\right)},$$

but invoking (a.1), this contradicts our presumption that $h'(w) > 0$. ■

Proof of Proposition 8. Consider a situation in which 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

$$U(x) \equiv \frac{1}{\int \pi_\theta(x)q(\theta)d\theta}\int_{-\infty}^{\infty} u(\theta)\frac{1}{\sigma(\theta)}\phi\left(\frac{x-\theta}{\sigma(\theta)}\right)q(\theta)d\theta. \tag{a.2}$$

Lemma 1. *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 sequence of cdfs. We claim that this sequence of probability measures 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 $x_n > \theta$,

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

as $n \to \infty$,[1] where the very last conclusion uses $\sigma_1 < \sigma_2$.

By symmetrically applying the same logic to the "other side" of θ^*, we must also conclude that $1 - H_n(\theta) \to 0$ for each $\theta > \theta^*$. It follows that $H_n(\theta) \to 1$ for each $\theta > \theta^*$. That completes the proof of convergence to the degenerate cdf placing all weight on θ^*.

By a standard characterization of weak convergence, and using the fact that $u(\theta)$ 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 2. *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

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

[1]The values σ_* and σ^* are the lowest and highest values that noise could optimally have, as per the discussion in the main text.

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

$$\phi\left(\frac{x^*-\theta}{\sigma}\right)\frac{x^*-\theta}{\sigma^2} - c'(\sigma) = 0. \tag{a.3}$$

By Condition U, (a.3) 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 maximum possible value of σ satisfying (a.3) 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

$$-\phi\left(\frac{x^*-\theta}{\sigma}\right)\frac{x^*-\theta}{\sigma^2} - c'(\sigma) = 0. \tag{a.4}$$

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.3) 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 3. *Let t^* stand for t_1 or t_2 as defined in Lemma 2. 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 1 and 2, $U(x)$ 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 4. $U(x^*) = V$.

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 4 combined with (a.2) 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 2. 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 3 pins down x^* uniquely:

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

so that combining these two inequalities, we conclude that

$$\int_{-\infty}^{\infty} h(\theta)q(\theta)d\theta = 0, \tag{a.5}$$

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 2. 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.5). It is obvious that $\mathcal{Q}-\mathcal{Q}^0$ is open and dense in $\mathcal{Q}$. ■

Section 7.3. Agents 1 and 2 simultaneously signal their types in a one-shot game:

$$x_i = \theta_{k(i)} + \sigma_{k(i)}\varepsilon_i,$$

where $i = 1, 2$, $k(i)$ denotes $i's$ type, and ε_1 and ε_2 are i.i.d. standard normal. When the principal observes a pair (x_1, x_1), her posterior probabilities are:

$$\Pr\left(k(1)=g|\left(x_1,x_2\right)\right) = \frac{\text{Density}\left(\left(x_1,x_2\right)|k(1)=g\right)}{\text{Density}\left(x_1,x_2\right)} = \frac{\text{Density}\left(\left(x_1,x_2\right)|k(1)=g, k(2)=b\right)}{\text{Density}\left(x_1,x_2\right)}.$$

So our principal will prefer agent 1 over agent 2 if (and modulo indifference, only if)

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

After some manipulation, the above inequality yields:

$$\left(\sigma_b^2-\sigma_g^2\right)x_1^2 - 2\left(\sigma_b^2\theta_g-\sigma_g^2\theta_b\right)x_1 \leq \left(\sigma_b^2-\sigma_g^2\right)x_2^2 - 2\left(\sigma_b^2\theta_g-\sigma_g^2\theta_b\right)x_2. \tag{a.6}$$

Proof of Proposition 9. We begin by ruling out the possibility that $\sigma_b = \sigma_g = \sigma$ in equilibrium. If that were the case, then (a.6) reduces to

$$x_1 \geq x_2;$$

that is, 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:

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

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

where we have used the property that $\int_{-\infty}^{\infty}\phi\left(w\right)\Phi\left(\frac{w-a}{b}\right)dw = \Phi\left(\frac{-a}{\sqrt{1+b^2}}\right)$. 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 soon see (but it is 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 the possibility that $\sigma_b<\sigma_g$. 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.6) becomes

$$\left|x_1-\hat{x}\right|\geq\left|x_2-\hat{x}\right|;$$

that is, the principal retains the agent whose signal is further away from $\hat{x}$. With these in hand, player i's retention probability, when his type is θ_i, is given by:

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

We want to evaluate 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:

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

The above equation has four terms on the right-hand side, which we will need to manipulate separately. The following expressions involving the normal density will be used repeatedly:

Lemma 5. *The normal density ϕ satisfies:*

$$\int_a^b \frac{w}{\sigma}\phi\left(\frac{w-\mu}{\sigma}\right)dw = \mu\left[\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)\right]-\sigma\left[\phi\left(\frac{b-\mu}{\sigma}\right)-\phi\left(\frac{a-\mu}{\sigma}\right)\right] \tag{a.7}$$

and

$$\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\omega}\phi\left(\frac{x-\tau}{\omega}\right) = \frac{1}{\frac{\sigma\omega}{\sqrt{\sigma^2+\omega^2}}}\phi\left(\frac{x-\frac{\omega^2\mu+\sigma^2\tau}{\sigma^2+\omega^2}}{\frac{\sigma\omega}{\sqrt{\sigma^2+\omega^2}}}\right)\frac{1}{\sqrt{\sigma^2+\omega^2}}\phi\left(\frac{\mu-\tau}{\sqrt{\sigma^2+\omega^2}}\right). \tag{a.8}$$

Proof. Equation (a.8) just need some standard algebra to be proven. As for (a.7), suppose that $w \sim N\left(\mu,\sigma^2\right)$. Then, integrating by parts, it is easy to see that:

$$\mathbb{E}\left(w|w\in[a,b]\right) = \frac{\int_a^b w\frac{1}{\sigma}\phi\left(\frac{w-\mu}{\sigma}\right)dw}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)} = \frac{b\Phi\left(\frac{b-\mu}{\sigma}\right)-a\Phi\left(\frac{a-\mu}{\sigma}\right)-\int_a^b\Phi\left(\frac{w-\mu}{\sigma}\right)dw}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}. \tag{a.9}$$

Observe that

$$\Phi\left(\frac{w-\mu}{\sigma}\right) = \frac{\partial}{\partial w}\left((w-\mu)\Phi\left(\frac{w-\mu}{\sigma}\right)+\sigma\phi\left(\frac{w-\mu}{\sigma}\right)\right).$$

Using this information in (a.9), we must conclude that

$$\begin{aligned}
\mathbb{E}\left(w\in[a,b]\right) &= \frac{b\Phi\left(\frac{b-\mu}{\sigma}\right)-a\Phi\left(\frac{a-\mu}{\sigma}\right)-\int_a^b\frac{\partial}{\partial w}\left((w-\mu)\Phi\left(\frac{w-\mu}{\sigma}\right)+\sigma\phi\left(\frac{w-\mu}{\sigma}\right)\right)dw}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\\
&= \frac{b\Phi\left(\frac{b-\mu}{\sigma}\right)-a\Phi\left(\frac{a-\mu}{\sigma}\right)-\left[\left((b-\mu)\Phi\left(\frac{b-\mu}{\sigma}\right)+\sigma\phi\left(\frac{b-\mu}{\sigma}\right)\right]-\left[(a-\mu)\Phi\left(\frac{a-\mu}{\sigma}\right)+\sigma\phi\left(\frac{a-\mu}{\sigma}\right)\right)\right]}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\\
&= \frac{\mu\Phi\left(\frac{b-\mu}{\sigma}\right)-\mu\Phi\left(\frac{a-\mu}{\sigma}\right)-\sigma\phi\left(\frac{b-\mu}{\sigma}\right)+\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\\
&= \mu-\sigma\frac{\phi\left(\frac{b-\mu}{\sigma}\right)-\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}.
\end{aligned}$$

With this expression in hand, we see that

$$\begin{aligned}
\int_a^b w\frac{1}{\sigma}\phi\left(\frac{w-\mu}{\sigma}\right)dw &= \mathbb{E}\left[w|w\in[a,b]\right]\cdot\left(\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)\right)\\
&= \mu\left(\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)\right)-\sigma\left(\phi\left(\frac{b-\mu}{\sigma}\right)-\phi\left(\frac{a-\mu}{\sigma}\right)\right)
\end{aligned}$$

■

We will use (a.7) and (a.8) repeatedly below to analyze the four terms in the expression for $\sigma_i\frac{\partial\Pi}{\partial\sigma_i}$. The first of these terms is given by:

$$
\begin{aligned}
&\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 \\
&\quad = (2\hat{x}-\theta_i)\int_{-\infty}^{\chi} \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)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)x_j dx_j \\
&\quad = (2\hat{x}-\theta_i)\frac{1}{\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)\int_{-\infty}^{\hat{x}} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)dx_j \\
&\quad -\frac{1}{\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)\int_{-\infty}^{\hat{x}} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)x_j dx_j \\
&\quad = (2\hat{x}-\theta_i)\frac{1}{\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)\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) \\
&\quad -\frac{1}{\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[\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^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)-\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^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)\right]
\end{aligned}
$$

In similar vein, and again using (a.8), the second term is given by:

$$
\begin{aligned}
&-\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_{-\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 dx_j + \theta_i\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)dx_j \\
&\quad = -\frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\int_{-\infty}^{\hat{x}} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)x_j dx_j + \theta_i\frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\int_{-\infty}^{\hat{x}} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)dx_j \\
&\quad = -\frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left[\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^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)-\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^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)\right] \\
&\quad +\theta_i\frac{1}{\sqrt{\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}
$$

The third term translates to:

$$
\begin{aligned}
&\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{x_j-\theta_i}{\sigma_i}\right)x_j dx_j - \theta_i\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)dx_j \\
&\quad = \frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\int_{\chi}^{\infty} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)x_j dx_j - \theta_i\frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\int_{\hat{x}}^{\infty} \frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)dx_j \\
&\quad = \frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left[\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}\left\{1-\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)\right\}+\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^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)\right] \\
&\quad -\theta_i\frac{1}{\sqrt{\sigma_i^2+\sigma_j^2}}\phi\left(\frac{\theta_j-\theta_i}{\sqrt{\sigma_i^2+\sigma_j^2}}\right)\left[1-\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)\right],
\end{aligned}
$$

and finally, the fourth term is given by

$$
\begin{aligned}
&-\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\hat{x}-x_j-\theta_i}{\sigma_i}\right)(2\hat{x}-x_j-\theta_i)\,dx_j \\
&=-(2\hat{x}-\theta_i)\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\hat{x}-x_j-\theta_i}{\sigma_i}\right)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\hat{x}-x_j-\theta_i}{\sigma_i}\right)x_j dx_j \\
&=-\frac{(2\hat{x}-\theta_i)}{\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)\int_{\hat{x}}^{\infty}\frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)dx_j+\frac{1}{\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)\int_{\hat{x}}^{\infty}\frac{1}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\phi\left(\frac{x_j-\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)x_j dx_j \\
&=-\frac{(2\hat{x}-\theta_i)}{\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[1-\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)\right] \\
&+\frac{1}{\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[\frac{\sigma_i^2\theta_j+\sigma_j^2(2\hat{x}-\theta_i)}{\sigma_i^2+\sigma_j^2}\left\{1-\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)\right\}+\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^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)\right]
\end{aligned}
$$

Summing all these terms up, we can conclude that:

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

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

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

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

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

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. However, notice that

$$\frac{\partial \Pi_b}{\partial \sigma_b}\sigma_b - \frac{\partial \Pi_g}{\partial \sigma_g}\sigma_g = -\frac{\sigma_b^2+\sigma_b^2}{\sqrt{\sigma_g^2+\sigma_b^2}}\phi\left(\frac{\theta_g-\theta_b}{\sqrt{\sigma_g^2+\sigma_b^2}}\right)\left[2\Phi\left(\frac{\hat{x}-\frac{\sigma_b^2\theta_g+\sigma_g^2\theta_b}{\sigma_g^2+\sigma_b^2}}{\frac{\sigma_b\sigma_g}{\sqrt{\sigma_g^2+\sigma_b^2}}}\right)-1\right]\left(\frac{\theta_g-\theta_b}{\sigma_g^2+\sigma_b^2}\right) > 0,$$

so that, because $\partial\Pi_k/\partial\sigma_k = c'(\sigma_k)$ for $k = g, b$, we have

$$c'(\sigma_b)\,\sigma_b > c'(\sigma_g)\,\sigma_g,$$

and moreover, $\sigma_b > \underline{\sigma}$ (because $\partial\Pi_b/\partial\sigma_b > 0$). Therefore the above inequality implies that $\sigma_b > \sigma_g$, a contradiction.

