Supplemental Information

"Biased Judgments without Biased Judges: How Legal Institutions Cause Errors"

Ryan Hübert
University of California, Davis

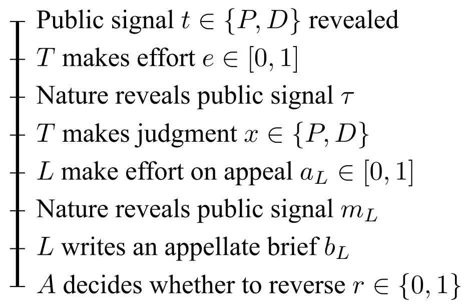
May 9, 2019

FOR ONLINE PUBLICATION

1 Formal Results

In the text of the paper, I describe the players, sequence, information structure and preferences of the players. Figure 1 depicts the sequence of the model.

Figure 1: The sequence of the model



I make the following assumptions.

Assumption A1 (no fabrication)**.** The case merits cannot be fabricated.

Assumption A2 (sufficient resource constraints)**.** For the litigants, $c_P > c_D > 1$. For the judge, $c_T > \bar{c}_T$, where $\bar{c}_T$ is defined in the Supplemental Information.

Assumption A3 (indifference)**.** When indifferent, the judge rules in favor of the defendant and the appellate court's reversal strategy favors the defendant.

Assumption A4 (damaging message)**.** L never writes a damaging brief. Formally, the action sets available for P and D are $b_P \in \{P, \phi\}$ and $b_D \in \{D, \phi\}$, respectively.

Lemma A1. Let μ_A be A's posterior belief that $\omega = P$ at its information set. In equilibrium, A's optimal reversal strategy is

$$r(\tau, x, b_L) = \begin{cases} 1 & \text{if } \left[\mu_A \leq \frac{1}{2} \text{ and } x = P\right] \text{ or } \left[\mu_A > \frac{1}{2} \text{ and } x = D\right] \\ 0 & \text{otherwise} \end{cases}$$

Proof of Lemma A1. This is directly from the requirement of perfect Bayesian equilibrium that players best respond to their beliefs. Since A prefers $x = \omega$ to $x \neq \omega$, then A reverses if and only if it believes $x \neq \omega$ is more probable than $x = \omega$. Since μ_A is A's equilibrium posterior that $\omega = P$, then it is optimal to reverse if and only if $\left[\mu_A \leq \frac{1}{2} \text{ and } x = P\right]$ or $\left[\mu_A > \frac{1}{2} \text{ and } x = D\right]$. □

Lemma A2. In equilibrium, $b_L = \omega$ if and only if $m_L = \omega \neq x$.

Proof of Lemma A2. Invoking Assumption A1 and Assumption A4, the only candidate brief is $m_L = \omega \neq x$. Now we verify if it is feasible and optimal. First, if $m_L = \phi$, then $m_L = \omega$ is not feasible since L has no information, and by Assumption A1, L cannot fabricate information. Next, if $m_L = \omega = L$, then $x \neq L$ and by Lemma A1, A reverses x. This makes L strictly better off. Therefore, $b_L^* = \omega$ if and only if $m_L = \omega \neq x$. □

Lemma A3. The litigant's optimal appeal effort is

$$a_L(\tau, x) = \begin{cases} \dfrac{\mu_P}{c_P} & \text{if } x = D \\ \dfrac{1-\mu_D}{c_D} & \text{if } x = P \\ 0 & \text{if } r(\tau, x, \phi) = 1 \end{cases} \tag{1}$$

where $\mu_L = \pi$ if $\tau = \phi$ and $\mu_L = \hat{\pi}$ if $\tau \in \{P, D\}$.

Proof of Lemma A3. It is straightforward to see that if $r(\tau, x, \phi) = 1$, then L secures reversal regardless of her brief and makes no costly effort to discover an error. At L's information set, let μ_L denote L's posterior belief that $\omega = P$.

Suppose $r(\tau, x, \phi) = 0$. Then by Lemma A1 and Lemma A2, L's interim expected utility is

$$U_P(a_P, \tau) = a_P\mu_P - \frac{1}{2}c_P a_P^2 \qquad U_D(a_D, \tau) = a_D(1-\mu_D) - \frac{1}{2}c_D a_D^2$$

Maximizing this with respect to a_P and a_D yields

$$a_P = \frac{\mu_P}{c_P} \qquad a_D = \frac{1-\mu_D}{c_D}$$

Note that the second derivative is negative, guaranteeing a_L is a maximum. Finally, since τ is a public signal, L's equilibrium belief is π if $\tau = \phi$ and $\hat{\pi}$ if $\tau \in \{P, D\}$. □

Lemma A4. Let μ_T be T's posterior belief that $\omega = P$ after observing τ. In equilibrium, T's judgment is as follows.

