Technical Appendix for
Getting Their Way: Bias and Deference to Trial Courts

Ryan Hübert
University of California, Davis

December 15, 2018

FOR ONLINE PUBLICATION

Contents

In the main text, I describe the set up of the model, including the players, the sequence, the strategy sets for each player and the payoffs for each player. I derive perfect Bayesian equilibria (PBE) of the model. A PBE is a profile of sequentially rational strategies with beliefs that are consistent with the equilibrium strategy profile and updated via Bayes' rule where possible.

Since a PBE requires all players' strategies to be sequentially rational, I use backward induction to solve the game. However, since backward induction only yields the sequentially rational strategy profile, I will also pin down consistent beliefs required for the sequentially rational strategies.

The formal analyses corresponding to Sections 3 to 5 of the article are presented in each of the following sections.

1 Formal Results for Rationalizing Deference

Two definitions will be useful for the analysis.

Definition A1. A disposition x is **consistent** with a legal rule ℓ_i if and only if $[x = p \iff \omega > \ell_i]$ or $[x = d \iff \omega \leq \ell_i]$. A disposition x is **compliant** if and only if it is consistent with ℓ_A.

Let $\hat{x}^c(\omega)$ denote a compliant disposition and $\hat{x}^n(\omega)$ denote a non-compliant disposition.

Definition A2. A's belief that judgment (x, o) is compliant is:

$$\xi(x, o) = \begin{cases} \Pr(\omega > \ell_A | x, o) & \text{if } x = p \\ \Pr(\omega \leq \ell_A | x, o) & \text{if } x = d \end{cases}$$

Assumption A1 (indifference)**.**

- A upholds when indifferent.
- T rules for the defendant when indifferent.
- T writes an informative opinion when indifferent.

Now, I derive the equilibrium strategy profiles using backward induction.

Lemma A1. Using Assumption A1, A's sequentially rational reversal strategy is

$$\rho^*(x,o) = \begin{cases} 0 & \text{if } \xi(x,o) \geq \frac{1}{2} \\ 1 & \text{if } \xi(x,o) < \frac{1}{2} \end{cases}$$

Proof. The result follows directly from A's preferences, the definition of $\xi(x,o)$ and Assumption A1. □

Before proceeding, I introduce an additional piece of notation that will be helpful in the following analysis. Let s_T be the signal that T receives about ω. That is: with probability $e\varepsilon$, T receives a signal $s_T = \omega$ and with probability $1 - e\varepsilon$, T receives a signal $s_T = \phi$.

Lemma A2 (Reversals When Informed)**.** If $o(s_T, \kappa, x) = \omega$, then $\xi(x,\omega) \in \{0,1\}$ and $\rho(x,\omega) = 1$ if and only if $\xi(x,\omega) = 0$.

Proof. Let $o(s_T, \kappa, x) = \omega$, and fix x. Since $o = \omega$, A's beliefs about ω—$\Pr(\omega \leq \ell_A)$ and $\Pr(\omega > \ell_A)$–are degenerate, and using Bayes' rule, it follows that $\xi(x, o = \omega) \in \{0,1\}$. Then, □

The appellate court never defers when it is fully informed.

I now characterize T's optimal disposition x and opinion o. Before proceeding, I adopt the following assumption.

Assumption A2. T cannot fabricate information. Formally, $o \in \{s_T, \phi\}$.

Lemma A3. T's sequentially rational opinion is

$$o^*(s_T,\kappa,x)=\begin{cases}\phi & \text{if}\begin{cases}s_T=\phi\\ s_T=\omega\in(\ell_A,\ell_T] \text{ and } x=\hat{x}^n(\omega) \text{ and } \kappa>q-(\alpha+k)(1-\rho(x,\phi))\\ s_T=\omega\in(\ell_A,\ell_T] \text{ and } x=\hat{x}^c(\omega) \text{ and } \kappa>q-(\alpha-k)\rho(x,\phi)\\ s_T=\omega\notin(\ell_A,\ell_T] \text{ and } x=\hat{x}^n(\omega) \text{ and } \kappa>q+(\alpha-k)(1-\rho(x,\phi))\\ s_T=\omega\notin(\ell_A,\ell_T] \text{ and } x=\hat{x}^c(\omega) \text{ and } \kappa>q+(\alpha+k)\rho(x,\phi)\end{cases}\\ \omega & \text{if}\begin{cases}s_T=\omega\in(\ell_A,\ell_T] \text{ and } x=\hat{x}^n(\omega) \text{ and } \kappa\le q-(\alpha+k)(1-\rho(x,\phi))\\ s_T=\omega\in(\ell_A,\ell_T] \text{ and } x=\hat{x}^c(\omega) \text{ and } \kappa\le q-(\alpha-k)\rho(x,\phi)\\ s_T=\omega\notin(\ell_A,\ell_T] \text{ and } x=\hat{x}^n(\omega) \text{ and } \kappa\le q+(\alpha-k)(1-\rho(x,\phi))\\ s_T=\omega\notin(\ell_A,\ell_T] \text{ and } x=\hat{x}^c(\omega) \text{ and } \kappa\le q+(\alpha+k)\rho(x,\phi)\end{cases}\end{cases}$$

Proof. T's chooses o at four information sets, after observing $x\in\{p,d\}$ and $s_T\in\{\omega,\phi\}$. By Assumption A2, it is trivial to see that for two of those information sets, T has no real choice to make:

$$o^*(s_T=\phi,x=p)=o^*(s_T=\phi,x=d)=\phi$$

Next, we consider the remaining two information sets. The proof proceeds by considering two mutually exclusive and exhaustive cases. Let $\rho(x,o)$ denote A's reversal strategy.