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

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

so the principal retains the agent whose signal is now closer to $\hat{x} := \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} > \theta_g$, which now maximizes the likelihood ratio $\left[\frac{1}{\sigma_g}\phi\left(\frac{x-\theta_g}{\sigma_g}\right)\right] \left[\frac{1}{\sigma_b}\phi\left(\frac{x-\theta_b}{\sigma_b}\right)\right]$. The objective function of i is:

$$\begin{aligned}\Pi_i\left(\sigma_i;\sigma_j,\hat{x}\right) &= \int_{-\infty}^{\chi} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(\Phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right) - \Phi\left(\frac{x_j-\theta_i}{\sigma_i}\right)\right)dx_j \\ &\quad + \int_{\hat{x}}^{\infty} \frac{1}{\sigma_j}\phi\left(\frac{x_j-\theta_j}{\sigma_j}\right)\left(\Phi\left(\frac{x_j-\theta_i}{\sigma_i}\right) - \Phi\left(\frac{2\hat{x}-x_j-\theta_i}{\sigma_i}\right)\right)dx_j.\end{aligned}$$

Momentarily ignoring the fact that $\hat{x}$ has a different value than before (because we have different values of σ), clearly, this probability and $i's$ previous probability of retention in the case where $\sigma_g > \sigma_b$ add up to 1: before, i was elected if, for a given x_j, his own signal fell outside a given interval, whereas now i is elected if the signal falls in the complementary set. Therefore, the derivative with respect to σ_i evaluated at $\hat{x} = \frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2}$ satisfies

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

Now $\hat{x} > \theta_g > \frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}$, so $\Phi\left(\frac{\hat{x}-\frac{\sigma_i^2\theta_j+\sigma_j^2\theta_i}{\sigma_i^2+\sigma_j^2}}{\frac{\sigma_j\sigma_i}{\sqrt{\sigma_i^2+\sigma_j^2}}}\right) > \frac{1}{2}$. Then,

$$\frac{\partial\Pi_g\left(\sigma_g;\sigma_b,\hat{x}\right)}{\partial\sigma_g} < 0,$$

whereas the sign of $\frac{\partial\Pi_b(\sigma_b;\sigma_g,\hat{x})}{\partial\sigma_b}$ is ambiguous. However, notice that

$$\begin{aligned}&\frac{\partial\Pi_b\left(\sigma_b;\sigma_g,\hat{x}\right)}{\partial\sigma_b}\sigma_b - \frac{\partial\Pi_g\left(\sigma_g;\sigma_b,\hat{x}\right)}{\partial\sigma_g}\sigma_g \\ &= \frac{\sigma_b^2+\sigma_b^2}{\sqrt{\sigma_g^2+\sigma_b^2}}\phi\left(\frac{\theta_g-\theta_b}{\sqrt{\sigma_g^2+\sigma_b^2}}\right)\left(2\Phi\left(\frac{\hat{x}-\frac{\sigma_b^2\theta_g+\sigma_g^2\theta_b}{\sigma_g^2+\sigma_b^2}}{\frac{\sigma_b\sigma_g}{\sqrt{\sigma_g^2+\sigma_b^2}}}\right)-1\right)\left(\frac{\theta_g-\theta_b}{\sigma_g^2+\sigma_b^2}\right) \\ &> 0.\end{aligned}$$

Then

$$c'\left(\sigma_b\right)\sigma_b > c'\left(\sigma_g\right)\sigma_g.$$

So, in principle there is no contradiction here. Moreover, if $c'\left(\sigma\right)\sigma$ is always increasing, this inequality is consistent with $\sigma_b > \sigma_g$. ■

Proof of the Claim in Section 7.4. In Section 7.4 of the main text we claim that $\theta_g > \theta_b$ even when these values are endogenously chosen.

We begin by eliminating the possibility that $\theta_b > \theta_g$. From the definition in (a.20) is it clear that a bounded retention regime is still 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

$$\frac{1}{\sigma_k}\phi\left(\frac{x_- - \theta_k}{\sigma_k}\right) - \frac{1}{\sigma_k}\phi\left(\frac{x_+ - \theta_k}{\sigma_k}\right) \leq d'\left(\theta_k - \underline{\theta}_k\right), \tag{a.10}$$

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.10) cannot hold. 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, there cannot be an equilibrium where $\theta_g = \theta_b = \theta$. For if there were, 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 $\theta_g \neq \theta_b$. ■

Analysis of the Model in Section 7.5. We begin by eliminating monotone equilibria. We already know that if the retention threshold, $x^*(\sigma)$ lies between θ_b and θ_g the good type will want to minimize the noise of his signal, while the bad type will want to maximize it, destroying the best response property for X. In the same negative vein, if the retention threshold that lies to the right of θ_g, both agents always respond by (costlessly) increasing noise, so a monotone equilibrium cannot exist in this case. So it remains to consider the case in which $x^* \leq \theta_b$. In this case, both the bad and the good type want to reduce the noise to its minimum: they both play $\sigma = \underline{\sigma}$. That is consistent with a monotone retention rule as long as $\underline{\sigma}$ is sufficiently large.[2] That is, $x^*(\underline{\sigma}) \leq \theta_b$ if, and only if

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

or

$$\ln(\beta) \geq \frac{\Delta^2}{2\underline{\sigma}^2}.$$

This is only possible if and only if (24) fails.

Next, we eliminate bounded replacement equilibria. It will be useful to recall the following results from the main text (see Proposition 2 and Lemma 1 in the Appendix):

Lemma 6. (i) *If $\sigma_b > \sigma_g$, then retention occurs in $X = [x_-, x_+]$, and $\theta_g < \frac{x_- + x_+}{2} < x_+$.*

(ii) *If $\sigma_b < \sigma_g$, then replacement occurs in $X = [x_+, x_-]$, and $\theta_b > \frac{x_- + x_+}{2} > x_+$.*

Lemma 6 tells us that in a bounded replacement equilibrium the principal replaces the agent when the signal falls inside $[x_+, x_-]$, and this regime is associated with $\sigma_b < \sigma_g$. Under this retention regime, it is easy to see that the retention probability of any type converges to 1 as $\sigma_k \to \infty$, and it is therefore clear that $\sigma_b < \sigma_g$ can never hold.

Proof of the Assertion in Footnote 14 of the Main Text. This assertion states that the non-existence of bounded replacement is robust to allowing for a finite upper bound to the choice of noise. In what follows, then, assume that there exists an upper bound on noise, so $\sigma_k \in [\underline{\sigma}, \bar{\sigma}]$ is costless, whereas the cost of going below $\underline{\sigma}$ or above $\bar{\sigma}$ is prohibitively high. Suppose on the contrary that a bounded replacement equilibrium exists.

Now, any type inside the replacement interval chooses $\bar{\sigma}$, trying to escape from the danger zone. That immediately tells us that $\theta_b \notin [x_+, x_-]$, otherwise σ_b cannot be lower than σ_g. Then, since $x_+ < \theta_b$ by Lemma 6, we need $x_- < \theta_b < \theta_g$, and therefore both types must be in the retention zone, X. Any type in X will choose $\underline{\sigma}$ or $\bar{\sigma}$, depending on which one of the two leads to a smaller probability mass in the replacement set. It follows that to maintain $\sigma_> \sigma_g$, the bad type must choose $\underline{\sigma}$ and the good type $\bar{\sigma}$, so that the first-order derivatives must satisfy

$$\text{(a.11)} \qquad -\phi\left(\frac{x_+ - \theta_g}{\bar{\sigma}}\right)\left(\frac{x_+ - \theta_g}{\bar{\sigma}^2}\right) + \phi\left(\frac{x_- - \theta_g}{\bar{\sigma}}\right)\left(\frac{x_- - \theta_g}{\bar{\sigma}^2}\right) \geq 0$$

for the good type, and

$$\text{(a.12)} \qquad -\phi\left(\frac{x_+ - \theta_b}{\underline{\sigma}}\right)\left(\frac{x_+ - \theta_b}{\underline{\sigma}^2}\right) + \phi\left(\frac{x_- - \theta_b}{\underline{\sigma}}\right)\left(\frac{x_- - \theta_b}{\underline{\sigma}^2}\right) \leq 0.$$

[2]For small $\underline{\sigma}$, there is a high degree of "separation" between the two types. To see this, consider the signal $x = \theta_b$. At this value, the likelihood of the bad type relative to the good type explodes as $\underline{\sigma}$ goes to 0. This makes x^* shift to the right, until it goes above θ_b. Now the equilibrium falls apart.

for the bad type. Combining (a.12) with the principal's indifference condition (9) in the main text, we have:

$$-\beta\frac{1}{\bar{\sigma}}\phi\left(\frac{x_+-\theta_g}{\bar{\sigma}}\right)\left(\frac{x_+-\theta_b}{\underline{\sigma}}\right)+\beta\frac{1}{\bar{\sigma}}\phi\left(\frac{x_--\theta_g}{\bar{\sigma}}\right)\left(\frac{x_--\theta_b}{\underline{\sigma}}\right)\leq 0.$$

But this, together with type-g's first-order condition (a.11), yields:

$$\left[\phi\left(\frac{x_+-\theta_g}{\bar{\sigma}}\right)-\phi\left(\frac{x_--\theta_g}{\bar{\sigma}}\right)\right](\theta_g-\theta_b)\geq 0.$$

Because $\theta_g > \theta_b$, this implies a contradiction since $\phi\left(\frac{x_+-\theta_g}{\bar{\sigma}}\right) < \phi\left(\frac{x_--\theta_g}{\bar{\sigma}}\right)$ (see Lemma 2 in the Appendix of the main text). ■

We now work towards a tighter description of bounded retention equilibrium and the proof of Proposition 10 in the main text. We begin with agent best responses:

Lemma 7. (i) *If $X = [x^*, \infty)$ and $\theta_k > x^*$, the agent 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) *Given a retention zone of the form $[x_-, x_+]$ with $x_- < x_+$, and $x_+ > \theta_g$, if $x_- \leq \theta_k$, then $\sigma_k = \underline{\sigma}$.*

(iii) *Given a retention zone of the form $[x_-, x_+]$ with $x_- < x_+$, and $x_+ > \theta_g$, if $x_- > \theta_k$ and $x_+ < \infty$, then for each k define*

$$d_k(\sigma_k) := \phi\left(\frac{x_--\theta_k}{\sigma_k}\right)(x_--\theta_k)-\phi\left(\frac{x_+-\theta_k}{\sigma_k}\right)(x_+-\theta_k) \text{ for all } \sigma_k > 0. \tag{a.13}$$

Then d_k is continuous, initially positive then negative, with a unique root to $d_k(\sigma_k) = 0$, given by

$$\sigma_k^* = \sqrt{\frac{(x_+-x_-)(\frac{x_-+x_+}{2}-\theta_k)}{\ln(x_+-\theta_k)-\ln(x_--\theta_k)}} \in (x_--\theta_k, x_+-\theta_k), \tag{a.14}$$

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

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

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

$$\Phi\left(\frac{x_+-\theta_k}{\sigma_k}\right)-\Phi\left(\frac{x_--\theta_k}{\sigma_k}\right), \tag{a.15}$$

where Φ is the cdf of the standard normal. The first-order derivative of the objective function with respect to σ_k is

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

where d_k is defined in (a.13). If $x_- \leq \theta_k$ and $x_+ > \theta_k$, then the sign of the derivative is always negative, so $\sigma_k = \underline{\sigma}$.