- If $r_D^\phi = 0$ and $r_P^\phi = 1$, then, $x(\tau) = D$.
- If $r_D^\phi = 1$ and $r_P^\phi = 0$, then, $x(\tau) = P$.
- If $r_D^\phi = r_P^\phi = 0$, then:

$$x(\tau) = \begin{cases} D & \text{if } \mu_T \leq \min\{\tilde{\mu}, \overline{\mu}\} \\ P & \text{if } \mu_T > \min\{\tilde{\mu}, \overline{\mu}\} \end{cases}$$

where $\tilde{\mu}$ and $\overline{\mu}$ are defined in the proof.

Proof of Lemma A4. To simplify notation, let $r_x^\phi := r(\mu_T, x, \phi)$. T rules for the defendant if and only if

$$(1-\mu_T)[\delta(1-r_D^\phi) - kr_D^\phi] + \mu_T(\delta - k)[a_P + (1-a_P)r_D^\phi]$$
$$\geq \mu_T[\delta(1-r_P^\phi) - kr_P^\phi] + (1-\mu_T)(\delta-k)[a_D + (1-a_D)r_P^\phi]$$

Consider each reversal strategy in turn:

If $r_D^\phi = r_P^\phi = 1$, then the condition collapses to $\mu_T \geq \frac{1}{2}$, a contradiction since $r(\mu_T, D, \phi) = 1$ implies $\mu_T < \frac{1}{2}$.

If $r_D^\phi = 1$ and $r_P^\phi = 0$, the condition collapses to $-k(1-(1-\mu_T)a_D) \geq (1-\mu_T)a_D\delta$, which never holds.

If $r_D^\phi = 0$ and $r_P^\phi = 1$, condition always holds.

Finally, consider $r_D^\phi = r_P^\phi = 0$. Using Lemma A3, this reduces to

$$(1-2\mu_T)\delta + (\delta - k)\left(\frac{\mu_T^2}{c_P} - \frac{(1-\mu_T)^2}{c_D}\right) \geq 0 \tag{2}$$

Note that the left hand side of (2) is strictly decreasing in μ_T and that it takes a positive value at $\mu_T = 0$ and a negative value at $\mu_T = 1$. To see that it is strictly decreasing, consider its derivative, which is strictly negative for all $c_P > c_D > 1$:

$$-2\delta\left[1 - \frac{(1-\mu_T)}{c_D} - \frac{\mu_T}{c_P}\right] - 2k\left[\frac{(1-\mu_T)}{c_D} + \frac{\mu_T}{c_P}\right] < 0$$

Then, there is a threshold $\tilde{\mu} \in (0,1)$ such that (2) holds with equality and where $x = D$ if and only if $\mu_T \leq \tilde{\mu}$.

$$\tilde{\mu} = \begin{cases} \dfrac{\sqrt{c_D c_P((c_D-1)\delta+k)((c_P-1)\delta+k)} - c_P(\delta(c_D-1)+k)}{(c_P-c_D)(\delta-k)} & \text{if } \delta > k \\ \dfrac{1}{2} & \text{if } \delta = k \\ \dfrac{-\sqrt{c_D c_P((c_D-1)\delta+k)((c_P-1)\delta+k)} + c_P(\delta(c_D-1)+k)}{(c_P-c_D)(k-\delta)} & \text{if } \delta < k \end{cases} \tag{3}$$

Next, we verify that it is sequentially rational for A to affirm T's decision when L does not provide an informative brief. Using Assumption A1 so that L cannot fabricate information, affirming is a best response to the following beliefs:

$$\underbrace{\frac{\mu_T}{\mu_T + (1-\mu_T)(1-a_D)}}_{\mu_A(\tau,P,\phi)} \geq \frac{1}{2} \qquad \underbrace{\frac{\mu_T(1-a_P)}{\mu_T(1-a_P)+1-\mu_T}}_{\mu_A(\tau,D,\phi)} \leq \frac{1}{2}$$

Substituting and rearranging yields

$$\mu_T \geq \sqrt{(c_D-1)c_D} - (c_D-1) =: \underline{\mu} < \frac{1}{2} \qquad \mu_T \leq c_P - \sqrt{(c_P-1)c_P} =: \overline{\mu} > \frac{1}{2}$$