1. Assume $s_T=\omega\le\ell_A$ or $s_T=\omega>\ell_T$. Lemma A2 implies that $\rho(x=\hat{x}^c(\omega),\omega)=0$ and $\rho(x=\hat{x}^n(\omega),\omega)=1$. We now consider the two possible choices of x.

   (a) If $x=\hat{x}^c(\omega)$, then T is better off with $o(\omega,x)=\omega$ if and only if

   $$\underbrace{q-\kappa+\alpha}_{U_T(o=\omega)}\ge\underbrace{\alpha(1-\rho(\hat{x}^c(\omega),\phi))-k\rho(\hat{x}^c(\omega),\phi)}_{U_T(o=\phi)}\Longleftrightarrow\kappa\le q+(\alpha+k)\rho(\hat{x}^c(\omega),\phi)$$

   (b) If $x=\hat{x}^n(\omega)$, then T is better off with $o(\omega,x)=\omega$ if and only if

   $$\underbrace{q-\kappa+\alpha-k}_{U_T(o=\omega)}\ge\underbrace{(\alpha-k)\rho(x,\phi)}_{U_T(o=\phi)}\Longleftrightarrow\kappa\le q+(\alpha-k)(1-\rho(x,\phi))$$

2. Assume $s_T=\omega\in(\ell_A,\ell_T]$. Lemma A2 implies that $\rho(x=d,\omega)=1$ and $\rho(x=p,\omega)=0$. We now consider the two possible choices of x.

(a) If $x = d$, then T is better off with $o(\omega, d) = \omega$ if and only if

$$\underbrace{q - \kappa - k}_{U_T(o=\omega)} \geq \underbrace{-k\rho(d,\phi) + \alpha(1 - \rho(d,\phi))}_{U_T(o=\phi)} \iff \kappa \leq q - (\alpha + k)(1 - \rho(d,\phi))$$

(b) If $x = p$, then T is better off with $o(\omega, x) = \omega$ if and only if

$$\underbrace{q - \kappa}_{U_T(o=\omega)} \geq \underbrace{(\alpha - k)\rho(p,\phi)}_{U_T(o=\phi)} \iff \kappa \leq q - (\alpha - k)\rho(p,\phi)$$

Taking these together gives $o^*(s_T, \kappa, x)$, as required. □

Lemma A4. T's sequentially rational disposition is characterized by:

$$x^*(s_T = \phi, \kappa) = \begin{cases} d & \text{if } \rho(d,\phi) = 0 \\ p & \text{if } \rho(d,\phi) = 1 \end{cases}$$

$$x^*(s_T \leq \ell_A, \kappa) = \begin{cases} d & \text{if } \begin{cases} (1-\rho(p,\phi))(\alpha-k) < \rho(d,\phi)(\alpha+k) \text{ and } \kappa \leq q + \alpha - \rho(p,\phi)(\alpha - k) \\ (1-\rho(p,\phi))(\alpha-k) > \rho(d,\phi)(\alpha+k) \text{ and } \kappa \geq q - k + \rho(d,\phi)(\alpha + k) \\ (1-\rho(p,\phi))(\alpha-k) = \rho(d,\phi)(\alpha+k) \text{ and } \kappa \leq q \end{cases} \\ p & \text{otherwise} \end{cases}$$

$$x^*(s_T > \ell_A, \kappa) = \begin{cases} p & \text{if } \begin{cases} (1-\rho(d,\phi))(\alpha-k) < \rho(p,\phi)(\alpha+k) \text{ and } \kappa < q + \alpha - \rho(d,\phi)(\alpha - k) \\ (1-\rho(d,\phi))(\alpha-k) > \rho(p,\phi)(\alpha+k) \text{ and } \kappa > q - k + \rho(p,\phi)(\alpha + k) \\ (1-\rho(d,\phi))(\alpha-k) = \rho(p,\phi)(\alpha+k) \text{ and } \kappa < q \end{cases} \\ d & \text{otherwise} \end{cases}$$

$$x^*(s_T \in (\ell_A, \ell_T], \kappa) = \begin{cases} d & \text{if } \begin{cases} (1-\rho(d,\phi))(\alpha+k) > \rho(p,\phi)(\alpha-k) \text{ and } \kappa \geq q - \alpha(1-\rho(d,\phi)) + k\rho(d,\phi) \\ (1-\rho(d,\phi))(\alpha+k) < \rho(p,\phi)(\alpha-k) \text{ and } \kappa \leq q - \alpha\rho(p,\phi) - k(1-\rho(p,\phi)) \end{cases} \\ p & \text{otherwise} \end{cases}$$