(iii) Since both x_- and x_+ are higher than θ_k, the sign of the derivative depends on the value of σ_k. After some elementary manipulation, we see that

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

The term inside the curly brackets is the only one that can change sign. Moreover, this term is continuous and strictly decreasing in σ_k, with limit $\frac{x_- - \theta_k}{x_+ - \theta_k} - 1 < 0$ when $\sigma_k \to \infty$, and ∞ as $\sigma_k \to 0$. So d_k has all the claimed properties, and there exists a unique σ_k^* that solves (a.15), given by setting the term within curly brackets equal to zero, which yields:

$$\sigma_k^* = \sqrt{\frac{(x_+ - x_-)\left(\frac{x_- + x_+}{2} - \theta_k\right)}{\ln(x_+ - \theta_k) - \ln(x_- - \theta_k)}}$$

Therefore, the agent will optimally choose $\sigma_k = \max\{\underline{\sigma}, \sigma_k^*\}$.

To show that $\sigma_k^* \in (x_- - \theta_k, x_+ - \theta_k)$, first define $\hat{x}_k := [(x_+ - \theta_k)/(x_- - \theta_k)]^2 \in (1, \infty)$. Provided $x_- > \theta_k$, we will have $\theta_k + \sigma_k^* > x_-$ if and only if $\hat{x}_k - 1 > \ln(\hat{x}_k)$, which is always true because equality holds at $\hat{x}_k = 1$ and then the left-hand side increases at a rate of 1, whereas the right-hand side increases at a rate of $1/\hat{x}_k < 1$. Similarly, $\theta_k + \sigma_k^* < x_+$ iff $1 - (1/\hat{x}_k) < \ln(\hat{x}_k)$. The condition holds with equality for $\hat{x}_k = 1$, and the derivatives of the left and right-hand sides are $1/\hat{x}_k^2$ and $1/\hat{x}_k$, respectively, making the condition valid for any $\hat{x}_k \in (1, \infty)$. ■

We now take note of a property of bounded retention equilibrium in the costless noise model:

Lemma 8. *If a bounded retention equilibrium exists, it can never be the case that $\underline{\sigma} < \sigma_g$.*

Proof. Suppose that $\underline{\sigma} < \sigma_g$. Then, since both choices of noise are interior solutions, agent optimality requires

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

Combining these equations with the principal's indifference condition (see (9) in the main text), we obtain

$$\phi\left(\frac{x_- - \theta_g}{\sigma_g}\right) = \phi\left(\frac{x_+ - \theta_g}{\sigma_g}\right),$$

which contradicts Lemma 2 in the Appendix of the main text. ■

Proof of Proposition 10. We proceed in a number of steps.

Lemma 9. *$\alpha(\beta)$ as defined by (22) is strictly decreasing in β for all $\beta \in (0,1)$.*

Proof. Multiplying both sides of (22) by α and defining

$$y := \frac{\alpha}{\alpha + \sqrt{1+\alpha^2}} \in [0,1), \tag{a.16}$$

we have

$$\alpha\beta = y \exp(-y). \tag{a.17}$$

If the assertion is false, then there is β such that α is locally nondecreasing in β. But that means that $\alpha\beta$ is *strictly* locally increasing at the very same β. Because (a.17) holds throughout and $y\exp(-y)$ is strictly increasing in y when $y \in [0,1)$,[3] it follows that y is also locally strictly increasing at that β. But from (a.16), it is easy to see that $d\alpha/dy < 0$. These last two observations contradict our presumption that α is locally nondecreasing in β. ■

Define $\alpha := (\theta_g - \theta_b)/2\underline{\sigma}$, and then let

$$\beta_l := \frac{1}{\alpha + \sqrt{1+\alpha^2}} \exp\left[-\frac{\alpha}{\alpha+\sqrt{1+\alpha^2}}\right], \tag{a.18}$$

and

$$\beta_h := \exp\left[2\alpha^2\right]. \tag{a.19}$$

Observation 1. *It is easy to see that β_l and β_h are defined by the requirement that, respectively, (23) and (24) hold with equality. Furthermore, it is clear that (24) is equivalent to asking for $\beta < \beta_h$, whereas Lemma 9 shows that (23), which can be also written as $\alpha(\beta) < \alpha$, is equivalent to asking for $\beta > \beta_l$.*

Next, define for any $\sigma \geq \underline{\sigma}$,

$$x_-(\sigma) := \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b - \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2} \text{ and } x_+(\sigma) := \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b + \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2}, \tag{a.20}$$

where

$$R(\sigma)^2 := (\theta_g - \theta_b)^2 + \left(\sigma^2 - \underline{\sigma}^2\right) 2\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right). \tag{a.21}$$

Clearly, x_- and x_+ are the roots to $\beta\frac{1}{\underline{\sigma}}\phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma}\phi\left(\frac{x-\theta_b}{\sigma}\right)$. So these are the bounds of the principal's retention regime X when she expects the bad type to choose σ and the good type to choose $\underline{\sigma}$. It is obvious that these two roots exist if and only if the right-hand side of (a.21) is greater or equal than 0. The following lemma determines the conditions under which this is true. First define

$$\sigma^* := \underline{\sigma}\left(\alpha + \sqrt{\alpha^2+1}\right). \tag{a.22}$$

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

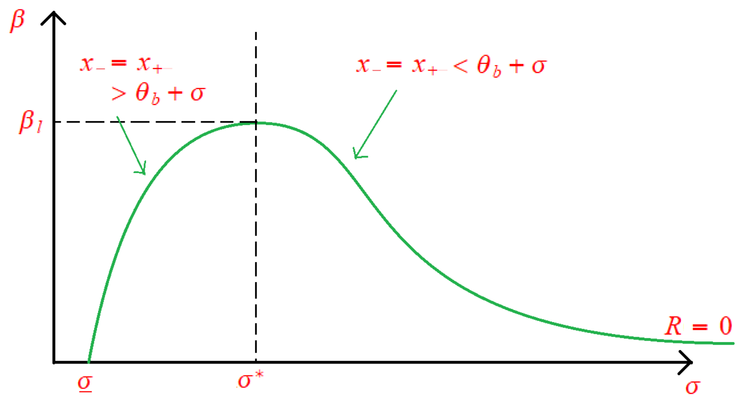


FIGURE A.1. The $R(\sigma, \beta) = 0$ locus.

Lemma 10. *Consider $\sigma_g = \underline{\sigma}$ and $\sigma_b = \sigma > \underline{\sigma}$.*

(i) *If (23) holds, there exist, for any σ, two distinct real roots $x_-(\sigma)$ and $x_+(\sigma)$ to $\beta\frac{1}{\underline{\sigma}}\phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma}\phi\left(\frac{x-\theta_b}{\sigma}\right)$, continuous in σ, given by (a.20).*

(ii) *When (23) holds with equality, $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$, and $x_-(\sigma) < x_+(\sigma)$ for $\sigma \neq \sigma^*$.*

Proof. (i) First recall Observation 1: (23) is equivalent to $\beta > \beta_l$. Now, consider the set of pairs (β, σ) such that $R = 0$. From (a.21) it is easy to see that these pairs satisfy

$$\beta = \frac{\underline{\sigma}}{\sigma}\exp\left[-\frac{\Delta^2}{2\left(\sigma^2 - \underline{\sigma}^2\right)}\right]. \tag{a.23}$$

Let us interpret (a.23) as β being a function of σ, depicted in Figure A.1. Any pair (β, σ) below the $R = 0$ locus (the green curve in the diagram) implies that the right-hand side in (a.21) is strictly negative, and therefore the functions $x_-(\sigma)$ and $x_+(\sigma)$ are not well-defined for such a 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 (β, σ) is above the locus, $R > 0$ and, therefore, two distinct real roots exist.

Now let us analyze the shape of the $R = 0$ locus. It is clear that $\beta \to 0$ as $\sigma \downarrow \underline{\sigma}$ and as $\sigma \to \infty$. By computing the derivative with respect to σ, we find that β in (a.23) 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 $\sigma^* = \underline{\sigma}\left(\alpha + \sqrt{\alpha^2+1}\right)$. Since the first one is negative, we have that β is strictly increasing in σ for $\sigma \in [\underline{\sigma}, \sigma^*)$, and it is strictly decreasing for $\sigma > \sigma^*$. At $\sigma = \sigma^*$ the derivative is zero, and therefore a global maximum is attained. By evaluating (a.23) at $\sigma = \sigma^*$ we find that this maximum value is equal to β_l, as defined in (a.18).

This means that any $\beta > \beta_l$ guarantees that $x_-(\sigma)$ and $x_+(\sigma)$ are well-defined for any value of σ and, moreover, these two real roots are always distinct.

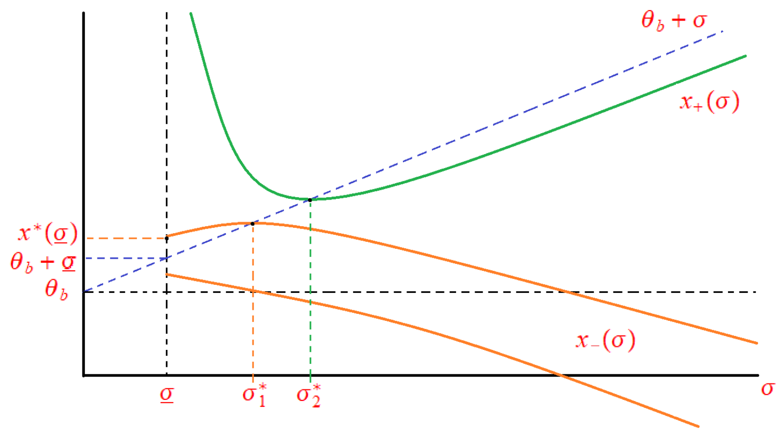


FIGURE A.2. Principal's Best Responses.

(ii) From (a.20) we can see that, along the $R = 0$ locus, $x_-(\sigma) = x_+(\sigma)$. On the other hand, for all those pairs (β, σ) above the locus, we have that $R > 0$ and, therefore, $x_-(\sigma) < x_+(\sigma)$. Since (23) with equality is equivalent to $\beta = \beta_l$, from the analysis just developed in part (i) it is clear that, when (23) holds with equality, $x_-(\sigma^*) = x_+(\sigma^*)$, and $x_-(\sigma) < x_+(\sigma)$ for $\sigma \neq \sigma^*$.

To establish that $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$, notice that, if $R = 0$, then $x_-(\sigma) = x_+(\sigma) = \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b}{\sigma^2 - \underline{\sigma}^2}$. If we impose $\frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b}{\sigma^2 - \underline{\sigma}^2} = \theta_b + \sigma$, we obtain a quadratic equation in σ: $-\sigma^2 + \sigma\Delta + \underline{\sigma}^2 = 0$. One of its roots is negative and therefore disregarded. The other one is $\sigma = \sigma^*$, so we conclude that $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$. ■

Figure A.2 depicts the two retention thresholds $x_-(\sigma)$ and $x_+(\sigma)$ for $\sigma_g = \underline{\sigma}$ and $\sigma_b = \sigma \geq \underline{\sigma}$, together with the function $\theta_b + \sigma$ which will be crucial to the analysis. The following Lemma serves as formal justification that the shapes of these objects are indeed as depicted in the diagram.

