Online Appendix

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

Ryan Hübert
University of California, Davis

November 5, 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

- Public signal $t \in \{P, D\}$ revealed
- T makes effort $e \in [0, 1]$
- Nature reveals public signal $\tau \in \{\hat{P}, \hat{D}, \phi\}$
- T makes judgment $x \in \{P, D\}$
- L make effort on appeal $a_L \in [0, 1]$
- Nature sends private signal $m_L \in \{\omega, \phi\}$ to L
- L writes an appellate brief $b_L \in \{m_L, \phi\}$
- A decides whether to reverse $r \in \{0, 1\}$

I make the following assumptions.

> **Assumption A1** (no fabrication)**.** Information about the case merits cannot be fabricated.

> **Assumption A2** (litigant costs)**.** The litigants' costs are sufficiently high. Formally, $c_P > c_D > 1$.

> **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 brief)**.** The appealing litigant never writes a brief that undermines its own case. Formally, $b_P \in \{P, \phi\}$ and $b_D \in \{D, \phi\}$.

Let π_τ be the players' posterior belief upon observing τ, which is a public signal. Then:

$$\pi_P = \frac{(1-\varepsilon)\pi}{(1-\varepsilon)\pi + \varepsilon(1-\pi)} > \pi \qquad \pi_D = \frac{\varepsilon(1-\pi)}{\varepsilon(1-\pi) + (1-\varepsilon)\pi} < \pi \qquad \pi_\phi = \pi$$

1.1 Appeals

A's reversal strategy depends on its belief induced by τ, x and b_L. To simplify notation in the following analysis, I denote A's reversal decision after observing τ, x and b_L as

$$r_x^{\phi}(\tau) \equiv r(\tau, x, b_L = \phi) \qquad\qquad r_x(\omega) \equiv r(\tau, x, b_L = \omega).$$

Lemma A1. Let $\mu(\tau, x, b_L)$ denote A's posterior belief that $\omega = P$. Then, A's best response is characterized by

$$r_P^{\phi}(\tau) = \begin{cases} 1 & \text{if } \mu(\tau, P, \phi) \leq 1/2 \\ 0 & \text{if } \mu(\tau, P, \phi) > 1/2 \end{cases}$$

$$r_D^{\phi}(\tau) = \begin{cases} 0 & \text{if } \mu(\tau, D, \phi) \leq 1/2 \\ 1 & \text{if } \mu(\tau, D, \phi) > 1/2 \end{cases}$$

$$r_x(\omega) = \begin{cases} 1 & \text{if } x \neq b_L = \omega \\ 0 & \text{otherwise} \end{cases}$$

On the equilibrium path, $\mu(\tau, x, b_L)$ is formed by Bayes' rule as follows. If $b_L = \phi$:

$$\mu(\tau, x, \phi) = \begin{cases} \dfrac{\pi_\tau}{\pi_\tau + (1 - \pi_\tau)(1 - a_D)} & \text{if } x = P \\ \dfrac{\pi_\tau(1 - a_P)}{\pi_\tau(1 - a_P) + 1 - \pi_\tau} & \text{if } x = D \end{cases}$$

And if $b_L = \omega$: $\mu(\tau, D, \omega) = 0$ and $\mu(\tau, P, \omega) = 1$.

Proof of Lemma A1. 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 $x = \omega$. We consider two cases:

Case 1. Suppose $b_L = \omega$. Then, A's belief over ω is degenerate: $\mu(\tau, D, \omega) = 0$ and $\mu(\tau, P, \omega) = 1$. It follows directly that A is strictly better off reversing if and only if $b_L = \omega \neq x$.

Case 2. Suppose $b_L = \phi$. Then, A's equilibrium posterior belief that $\omega = P$ is formed by

Bayes' rule:

$$\mu(\tau, x, \phi) = \begin{cases} \dfrac{\pi_\tau}{\pi_\tau + (1-\pi_\tau)(1-a_D)} & \text{if } x = P \\ \dfrac{\pi_\tau(1-a_P)}{\pi_\tau(1-a_P) + 1 - \pi_\tau} & \text{if } x = D \end{cases}$$

Using Assumption A3, reversal is a best response for A if and only if $\mu(\tau, P, \phi) \leq 1/2$ and $\mu(\tau, D, \phi) > 1/2$. □

Lemma A2. In equilibrium, L's brief is

$$b_L(m_L) = \begin{cases} \omega & \text{if } m_L = \omega = L \\ \phi & \text{otherwise} \end{cases}$$

Proof of Lemma A2. Consider the three possible cases, $m_L = \phi$, $m_L = L$ and $m_L \neq L$. First, if $m_L = \phi$, then $b_L = \phi$ since by Assumption A1, L cannot fabricate information. Second, if $m_L = \omega \neq L$, then by Assumption A4, $b_L = \phi$ since L never writes a damaging brief. Finally, if $m_L = \omega = L$, then by Lemma A1, A is strictly better off reversing if $b_L = m_L$. This also makes L strictly better off. Therefore, $b_L = \omega$. □