It is straightforward to verify that $\tilde{\mu} > \underline{\mu}$ for all c_D, c_P. However, there may exist $\tilde{\mu} \geq \overline{\mu}$ for some c_D, c_P. Assuming that T finds it optimal to avoid reversal, then $x(\tau) = D$ if and only if $\mu_T \leq \tilde{\mu}$ and $\mu_T \leq \overline{\mu}$. Finally, we verify that for all $\mu_T \in [\overline{\mu}, \tilde{\mu}]$ (if such an interval exists), it is indeed optimal that T sets $x = P$ in order to avoid reversal. Suppose, by contradiction that there is a profitable deviation to $x' = D$. Then, A reverses in the absence of information, and T suffers a cost k and ends up with an outcome P. If instead she sets $x = P$ as the equilibrium requires, then she achieves the same outcome without the reversal cost. □

Lemma A5. In equilibrium, T exerts effort according to the following. If $r(\tau, P, \phi) = r(\tau, D, \phi) = 0$ and $\varepsilon < \overline{\varepsilon}$ (where $\overline{\varepsilon}$ is defined in the proof), then:

$$e(x) = \begin{cases} \frac{(1-\pi)\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{(1-\pi)^2}{c_D}\right) - \varepsilon\left(\frac{\pi^2}{c_P}\right)\right] & \text{if } x_\phi = P \\ \frac{\pi\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{\pi^2}{c_P}\right) - \varepsilon\left(\frac{(1-\pi)^2}{c_D}\right)\right] & \text{if } x_\phi = D \end{cases} \tag{4}$$

Otherwise, $e(x) = 0$ for $x \in \{P, D\}$.

Proof of Lemma A5. First, note that if x does not depend on τ, there is no incentive to exert effort since e only improves the informativeness of τ. Thus, if $r_x^\phi = 1$ for some x or if $x = P, D$ for all τ, then $e = 0$.

Consider the remaining case where $x = \tau$ and $r_P^\phi = r_D^\phi = 0$. Then, T's ex ante expected utility when $x = \tau = P$ is given by

$$e[(1-\varepsilon)\delta + \varepsilon(\delta - k)(\pi a_P + (1-\pi)a_D)] + (1-e)\left[\pi\delta - (1-\pi)a_D(k-\delta)\right] - \frac{c_T}{2}e^2$$

Substituting and optimizing yields:

$$e(x = P; \cdot) = \frac{(1-\pi)\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{(1-\pi)^2}{c_D}\right) - \varepsilon\left(\frac{\pi^2}{c_P}\right)\right] \tag{5}$$

When $x = \tau = D$, the corresponding ex ante expected utility is

$$e[(1-\varepsilon)\delta + \varepsilon(\delta - k)(\pi a_P + (1-\pi)a_D)] + (1-e)\left[(1-\pi)\delta - \pi a_P(k-\delta)\right] - \frac{c_T}{2}e^2$$

Substituting and optimizing yields:

$$e(x = D; \cdot) = \frac{\pi\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{\pi^2}{c_P}\right) - \varepsilon\left(\frac{(1-\pi)^2}{c_D}\right)\right] \tag{6}$$

Finally, we characterize the condition required for x to depend on τ when $r_P^\phi = r_D^\phi = 0$. From Lemma A4, for τ to change T's judgment:

$$\Pr(\omega = P|\tau = P) = \frac{(1-\varepsilon)\pi}{(1-\varepsilon)\pi + \varepsilon(1-\pi)} > \tilde{\mu} \iff \varepsilon < \frac{\pi(1-\tilde{\mu})}{\tilde{\mu} - \pi(2\tilde{\mu}-1)}$$
$$\Pr(\omega = P|\tau = D) = \frac{\varepsilon\pi}{\varepsilon\pi + (1-\varepsilon)(1-\pi)} \leq \tilde{\mu} \iff \varepsilon \leq \frac{\tilde{\mu}(1-\pi)}{\tilde{\mu} - \pi(2\tilde{\mu}-1)}$$

Let $\overline{\varepsilon} := \min\left\{\frac{\tilde{\mu}(1-\pi)}{\tilde{\mu}-\pi(2\tilde{\mu}-1)}, \frac{\pi(1-\tilde{\mu})}{\tilde{\mu}-\pi(2\tilde{\mu}-1)}\right\}$. Then, setting aside knife edge conditions, the equilibrium effort is defined by (5) and (6) if and only if $r_P^\phi = r_D^\phi = 0$ and $\varepsilon < \overline{\varepsilon}$. This completes the proof. □

The previous result describes two scenarios where the trial judge declines to exert effort. First, if the appellate court uses a reversal strategy that requires a specific judgment, then there is no value to acquiring additional information. However, when the appellate court uses a fully deferential reversal strategy, there may still be scenarios where the trial judge does not exert effort. Specifically, in some region of the parameter space (e.g. very high or very low π), no effort will be made because the public signal τ does not provide accurate enough information that the trial judge would change her judgment even if $\tau \in \{P, D\}$.