Proof. Let $\rho(x, o)$ denote A's reversal strategy. Chooses x at information set s_T. If $s_T = \phi$, then by Lemma A3, $o^*(\phi, x) = \phi$ for $x \in \{p, d\}$. Then, $x = d$ if and only if

$$\underbrace{(1-\rho(d,\phi))\delta_T\alpha + \rho(d,\phi)((1-\delta_T)\alpha - k)}_{U_T(x=d)} \geq \underbrace{(1-\rho(p,\phi))(1-\delta_T)\alpha + \rho(p,\phi)(\delta_T\alpha - k)}_{U_T(x=p)}$$

This holds if $\rho(d,\phi) = \rho(p,\phi) = 0$ or $\rho(d,\phi) = 0, \rho(p,\phi) = 1$, and fails otherwise.

Next, suppose $s_T = \omega$. We consider several mutually exclusive and exhaustive cases, based on the sequentially rational strategies of T and A in Lemmas A1 and A3.

1. Assume $s_T = \omega \leq \ell_A$ or $s_T = \omega > \ell_T$.

(a) Suppose $\kappa \leq \min\{q + (1-\rho(\hat{x}^n(\omega), \phi))(\alpha - k), q + \rho(\hat{x}^c(\omega), \phi)(\alpha + k)\}$. Then $o(s_T, \kappa, x) = \omega$ for all x. Then $x = \hat{x}^c(\omega)$ if and only if

$$\underbrace{\alpha + q - \kappa}_{U_T(x=\hat{x}^c(\omega))} \geq \underbrace{\alpha - k + q - \kappa}_{U_T(x=\hat{x}^n(\omega))}$$

This holds strictly, and $x = \hat{x}^c(\omega)$.

(b) Suppose $\kappa \in (q + (1-\rho(\hat{x}^n(\omega), \phi))(\alpha - k), q + \rho(\hat{x}^c(\omega), \phi)(\alpha + k)]$. Then $o(s_T, \kappa, x = \hat{x}^c(\omega)) = \omega$ and $o(s_T, \kappa, x = \hat{x}^n(\omega)) = \phi$. Then $x = \hat{x}^c(\omega)$ if and only if

$$\underbrace{\alpha + q - \kappa}_{U_T(x=\hat{x}^c(\omega))} \geq \underbrace{(\alpha - k)\rho(\hat{x}^n(\omega), \phi) + 0(1-\rho(\hat{x}^n(\omega), \phi))}_{U_T(x=\hat{x}^n(\omega))}$$
$$\iff \kappa \leq \alpha(1-\rho(\hat{x}^n(\omega), \phi)) + q + k\rho(\hat{x}^n(\omega), \phi)$$

(c) Suppose $\kappa \in (q + \rho(\hat{x}^c(\omega), \phi)(\alpha + k), q + (1-\rho(\hat{x}^n(\omega), \phi))(\alpha - k)]$. Then $o(s_T, \kappa, x = \hat{x}^c(\omega)) = \phi$ and $o(s_T, \kappa, x = \hat{x}^n(\omega)) = \omega$. Then $x = \hat{x}^c(\omega)$ if and only if

$$\underbrace{\alpha(1-\rho(\hat{x}^c(\omega), \phi)) - k\rho(\hat{x}^c(\omega), \phi)}_{U_T(x=\hat{x}^c(\omega))} \geq \underbrace{q - \kappa + \alpha - k}_{U_T(x=\hat{x}^n(\omega))}$$
$$\iff \kappa \geq q + \alpha\rho(\hat{x}^c(\omega), \phi) - k(1-\rho(\hat{x}^c(\omega), \phi))$$

(d) Suppose $\kappa > \max\{q + (1-\rho(\hat{x}^n(\omega), \phi))(\alpha - k), q + \rho(\hat{x}^c(\omega), \phi)(\alpha + k)\}$. Then $o(s_T, \kappa, x) = \phi$ for all x. Then $x = \hat{x}^c(\omega)$ if and only if

$$\underbrace{-k\rho(\hat{x}^c(\omega), \phi) + \alpha(1-\rho(\hat{x}^c(\omega), \phi))}_{U_T(x=\hat{x}^c(\omega))} \geq \underbrace{(\alpha - k)\rho(\hat{x}^n(\omega), \phi)}_{U_T(x=\hat{x}^n(\omega))}$$
$$\iff \alpha - \rho(\hat{x}^c(\omega), \phi)(\alpha + k) \geq \rho(\hat{x}^n(\omega), \phi)(\alpha - k)$$

2. Assume $s_T \in (\ell_A, \ell_T]$.

(a) Suppose $\kappa \leq \min\{q - (1-\rho(d, \phi))(k + \alpha), q - \rho(p, \phi)(\alpha - k)\}$. Then $o(s_T, \kappa, x) = \omega$ for all x. Then $x = d$ if and only if

$$\underbrace{-k + q - \kappa}_{U_T(x=d)} \geq \underbrace{q - \kappa}_{U_T(x=p)}$$

This never holds, and $x = p$.