Lemma A3. The litigant's optimal appeal effort is

$$a_D(\tau) = \begin{cases} \dfrac{1-\pi_\tau}{c_D} & \text{if } r_P^\phi(\tau) = 0 \\ 0 & \text{if } r_P^\phi(\tau) = 1 \end{cases} \qquad a_P(\tau) = \begin{cases} \dfrac{\pi_\tau}{c_P} & \text{if } r_D^\phi(\tau) = 0 \\ 0 & \text{if } r_D^\phi(\tau) = 1 \end{cases}$$

Proof of Lemma A3. It is straightforward to see that if $r_x^\phi(\tau) = 1$, then L secures reversal regardless of her brief and makes no costly effort to discover an error.

Suppose $r_x^\phi(\tau) = 0$. A reverses if $b_L = \omega \neq x$ (by Lemma A1), and L writes an informative brief if $m_L = \omega \neq x$ (by Lemma A2). Then, L's expected utility is

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

Maximizing with respect to a_P and a_D yields the following:

$$a_P(\tau) = \frac{\pi_\tau}{c_P} \qquad\qquad a_D(\tau) = \frac{1-\pi_\tau}{c_D} \tag{1}$$

Since the second derivatives are negative, this guarantees $a_P(\tau)$ and $a_D(\tau)$ maximize U_P and U_D, respectively. □

1.2 Trials

Every realization of τ induces an isomorphic continuation game. Next, I characterize the players' equilibrium strategies for a generic continuation game induced by τ. Following this, I will characterize T's optimal effort choice given the equilibria of the continuation games induced by τ.

Lemma A4. Fix τ, and let $\underline{\pi}$, $\overline{\pi}$ and $\tilde{\pi}$ be defined as in the proof. Then, in every equilibrium, $r_x(\omega)$ and b_L are as in Lemmas A1 and A2. Moreover:

- If $\pi_\tau > 1/2$, then there is an equilibrium in which $r_D^\phi(\tau) = 1$ and $r_P^\phi(\tau) = 0$, $a_D = (1-\pi_\tau)/c_D$, and $a_P = 0$ and $x(\tau) = P$ for all τ. A's equilibrium belief is as in Lemma A1, and its belief off the equilibrium path is $\mu^{\text{off}}(\tau, D, \phi) > \frac{1}{2}$.
- If $\pi_\tau \leq 1/2$, then there is an equilibrium in which $r_D^\phi(\tau) = 0$ and $r_P^\phi(\tau) = 1$, $a_D = 0$, $a_P = \pi_\tau/c_P$ and $x(\tau) = D$ for all τ. A's equilibrium belief is as in Lemma A1, and its belief off the equilibrium path is $\mu^{\text{off}}(\tau, P, \phi) \leq \frac{1}{2}$.
- If $\underline{\pi} < \pi_\tau \leq \overline{\pi}$, then there is an equilibrium in which $r_D^\phi(\tau) = r_P^\phi(\tau) = 0$, $a_D = (1-\pi_\tau)/c_D$, $a_P = \pi_\tau/c_P$ and

$$x(\tau) = \begin{cases} D & \text{if } \pi_\tau \leq \tilde{\pi} \\ P & \text{if } \pi_\tau > \tilde{\pi} \end{cases}$$

  A's equilibrium belief is as in Lemma A1, and its belief off the equilibrium path is $\mu^{\text{off}}(\tau, D, \phi) \leq \frac{1}{2}$ if $\pi_\tau > \tilde{\pi}$ and $\mu^{\text{off}}(\tau, P, \phi) > \frac{1}{2}$ if $\pi_\tau \leq \tilde{\pi}$.

For L and T, relevant beliefs at every information set are π_τ.

Proof of Lemma A4. Suppose $b_L = \omega$. It is straight forward to see from Lemmas A1 and A2 that $r_x(\omega)$ and b_L are mutual best responses.

Now suppose $b_L = \phi$. We characterize all the equilibria by considering each of A's possible strategies. Before proceeding, however, note that because τ is a public signal, the equilibrium beliefs of L and T are simply π_τ. Moreover, conditional on observing τ, we do not need to consider off equilibrium path beliefs for either player, since π_τ is direct from observation of τ.

Case 1. Suppose that $r_D^\phi(\tau) = 1$ and $r_P^\phi(\tau) = 0$. From Lemma A1 consistent beliefs require

$$\underbrace{\frac{\pi_\tau}{\pi_\tau + (1-\pi_\tau)(1-a_D)} > \frac{1}{2}}_{\text{if } x=P} \qquad \underbrace{\frac{\pi_\tau(1-a_P)}{\pi_\tau(1-a_P) + 1 - \pi_\tau} > \frac{1}{2}}_{\text{if } x=D}$$

If a_P and a_D are best responses, then we can substitute values from Lemma A3 and see that these collapse to $\pi_\tau > \sqrt{(c_D-1)c_D} - (c_D - 1) \equiv \underline{\pi}$ and $\pi_\tau > \frac{1}{2}$. And since $\underline{\pi} < 1/2$, this further reduces to $\pi_\tau > \frac{1}{2}$.