The following assumption rules out equilibria where the appellate court's reversal strategy discourages effort by the trial judge in all situations. Such equilibria are substantively implausible and they discourage information acquisition by trial judges.

Assumption A5 (deference to trial judge)**.** If it is sequentially rational for the appellate court to defer to the trial judge's decision, it does so. Formally, if $\mu_A > \frac{1}{2}$ and $x = P$ or if $\mu_A \leq \frac{1}{2}$ and $x = D$, then $r = 0$.

Proposition A1. There is a unique perfect Bayesian equilibrium that satisfies Assumption A3 and Assumption A5. It is characterized by the equilibrium strategies and equilibrium beliefs in Lemmas A1 to A5. Where relevant, L and A's off-equilibrium path beliefs are formed by application of Bayes' rule to public signals t and τ.

Proof of Proposition A1. A perfect Bayesian equilibrium is a sequentially rational strategy profile and consistent beliefs. Existence of a sequentially rational profile follows directly from Lemmas A1 to A5, which characterize the player's equilibrium strategies. Relevant beliefs are over the merits, ω. Whenever ω is revealed, all player's beliefs become degenerate and subsequent subgames no longer feature imperfect information. The equilibrium beliefs of T and L are straight forward since T's information is publicly revealed. Specifically, they collapse to either the prior π (if $\tau = \phi$) or to $\hat{\pi}$ (if $\tau \in \{P, D\}$). A's belief when $\tau = b_L = \phi$ is more complicated, since L may be concealing information. The proof of Lemma A4 characterizes the equilibrium belief, which is consistent since it is formed by Bayes' rule using the players' equilibrium strategies.

We now consider off-equilibrium path beliefs. First, consider deviations by L. If $m_L = \phi$, then there is no deviation possible. If $m_L = \omega$, there are two possible deviations: $b_L = \omega = x$ or $b_L = \phi$. Deviations of the former type do not require we specify off-equilibrium path beliefs since all uncertainty over ω is resolved. Deviations of the latter type are not off the equilibrium path since $b_L = \phi$ is a best response to $m_L = \omega = x$.

Next, consider deviations by T. If $x^* = P$ but $x = D$ or $x^* = D$ but $x = P$, then L and A must form beliefs about ω. Since T's information is public, deviations by T do not affect what L or A know about τ. Since it is knowledge of τ that is relevant and both $\tau = P$ and $\tau = D$ are on the equilibrium path, I assume that L and A's belief after a deviation by T are formed by application of Bayes' rule to t and τ.

Finally, uniqueness follows directly from application of Assumption A3 and Assumption A5 since Assumption A3 rules out multiple equilibria induced by indifference and Assumption A5 rules out multiple equilibria driven by the fact that A has multiple sequentially rational review strategies. □

Lemma A6. If $\varepsilon < \bar{\varepsilon}$, then T's prior belief about ω changes, her level of effort changes according to:

$$\frac{\partial e(x=P)}{\partial \pi} = -\frac{\delta}{c_T} - \frac{2(k-\delta)}{c_T}\left(\frac{\varepsilon\pi}{c_P} + \frac{(1-\varepsilon)(1-\pi)}{c_D}\right) < 0$$
$$\frac{\partial e(x=D)}{\partial \pi} = \frac{\delta}{c_T} + \frac{2(k-\delta)}{c_T}\left(\frac{(1-\varepsilon)\pi}{c_P} + \frac{\varepsilon(1-\pi)}{c_D}\right) > 0$$

Otherwise if $\varepsilon \geq \bar{\varepsilon}$, $e(x) = 0$ and $\frac{\partial e(x)}{\partial \pi} = 0$.