(b) Suppose $\kappa \in (q - (1-\rho(x,\phi))(k+\alpha), q - \rho(x,\phi)(\alpha - k)]$. Then $o(s_T, \kappa, x = d) = \phi$ and $o(s_T, \kappa, x = p) = \omega$. Then $x = d$ if and only if

$$\underbrace{\alpha(1-\rho(d,\phi)) - k\rho(d,\phi)}_{U_T(x=d)} \geq \underbrace{q-\kappa}_{U_T(x=p)} \iff \kappa \geq q - \alpha(1-\rho(d,\phi)) + k\rho(d,\phi)$$

(c) Suppose $\kappa \in (q - \rho(x,\phi)(\alpha - k), q - (1-\rho(x,\phi))(k+\alpha)]$. Then $o(s_T, \kappa, x = d) = \omega$ and $o(s_T, \kappa, x = p) = \phi$. Then $x = d$ if and only if

$$\underbrace{-k + q - \kappa}_{U_T(x=d)} \geq \underbrace{(\alpha - k)\rho(p,\phi)}_{U_T(x=p)} \iff \kappa \leq q - \alpha\rho(p,\phi) - k(1-\rho(p,\phi))$$

(d) Suppose $\kappa > \max\{q - (1-\rho(x,\phi))(k+\alpha), q - \rho(x,\phi)(\alpha - k)\}$. Then $o(s_T, \kappa, x) = \phi$ for all x. Then $x = d$ if and only if

$$\underbrace{\alpha(1-\rho(d,\phi)) - k\rho(d,\phi)}_{U_T(x=d)} \geq \underbrace{(\alpha - k)\rho(p,\phi)}_{U_T(x=p)}$$
$$\iff \alpha - \rho(d,\phi)(\alpha + k) - \rho(p,\phi)(\alpha - k) \geq 0$$

Taking these together gives $x^*(s_T, \kappa)$, as required. □

Lemmas A3 and A4 characterize T's sequentially rational actions, which are a function of A's reversal strategy, $\rho(x,o)$. To proceed, we now determine whether equilibria exist for each of A's four pure strategies when $o = \phi$:

$$\rho(p,\phi) = \rho(d,\phi) = 1 \qquad \rho(p,\phi) = 1, \quad \rho(d,\phi) = 0$$
$$\rho(p,\phi) = \rho(d,\phi) = 0 \qquad \rho(p,\phi) = 0, \quad \rho(d,\phi) = 1$$

In the analysis that follows, we characterize T's optimal optimal opinion as a function of s_T and κ, under the presumption that T does not have to consider her own deviations from her optimal disposition, x. This is a technicality that makes the analysis less cumbersome, but does not alter the derivations. It means, however, that characterizations of $o(s_T, \kappa)$ in the following results are on the equilibrium path. Consult Lemma A3 for a full specification of T's sequentially rational strategy.

We adopt the following assumption so that T has "dispositional" preferences. That is, we do not wish for results to be driven by T's desire to write opinions independent of case outcomes. To see why such behavior would be unreasonable, note that if the following assumption is violated, then T would sometimes issue judgments with dispositions it does not prefer in order to write an opinion justifying that disposition.

Assumption A3. The trial judge has dispositional preferences. Formally, $q + k - \alpha < \kappa_L < q < \kappa_H = \infty$.

Lemma A5. There is no equilibrium where $\rho(d, \phi) = 0$ and $\rho(p, \phi) = 1$.

Proof. We prove this by contradiction. Specifically, we will show that the sequentially rational strategies required by the conjectured equilibrium require inconsistent beliefs. Suppose that the conjectured equilibrium exists. By Lemmas A3 and A4, T's sequentially rational strategy can be characterized as follows:

- If $s_T = \phi$, then $x(s_T, \kappa) = p$ and $o(s_T, \kappa) = \phi$.
- If $s_T \leq \ell_A$, then $x(s_T, \kappa) = d$ and

$$o(s_T, \kappa) = \begin{cases} \omega & \text{if } \kappa \leq q \\ \phi & \text{if } \kappa > q \end{cases}$$

- If $s_T > \ell_T$, then

$$x(s_T, \kappa) = \begin{cases} p & \text{if } \kappa \leq q \\ d & \text{if } \kappa > q \end{cases} \qquad o(s_T, \kappa) = \begin{cases} \omega & \text{if } \kappa \leq q \\ \phi & \text{if } \kappa > q \end{cases}$$

- If $s_T \in (\ell_A, \ell_T]$, then

$$x(s_T, \kappa) = \begin{cases} p & \text{if } \kappa \leq q - \alpha \\ d & \text{if } \kappa > q - \alpha \end{cases} \qquad o(s_T, \kappa) = \begin{cases} \omega & \text{if } \kappa \leq q - \alpha \\ \phi & \text{if } \kappa > q - \alpha \end{cases}$$

  Using Assumption A3, this reduces to $x(s_T, \kappa) = d$ and $o(s_T, \kappa) = \phi$.