Lemma 11. *If (23) holds, $\sigma_b = \sigma > \underline{\sigma}$ and $\sigma_g = \underline{\sigma}$:*

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

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

(ii) $x_+(\sigma)$ *finds its minimum at σ_2^*, where $x_+(\sigma_2^*) = \theta_b + \sigma_2^*$, and $x_-(\sigma)$ finds its maximum at σ_1^* where $x_-(\sigma_1^*) = \theta_b + \sigma_1^*$ if $x^*(\underline{\sigma}) \geq \theta_b + \underline{\sigma}$ or at $\sigma = \underline{\sigma}$ if $x^*(\underline{\sigma}) < \theta_b + \underline{\sigma}$.*

Furthermore, $\sigma_2^* > \sigma_1^*$ when (23) holds; and $\sigma_2^* = \sigma_1^* = \sigma^*$ when (23) holds with equality.

(iii) *If (24) fails, then $x_-(\sigma, \beta) < \theta_b$ for all $\sigma > \underline{\sigma}$.*

(iv) *If (24) holds, then there exists $\hat{\sigma} > \underline{\sigma}$ such that $x_-(\sigma, \beta) > \theta_b$ for all $\sigma \in (\underline{\sigma}, \hat{\sigma})$ and $x_-(\sigma, \beta) < \theta_b$ for all $\sigma > \hat{\sigma}$.*

Proof. (i) Inspection of (a.20) immediately reveals that $\lim_{\sigma\to\underline{\sigma}^+} x_+(\sigma) = \infty$. For the corresponding limit of $x_-(\sigma)$, apply L'Hôpital's rule to see that

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

As for the limits of $x_-(\sigma)$ and $x_+(\sigma)$ when $\sigma\to\infty$, first notice that

$$\begin{aligned}
\lim_{\sigma\to\infty}\left(\frac{R(\sigma)}{\sigma}\right)^2 &= \lim_{\sigma\to\infty}\frac{(\theta_g-\theta_b)^2}{\sigma^2} + \left(1-\frac{\underline{\sigma}^2}{\sigma^2}\right)2\ln\left(\beta\frac{\sigma}{\underline{\sigma}}\right) \\
&= \infty.
\end{aligned}$$

Then, since $x_-(\sigma)$ and $x_+(\sigma)$ can be written, respectively, as

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

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

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

$$\begin{aligned}
\frac{\partial x_-(\sigma)}{\partial\sigma} &= \frac{\underline{\sigma}}{R(\sigma)}\left[\left(\frac{x_-(\sigma)-\theta_b}{\sigma}\right)^2 - 1\right] \quad \text{and} \\
\frac{\partial x_+(\sigma)}{\partial\sigma} &= \frac{\underline{\sigma}}{R(\sigma)}\left[1-\left(\frac{x_+(\sigma)-\theta_b}{\sigma}\right)^2\right]
\end{aligned} \tag{a.24}$$

respectively. In what follows, let us use the notation x' to denote these derivatives.

By part (i) of this Lemma, we know that $x_+(\sigma)\to\infty$ as $\sigma\to\underline{\sigma}^+$. This means that $x'_+(\sigma)<0$ for σ in a neighborhood of $\underline{\sigma}$. The derivative remains strictly negative as long as $x_+(\sigma)>\theta_b+\sigma$, as we can see from (a.24). $x_+(\sigma)$ keeps decreasing with σ, whereas $\theta_b+\sigma$ increases. This means that there exists a point σ_2^* such that $x_+(\sigma_2^*)=\theta_b+\sigma_2^*$, and $x'_+(\sigma_2^*)=0$ (see Figure A.2). It is clear then that for $\sigma=\sigma_2^*+\delta$ with $\delta>0$ and small, $x_+(\sigma)<\theta_b+\sigma$ and therefore $x'_+(\sigma)>0$. Now both functions, $x_+(\sigma)$ and $\theta_b+\sigma$ are increasing in σ, for $\sigma=\sigma_2^*+\delta$ with $\delta>0$ and small. Moreover, $x_+(\sigma)$ can never exceed $\theta_b+\sigma$ again. If it did, another intersection point exists, call it σ_2^{**}, and we would need $x'_+(\sigma_2^{**})\geq 1$, but the fact that $x_+(\sigma_2^{**})=\theta_b+\sigma_2^{**}$ says that $x'_+ (\sigma_2^{**}) = 0$. We therefore conclude that $x_+ (\sigma)$ is strictly decreasing for $\sigma < \sigma_2^*$, and it is strictly increasing for $\sigma > \sigma_2^*$. Its unique minimum is therefore reached at σ_2^*.

As for $x_- (\sigma)$, by part (i) of this Lemma we know that $x_- (\sigma) \to x^* (\underline{\sigma})$ as $\sigma \to \underline{\sigma}^+$.

Suppose first that $x^* (\underline{\sigma}) > \theta_b + \underline{\sigma}$, which by (a.24) says that $x'_- (\sigma) > 0$ for σ in a neighborhood of $\underline{\sigma}$. Then, both functions $x_- (\sigma)$ and $\theta_b + \sigma$ increase with σ at $\sigma = \underline{\sigma}$. However, $x_- (\sigma)$ must eventually stop increasing, since $x_- (\sigma) \to -\infty$ as $\sigma \to \infty$. So there exists σ_1^* such that $x_- (\sigma_1^*) = \theta_b + \sigma_1^*$, and therefore $x'_- (\sigma_1^*) = 0$ by (a.24). Then, for $\sigma = \sigma_1^* + \delta'$ with $\delta' > 0$ and small, $x_- (\sigma) < \theta_b + \sigma$ and therefore $x_- (\sigma)$ decreases at such σ, staying thus always below $\theta_b + \sigma$. What if $x_- (\sigma)$ goes below $\theta_b - \sigma$, in which case $x'_- (\sigma)$ would change signs again (take a look at (a.24) once again)? This cannot happen either: if an intersection point between $x_- (\sigma)$ and $\theta_b - \sigma$ exists — call it σ_1^{**} — we would need $x'_- (\sigma_1^{**}) \leq 1$ (because $x_- (\sigma)$ is crossing from above), but $x_- (\sigma_1^{**}) = \theta_b - \sigma_1^{**}$ means $x'_- (\sigma_1^{**}) = 0$. We must therefore conclude that, if $x^* (\underline{\sigma}) > \theta_b + \underline{\sigma}$, then $x_- (\sigma)$ is strictly increasing for $\sigma < \sigma_1^*$ and it is strictly decreasing for $\sigma > \sigma_1^*$. It therefore finds its maximum at $\sigma = \sigma_1^*$.

Suppose now that $x^* (\underline{\sigma}) \leq \theta_b + \underline{\sigma}$. This means that $x'_- (\underline{\sigma}^+) \leq 0$. Since $\frac{\partial(\theta_b+\sigma)}{\partial\sigma}|_{\sigma=\underline{\sigma}} = 1 > 0$, we have that $x_- (\sigma) < \theta_b + \sigma$ for σ close enough to $\underline{\sigma}$. This says that $x'_- (\sigma) < 0$ for such σ, and we already know that $x_- (\sigma)$ decreases thereafter. It therefore finds its maximum at $\sigma = \underline{\sigma}$.

It is clear that $\sigma_2^* > \sigma_1^*$ when (23) holds (or, equivalently, when $\beta > \beta_l$): $x_- (\sigma_1^*) = \theta_b + \sigma_1^*$ and $x_+ (\sigma_1^*) > x_- (\sigma_1^*)$, so $x_+ (\sigma_1^*) > \theta_b + \sigma_1^*$ which implies that $x'_2 (\sigma_1^*) < 0$ and, therefore, $\sigma_2^* > \sigma_1^*$. That is, the value of σ that maximizes $x_-(\sigma)$ is to the left of the one that minimizes $x_+(\sigma)$.

Now we will see that, as we decrease β, σ_1^* increases and σ_2^* decreases, thus getting closer to each other (see Figure A.2).

σ_1^* satisfies $\theta_b + \sigma_1^* = x_- (\sigma_1^*)$, so

$$\frac{\partial \sigma_1^*}{\partial \beta} = \frac{\partial x_- (\sigma)}{\partial \sigma}|_{\sigma=\sigma_1^*} \frac{\partial \sigma_1^*}{\partial \beta} + \frac{\partial x_- (\sigma)}{\partial \beta}|_{\sigma=\sigma_1^*}.$$

Since σ_1^* maximizes $x_- (\sigma)$, $\frac{\partial x_-(\sigma)}{\partial \sigma}|_{\sigma=\sigma_1^*} = 0$, so

$$\frac{\partial \sigma_1^*}{\partial \beta} = \frac{\partial x_- (\sigma)}{\partial \beta}|_{\sigma=\sigma_1^*} < 0.$$

Similarly, for σ_2^* we have

$$\begin{aligned} \frac{\partial \sigma_2^*}{\partial \beta} &= \frac{\partial x_+ (\sigma)}{\partial \sigma}|_{\sigma=\sigma_2^*} \frac{\partial \sigma_2^*}{\partial \beta} + \frac{\partial x_+ (\sigma)}{\partial \beta}|_{\sigma=\sigma_2^*} \\ &= \frac{\partial x_+ (\sigma)}{\partial \beta}|_{\sigma=\sigma_2^*} > 0. \end{aligned}$$

We are interested in the value of β such that $\sigma_1^* = \sigma_2^* = \sigma^*$. By the definitions of σ_1^* and σ_2^*, this value of β says, therefore, that

$$x_- (\sigma^*) = \theta_b + \sigma^* = x_+ (\sigma^*),$$

which is true only if $R(\sigma^*) = 0$, in which case $x_-(\sigma^*) = x_+(\sigma^*) = \frac{\sigma^{*2}\theta_g - \underline{\sigma}^2\theta_b}{\sigma^{*2} - \underline{\sigma}^2}$ and the above condition reads

$$\frac{\sigma^{*2}\theta_g - \underline{\sigma}^2\theta_b}{\sigma^{*2} - \underline{\sigma}^2} = \theta_b + \sigma^*,$$

so

$$\sigma^* = \underline{\sigma}\left(\alpha + \sqrt{\alpha^2 + 1}\right). \tag{a.25}$$

Also, $R(\sigma^*) = 0$ means

$$(\theta_g - \theta_b)^2 + \left(\sigma^{*2} - \underline{\sigma}^2\right) 2 \ln\left(\beta \frac{\sigma^*}{\underline{\sigma}}\right) = 0. \tag{a.26}$$

Then, by combining (a.25) and (a.26) we obtain

$$\beta = \beta_l = \frac{1}{\alpha + \sqrt{1 + \alpha^2}} \exp\left[-\frac{\alpha}{\alpha + \sqrt{1 + \alpha^2}}\right], \tag{a.27}$$

which is equivalent to (23) holding with equality (recall Observation 1).

(iii) Suppose that (24) fails. Using (a.20), we see that

$$x_-(\sigma, \beta) - \theta_b = \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b - \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2} - \theta_b = \frac{\sigma^2(\theta_g - \theta_b) - \sigma\underline{\sigma}R(\sigma)}{\sigma^2 - \underline{\sigma}^2}. \tag{a.28}$$

So the claim is established if the last term in (a.28) 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.21), 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 (24), because $\sigma \geq \underline{\sigma}$.