Next, consider T's judgment. It is a best response for T to rule for the defendant if and only if doing so yields a weakly higher expected utility than ruling for the plaintiff. Formally:

$$\begin{aligned}(1-\pi_\tau)[\delta(1-r_D^\phi(\tau)) - kr_D^\phi(\tau)] + \pi_\tau(\delta-k)[a_P + (1-a_P)r_D^\phi(\tau)] \\ \geq \pi_\tau[\delta(1-r_P^\phi(\tau)) - kr_P^\phi(\tau)] + (1-\pi_\tau)(\delta-k)[a_D + (1-a_D)r_P^\phi(\tau)]\end{aligned} \tag{2}$$

Note that this condition is weak due to Assumption A3.

Since $r_D^\phi(\tau) = 1$ and $r_P^\phi(\tau) = 0$, condition (2) collapses to

$$-(1-\pi_\tau)k + \pi_\tau(\delta - k) \geq \pi_\tau\delta + (1-\pi_\tau)(\delta-k)a_D$$

which fails for any π_τ. It follows that $x = P$ for all τ.

Finally, to support the equilibrium, A's off equilibrium belief must be $\mu^{\text{off}}(\tau, D, \phi) > 1/2$.

Case 2. Suppose that $r_D^\phi(\tau) = 0$ and $r_P^\phi(\tau) = 1$. From Lemma A1 consistent beliefs require

$$\underbrace{\frac{\pi_\tau}{\pi_\tau + (1-\pi_\tau)(1-a_D)} \leq \frac{1}{2}}_{\text{if } x=P} \qquad \underbrace{\frac{\pi_\tau(1-a_P)}{\pi_\tau(1-a_P) + 1 - \pi_\tau} \leq \frac{1}{2}}_{\text{if } x=D}$$

Again, if a_P and a_D are best responses, these collapse to $\pi_\tau \leq \frac{1}{2}$ and $\pi_\tau \leq c_P - \sqrt{(c_P-1)c_P} \equiv \overline{\pi}$. Since $\overline{\pi} > 1/2$, this further reduces to $\pi_\tau \leq 1/2$.

Next, consider T's judgment. Since $r_D^\phi(\tau) = 0$ and $r_P^\phi(\tau) = 1$, condition (2) collapses to

$$(1-\pi_\tau)\delta + \pi_\tau(\delta - k)a_P \geq -\pi_\tau k + (1-\pi_\tau)(\delta - k)$$

which holds for any π_τ. It follows that $x = D$ for all τ.

Finally, to support the equilibrium, A's off equilibrium belief must be $\mu^{\text{off}}(\tau, P, \phi) \leq 1/2$.

Case 3. Suppose that $r_D^\phi(\tau) = r_P^\phi(\tau) = 0$. Then if $x = P$, consistent beliefs require

$$\underbrace{\frac{\pi_\tau}{\pi_\tau + (1-\pi_\tau)(1-a_D)} > \frac{1}{2}}_{\text{if } x=P} \qquad \underbrace{\frac{\pi_\tau(1-a_P)}{\pi_\tau(1-a_P) + 1 - \pi_\tau} \leq \frac{1}{2}}_{\text{if } x=D}$$

Again, if a_P and a_D are best responses, these collapse to $\pi_\tau > \underline{\pi}$ and $\pi_\tau \leq \overline{\pi}$

Next, consider T's judgment. Since $r_D^\phi(\tau) = r_P^\phi(\tau) = 0$, condition (2) collapses to

$$(1-2\pi_\tau)\delta + (\delta - k)\left(\frac{\pi_\tau^2}{c_P} - \frac{(1-\pi_\tau)^2}{c_D}\right) \geq 0 \tag{3}$$

Note that the left hand side of (3) is strictly decreasing in π_τ and that it takes a positive value at $\pi_\tau = 0$ and a negative value at $\pi_\tau = 1$. To see that it is strictly decreasing, consider its derivative, which is strictly negative for all $c_P, c_D > 1$:

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

By Assumption A2, $c_P, c_D > 1$, indicating that there is a threshold $\tilde{\pi} \in (0,1)$ such that (3) holds with equality. Formally, this value is defined by:

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

Then $x = D$ is a best response if $\pi_\tau \leq \tilde{\pi}$ and $x = P$ is a best response if $\pi_\tau > \tilde{\pi}$.

Finally, to support the equilibrium, A's off equilibrium belief must be $\mu^{\text{off}}(\tau, D, \phi) \leq 1/2$ if $\pi_\tau > \tilde{\pi}$ and $\mu^{\text{off}}(\tau, P, \phi) < 1/2$ if $\pi_\tau \leq \tilde{\pi}$.