Given T's sequentially rational actions when $\rho(d, \phi) = 0$ and $\rho(p, \phi) = 1$, A's consistent belief at its information set is

$$\xi(d, \phi) = \frac{\delta_A(1-\eta)\varepsilon}{\delta_A(1-\eta)\varepsilon + (\delta_T - \delta_A)\varepsilon + (1-\delta_T)(1-\eta)\varepsilon}$$

By Lemma A1, in order for $\rho(d, \phi) = 0$ to be a best response:

$$\xi(d, \phi) \geq \frac{1}{2}$$

Substituting, this simplifies to

$$\delta_A(1-\eta) \geq (\delta_T - \delta_A) + (1-\delta_T)(1-\eta)$$

This condition never holds for $\delta < \frac{1}{2} < \delta_T$ and $0 < \eta < 1$, a contradiction. We have shown that A's conjectured sequentially rational action $\rho(d,\phi) = 0$ is not a best response to beliefs induced by the conjectured equilibrium strategy of T. This is sufficient to demonstrate there is no equilibrium with $\rho(d,\phi) = 0$ and $\rho(p,\phi) = 1$. □

Lemma A6. There exists an equilibrium in which $\rho(x,\omega)$ is given by Lemma A2, $\rho(d,\phi) = 1$, $\rho(p,\phi) = 0$, and where:

$$x(s_T,\kappa) = \begin{cases} d & \text{if } s_T = \omega \leq \ell_A \text{ and } \kappa \leq q+\alpha \\ p & \text{otherwise} \end{cases}$$

$$o(s_T,\kappa,x) = \begin{cases} \omega & \text{if } \begin{cases} s_T = \omega \leq \ell_A \text{ and } \kappa \leq q+\alpha \\ s_T = \omega > \ell_A \text{ and } \kappa \leq q \end{cases} \\ \phi & \text{otherwise} \end{cases}$$

A's belief off the equilibrium path is $\xi(d,\phi) < \frac{1}{2}$, and A's belief on the equilibrium path is defined in the proof.

Proof. Our task is to demonstrate the existence of an equilibrium as conjectured. We do so by contradiction. Suppose that there does not exist an equilibrium. First, by Lemmas A3 and A4, T's sequentially rational strategy can be characterized as follows:

- If $s_T = \phi$, then $x = p$ and $o = \phi$.
- If $s_T \leq \ell_A$, then $x = d$ and $o = \omega$ if $\kappa \leq \alpha + q$ and $x = p$ and $o = \phi$ if $\kappa > \alpha + q$.
- If $s_T > \ell_A$, then $x = p$; and $o = \omega$ if $\kappa \leq q$ and $o = \phi$ if $\kappa > q$.

Now we consider A's belief given T's sequentially rational strategy characterized above. First, note that $x = d$ and $o = \phi$ is off the equilibrium path. Because Bayes' rule does not pin down a belief at this information set, we must construct beliefs that rationalize A's conjectured equilibirum action at this information set. By Lemma A1, $\rho(d,\phi) = 1$ is optimal if $\xi(d,\phi) < \frac{1}{2}$. Then, we can write A's belief that x is compliant (conditional on $o = \phi$) as

$$\xi(x, \phi) = \begin{cases} \dfrac{(1-\varepsilon\eta)(1-\delta_A)}{(1-\varepsilon\eta)(1-\delta_A) + (1-\varepsilon\eta)\delta_A} & \text{if } x = p \\ z \in \left[0, \dfrac{1}{2}\right) & \text{if } x = d \end{cases}$$

We must also check that $\rho(p, \phi) = 0$ is optimal for A given A's beliefs:

$$\frac{(1-e\varepsilon\eta)(1-\delta_A)}{(1-e\varepsilon\eta)(1-\delta_A) + (1-e\varepsilon\eta)\delta_A} \geq \frac{1}{2}$$

This is always true, so A has no incentive to deviate, a contradiction. We have therefore proven that the conjectured strategy profile is an equilibrium. □

Lemma A7. If and only if $\eta \leq \tilde{\eta}$ and $e = 1$, there exists an equilibrium in which $\rho(x, \omega)$ is given by Lemma A2, $\rho(d, \phi) = \rho(p, \phi) = 0$, and where:

$$x(s_T, \kappa) = \begin{cases} d & \text{if } s_T = \omega \leq \ell_T \text{ or } s_T = \phi \\ p & \text{if } s_T = \omega > \ell_T \end{cases}$$

$$o(s_T, \kappa) = \begin{cases} \omega & \text{if } s_T \in \{\omega : \omega \leq \ell_A \text{ or } \omega > \ell_T\} \text{ and } \kappa \leq q \\ \phi & \text{otherwise} \end{cases}$$

A's equilibrium belief and $\tilde{\eta}$ are defined in the proof.

Proof. ($\Longrightarrow$) Suppose $e = 1$, and $\eta \leq \tilde{\eta} \equiv \frac{1}{\varepsilon\delta_A}(\varepsilon(1-\delta_T) + 2\delta_A - 1)$. Our task is to demonstrate the existence of an equilibrium as conjectured. We do so by showing that no player has an incentive to deviate. Suppose A plays $\rho(x, o)$ as conjectured. By Lemmas A3 and A4, T's sequentially rational strategy can be characterized as follows:

- If $s_T = \phi$, then $x = d$ and $o = \phi$.
- If $s_T \leq \ell_A$, then $x = d$ and $o = \omega$ if $\kappa \leq q$ and $o = \phi$ if $\kappa > q$.
- If $s_T > \ell_T$, then $x = p$ and $o = \omega$ if $\kappa \leq q$ and $o = \phi$ if $\kappa > q$.
- If $s_T \in (\ell_A, \ell_T]$, then using Assumption A3, $x = d$ and $o = \phi$.