(iv) From the previous point, we know that $x_-(\sigma) \leq \theta_b$ if and only if

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

The right-hand side is strictly increasing in σ. Then, if (24) holds, the condition is violated for σ close enough to $\underline{\sigma}$. In other words, there exists $\hat{\sigma} > \underline{\sigma}$ such that $x_-(\sigma) > \theta_b$ for all $\sigma \in (\underline{\sigma}, \hat{\sigma})$, and $x_-(\sigma) < \theta_b$ for all $\sigma > \hat{\sigma}$. ■

Let us construct a self-map on $(\underline{\sigma}, \infty)$, with domain to be interpreted as the principal's conjecture about the noise used by the low type, and range as the subsequent optimal choice of noise by the

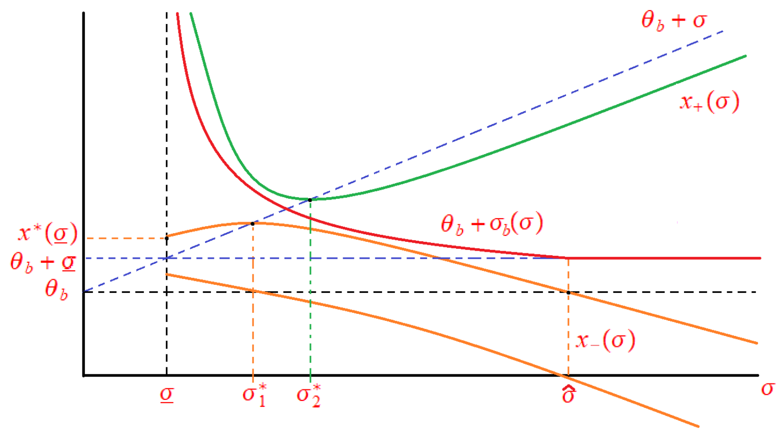


FIGURE A.3. Principal's best responses $x_-(\sigma)$ and $x_+(\sigma)$ and b's counter-response.

bad type, in response to the retention decision. (Throughout, we presume that $\sigma_g = \underline{\sigma}$.) Guided by Lemma 7iii, our self-map is:

$$\Psi(\sigma) \equiv \max\left\{\sqrt{\frac{[x_+(\sigma) - x_-(\sigma)](\frac{x_-(\sigma)+x_+(\sigma)}{2} - \theta_b)}{[\ln(x_+(\sigma) - \theta_b) - \ln(x_-(\sigma) - \theta_b)]}}, \underline{\sigma}\right\}. \tag{a.29}$$

Take a look at Figure A.3 which will guide you through the following Lemma. Keep also in mind equation (a.14) in Lemma 7iii, which says that an unrestricted maximizer when retention zone is $X = [x_-, x_+]$ with $\theta_k < x_- < x_+ < \infty$ satisfies $\sigma_k \in (x_- - \theta_k, x_+ - \theta_k)$ or equivalently, $x_- < \theta_k + \sigma_k < x_+$.

Lemma 12. *If (23) holds, there exists a unique fixed-point σ^* of $\Psi(\sigma)$;*

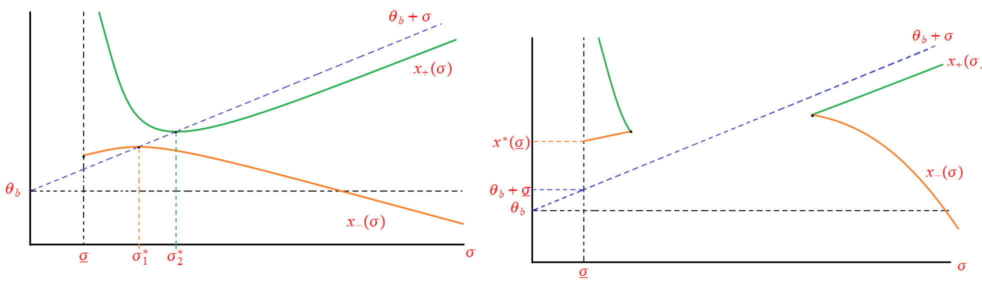
(i) If (24) holds, then $\sigma^* > \underline{\sigma}$;

(ii) If (24) fails, then $\sigma^* = \underline{\sigma}$.

Proof. (i) By Lemma 11i, $\lim_{\sigma\to\underline{\sigma}^+} x_-(\sigma) = x^*(\underline{\sigma})$, and if (24) holds, then $x^*(\underline{\sigma}) > \theta_b$. Also, $\lim_{\sigma\to\underline{\sigma}^+} x_+(\sigma) = \infty$. By inspecting (a.29), we can see that $\lim_{\sigma\to\underline{\sigma}^+} \Psi(\sigma) = \infty$. By Lemma 11iv, the interval $(x_-(\sigma), x_+(\sigma))$ contains θ_b for σ large, so that by Lemma 7i, $\Psi(\sigma) = \underline{\sigma}$ for all such σ. By Lemma 10i, $x_-(\sigma)$ and $x_+(\sigma)$ are well-defined and distinct for every $\sigma > \underline{\sigma}$, so these values move continuously with σ. Consequently, so does $\Psi(\sigma)$. Therefore Ψ has at least one fixed point.

At any such fixed point σ, we have $\underline{\sigma} < \sigma = \Psi(\sigma)$. Consequently, the first term on the right hand side of (a.29) must bind. It follows that $\Psi(\sigma)$ solves $d_b(\Psi(\sigma)) = 0$, where d_b is defined in (a.13), so that

$$\phi\left(\frac{x_+(\sigma) - \theta_b}{\Psi(\sigma)}\right)(x_+(\sigma) - \theta_b) = \phi\left(\frac{x_-(\sigma) - \theta_b}{\Psi(\sigma)}\right)(x_-(\sigma) - \theta_b). \tag{a.30}$$



(A) When (23) holds. (B) When (23) fails.

FIGURE A.4. $x_-(\sigma)$ and $x_+(\sigma)$ are always well-defined if and only if (23) holds.

We compute $\Psi'(\sigma)$ from (a.30). This is routine but tedious and we only outline the steps. First, consider the derivative of the left-hand side of (a.30) with respect to σ; this can be expressed 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 the manipulations used to obtain this expression 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 (a.30). Rearranging terms, we obtain

$$\Psi'(\sigma)=\frac{\Psi(\sigma)^3}{2}\frac{\frac{x_-'(\sigma)}{x_-(\sigma)-\theta_b}\left(1-\frac{(x_-(\sigma)-\theta_b)^2}{\Psi(\sigma)^2}\right)+\frac{x_+'(\sigma)}{(x_+(\sigma)-\theta_b)}\left(\frac{(x_+(\sigma)-\theta_b)^2}{\Psi(\sigma)^2}-1\right)}{(x_+(\sigma)-x_-(\sigma))\left(\frac{x_+(\sigma)+x_-(\sigma)}{2}-\theta_b\right)}. \tag{a.31}$$

Using the expressions for $x_-'(\sigma)$ and $x_+'(\sigma)$ in equations (a.24) and plugging them in (a.31) and evaluating the resulting derivative at $\sigma=\Psi(\sigma)$, we see that

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

This says that the self-map $\Psi(\sigma)$ is strictly decreasing at any fixed point. Consequently, there can only be one fixed point σ^*.

(ii) If (24) fails, then $x_-(\sigma)<\theta_b$ for all $\sigma>\underline{\sigma}$ by Lemma 11iii. At the same time, by Lemma 6, $\theta_b<x_+(\sigma)$, so $\theta_b\in(x_-(\sigma),x_+(\sigma))$ for all $\sigma\geq\underline{\sigma}$. Then, by Lemma 7ii, $\Psi(\sigma)=\underline{\sigma}$ for all $\sigma\geq\underline{\sigma}$. Therefore, there is a unique fixed-point given by $\underline{\sigma}$. ■

Lemma 13. *If (23) fails, a non-trivial equilibrium does not exist and trivial equilibria exist.*

Proof. First, remember what is established in Observation 1: the failure of (23) is equivalent to the statement $\beta \leq \beta_l$.

Recall also equation (a.23) from Lemma 10, which established the set of pairs (β, σ) such that $R(\sigma)$, as defined in (a.21), equals 0. This locus is depicted in Figure A.1. Any pair below the $R = 0$ locus implies that the right-hand side of (a.21) is strictly negative, and therefore there are no real roots to $\beta \frac{1}{\underline{\sigma}} \phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma} \phi\left(\frac{x-\theta_b}{\sigma}\right)$: the principal will follow a trivial retention rule.

From (a.20), we have that, along the locus, $x_-(\sigma) = x_+(\sigma) = \frac{\sigma^2\theta_g - \underline{\sigma}^2\theta_b}{\sigma^2 - \underline{\sigma}^2}$. From this, we obtain that $x_-(\sigma) = x_+(\sigma) > \theta_b + \sigma$ if, and only if, $\sigma \in [\underline{\sigma}, \sigma^*)$, and $x_-(\sigma) = x_+(\sigma) < \theta_b + \sigma$ if and only if $\sigma > \sigma^*$. That is, along the locus, if $\sigma > \sigma^*$ then $x_-(\sigma) = x_+(\sigma) < \theta_b + \sigma$, and if $\sigma < \sigma^*$ then $x_-(\sigma) = x_+(\sigma) > \theta_b + \sigma$. See Figure A.4.

Then, for any given $\beta < \beta_l$, there exists no value of σ such that $x_-(\sigma) < \sigma + \theta_b < x_+(\sigma)$, and therefore by Lemma 7iii there cannot be a bounded retention equilibrium with $\sigma_b > \sigma_g = \underline{\sigma}$. On the other hand, when $\beta = \beta_l$, by Lemma 10ii we have that $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$, and by Lemma 11ii, $\sigma_1^* = \sigma_2^* = \sigma^*$, so the equilibrium is $\sigma_g = \underline{\sigma}$, $\sigma_b = \sigma^*$, and the principal retains inside $X = [x_-, x_+] = \{\theta_b + \sigma^*\}$: she almost never retains.

Notice also that a monotone equilibrium $(\sigma_g, \sigma_g) = (\underline{\sigma}, \underline{\sigma})$ cannot exist either: $\beta \leq \beta_l < 1$, so it can never be the case that $x^* < \theta_b$.

For any value of $\beta < \beta_l$, there can only be trivial equilibria, and these are of the type "never retain". To obtain such an equilibrium, consider $\sigma_g = \underline{\sigma}$ and σ_b such that the (β, σ_b) pair is below the $R = 0$ locus. Since for any such pair the $\{x_-, x_+\}$ roots do not exist, the principal will never retain when facing this pair (σ_g, σ_b). ■

Lemma 14. *If σ_b satisfies $d_b(\sigma_b) = 0$ and $\{x_-, x_+\}$ each solve $\beta \frac{1}{\underline{\sigma}} \phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) = \frac{1}{\sigma_b} \phi\left(\frac{x-\theta_b}{\sigma_b}\right)$, then the good type optimally chooses $\sigma_g = \underline{\sigma}$.*

Proof. By Lemma 6 we know that it is always the case that $x_+ > \theta_g$. If, in addition, $x_- \leq \theta_g$, then by Lemma 7ii g optimally chooses $\sigma_g = \underline{\sigma}$, and we are done.

If $x_- > \theta_g$, we are in Case iii of Lemma 7, which tells us that there is a unique value of σ_g (not worrying about the lower bound $\underline{\sigma}$) maximizing g's probability of retention. This value is the one that solves $d_g(\sigma_g) = 0$, where d_g is defined in (a.13). We now claim that this value is smaller than $\underline{\sigma}$. By Lemma 7iii, it will suffice to show that $d_g(\underline{\sigma}) < 0$.

We know that the triplet (σ, x_-, x_+) satisfies (9) with equality, which is

$$\frac{1}{\sigma} \phi\left(\frac{x-\theta_b}{\sigma}\right) = \beta \frac{1}{\underline{\sigma}} \phi\left(\frac{x-\theta_g}{\underline{\sigma}}\right) \tag{a.32}$$

for $x = x_-, x_+$, while by the optimality of σ and Lemma 7iii,

$$\phi\left(\frac{x_+ - \theta_b}{\sigma}\right)(x_+ - \theta_b) = \phi\left(\frac{x_- - \theta_b}{\sigma}\right)(x_- - \theta_b). \tag{a.33}$$

It follows that

$$\begin{aligned} d_g(\underline{\sigma}) &= \phi\left(\frac{x_- - \theta_g}{\underline{\sigma}}\right)(x_- - \theta_g) - \phi\left(\frac{x_+ - \theta_g}{\underline{\sigma}}\right)(x_+ - \theta_g) \\ &= \frac{\underline{\sigma}}{\beta}\left[\phi\left(\frac{x_- - \theta_b}{\sigma}\right)\frac{x_- - \theta_g}{\sigma} - \phi\left(\frac{x_+ - \theta_b}{\sigma}\right)\frac{x_+ - \theta_g}{\sigma}\right] \\ &= \frac{\underline{\sigma}}{\beta}\frac{x_- - \theta_b}{\sigma}\phi\left(\frac{x_- - \theta_b}{\sigma}\right)\left[\frac{x_- - \theta_g}{x_- - \theta_b} - \frac{x_+ - \theta_g}{x_+ - \theta_b}\right] < 0, \end{aligned}$$

where the second and the third lines follow from (a.32) and (a.33), respectively, and the last inequality from $\theta_g > \theta_b$ and $x_+ > x_-$. ■

Proof of Proposition 10.