Case 4. Suppose that $r_D^\phi(\tau) = r_P^\phi(\tau) = 1$. From Lemma A1 consistent beliefs require

$$\underbrace{\frac{\pi_\tau}{\pi_\tau + (1-\pi_\tau)(1-a_D)} \leq \frac{1}{2}}_{\text{if } x=P} \underbrace{\frac{\pi_\tau(1-a_P)}{\pi_\tau(1-a_P) + 1 - \pi_\tau} > \frac{1}{2}}_{\text{if } x=D}$$

Again, if a_P and a_D are best responses, these collapse to $\pi_\tau \leq \frac{1}{2}$ and $\pi_\tau > \frac{1}{2}$. Since either $\pi_\tau \leq 1/2$ or $\pi_\tau > 1/2$ (but not both), this presents a contradiction. Therefore, there is no equilibrium in which $r_P^\phi(\tau) = r_D^\phi(\tau) = 1$. □

Lemma A4 characterizes the equilibria of the continuation games induced by a realization of τ. There is the possibility of multiple equilibria for all τ such that $\pi_\tau \in (\underline{\pi}, \overline{\pi}]$. In one equilibrium, A affirms T's judgment off the equilibrium path and in another A reverses T's judgment off the equilibrium path.

While either equilibria are possible, the latter equilibrium features no deference to T's decision. This stands in contrast to the widespread deference to trial courts that is observed in U.S. courts, and even enshrined in federal rules on court procedure and evidence. In the subsequent analysis, I make the following assumption to rule out these equilibria. However, I note that in the absence of strong institutional rules (or norms) granting wide deference to trial judge decisions, this assumption may be violated. As a result, the analysis in this paper is best understood as corresponding to a situation in which courts coordinate on a system of deference to trial court decisions.

> **Assumption A5** (deference to trial judge)**.** If there is an equilibrium in which the appellate court defers to any decision by the trial judge, then it does so.

To analyze T's optimal effort choice, it will be useful to distinguish between situations where T's judgment depends on what is learned (i.e. τ), and those where it does not. The following result demonstrates that the former situation arises if and only if τ is sufficiently informative.

> **Lemma A5.** In an equilibrium satisfying Assumption A5, $x(\tau \neq \phi) = \tau$ if and only if $\pi_D \leq \min\{\tilde{\pi}, \overline{\pi}\} < \pi_P$.

Proof of Lemma A5. Suppose $\tau \in \{D, P\}$. We prove this by considering several mutually exclusive and exhaustive cases. Preliminarily, it is straight forward to verify that $\tilde{\pi} > \underline{\pi}$. Then, there are three cases to consider.

Case 1. Suppose $\pi_D < \pi_P \leq \min\{\tilde{\pi}, \overline{\pi}\}$. Then, $\pi_\tau \leq \tilde{\pi}$ and $\pi_\tau \leq \overline{\pi}$ for all τ. By Lemma A4 and using Assumption A5, $x(\tau) = D$ for all τ.

Case 2. Suppose $\min\{\tilde{\pi}, \overline{\pi}\} < \pi_D < \pi_P$. Then, $\pi_\tau > \tilde{\pi}$ or $\pi_\tau > \overline{\pi}$ for all τ. By Lemma A4 and using Assumption A5, $x(\tau) = P$ for all τ.

Case 3. Suppose $\pi_D \leq \min\{\tilde{\pi}, \overline{\pi}\} < \pi_P$. Since $\pi_D \leq \tilde{\pi}$ and $\pi_D \leq \overline{\pi}$, by Lemma A4 and using Assumption A5, $x(\tau = D) = D$. Since $\pi_P > \tilde{\pi}$ or $\pi_P > \overline{\pi}$, by Lemma A4 and using Assumption A5, $x(\tau = P) = P$. Combining these two facts, $x(\tau) = \tau$. □

Let $\pi_M = \min\{\tilde{\pi}, \overline{\pi}\}$ and define

$$\pi_D \leq \pi_M \iff \varepsilon \leq \frac{\pi_M(1-\pi)}{\pi_M - \pi(2\pi_M - 1)} \equiv \varepsilon_D \quad \pi_P > \pi_M \iff \varepsilon < \frac{\pi(1-\pi_M)}{\pi_M - \pi(2\pi_M - 1)} \equiv \varepsilon_P$$

A useful restatement of the previous result is

Corollary A1. In an equilibrium satisfying Assumption A5, $x(\tau \neq \phi) = \tau$ if and only if $\varepsilon < \varepsilon_P \leq \varepsilon_D$ or $\varepsilon \leq \varepsilon_D < \varepsilon_P$.

I will set aside the possibility that $\varepsilon = \varepsilon_D$, which is a substantively uninteresting knife-edge case driven by exogenous parameters. Then, define $\overline{\varepsilon} \equiv \min\{\varepsilon_D, \varepsilon_P\}$. I can now characterize the trial judge's optimal effort on the case.