Given the strategy profile, A's belief that x is compliant (conditional on $o = \phi$) is

$$\xi^0(x,\phi) = \begin{cases} 1 & \text{if } x = p \\ \dfrac{(1-\varepsilon\eta)\delta_A}{(1-\varepsilon\eta)\delta_A + (\delta_T - \delta_A) + (1-\varepsilon)(1-\delta_T)} & \text{if } x = d \end{cases}$$

Then, by Lemma A1, $\rho(p,\phi) = 0$ is optimal and $\rho(d,\phi) = 0$ is optimal for A if

$$\frac{(1-\varepsilon\eta)\delta_A}{(1-\varepsilon\eta)\delta_A + (\delta_T - \delta_A) + (1-\varepsilon)(1-\delta_T)} \geq \frac{1}{2}$$

which reduces to

$$\eta \leq \frac{1}{\varepsilon\delta_A}\left(\varepsilon(1-\delta_T) + 2\delta_A - 1\right) = \tilde{\eta}$$

Since we postulated $\eta \leq \tilde{\eta}$, then A has no incentive to deviate. We have therefore proven that the existence of the conjectured equilibrium.

($\Longleftarrow$) Suppose that the conjectured equilibrium exists. Our task is to show that $e = 1$ and $\eta \leq \tilde{\eta}$, and we prove each by contradiction. First, suppose $e = 0$. Then $\xi^0(x = d, \phi) = \delta_A < \frac{1}{2}$. By Lemma A1, this implies that $\rho(x = d, \phi) = 1$, a contradiction since in the conjectured equilibrium, $\rho(x = d, \phi) = 0$. Second, suppose $\eta > \tilde{\eta}$. Then, using the arguments above, $\xi^0(x = d, \phi) < \frac{1}{2}$. Again, by Lemma A1, this implies that $\rho(x = d, \phi) = 1$, a contradiction since in the conjectured equilibrium, $\rho(x = d, \phi) = 0$. □

Lemma A8. If and only if $\eta < \tilde{\eta}$ and $e = 1$, there exists an equilibrium in which $\rho(x,\omega)$ is given by Lemma A2, $\rho(d,\phi) = \rho(p,\phi) = 1$, and where:

$$x(s_T,\kappa) = \begin{cases} d & \text{if } \begin{cases} s_T = \omega \leq \ell_A \text{ and } \kappa \leq q + k \\ s_T = \omega > \ell_T \text{ and } \kappa > q + k \end{cases} \\ p & \text{if } \begin{cases} s_T \in \{\phi\} \cup \{\omega : \omega \in (\ell_A, \ell_T]\} \\ s_T = \omega > \ell_T \text{ and } \kappa \leq q + k \\ s_T = \omega \leq \ell_A \text{ and } \kappa > q + k \end{cases} \end{cases}$$

$$o(s_T,\kappa) = \begin{cases} \omega & \text{if } \begin{cases} s_T = \omega \leq \ell_A \text{ and } \kappa \leq q + k \\ s_T = \omega > \ell_T \text{ and } \kappa \leq q + k \end{cases} \\ \phi & \text{otherwise} \end{cases}$$

A's equilibrium belief is defined in the proof.

Proof. The proof proceeds by the same arguments as the proof for Lemma A7 with the added caveats that (1) A's equilibrium belief is

$$\xi^1(x,\phi) = \begin{cases} \dfrac{(1-\varepsilon)(1-\delta_T)+(\delta_T-\delta_A)}{(1-\varepsilon)(1-\delta_T)+(\delta_T-\delta_A)+(1-\varepsilon\eta)\delta_A} & \text{if } x=p \\ 0 & \text{if } x=d \end{cases}$$

and (2) Assumption A1 ensures that $\eta \leq \tilde{\eta}$ must hold strictly in the conjectured equilibrium. □

Next we derive T's equilibrium effort in each of the candidate equilibria characterized above. Before proceeding, we adopt the following assumption, as described in the text of the paper.

Assumption A4. $\alpha \geq \frac{2c}{\varepsilon}$.