(i) The necessity of condition (23) for the existence of a nontrivial equilibrium is given by Lemma 13. Its sufficiency is given by Lemmas 12 and 14.

(ii) The statement is proven by combining Lemmas 6, 12 and 14.

(iii) Trivial.

(iv) If (24) fails, then by Lemma 12 the bad type is willing to play $\sigma_b = \underline{\sigma}$ if the good type is playing $\sigma_g = \underline{\sigma}$. For $(\sigma_g, \sigma_b) = (\underline{\sigma}, \underline{\sigma})$ to be compatible with a monotone retention regime, it must be the case that $x^*(\underline{\sigma}) \leq \theta_b$, which simple algebra shows to be equivalent to the failure of (24). ■

Section 7.6. Recall the formula (26) for the equilibrium ratio β in the main text:

$$\beta = \frac{q}{1-q}\frac{1-p}{p} = \frac{1+\delta\Pi_b}{1+\delta\Pi_g}. \tag{a.34}$$

Also recall that we are now working with the costless noise model from Section 7.5: any choice of noise above $\underline{\sigma}$ is costless, whereas going below $\underline{\sigma}$ is impossible.

Proof of Proposition 11. For some (provisionally given) value of β, obtain the baseline static model and then use Proposition 10 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 (a.34).

As argued in the main text, it has to be the case that $\Pi_g \geq \Pi_b$, because the principal will choose a retention zone that retains the high type at least as often than the low type. This says that, in the dynamic model, $\beta \leq 1$, and therefore condition (24) trivially holds. Then, following Proposition 10 and Observation 1, we can separate the analysis into two cases: either (23) fails and $\beta < \beta_l < 1$, or (23) holds and $\beta \in (\beta_l, 1]$.

In the former case, Proposition 10 tells us that in the static model, only a trivial equilibrium exists (see Lemma 13). Then, $\Pi_b = \Pi_g$. But equilibrium condition (a.34) 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 be a bounded retention equilibrium. We now prove its existence.

For any given $\beta \in (\beta_l, 1]$, by Proposition 10, there is a unique equilibrium in the static model, and it involves 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 $k = b, g$:

$$\Pi_k(\beta) = \int_X \pi_k(x)dx = \frac{1}{\sigma_k(\beta)} \int_{x_-(\beta)}^{x_+(\beta)} \phi\left(\frac{x - \theta_k}{\sigma_k(\beta)}\right) dx. \tag{a.35}$$

Now, in line with (a.34), define a mapping $\beta' = \psi(\beta)$ by

$$\beta' = \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)} \tag{a.36}$$

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 $\beta \downarrow \beta_l$, the boundaries of the static equilibrium retention thresholds x_-^* and x_+^* converge to each other (see Lemma 15), so that $\lim_{\beta\downarrow\beta_l} \Pi_g = \lim_{\beta\downarrow\beta_l} \Pi_b = 0$, and therefore

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

as $\beta \downarrow \beta_l$. This verifies a second end-point condition $\lim_{\beta\downarrow\beta_l} \psi(\beta) > \beta_l$. By the intermediate value theorem, there is at least one value of 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.35) 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:

$$\frac{\partial\Pi_k(\beta)}{\partial\beta} = \frac{1}{\sigma_k(\beta)}\left[\phi\left(\frac{x_+(\beta) - \theta_k}{\sigma_k(\beta)}\right) x'_+(\beta) - \phi\left(\frac{x_-(\beta) - \theta_k}{\sigma_k(\beta)}\right) x'_-(\beta)\right]. \tag{a.37}$$

Next, observe that

$$\frac{\partial}{\partial\beta}\frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)} = \delta\frac{\frac{\partial\Pi_b(\beta)}{\partial\beta}(1 + \delta\Pi_g(\beta)) - (1 + \delta\Pi_b(\beta))\frac{\partial\Pi_g(\beta)}{\partial\beta}}{(1 + \delta\Pi_g(\beta))^2}. \tag{a.38}$$

Substitute (a.37) in (a.38) and use the fact that $x_-(\beta)$ and $x_+(\beta)$ solve (9) with equality to obtain (after some manipulation)

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

Because $x'_+(\beta) > 0$ and $x'_-(\beta) < 0$ (Lemma 15i), we must conclude that

$$\text{Sign}\left\{\frac{\partial}{\partial\beta}\frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)}\right\} = \text{Sign}\left\{\beta - \frac{1 + \delta\Pi_b(\beta)}{1 + \delta\Pi_g(\beta)}\right\}. \tag{a.39}$$

This last observation, along with the end-point condition $\lim_{\beta\to\beta_l}\phi(\beta) > \beta_l$, eliminates two solutions to the equation

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

for that would require the sign equality (a.39) to be violated for some β. ■

Proof of Proposition 12. Recall Observation 1: β_l in (a.18) and β_h in (a.19) are defined, respectively, by the requirement that (23) and (24) hold with equality. Therefore, conditions (23) and (24) together are equivalent to $\beta \in (\beta_l, \beta_h)$.

We proceed in a number of steps.

Lemma 15. *Assume* $\beta \in (\beta_l, \beta_h)$*, then*

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

(ii) $\lim_{\beta\to\beta_l^+} x_-^* = \lim_{\beta\to\beta_l^+} x_+^*$ *and* $\lim_{\beta\to\beta_l^+}\sigma_b^* = \underline{\sigma}\left(\alpha+\sqrt{\alpha^2+1}\right)$

Proof. (i) When $\beta \in (\beta_l, \beta_h)$, Proposition 10 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 $X = [x_-, x_+]$. The equilibrium values $\left(\sigma_b^*, x_-^*, x_+^*\right)$ are determined by

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

and

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

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

$$\text{(a.40)}\qquad \begin{aligned} \frac{\sigma_b'}{\sigma_b}\left(\left(\frac{x_- -\theta_b}{\sigma_b}\right)^2-1\right) &= \frac{R(\sigma_b)}{\sigma_b\underline{\sigma}}x_-' + \frac{1}{\beta},\\ \frac{\sigma_b'}{\sigma_b}\left(1-\left(\frac{x_+ -\theta_b}{\sigma_b}\right)^2\right) &= \frac{R(\sigma_b)}{\sigma_b\underline{\sigma}}x_+' - \frac{1}{\beta}. \end{aligned}$$

In the case of the second equilibrium equation we obtain the same expression as in (a.31), where $\Psi(\sigma)$ is now σ_b, and the derivatives are those with respect to β. By combining it with (a.40), and after some heavy algebra, we obtain:

(a.41)

$$x_-' = -\frac{y_-(y_+ + y_-)}{\beta}\frac{\underline{\sigma}^2}{(\theta_g-\theta_b)}\left(\frac{\frac{\theta_g-\theta_b}{\underline{\sigma}}\frac{\sigma_b}{\underline{\sigma}}(y_+ - y_-)y_+ + (y_- + y_+)\left(y_+^2-1\right)}{\frac{\theta_g-\theta_b}{\underline{\sigma}}\frac{\sigma_b}{\underline{\sigma}}y_-y_+(y_+ - y_-)^2 + \left(1-y_-^2\right)^2 y_+ + \left(1-y_+^2\right)^2 y_-}\right)$$

$$x_+' = \frac{y_+(y_+ + y_-)}{\beta}\frac{\underline{\sigma}^2}{(\theta_g-\theta_b)}\left(\frac{\frac{\theta_g-\theta_b}{\underline{\sigma}}\frac{\sigma_b}{\underline{\sigma}}(y_+ - y_-)y_- + (y_+ + y_-)\left(1-y_-^2\right)}{\frac{\theta_g-\theta_b}{\underline{\sigma}}\frac{\sigma_b}{\underline{\sigma}}y_-y_+(y_+ - y_-)^2 + \left(1-y_-^2\right)^2 y_+ + \left(1-y_+^2\right)^2 y_-}\right)$$

where the notation x' means $\frac{\partial x}{\partial \beta}$, and $y_i := \frac{x_i-\theta_b}{\sigma_b}$, for $i = -, +$.

Since $\beta < \beta_h$ condition (24) holds, and therefore $x_- > \theta_b$. This says that $y_+ > y_- > 0$. Also, from Lemma 7iii, $y_+ > 1 > y_-$. Then, from the above expressions we can see that $x'_- < 0$ and $x'_+ > 0$, which means that the interval shrinks as β decreases.

(ii) For $\beta > \beta_l$, Proposition 10 tells us that there exists a unique equilibrium, which exhibits bounded retention. By Lemma 7iii, equilibrium σ_b satisfies $x_-(\sigma_b) < \theta_b + \sigma_b < x_+(\sigma_b)$. In the limit as $\beta \to \beta_l$, condition (23) holds with equality, and by Lemma 10ii, $x_-(\sigma^*) = x_+(\sigma^*) = \theta_b + \sigma^*$ where $\sigma^* = \underline{\sigma}\left(\alpha + \sqrt{\alpha^2+1}\right)$. Then, $\sigma_b \to \sigma^*$, and the result is proven.

■

Lemma 16. $\frac{\partial \sigma_b^*}{\partial \beta} = 0$ *at* $\beta = \beta_l, \beta_h$ *and* $\frac{\partial \sigma_b^*}{\partial \beta} < 0$ *for* $\beta \in (\beta_l, \beta_h)$.

Proof. By combining the derivatives in (a.41) and (a.40) we obtain

$$\frac{\sigma_b'}{\sigma_b} = \frac{1}{\beta}\frac{(y_- + y_+)(y_-y_+ - 1)}{\frac{\theta_g - \theta_b}{\underline{\sigma}}\frac{\sigma_b}{\underline{\sigma}}y_-y_+(y_+ - y_-)^2 + \left(1-y_-^2\right)^2 y_+ + \left(1-y_+^2\right)^2 y_-},$$

where $y_i := \frac{x_i - \theta_b}{\sigma_b}, i = -, +$. So, $\sigma_b' \le 0$ if and only if $y_-y_+ \le 1$, which, from the expressions for x_- and x_+ in (a.20), is equivalent to

$$\beta \ge \frac{\exp\left[2\alpha^2\right]\exp\left[\frac{1}{2}\left(1-\left(\frac{\sigma_b(\beta)}{\underline{\sigma}}\right)^2\right)\right]}{\frac{\sigma_b(\beta)}{\underline{\sigma}}} =: f(\beta). \tag{a.42}$$

Notice that, since by Lemma 15 $\lim_{\beta\to\beta_l}\sigma_b = \underline{\sigma}\left(\alpha + \sqrt{\alpha^2+1}\right)$,

$$\begin{aligned}
\lim_{\beta\to\beta_l} f(\beta) &= \frac{\exp\left[2\alpha^2\right]\exp\left[\frac{1}{2}\left(1-\left(\alpha+\sqrt{\alpha^2+1}\right)^2\right)\right]}{\alpha+\sqrt{\alpha^2+1}} \\
&= \frac{\exp\left[\alpha\left(\alpha-\sqrt{\alpha^2+1}\right)\right]}{\alpha+\sqrt{\alpha^2+1}} \\
&= \frac{\exp\left[-\frac{\alpha\left(\sqrt{\alpha^2+1}-\alpha\right)\left(\alpha+\sqrt{\alpha^2+1}\right)}{\left(\alpha+\sqrt{\alpha^2+1}\right)}\right]}{\alpha+\sqrt{\alpha^2+1}} \\
&= \frac{\exp\left[-\frac{\alpha}{\alpha+\sqrt{\alpha^2+1}}\right]}{\alpha+\sqrt{\alpha^2+1}} \\
&= \beta_l.
\end{aligned}$$