Lemma A6. Let e_P and e_D be defined as in the proof. In an equilibrium satisfying Assumption A5, T exerts effort according to the following. If $\varepsilon < \overline{\varepsilon}$, then:

$$e^* = \begin{cases} \min\{\max\{0, e_P\}, 1\} & \text{if } x(\phi) = P \\ \min\{\max\{0, e_D\}, 1\} & \text{if } x(\phi) = D \end{cases} \tag{4}$$

Otherwise, $e^* = 0$.

Proof of Lemma A6. We first characterize T's continuation value from the equilibrium strategies induced by a realization of τ. From Lemma A4, T's interim expected utility from judgment x as follows:

$$\tilde{U}_T(x,\tau) = \begin{cases} \pi_\tau[\delta(1-r_P^\phi(\tau)) - kr_P^\phi(\tau)] + (1-\pi_\tau)(\delta-k)[a_D + (1-a_D)r_P^\phi(\tau)] & \text{if } x = P \\ (1-\pi_\tau)[\delta(1-r_D^\phi(\tau)) - kr_D^\phi(\tau)] + \pi_\tau(\delta-k)[a_P + (1-a_P)r_D^\phi(\tau)] & \text{if } x = D \end{cases}$$

Also from Lemma A4, on the equilibrium path, $r_x^\phi(\tau) = 0$, and $a_L > 0$. Then, we can rewrite $\tilde{U}_T(x,\tau)$ as

$$\tilde{U}_T(x^*,\tau) = \begin{cases} \pi_\tau\delta + \dfrac{(1-\pi_\tau)^2}{c_D}(\delta-k) & \text{if } x^* = P \\ (1-\pi_\tau)\delta + \dfrac{\pi_\tau^2}{c_P}(\delta-k) & \text{if } x^* = D \end{cases}$$

When T decides whether to make effort, she picks e to maximize her ex ante expected utility, which we denote by U_T. By Corollary A1, there are three cases to consider.

Case 1. If $\varepsilon_P < \varepsilon \leq \varepsilon_D$, then $x(\tau) = D$ for all τ. T's ex ante expected utility is:

$$U_T = e\tilde{U}_T(D,\phi) + (1-e)\tilde{U}_T(D,\phi) - \frac{c_T}{2}e^2$$

This collapses to

$$U_T = \tilde{U}_T(D,\phi) - \frac{c_T}{2}e^2$$

Maximizing expected utility, T sets $e^* = 0$.

Case 2. If $\varepsilon_D \leq \varepsilon < \varepsilon_P$, then $x(\tau) = P$ for all τ. By the same arguments as Case 1, $e^* = 0$.

Case 3. If $\varepsilon < \overline{\varepsilon}$, then $x(\tau \neq \phi) = \tau$. T's ex ante expected utility is:

$$U_T = e\left[(1-\varepsilon)\delta + \varepsilon(\delta-k)(\pi a_P + (1-\pi)a_D)\right] + (1-e)\tilde{U}_T(x(\phi),\phi) - \frac{c_T}{2}e^2$$

Maximizing expected utility, T sets

$$e_{x(\phi)} = \frac{1}{c_T}\left[(1-\varepsilon)\delta + \varepsilon(\delta-k)(\pi a_P + (1-\pi)a_D) - \tilde{U}_T(x(\phi),\phi)\right]$$

We now substitute equilibrium values. If $x(\phi) = P$,

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

and if $x(\phi) = D$,

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

Finally, since $e \in [0, 1]$, $e^* = \min\{\max\{0, e_D\}, 1\}$ if $x(\phi) = D$ and $e^* = \min\{\max\{0, e_P\}, 1\}$ if $x(\phi) = P$. □

1.3 Equilibrium

I now summarize the analysis in the following proposition.

Proposition A1. There is a unique perfect Bayesian equilibrium that satisfies Assumption A3 and Assumption A5. It is characterized by the equilibrium strategies and beliefs (where necessary) in Lemmas A4 and A6.

Proof of Proposition A1. A perfect Bayesian equilibrium is a sequentially rational strategy profile and consistent beliefs. In Lemma A4, we characterize the sequentially rational strategy profiles and consistent beliefs after τ is realized. Using Assumption A5, we rule out one class of equilibria, and using Assumption A3 we rule out multiple equilibria induced by players' indifference. Together, these assumptions guarantee there is a unique equilibrium in the continuation game induced by the realization any τ. In Lemma A6 we characterize T's equilibrium choice of effort, which completes the equilibrium analysis. □

1.4 Comparative Statics

The following results describe how the equilibrium changes as the exogenous parameters change. Of particular interest is the way that $\tilde{\pi}$ and e^* change as the judge's objectives (as captured by δ

and k) change.

Definition A1. A trial judge's decision rule is **impartial** if and only if $x = D \Leftrightarrow \pi_\tau \leq \frac{1}{2}$. A trial judge's decision rule is **biased in favor of litigant** L if and only if there exists some π_τ such that T chooses $x = L$ even though an impartial judge would choose $x \neq L$.