Lemma A9. In an equilibrium described in Lemma A7, T exerts effort, $e^0 = 1$. And in an equilibrium described in Lemma A8, T exerts effort, $e^1 = 1$, if and only if

$$\eta \geq \frac{c+k-\alpha(\delta_T-(1-\delta_T)(1-\varepsilon))}{\varepsilon(\delta_A+1-\delta_T)(q-\kappa_L+k)}$$

Proof. First, assume players use judgment and reversal strategies as in Lemma A7. Then, T's ex ante expected utility from exerting effort is:

$$U_T^0(e=1) = \alpha(\delta_T+\varepsilon(1-\delta_T)) + \varepsilon(\delta_A+1-\delta_T)\eta(q-\kappa_L) - c \tag{1}$$

If T does not exert effort in equilibrium, she receives no deference and her expected utility is $(1-\delta_T)\alpha$. Then, it is a best response to exert effort if

$$\alpha(\delta_T+\varepsilon(1-\delta_T)) + \varepsilon(\delta_A+1-\delta_T)\eta(q-\kappa_L) - c \geq (1-\delta_T)\alpha$$

which reduces to

$$\eta \geq \frac{c-\alpha(\delta_T-(1-\delta_T)(1-\varepsilon))}{\varepsilon(\delta_A+1-\delta_T)(q-\kappa_L)}$$

By Assumption A4, the right hand side is negative, implying the condition always holds and T exerts effort in the equilibrium described in Lemma A7.

Next, assume players use judgment and reversal strategies as in Lemma A8. Then, T's ex ante expected utility is:

$$U_T^1(e=1) = \alpha(\delta_T + \varepsilon(1-\delta_T)) - k(1-\varepsilon+\varepsilon(\delta_T-\delta_A)) \\ + \varepsilon(\delta_A + 1 - \delta_T)\big[(q-\kappa_L)\eta - k(1-\eta)\big] - c$$

This simplifies to

$$U_T^1(e=1) = \alpha(\delta_T + \varepsilon(1-\delta_T)) - k + \varepsilon(\delta_A + 1 - \delta_T)(q-\kappa_L+k)\eta - c \tag{2}$$

If T does not exert effort in equilibrium, she receives no deference and her expected utility is $(1-\delta_T)\alpha$. Then, it is a best response to exert effort if

$$\alpha(\delta_T + \varepsilon(1-\delta_T)) - k + \varepsilon(\delta_A + 1 - \delta_T)(q-\kappa_L+k)\eta - c \geq (1-\delta_T)\alpha$$

which reduces to

$$\eta \geq \frac{c + k - \alpha(\delta_T - (1-\delta_T)(1-\varepsilon))}{\varepsilon(\delta_A + 1 - \delta_T)(q-\kappa_L+k)}$$

We have directly proven the statement. □

Lemma A10. In an equilibrium described in Lemma A6, T exerts effort, $e^n = 1$, if and only if

$$\eta \geq \frac{c}{\varepsilon(\delta_A\alpha + q - \kappa_L)} \equiv \hat{\eta} \tag{3}$$

Proof. Assume players use judgment and reversal strategies as in Lemma A6. Then, T's ex ante expected utility is:

$$U_T^n(e=1) = (1-\delta_T)\alpha + \varepsilon\eta(\delta_A\alpha + (q-\kappa_L)) - c \tag{4}$$

Then, it is a best response to exert effort if

$$(1-\delta_T)\alpha + \varepsilon\eta(\delta_A\alpha + (q-\kappa_L)) - c \geq (1-\delta_T)\alpha$$

which reduces to

$$\eta \geq \frac{c}{\varepsilon(\delta_A\alpha + q - \kappa_L)} \equiv \hat{\eta}$$

We have directly proven the statement. □

There are two equilibria of the model which allow T to achieve their preferred disposition (see Lemmas A7 and A8). I do not use a formal equilibrium selection criterion to adjudicate between them, but I note that one of them features an appellate court that *always* reverses a trial judge who does not write an informative opinion. Because this is substantively implausible, we adopt the following assumption to streamline the analysis.

Assumption A5. The players do not play the equilibrium described in Lemma A8.

Since the two equilibria provide *qualitatively* similar outcomes (as evidenced by the results above), the main substantive lessons are unchanged with this assumption.

Proposition A1 (Non-Deferential Equilibrium)**.** There always exists a non-deferential equilibrium σ^n with equilibrium strategies and beliefs characterized by Lemmas A6 and A10.

Proof. The proof follows directly from Lemmas A6 and A10. □

Proposition A2 (Deferential Equilibrium)**.** If and only if $\eta \leq \tilde{\eta}$, there exists a deferential equilibrium σ^0 with equilibrium strategies and beliefs characterized by Lemmas A6 and A10.

Proof. This follows directly from Lemmas A7 and A9. □

The previous result hinges on T exerting effort (see Lemma A7). Therefore, a necessary condition for A to defer in any equilibrium is that T engages in fact-finding effort.

Corollary A1. If $e = 0$, the unique equilibrium is a non-deferential equilibrium, σ^n.

Moreover, no deference is possible if T is too biased.

Corollary A2. If $\delta_T > \overline{\delta} \equiv 1 - \frac{1}{\varepsilon}(1 - 2\delta_A)$, then there is no deferential equilibrium for any $\eta \in (0, 1)$.

We now examine the welfare implications of the model's equilibria. If $\eta \leq \tilde{\eta}$, then there are multiple equilibria: a deferential equilibrium σ^0 and a non-deferential equilibrium σ^n. The next result demonstrates that the non-deferential equilibrium is inefficient when both exist.

Lemma A11. From an ex ante perspective, T is strictly better off under σ^0 than under σ^n for all $\eta \in [0,1]$.