Also, since $\sigma_b \to \underline{\sigma}$ as $\beta \to \beta_h^-$ by Proposition 10,

$$\begin{aligned}
\lim_{\beta\to\beta_h} f(\beta) &= \exp\left[2\alpha^2\right] \\
&= \beta_h.
\end{aligned}$$

So, for the extreme values of β, inequality (a.42) holds with equality, and therefore $\sigma_b'(\beta) = 0$. The derivative of f with respect to β is

$$f'(\beta) = -\underline{\sigma} \exp\left[2\alpha^2\right] \exp\left[\frac{1}{2}\left(1 - \left(\frac{\sigma_b}{\underline{\sigma}}\right)^2\right)\right]\left(\frac{1}{\underline{\sigma}^2} + \frac{1}{\sigma_b^2}\right)\sigma_b'.$$

Assume that $\lim_{\beta\to\beta_l^+} f'(\beta) \geq 1$. This means that in a neighborhood of β_l, $f(\beta) \geq \beta$, which implies that $\sigma_b' \geq 0$. But then $f'(\beta) \leq 0$ from the above expression. We conclude that $\lim_{\beta\to\beta_l^+} f'(\beta) < 1$ and therefore $f(\beta) < \beta$ for β close enough to β_l. Can $f(\beta)$ go above β? If such a point exists, if has to be the case that $f'(\beta) > 1 > 0$, but at the same time $f(\beta) = \beta$ implies $\sigma_b' = 0$, which implies that $f'(\beta) = 0$ from the above expression. We conclude that $f(\beta) < \beta$ at any $\beta \in (\beta_l, \beta_h)$ and therefore $\frac{\partial\sigma_b}{\partial\beta} < 0$ for $\beta \in (\beta_l, \beta_h)$, with $\frac{\partial\sigma_b}{\partial\beta} = 0$ at the end-points $\beta = \beta_l, \beta_h$. ■

2. Examples

2.1. **(Non-Generic) Example of a Monotone Equilibrium.** In the main text we have argued that with a monotone regime $X = [x^*, \infty)$ with $x^* \in [\theta_b, \theta_g]$ the bad type will optimally respond by choosing $\sigma_b > \underline{\sigma}$, whereas the good type will play $\sigma_g < \underline{\sigma}$, so this cannot be an equilibrium. Consider then the case $x^* \notin [\theta_b, \theta_g]$. Type-$k$ agent's objective function is:

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

and the corresponding first order condition is

$$\phi\left(\frac{x^* - \theta_k}{\sigma_k}\right)\frac{x^* - \theta_k}{\sigma_k^2} - c'(\sigma_k) = 0$$

If $\sigma_g = \sigma_b = \sigma$, the two first-order condition together imply that

$$\phi\left(\frac{x^* - \theta_g}{\sigma}\right)(x^* - \theta_g) = \phi\left(\frac{x^* - \theta_b}{\sigma}\right)(x^* - \theta_b). \tag{a.43}$$

Furthermore, from the principal's indifference condition, we have that

$$\beta\phi\left(\frac{x^* - \theta_g}{\sigma}\right) = \phi\left(\frac{x^* - \theta_b}{\sigma}\right),$$

which determines the value of x^* :

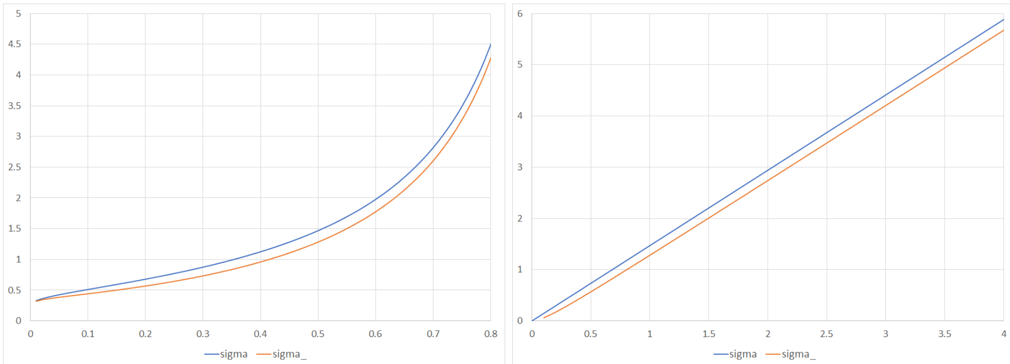
$$x^* = \frac{\theta_g + \theta_b}{2} - \frac{\sigma^2}{\theta_g - \theta_b}\ln(\beta). \tag{a.44}$$

If we use this indifference condition in equation ($a.43$) we obtain the following:

$$(x^* - \theta_g) = \beta(x^* - \theta_b),$$

or

$$x^* = \frac{\theta_g - \beta\theta_b}{1 - \beta}.$$



(A) Equilibrium Noise and Ambient Noise as functions of β

(B) Equilibrium Noise and Ambient Noise as functions of Δ

FIGURE A.5. Monotone Equilibria

Now, combining this expression for x^* with $(a.44)$ the equilibrium value of σ is fully determined:

$$\sigma = \Delta\sqrt{\frac{1}{2}\frac{1}{\ln(\beta)}\frac{1+\beta}{\beta-1}},$$

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

Assume $\beta < 1$, so $x^* > \theta_k\ \forall k$. The good type's first-order condition is

$$\begin{aligned} c'(\sigma) &= \phi\left(\frac{\beta}{1-\beta}\frac{\Delta}{\sigma}\right)\frac{\Delta}{\sigma^2}\frac{\beta}{1-\beta} \\ &= \phi\left(\sqrt{2\ln(\beta)\frac{\beta-1}{1+\beta}}\frac{\beta}{1-\beta}\right)\frac{2}{\Delta}\ln\left(\frac{1}{\beta}\right)\frac{\beta}{1+\beta} > 0 \end{aligned}$$

Let $c(\sigma) = \frac{1}{2\underline{\sigma}}(\sigma - \underline{\sigma})^2\ \forall\sigma \geq \underline{\sigma}$. Then the condition reads

$$\frac{\sigma-\underline{\sigma}}{\underline{\sigma}} = \phi\left(\sqrt{2\ln(\beta)\frac{\beta-1}{1+\beta}}\frac{\beta}{1-\beta}\right)\frac{2}{\Delta}\ln\left(\frac{1}{\beta}\right)\frac{\beta}{1+\beta},$$

or

$$\underline{\sigma} = \frac{\Delta\sqrt{\frac{1}{2}\frac{1}{\ln(\beta)}\frac{1+\beta}{\beta-1}}}{1+\phi\left(\sqrt{2\ln(\beta)\frac{\beta-1}{1+\beta}}\frac{\beta}{1-\beta}\right)\frac{2}{\Delta}\ln\left(\frac{1}{\beta}\right)\frac{\beta}{1+\beta}}.$$

Figure A.5 plots σ and $\underline{\sigma}$ as functions of $\beta \in \left(0, \frac{4}{5}\right)$ for $\Delta = 1$ (Panel a) and as functions of Δ for $\beta = \frac{1}{2}$ (Panel b):

2.2. **Examples of Bounded Replacement Equilibria.**

Lemma 17. *Suppose the principal retains and replaces according to some rule such that* $x_- \in [\theta_b, \theta_g]$. *Then, the agents' best responses satisfy* $\sigma_b > \sigma_g$.

The proof of Lemma 17 is identical to the one of Proposition 5, which is developed in the Appendix. If we define $B_k(\sigma)$ to be type-k's marginal benefit of noise, under the premise $x_- \in [\theta_b, \theta_g]$, the marginal benefit of noise for the bad type strictly exceeds that for the good type at *every* noise level. Then, by a simple single-crossing argument, we must have $\sigma_b > \sigma_g$.

By combining this Lemma with Proposition 4, we have the following

Observation 2. *Consider a bounded replacement equilibrium, in which $\sigma_g > \sigma_b$. Then, either*

(i) $\beta > 1$ *and large enough so that* $x_+ < x_- < \theta_b < \theta_g$; *or*

(ii) $\beta < 1$ *and small enough so that* $x_+ < \theta_b < \theta_g < x_-$.

We will now construct two examples of bounded replacement, one for $\beta < 1$ and another for $\beta > 1$. For the case $\beta < 1$, Observation 2 says that both choices of noise must be above the ambient noise. Let us then take $\theta_b = 1$, $\theta_g = 2$ and $x_- = 2.3$. The idea is that both types are inside the replacement zone, as Observation 2 determines, but the bad type is deep inside 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 $\sigma_g > \sigma_b$.

Let us impose $\sigma_g = 0.42$ and $\sigma_b = 0.250001$. We have now pinned down the value of x_+ :

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

so that

$$x_+ \quad = \quad 2\frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} - x_- \approx -1.38$$

The value of $\beta < 1$ is now also determined:

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

Finally, the two first-order conditions need to be satisfied. Both types' marginal benefits are always positive, so we just have to care about marginal cost for values of σ above $\underline{\sigma}$. We choose

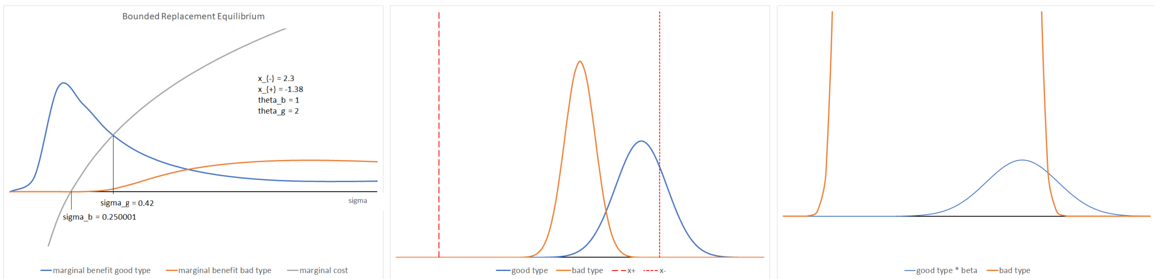
$$c'(\sigma) = A\ln(\sigma) + B.$$

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:

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



(A) Marginal Benefits and Costs (B) Equilibrium Likelihoods (C) Adjusted Eq. Likelihoods

FIGURE A.6. A Bounded Replacement Equilibrium for β Small.

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 the ambient noise ($c'(\sigma) = 0$) is $\underline{\sigma} \approx \frac{1}{4}$.

Figure A.6 depicts the equilibrium.

Now we find an example of a bounded replacement equilibrium for the case $\beta > 1$. By Observation 2 it must be the case that $x_+ < x_- < \theta_b < \theta_g$. Both agents are now in the retention zone so they want to stay there: $\sigma_b, \sigma_g < \underline{\sigma}$. The bad type is closer to the replacement zone, though, so he will make a bigger effort than the good type to stay safe: $\sigma_b < \sigma_g < \underline{\sigma}$.

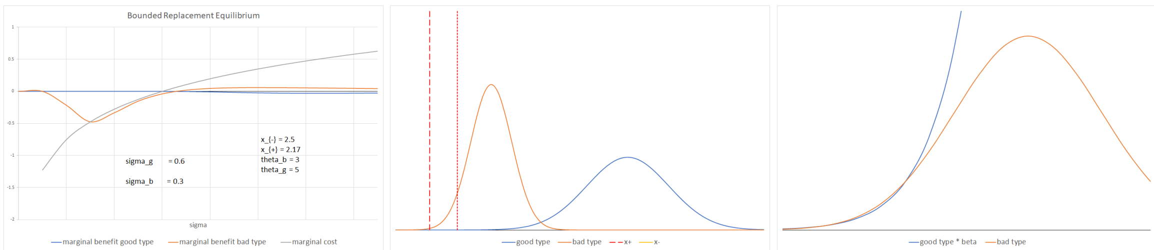
Let us then choose $\theta_b = 3$, $\theta_g = 5$ and $x_- = 2.5$. For the choices of noise, let's take $\sigma_b = 0.3$ and $\sigma_g = 0.6$. All this again pins down the value of x_+:

$$x_+ = 2\frac{\sigma_b^2\theta_g - \sigma_g^2\theta_b}{\sigma_b^2 - \sigma_g^2} - x_- \approx 2.17.$$

For β we have:

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

β is now immensely big.