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

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

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 $\pi_\tau \leq \tilde{\pi}$. It follows directly from the definition of $\tilde{\pi}$ in the proof of Lemma A4 that $\tilde{\pi} < \frac{1}{2}$ if $\delta > k$, $\tilde{\pi} > \frac{1}{2}$ if $k > \delta$ and $\tilde{\pi} = \frac{1}{2}$ if $k = \delta$. Note that $\tilde{\pi} < \frac{1}{2}$ implies that T rules for the plaintiff (the less powerful litigant) more often than under an impartial decision rule and $\tilde{\pi} > \frac{1}{2}$ 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 $\tilde{\pi} > \overline{\pi}$, then she is constrained in her decision making by Lemma A4 and uses a decision rule of the form $\pi_\tau \leq \overline{\pi} \Leftrightarrow x = D$. Since $\overline{\pi}$ does not depend on δ or k, then it is not affected as $|\delta - k|$ increases.

Now suppose $\tilde{\pi} \leq \overline{\pi}$. Recall the condition that defines $\tilde{\pi}$ from Lemma A4

$$(1 - 2\tilde{\pi})\delta + (\delta - k)\left(\frac{\tilde{\pi}^2}{c_P} - \frac{(1-\tilde{\pi})^2}{c_D}\right) = 0 \tag{5}$$

Case 1. Suppose $\delta > k$. From above, $\tilde{\pi} < \frac{1}{2}$, so the left hand side of (5) is strictly increasing in k and strictly decreasing in $\tilde{\pi}$. Then as k decreases (and $|\delta - k|$ increases), $\tilde{\pi}$ 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 2. Suppose $k > \delta$. From above, $\tilde{\pi} > \frac{1}{2}$, so the left hand side of (5) is strictly decreasing in both δ and $\tilde{\pi}$. Then as δ decreases (and $|\delta - k|$ increases), $\tilde{\pi}$ 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 $\pi \in (\max\{\frac{1}{2}, \tilde{\pi}\}, \min\{\frac{1}{2}, \tilde{\pi}\}]$.

Proof of Proposition A3. We prove the claim directly by considering three cases.

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

Case 2. If $\delta > k$ and $\tilde{\pi} < \pi \leq \frac{1}{2}$, then T expects to rule for P whereas an impartial judgment would find for D. Moreover, it is straight forward to verify that $e_D > e_P$ when $\tilde{\pi} < \pi \leq \frac{1}{2}$.

Case 3. If $\delta < k$ and $\frac{1}{2} < \pi \leq \tilde{\pi}$, then T's expects to rule for D whereas an impartial judgment would find for P. Moreover, it is straight forward to verify that $e_P > e_D$ when $\frac{1}{2} < \pi < \tilde{\pi}$ and $e_D = e_P$ when $\pi = \tilde{\pi}$. □

Proposition A4. If ε is sufficiently low, then 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.
- If $k = \delta$, then litigant-driven appellate review has no effect on the trial judge's equilibrium effort.

Proof of Proposition A4. If not subjected to appellate review, then $x = D$ if and only if $\pi_\tau \leq \frac{1}{2}$, and effort is determined by maximizing

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

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

$$e^{\text{no}} = \begin{cases} \frac{(1-\pi-\varepsilon)\delta}{c_T} & \text{if } x(\phi) = P \\ \frac{(\pi-\varepsilon)\delta}{c_T} & \text{if } x(\phi) = D \end{cases} \tag{6}$$

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 if ε is sufficiently close to zero, then (6) is strictly larger (smaller) than (4) for all δ, k such that $\delta > k$ ($k > \delta$). □

1.5 Diversity on the Bench

We now assume that $\varepsilon = 0$, and write several endogenous components of the model explicitly in terms of δ and c_T. As in the main text, I now consider two hypothetical judges, each of whom varies with respect to their δ and c_T parameters: (1) $\delta = \bar{\delta} > 0$ and $c_T = c_L$ and (2) $\delta = 0$ and $c_T = c_H > c_L$.

Lemma A7. Let $e^*(\pi, \delta, c_T)$ be the equilibrium effort of T as a function of π, δ and c_T. Then, $e^*(\pi, \bar{\delta}, c_L) \geq e^*(\pi, 0, c_H)$. Moreover, if $c_H > \bar{c}_T$, where $\bar{c}_T$ is defined in the proof, then the inequality is strict: $e^*(\pi, \bar{\delta}, c_L) > e^*(\pi, 0, c_H)$.

Proof of Lemma A7. Since $\tilde{\pi}$ is a function of δ, we will explicitly write it as a function: $\tilde{\pi}(\delta)$. Note that $\tilde{\pi}(\delta)$ decreases in δ. So, $\tilde{\pi}(\bar{\delta}) < \tilde{\pi}(0)$. Then there are two cases to consider: $x(\phi; \delta = \bar{\delta}) = x(\phi; \delta = 0)$ and $x(\phi; \bar{\delta}) = P$ and $x(\phi; 0) = D$. (Note that $x(\phi; \bar{\delta}) = D$ and $x(\phi; 0) = P$ is not possible since $\tilde{\pi}(\bar{\delta}) < \tilde{\pi}(0)$.)