Proof. First, let $\varepsilon < \bar{\varepsilon}$. Then using Assumption A5 and Lemma A5, $\frac{\partial e(x)}{\partial \pi}$ follows directly from the first derivative of (4). Next, note the following

- If $k \geq \delta$, then $\frac{\partial e(x=P)}{\partial \pi} < 0$.
- If $k \geq \delta$, then $\frac{\partial e(x=D)}{\partial \pi} > 0$.
- If $\delta > k$ and $\varepsilon < \bar{\varepsilon}$, then $\frac{\partial e(x=P)}{\partial \pi} < 0$.
- If $\delta > k$ and $\pi \leq \tilde{\mu}$, then $\frac{\partial e(x=D)}{\partial \pi} > 0$.

Next, let $\varepsilon \geq \bar{\varepsilon}$. Then, by Lemma A5, $e(x) = 0$ and trivially, $\frac{\partial e(x)}{\partial \pi} = 0$. This proves the result. □

Definition A1. A trial judge's decision rule is **impartial** if and only if $x = D \Leftrightarrow \mu_T \leq \frac{1}{2}$.

Proposition A2. The trial judge's decision rule has the following properties:

- If $k < \delta$, then her judgments are biased in favor of the less powerful litigant, $\tilde{\mu} < \frac{1}{2}$.
- If $k > \delta$, then her judgments are biased in favor of the more powerful litigant, $\tilde{\mu} > \frac{1}{2}$.
- If $k = \delta$, then her decision rule is impartial, $\tilde{\mu} = \frac{1}{2}$.

Moreover, the bias in her decision rule becomes weakly larger as $|\delta - k|$ increases.

Proof of Proposition A2. Recall from Lemma A4 (and using Assumption A5) that T's decision rule is to rule for the defendant if and only if $\mu_T \geq \tilde{\mu}$. It follows directly from the definition of $\tilde{\mu}$ in the proof of Lemma A4 that $\tilde{\mu} < \frac{1}{2}$ if $\delta > k$, $\tilde{\mu} > \frac{1}{2}$ if $k > \delta$ and $\tilde{\mu} = \frac{1}{2}$ if

$k = \delta$. Note that $\tilde{\mu} < \frac{1}{2}$ implies that T rules for the plaintiff (the less powerful litigant) more often than under an impartial decision rule and $\tilde{\mu} > \frac{1}{2}$ implies that T rules for the defendant (the more powerful litigant) more often than an impartial judge.

Finally, we show that this bias increases as $|\delta - k|$ increases. First note that if $\tilde{\mu} \notin [\underline{\mu}, \overline{\mu}]$, then she is constrained in her decision making by Lemma A4 and uses a decision rule of the form $\mu_T \leq \overline{\mu} \Leftrightarrow x = D$. Since $\overline{\mu}$ does not depend on δ or k, then it is not affected as $|\delta - k|$ increases.

Now suppose $\tilde{\mu} \in [\underline{\mu}, \overline{\mu}]$. Recall the condition that defines $\tilde{\mu}$ from Lemma A4

$$(1 - 2\tilde{\mu})\delta + (\delta - k)\left(\frac{\tilde{\mu}^2}{c_P} - \frac{(1-\tilde{\mu})^2}{c_D}\right) = 0 \tag{7}$$

Case I. Suppose $\delta > k$. From above, $\tilde{\mu} < \frac{1}{2}$, so the left hand side of (7) is strictly increasing in k and strictly decreasing in $\tilde{\mu}$. Then as k decreases (and $|\delta - k|$ increases), $\tilde{\mu}$ must decrease in order for the condition to continue to hold. Therefore the decision rule is becoming more biased in favor of the less powerful litigant.

Case II. Suppose $k > \delta$. From above, $\tilde{\mu} > \frac{1}{2}$, so the left hand side of (7) is strictly decreasing in both δ and $\tilde{\mu}$. Then as δ decreases (and $|\delta - k|$ increases), $\tilde{\mu}$ must increase in order for the condition to continue to hold. Therefore the decision rule is becoming more biased in favor of the more powerful litigant. □

Proposition A3. If $\delta \neq k$, then the trial judge's equilibrium effort is weakly lower than if she used an impartial decision rule. Moreover, it is strictly lower for all $\mu_T \in (\max\{\frac{1}{2}, \tilde{\mu}\}, \min\{\frac{1}{2}, \tilde{\mu}\}]$.

Proof of Proposition A3. Recall from Lemma A5 that T's effort (when non-zero) is given by:

$$e(x = P) = \frac{(1-\pi)\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{(1-\pi)^2}{c_D}\right) - \varepsilon\left(\frac{\pi^2}{c_P}\right)\right]$$

and

$$e(x = D) = \frac{\pi\delta}{c_T} - \frac{\varepsilon\delta}{c_T} + \frac{(k-\delta)}{c_T}\left[(1-\varepsilon)\left(\frac{\pi^2}{c_P}\right) - \varepsilon\left(\frac{(1-\pi)^2}{c_D}\right)\right]$$

Moreover, recall from Lemma A4 that for all $\tilde{\mu} < \mu_T < 1$, $x = P$ and for all $0 < \mu_T \leq \tilde{\mu}$, $x = D$.