(A) Marginal Benefits and Costs (B) Equilibrium Likelihoods (C) Adjusted Eq. Likelihoods

FIGURE A.7. A Bounded Replacement Equilibrium for β Large.

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

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

Ambien noise is now $\underline{\sigma} \approx 0.6004$. Figure A.7 depicts the equilibrium.

3. MISCELLANEOUS DETAILS

3.1. **The Behavior of** $\sigma(\theta)$**.** We will analyze the behavior of $\sigma(\theta)$ when the agents face a retention rule $X = [x_-, x_+]$, where x_+ can be equal to infinity (monotone regime). See Figure 5 in the main text. Even though the analysis will not be complete in the case of bounded retention, it will shed some light on the way σ changes with θ. The first-order condition of an agent of type θ is

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

By differentiating the first-order condition at $\sigma(\theta)$, we can find an expression for $\sigma'(\theta)$:

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

where

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

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

$$\text{Sign}\left\{\frac{\partial\sigma(\theta)}{\partial\theta}\right\} = \text{Sign}\{h(\theta)\}.$$

Let us study the case of monotone retention first (see Panel A of Figure 5 in the main text). In this case the term involving x_+ in $h(\theta)$ disappears, and therefore

$$\text{Sign}\left\{\frac{\partial\sigma(\theta)}{\partial\theta}\right\} = \text{Sign}\left\{|x_- - \theta| - \sigma(\theta)\right\}.$$

Then, all we have to do is to compare $\sigma(\theta)$ and $|x_- - \theta|$. Remember that at $\theta = x_-$, $\sigma(\theta) = \underline{\sigma} > 0$, 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[4]. However, $\sigma(\theta)$ cannot grow unboundedly, since this would mean $c'(\sigma) \to \infty$, but $\phi(z)z \to 0$ as $|z| \to \infty$, so $\sigma(\theta)$ approaches $\underline{\sigma}$ from below as $\theta \to \infty$.

If we move away from x_- but in the opposite direction; that is, as we decrease θ, $\sigma(\theta)$ cannot always stay above $x_- - \theta$ because this would mean $\sigma(\theta) \to \infty$ as $\theta \to -\infty$, and we have just argued that this is inconsistent with optimality. 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 turn to the case of bounded retention, depicted in Panel B of Figure 5 in the main text. The symmetry of $\sigma(\theta)$ around the midpoint $\frac{x_- + x_+}{2}$ of the retention interval is evident from the first-order condition. If $x_+ < \infty$, $\sigma(\theta) = \sigma(x_- + x_+ - \theta)$ for all $\theta \leq \frac{x_- + x_+}{2}$. Let us therefore study the behavior of $\sigma(\theta)$ for $\theta \geq \frac{x_- + x_+}{2}$.

First, consider type $\theta = \frac{x_- + x_+}{2}$. By the symmetry and the existence of $\sigma'(\theta)$, at this point $\sigma'(\theta) = 0$. Effectively:

$$h\left(\frac{x_- + x_+}{2}\right) = 0.$$

For the type at the edge, $\theta = x_+$, the first-order condition is

$$-\phi\left(\frac{x_+ - x_-}{\sigma}\right)\frac{x_+ - x_-}{\sigma^2} = c'(\sigma), \tag{a.45}$$

so $\sigma(\theta) < \underline{\sigma}$. But his chosen σ is bigger than the one of the midpoint type: the marginal benefit at $\theta = \frac{x_+ + x_-}{2}$ is equal to

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

and it is smaller than the left-hand side of (a.45). Also, $\sigma'(\theta)$ is positive at $\theta = x_+$:

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

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

What about types above x_+? There exists a type $\underline{\theta} > x_+$ such that $\sigma(\underline{\theta}) = \underline{\sigma}$. For such a type, the value of σ that maximizes his probability of retention is equal to $\underline{\sigma}$, that is,

$$\phi\left(\frac{x_- - \underline{\theta}}{\underline{\sigma}}\right)\frac{x_- - \underline{\theta}}{\underline{\sigma}} - \phi\left(\frac{x_+ - \underline{\theta}}{\underline{\sigma}}\right)\frac{x_+ - \underline{\theta}}{\underline{\sigma}} = 0.$$

Observe that for type $\underline{\theta}$ the solution is represented in Panel B of Figure 3 by those z_1 and z_2 such that $\phi(z_1) z_1 = \phi(z_2) z_1$ and $z_2 - z_1 = \frac{x_+ - x_-}{\underline{\sigma}}$.

Now consider any type θ above $\underline{\theta}$ who considers playing $\underline{\sigma}$. His first-order derivative would be

$$\phi\left(\frac{\theta - x_+}{\underline{\sigma}}\right)\frac{\theta - x_+}{\underline{\sigma}} - \phi\left(\frac{\theta - x_-}{\underline{\sigma}}\right)\frac{\theta - x_-}{\underline{\sigma}}.$$

But Panel B of Figure 3 in the main text reveals that the sign of this expression will be positive, since by increasing θ we are considering bigger values of both z_1 and z_2. This means that $\sigma(\theta) > \underline{\sigma}\ \forall\theta > \underline{\theta}$. Similarly, for any type $\theta \in (x_+, \underline{\theta})$, the first-order derivative at $\underline{\sigma}$ will be negative, so $\sigma(\theta) < \underline{\sigma}$ for such θ.

Let us focus on types $\theta > \underline{\theta}$. For such types, since $\sigma(\theta) > \underline{\sigma}$, $\sigma c'(\sigma)$ is increasing in σ. This means that σ cannot grow unboundedly with θ, since that would mean $\sigma c'(\sigma) \to \infty$, whereas $\phi\left(\frac{\theta - x_+}{\sigma}\right)\frac{\theta - x_+}{\sigma} - \phi\left(\frac{\theta - x_-}{\sigma}\right)\frac{\theta - x_-}{\sigma}$ is a bounded function (each term is between 0 and 1).

Depart from $\theta = \underline{\theta}$. Notice that type $\underline{\theta}$ satisfies that $\underline{\sigma} \in [\theta - x_+, \theta - x_-]$ (take a look at Panel B of Figure 3 again to convince yourself: $z_1 < 1 < z_2$). This says that $h(\underline{\theta}) > 0$ so $\sigma'(\underline{\theta}) > 0$. Notice then that the distance $\frac{\theta - x_-}{\sigma(\theta)} - \frac{\theta - x_+}{\sigma(\theta)} = \frac{x_+ - x_-}{\sigma(\theta)}$ is decreasing in θ for values close to $\underline{\theta}$. Furthermore, it has to be the case that $\frac{\theta - x_+}{\sigma(\theta)}$ is increasing in θ, since otherwise $\frac{\theta - x_-}{\sigma(\theta)}$ would be decreasing (remember their distance decreases), and therefore $\phi\left(\frac{\theta - x_+}{\sigma(\theta)}\right)\frac{\theta - x_+}{\sigma(\theta)} - \phi\left(\frac{\theta - x_-}{\sigma(\theta)}\right)\frac{\theta - x_-}{\sigma(\theta)}$ would decrease (remember that $\frac{\theta - x_+}{\sigma(\theta)} < 1 < \frac{\theta - x_-}{\sigma(\theta)}$) at the same time that $\sigma(\theta) c'(\sigma(\theta))$ increases.

So as long as $\frac{\theta - x_+}{\sigma(\theta)} < 1$, $\frac{\theta - x_+}{\sigma(\theta)}$ is increasing and so it $\sigma(\theta)$ (see function $h(\theta)$). Furthermore, since $\sigma(\theta)$ is bounded, this means that $\frac{\theta - x_+}{\sigma(\theta)}$ eventually goes above 1. Function $h(\theta)$ indicates that as soon as this happens ($\frac{\theta - x_+}{\sigma(\theta)} = 1$), $\sigma(\theta)$ is still increasing with θ. Since we know $\sigma(\theta)$ is bounded and it converges to $\underline{\sigma}$ as $\theta \to \infty$, there exists a point θ at which $\sigma'(\theta) = h(\theta) = 0$. Now take a look at Figure A.8 in this document, which plots functions $\phi(z) z$ (the orange curve, related to the first-order derivative) and $\phi(z)\left(z^2 - 1\right)$ (the blue curve, related to function $h(\theta)$). $h(\theta) = 0$ means that there are two points on the x axis that reach the same height on the blue curve. The smaller point corresponds to $\frac{\theta - x_+}{\sigma(\theta)}$, and the larger point to $\frac{\theta - x_-}{\sigma(\theta)}$ But then, both $\frac{\theta - x_+}{\sigma(\theta)}$ and $\frac{\theta - x_-}{\sigma(\theta)}$ are increasing in θ at such a point (because $\sigma'(\theta)$ but θ increases), and therefore $h(\theta) < 0$ forever after: both $\frac{\theta - x_+}{\sigma(\theta)}$ and $\frac{\theta - x_-}{\sigma(\theta)}$ will always be increasing forever after because $\sigma(\theta)$ is decreasing and the numerators increase with θ.

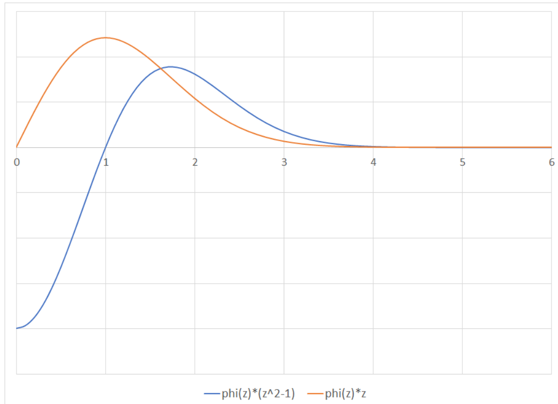


FIGURE A.8. Functions $\phi(z)\, z$ and $\phi(z)\left(z^2-1\right)$

3.2. **Sufficient Conditions for Uniqueness.** Condition U in the main text states that for every monotone or bounded retention zone, and for every agent type, the optimal choice of noise is unique. Here we show a condition on the cost function $c(\sigma)$ that guarantees the desired uniqueness.

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 :

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

This expression is related to the function $\phi(z)\, z\left(2-z^2\right)$, 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 $c''(\sigma) \geq \frac{\kappa}{\sigma^2}\ \forall\sigma$. Let us plot the $\phi(z)\, z\left(2-z^2\right)$ function in Figure A.9. 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:

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

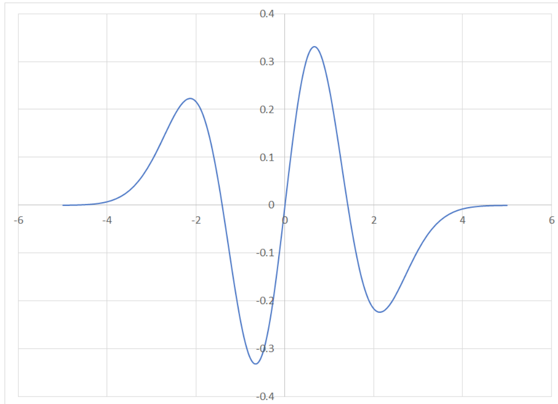


FIGURE A.9. Function $\phi(z) z \left(2 - z^2\right)$

Finally, to find the maximum value of $\phi(z_2) z_2 \left(2 - z_2^2\right) - \phi(z_1) z_1 \left(2 - z_1^2\right)$ 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

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