Case 1. Suppose that $x(\phi; \bar{\delta}) = x(\phi; 0)$. If T exerts an interior level of effort, and given that

$\varepsilon = 0$, from Lemma A6, we can write:

$$e_P(\pi, \delta, c_T) = \frac{1}{c_T}\left[(1-\pi)\delta + \frac{(1-\pi)^2}{c_D}(k-\delta)\right]$$
$$e_D(\pi, \delta, c_T) = \frac{1}{c_T}\left[\pi\delta + \frac{\pi^2}{c_P}(k-\delta)\right]$$

Then, note that $e_P(\pi, \bar{\delta}, c_L) > e_P(\pi, 0, c_H) > 0$ and $e_D(\pi, \bar{\delta}, c_L) > e_D(\pi, 0, c_H) > 0$. However, we must also consider the possible corner solution: $e^*(\pi, \bar{\delta}, c_L) = e^*(\pi, 0, c_H) = 1$. This is possible if

$$e^* = e_P \text{ and } \left(\frac{(1-\pi)^2}{c_D}\right) k \geq c_H \qquad \text{or} \qquad e^* = e_D \text{ and } \left(\frac{\pi^2}{c_P}\right) k \geq c_H$$

Define $\underline{c}_T$ and $\bar{c}_T$ as follows:

$$\underline{c}_T = \min\left\{\left(\frac{(1-\pi)^2}{c_D}\right) k, \left(\frac{\pi^2}{c_P}\right) k\right\} \qquad \bar{c}_T = \max\left\{\left(\frac{(1-\pi)^2}{c_D}\right) k, \left(\frac{\pi^2}{c_P}\right) k\right\}$$

We have now shown that if $x(\phi; \bar{\delta}) = x(\phi; 0)$, then $e^*(\pi, \bar{\delta}, c_L) \geq e^*(\pi, 0, c_H)$, and $e^*(\pi, \bar{\delta}, c_L) > e^*(\pi, 0, c_H)$ if $c_H > \bar{c}_T$.

Case 2. Suppose that $x(\phi; \bar{\delta}) = P$ and $x(\phi; 0) = D$. We begin again by assuming an interior level of effort. Then, $e^*(\pi, \bar{\delta}, c_L) = e_P$ and $e^*(\pi, 0, c_H) = e_D$. By contradiction, suppose that there exists some region of the parameter space such that $e^*(\pi, \bar{\delta}, c_L) \leq e^*(\pi, 0, c_H)$. Then,

$$\frac{\bar{\delta}(1-\pi)(c_D - (1-\pi))}{c_D c_L} + \frac{k(1-\pi)^2}{c_D c_L} \leq \frac{k\pi^2}{c_P c_H} \tag{7}$$

Moving all the terms to the left hand side, we can define an expression $E(\pi) \equiv e^*(\pi, \bar{\delta}, c_L) - e^*(\pi, 0, c_H)$. We then seek to show that there exist some parameters such that $E \leq 0$. Note that $E(\pi)$ is strictly concave in π with a maximum at

$$\pi' = \frac{c_H c_P((2-c_D)\delta - 2k)}{2(c_D c_L k + c_H c_P(\delta - k))}$$

Note that either $\pi' < \tilde{\pi}(\bar{\delta})$ or $\pi' > \tilde{\pi}(0)$ or $\pi' \in [\tilde{\pi}(\bar{\delta}), \tilde{\pi}(0)]$. In any of these cases, $E(\pi)$ takes a minimum value at either $\pi = \tilde{\pi}(\bar{\delta})$ or $\pi = \tilde{\pi}(0)$. First if $\pi = \tilde{\pi}(0)$, then:

$$E(\tilde{\pi}(0)) = \frac{1}{c_H c_L (c_D - c_P)^2} \Big[c_H \bar{\delta} \Big(c_D^2 + 2\sqrt{c_D c_P} + c_P \left(\sqrt{c_D c_P} - 1\right) - c_D \left(\sqrt{c_D c_P} + c_P + 1\right) \Big) + k \Big(c_D(c_H - c_L) + c_H c_P - 2c_H \sqrt{c_D c_P} + c_L(2\sqrt{c_D c_P} - c_P) \Big) \Big]$$

Note that the first term inside the large square brackets is weakly positive, and that the expression is strictly increasing in k. So in order for $E \leq 0$, then $\bar{\delta} = k = 0$. However, since we require $\bar{\delta} > 0$, this is a contradiction.