Case I. If $\mu_T \leq \min\{\frac{1}{2}, \tilde{\mu}\}$, $\mu_T > \max\{\frac{1}{2}, \tilde{\mu}\}$, or $\delta = k$, then T's equilibrium judgment $x(\tau)$ is the same as the judgment generated by an impartial decision rule. Then, effort is identical under either decision rule.

Case II. If $\delta > k$ and $\tilde{\mu} < \mu_T \leq \frac{1}{2}$, then T's equilibrium judgment is $x(\tau) = P$ whereas an impartial judgment would be $x = D$. Moreover, it is straight forward to verify that $e(x = D) > e(x = P)$ when $\tilde{\mu} < \mu_T \leq \frac{1}{2}$.

Case III. If $\delta < k$ and $\frac{1}{2} < \mu_T \leq \tilde{\mu}$, then T's equilibrium judgment is $x(\tau) = D$ whereas an impartial judgment would be $x = P$. Moreover, it is straight forward to verify that $e(x = P) > e(x = D)$ when $\frac{1}{2} < \mu_T < \tilde{\mu}$ and $e(x = D) = e(x = P)$ when $\mu_T = \tilde{\mu}$. □

Proposition A4. Litigant-driven appellate review has the following effect on the trial judge's equilibrium effort:

- If $k < \delta$, then her effort is strictly lower than without litigant-driven appellate review.

- If $k > \delta$, then her effort is strictly higher than without litigant-driven appellate review.

Proof. If not subjected to appellate review, then $x = D$ if and only if $\mu_T < \frac{1}{2}$, and effort is determined by maximizing

$$U_T^{\text{no}}(e, \cdot) = \begin{cases} e((1-\varepsilon)\delta) + (1-e)(\pi\delta) - \dfrac{c_T}{2}e^2 & \text{if } x_\phi = P \\ e((1-\varepsilon)\delta) + (1-e)((1-\pi)\delta) - \dfrac{c_T}{2}e^2 & \text{if } x_\phi = D \end{cases}$$

This yields an optimal level of effort for T when not subjected to review:

$$e^{\text{no}} = \begin{cases} \dfrac{(1-\pi-\varepsilon)\delta}{c_T} & \text{if } x_\phi = P \\ \dfrac{(\pi-\varepsilon)\delta}{c_T} & \text{if } x_\phi = D \end{cases} \tag{8}$$

Comparing this with (4) and given the assumptions on ε, c_P, c_D and c_T, it is it is immediate to see that (8) is strictly larger than (4) for all δ, k such that $\delta > k$. $\square$

Lemma A7. Let $e(\delta)$ be the equilibrium effort of T as a function of δ. If $\varepsilon < \tilde{\varepsilon}$ (as defined in the proof), then $e(\bar{\delta}) > e(0)$.

Proof of Lemma A7. There are two cases to consider.

Case I. Suppose $\mu_T \leq \min\{\frac{1}{2}, \tilde{\mu}\}$ or $\mu_T > \max\{\frac{1}{2}, \tilde{\mu}\}$. Then both kinds of judge issue the same judgment. First, suppose $e(x) > 0$, then the derivative with respect to δ is

$$\frac{\partial e(x)}{\partial \delta} = \begin{cases} \dfrac{1}{c_T}\left(1 - \pi - \varepsilon - (1-\varepsilon)\dfrac{(1-\pi)^2}{c_D} + \varepsilon\dfrac{\pi^2}{c_P}\right) & \text{if } x_\phi = P \\ \dfrac{1}{c_T}\left(\pi - \varepsilon - (1-\varepsilon)\dfrac{\pi^2}{c_P} + \varepsilon\dfrac{(1-\pi)^2}{c_D}\right) & \text{if } x_\phi = D \end{cases}$$

Both of these indicate that $e(x)$ linearly increases or decreases in δ. Next, note that $\frac{\partial e(x)}{\partial \delta} > 0$ when $\varepsilon = 0$ and that

$$\frac{\partial}{\partial \varepsilon}\left[\frac{\partial e(x)}{\partial \delta}\right] = \frac{1}{c_T}\left(-1 + \frac{(1-\pi)^2}{c_D} + \frac{\pi^2}{c_P}\right) < 0$$

Therefore, the positive relationship between δ and $e(x)$ is decreasing as ε increases away from zero. Then, there exists thresholds $\tilde{\varepsilon}_1(x)$ such that for all $\varepsilon < \min\{\tilde{\varepsilon}_1(P), \tilde{\varepsilon}_1(D)\}$, $\frac{\partial e(x)}{\partial \delta} > 0$.

Finally notice that since $\frac{\partial e(x)}{\partial c_T} < 0$ if $e > 0$. Then, if $\varepsilon < \min\{\tilde{\varepsilon}_1(P), \tilde{\varepsilon}_1(D)\}$, the decrease in c_T associated with an increase in δ simply exacerbates the positive relationship between δ and e.

Case II. Suppose $\min\{\frac{1}{2}, \tilde{\mu}\} < \mu_T \leq \max\{\frac{1}{2}, \tilde{\mu}\}$. By Lemma A6, $e^*(\pi, \delta_i = \delta)$ is strictly decreasing and $e^*(\pi, \delta_i = 0)$ is strictly increasing for all $\pi \in (\tilde{\mu}, \tilde{\mu}^0)$. Then to show that $e^*(\pi, \delta_i = \delta) > e^*(\pi, \delta_i = 0)$ in this interval, it suffices to show that $e^*(\tilde{\mu}^0, \delta_i = \delta) > e^*(\tilde{\mu}^0, \delta_i = 0)$:

$$\frac{(1-\tilde{\mu}^0-\varepsilon)\delta}{c_L} + \frac{(k-\delta)}{c_L}\left[(1-\varepsilon)\left(\frac{(1-\tilde{\mu}^0)^2}{c_D}\right) - \varepsilon\left(\frac{(\tilde{\mu}^0)^2}{c_P}\right)\right]$$
$$> \frac{k}{c_H}\left[(1-\varepsilon)\left(\frac{(\tilde{\mu}^0)^2}{c_P}\right) - \varepsilon\left(\frac{(1-\tilde{\mu}^0)^2}{c_D}\right)\right]$$

The condition holds for all $\varepsilon < \frac{c_P - \pi c_P - \pi^2}{c_P - \pi^2 - \pi^2} := \tilde{\varepsilon}_2$.

Finally define $\tilde{\varepsilon} := \min\{\tilde{\varepsilon}_1(P), \tilde{\varepsilon}_1(D), \tilde{\varepsilon}_2\}$. Then, we have established that for all $\varepsilon < \tilde{\varepsilon}$, $e(\bar{\delta}) > e(0)$. $\square$

We now assume that $\varepsilon = 0$. Then, define the accuracy of outcomes by

$$\xi(x_\phi, \delta, c_T) = \begin{cases} e(x_\phi, \delta, c_T) + [1 - e(x_\phi, \delta, c_T)][1 - \pi(1 - a_P)] & \text{if } x_\phi = D \\ e(x_\phi, \delta, c_T) + [1 - e(x_\phi, \delta, c_T)][1 - (1 - \pi)(1 - a_D)] & \text{if } x_\phi = P \end{cases} \tag{9}$$

Lemma A8. There are more errors when issue 2 judges hear issue 1 cases than when issue 1 judges hear issue 1 cases, and vice versa. Formally, $\xi(x_\phi, \bar{\delta}, c_L) > \xi(x_\phi, 0, c_H)$.

Proof of Lemma A8. Again, there are two cases to consider.

Case I. Suppose $\mu_T \leq \min\{\frac{1}{2}, \tilde{\mu}\}$ or $\mu_T > \max\{\frac{1}{2}, \tilde{\mu}\}$. Both types of judges make the same judgment. From lemma A7 and assuming $\varepsilon = 0$, $e(\bar{\delta}) > e(0)$ and it is direct to see from (9) that $\xi(x_\phi, \bar{\delta}, c_L) > \xi(x_\phi, 0, c_H)$.