Second if $\pi = \tilde{\pi}(\bar{\delta})$, then:

$$E(\tilde{\pi}(\bar{\delta})) = \frac{1}{c_D c_H c_L c_P (c_D - c_P)^2 (\bar{\delta} - k)^2} \Bigg[c_H c_P (k - \bar{\delta}) \left(c_D((c_D - 1)\bar{\delta} + k) - \sqrt{c_D c_P ((c_D - 1)\bar{\delta} + k)((c_P - 1)\bar{\delta} + k)} \right) \left(c_D((c_P - 1)\bar{\delta} + k) - \sqrt{c_D c_P ((c_D - 1)\bar{\delta} + k)((c_P - 1)\bar{\delta} + k)} \right) - c_D c_L k \left(c_P((c_D - 1)\bar{\delta} + k) - \sqrt{c_D c_P ((c_D - 1)\bar{\delta} + k)((c_P - 1)\bar{\delta} + k)} \right)^2 \Bigg]$$

This term is strictly positive, given the restrictions on $\bar{\delta}, k, c_D, c_P, c_L$ and c_H. This is again another contradiction. We have now shown that if $x(\phi; \bar{\delta}) = P$, $x(\phi; 0) = D$ and e^* is interior, then $E > 0$ for all $\pi \in [\tilde{\pi}(\bar{\delta}, \tilde{\pi}(0)]$. To conclude, we now consider the possible corner solution in which $e^*(\pi, \bar{\delta}, c_L) = e^*(\pi, 0, c_H) = 1$. From above, this occurs if $c_H \leq k\pi^2 / c_P$. □

I now define the accuracy of case outcomes as follows:

$$\xi(\pi, \delta, c_T) = \begin{cases} e^*(\pi, \delta, c_T) + (1 - e^*(\pi, \delta, c_T))(1 - \pi(1 - a_P^*)) & \text{if } x(\phi) = D \\ e^*(\pi, \delta, c_T) + (1 - e^*(\pi, \delta, c_T))(1 - (1 - \pi)(1 - a_D^*)) & \text{if } x(\phi) = P \end{cases} \tag{8}$$

Because we consider two possible judges, we can further define the following:

$$\overline{\xi} = \xi(\pi, \bar{\delta}, c_L) \qquad \underline{\xi} = \xi(\pi, 0, c_H)$$

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, $\overline{\xi} \geq \underline{\xi}$, where the condition holds strictly if $c_H > \overline{c}_T$ (with $\overline{c}_T$ defined in the proof).

Proof of Lemma A8. We consider the same cases as in the proof of Lemma A7. Let $\overline{c}_T \equiv \max\{k(1-\pi)^2/c_D, k\pi^2/c_P\}$.

Case 1. From Lemma A7, if $c_H > \overline{c}_T$, then $e^*(\pi, \bar{\delta}, c_L) > e^*(\pi, 0, c_H)$ and it is direct to see from (8) that $\overline{\xi} > \underline{\xi}$. If, however, $c_H \leq \overline{c}_T$, we may have a corner solution in which $e^*(\pi, \bar{\delta}, c_L) = e^*(\pi, 0, c_H) = 1$. If so, then $\overline{\xi} = \underline{\xi}$.

Case 2. In this case, $x(\phi; \bar{\delta}) = P$ and $x(\phi; 0) = D$. By contradiction, suppose that $\overline{\xi}$ were weakly less than $\underline{\xi}$ at some point in the interval $[\tilde{\pi}(\bar{\delta}), \min\{\tilde{\pi}(0), \overline{\pi}\}]$. First, suppose $c_H > \overline{c}_T$. Since $x(\phi; \bar{\delta}) = P$ and $x(\phi; 0) = D$, we can rewrite the ex ante probability as follows:

$$\xi(\pi, \delta, c_T) = \begin{cases} e_D(\pi, 0, c_H) + (1 - e_D(\pi, 0, c_H))(1 - \pi(1 - a_P)) & \text{if } \delta = 0 \\ e_P(\pi, \bar{\delta}, c_L) + (1 - e_P(\pi, \bar{\delta}, c_L))(1 - (1 - \pi)(1 - a_D)) & \text{if } \delta = \bar{\delta} \end{cases}$$

By Lemma A7, $e(\pi, \bar{\delta}, c_L) > e(\pi, 0, c_H)$. In order for $\overline{\xi} \leq \underline{\xi}$ for some $\pi \in [\tilde{\pi}(\bar{\delta}), \tilde{\pi}(0)]$ 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 $\pi = \tilde{\pi}(\bar{\delta})$:

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

However, the right hand side is greater than one for any $\tilde{\pi}(\bar{\delta}) \in [0, 1]$, a contradiction. Next, suppose $c_H \leq k\pi^2/c_P$, then $e(\pi, \bar{\delta}, c_L) = e(\pi, 0, c_H) = 1$, and $\overline{\xi} = \underline{\xi}$. □

Proposition A5. Random assignment of judges to cases leads to fewer accurate decisions than voluntary assignment. Formally, $R \leq V$, and $R < V$ if $c_H > \bar{c}_T$.

Proof of Proposition A5. This is direct:

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

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

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

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

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

Thus, we have shown that $R \leq V$. Moreover, if $c_H > \bar{c}_T$, then $\overline{\xi} > \underline{\xi}$ and these inequalities are strict. Then $R < V$. $\square$