Case II. By contradiction, suppose that $\xi(x_\phi = P, \delta)$ were weakly less than $\xi(x_\phi = D, 0)$ at some point in the interval $[\tilde{\mu}, \tilde{\mu}^0]$. Using the fact that, when uninformed, a $\delta_i = 0$ judge rules for the defendant and a $\delta_i = \delta$ judge rules for the plaintiff, we can rewrite the *ex ante* probability as follows:

$$\xi(x_\phi, \delta_i) = \begin{cases} e(x = D, 0) + [1 - e(x = D, 0)][1 - \pi(1 - a_P)] & \text{if } \delta_i = 0 \\ e(x = P, \bar{\delta}) + [1 - e(x = P, \bar{\delta})][1 - (1 - \pi)(1 - a_D)] & \text{if } \delta_i = \delta \end{cases}$$

By Lemma A7 and assuming $\varepsilon = 0$, $e(x_\phi = P, \delta) > e(x_\phi = D, 0)$. In order for $\xi(x_\phi = P, \bar{\delta}) \leq \xi(x_\phi = D, 0)$ for some $\pi \in [\tilde{\mu}, \tilde{\mu}^0]$ as conjectured, it must be that:

$$1 - \pi(1 - a_P) > 1 - (1 - \pi)(1 - a_D)$$

Substituting equilibrium values and simplifying yields:

$$1 > \frac{(1-\pi)^2}{c_D} - \frac{\pi^2}{c_P} + 2\pi$$

The right hand side is increasing in π, and thus is at its smallest in the relevant interval when $\pi = \tilde{\mu}$:

$$1 > \frac{(1-\tilde{\mu})^2}{c_D} - \frac{\tilde{\mu}^2}{c_P} + 2\tilde{\mu}$$

However, the right hand side is strictly greater than one and the condition fails, a contradiction.

□

In the following, define:

$$\overline{\xi} \coloneqq \xi(x_\phi, \overline{\delta}, c_L) \qquad\qquad \underline{\xi} \coloneqq \xi(x_\phi, 0, c_H)$$

Proposition A5. Random assignment of judges to cases leads to strictly fewer accurate decisions than voluntary assignment. Formally, $R < V$.

Proof of Proposition A5. This is direct:

$$\begin{aligned} & R < V \\ & \iff pq\overline{\xi} + p(1-q)\underline{\xi} + (1-p)q\underline{\xi} + (1-p)(1-q)\overline{\xi} \\ & \qquad < (1-m)\overline{\xi} + m\underline{\xi} \\ & \iff (1-p-q+2pq)\overline{\xi} + (q+p-2pq)\underline{\xi} < (1-m)\overline{\xi} + m\underline{\xi} \end{aligned}$$

If $m = q - p$, this reduces to

$$\underbrace{2p(1-q)}_{+}\underbrace{(\underline{\xi} - \overline{\xi})}_{-} < 0$$

If $m = p - q$, this reduces to

$$\underbrace{2q(1-p)}_{+}\underbrace{(\underline{\xi} - \overline{\xi})}_{-} < 0$$

Thus, we have shown that $R < V$. □

2 Trial Judge's Interior Effort

For T to exert interior effort in equilibrium, the following must hold:

$$c_T > \begin{cases} \dfrac{1}{c_P}\left[\pi(c_P - \pi)\delta + \pi^2 k\right] & \text{if } x_\phi = D \\ \dfrac{1}{c_D}\left[(1-\pi)(c_D - (1-\pi))\delta + (1-\pi)^2 k\right] & \text{if } x_\phi = P \end{cases} \tag{10}$$

We define a function $\tilde{c}_T(\pi)$, which returns a threshold value of c_T such that equation (10) holds with equality:

$$\tilde{c}_T(\pi) = \begin{cases} \dfrac{\pi(c_P - \pi)\delta + \pi^2 k}{c_P} & \text{if } \pi \leq \tilde{\mu} \\ \dfrac{(1-\pi)(c_D - (1-\pi))\delta + (1-\pi)^2 k}{c_D} & \text{if } \pi \geq \tilde{\mu} \end{cases}$$

There are three properties of this function worth noticing. First, the inequalities in the conditionals are weak. This is because the two components of the piecewise function are equal when evaluated at $\tilde{\mu}$. Second, $\tilde{c}_T(\pi)$ is strictly increasing when $\pi < \tilde{\mu}$ and strictly decreasing when $\pi > \tilde{\mu}$. Thus, $\tilde{c}_T(\pi)$ achieves its global maximum at $\tilde{\mu}$. Using these facts, in order for T's effort to be interior in in any equilibrium, $c_T > \bar{c}_T$, where:

$$\bar{c}_T \equiv \tilde{c}_T(\pi = \tilde{\mu}) = \frac{c_P \delta \tilde{\mu} - (\delta - k)\tilde{\mu}^2}{c_P} \tag{11}$$

and $\tilde{\mu}$ is given by equation (3). For the case of a biased judge, replace δ with β in Equation (11).