Proof. Let T's ex ante utility from effort be given by (1) and (4). Since σ^0 does not exist if $e^0 = 0$, it suffices to show that $U_T^0(e=1) > U_T^n(e=0)$ and $U_T^0(e=1) > U_T^n(e=1)$. It is immediately clear that $U_T^0(e=1) > U_T^n(e=0)$. Finally, we show directly that $U_T^0(e=1) > U_T^n(e=1)$. Using (1) and (4) and simplifying gives the following condition

$$\alpha(\delta_T - (1-\varepsilon)(1-\delta_T) - \varepsilon\eta\delta_A) > \varepsilon\eta(\delta_T - \delta_A)(q - \kappa_L)$$

Using Assumption A3, $\alpha > q - \kappa_L$. So, the left hand side is minimized by substituting $\alpha = q - \kappa_L$. This yields

$$\delta_T > \frac{1}{2} + \frac{\varepsilon}{2}(\delta_T(1+\eta) - 1)$$

which is true for all parameter values. We have proven $U_T^0(e=1) > U_T^n(e=1)$. □

Proposition A3 (Deference Pareto Dominates Non-Deference)**.** If σ^0 and σ^n both exist, then σ^n is weakly Pareto dominated by σ^0.

Proof. Suppose σ^0 and σ^n both exist. Lemma A11 establishes that T is strictly better off under σ^0 than under σ^n. We now demonstrate that A is strictly better off under σ^0 than under σ^n. First, note that A's ex ante expected utility from each equilibrium is given by:

$$U_A(\sigma^0) = \delta_A + \varepsilon(1-\delta_T) \qquad\qquad U_A(\sigma^n) = 1 - \delta_A + e^n\varepsilon\delta_A\eta$$

It suffices to show $U_A(\sigma^0) \geq U_A(\sigma^n)$ if $e^n = 1$. This condition reduces to

$$\eta \leq \frac{\varepsilon(1-\delta_T) - (1-2\delta_A)}{\varepsilon\delta_A} = \tilde{\eta}$$

The condition holds since $\eta \leq \tilde{\eta}$ by the supposition that σ^0 exists. □

However a more troubling kind of inefficiency emerges if $\eta \in (\tilde{\eta}, \hat{\eta})$. In this situation, A does not defer despite the fact that T would exert effort if A did defer but does not exert effort under non-deference.

Proposition A4 (Inefficient Non-Deference)**.** If $\tilde{\eta} < \eta < \hat{\eta}$ and $\delta_T \leq \overline{\delta}$, there is a unique non-deferential equilibrium that is inefficient.

Proof. Let $\tilde{\eta} < \eta < \hat{\eta}$ and $\delta_T \leq \overline{\delta}$. To prove the statement, we show the equilibrium is unique, then that it is inefficient. First, by Proposition A2, since $\eta > \tilde{\eta}$, there is no deference equilibrium. And by Proposition A1, there is a non-deference equilibrium. Therefore, by Lemma A5 and Assumption A5 there is a unique non-deference equilibrium, σ^n.

Next, to show that the equilibrium is inefficient, it suffices to identify a strategy profile that would strictly increase the payoff of one player without decreasing the payoff of another player. By Lemma A11, T is always strictly better off in equilibrium σ^0 than in equilibrium σ^n.

Under the current equilibrium σ^n, $e^n = 0$ since $\eta < \hat{\eta}$ (see Lemma A10). A's expected payoff is

$$U_A(\sigma^n) = 1 - \delta_A$$

Suppose that the players instead played the strategy profiles in σ^0, but that A's equilibrium belief supported its deference. Then, Lemma A9 implies $e^0 = 1$, and given T's optimal judgment, the players' expected payoffs are

$$U_A(\sigma^0) = \delta_A + \varepsilon(1 - \delta_T)$$

$U_A(\sigma^0) \geq U_A(\sigma^n)$ if and only if

$$\delta_A + \varepsilon(1 - \delta_T) \geq 1 - \delta_A \iff \delta_T \leq 1 - \frac{1}{\varepsilon}(1 - 2\delta_A) = \overline{\delta}$$

This is true by assumption, proving the claim. □

2 Formal Results for Procedural Discretion

The following ensures the players play deferential equilibrium when it exists.

Assumption A6. In the extended game with procedural discretion, players do not play Pareto dominated equilibria.

Since the subgame after remand looks identical to the baseline model, the unique equilibrium (with Assumption A5 and Assumption A6) is as in Propositions A1 and A2.

Definition A3. T **exercises procedural discretion** if and only if $\pi \neq \eta$.

Lemma A12. If procedural discretion does not alter A's review posture, then T's ex ante utility is increasing in π and optimally sets $\pi = \eta$.

Proof. First note that T's ex ante utility is weakly increasing in η, holding fixed $\rho(x, o)$. This follows from equations (1), (2) and (4), and the fact that T's ex ante utility is $(1 - \delta_T)\alpha$ when $e = 0$. Then, if $\pi < \eta$ does not induce A's optimal $\rho(x, o)$ to change, then it is utility maximizing to set $\pi = \eta$. □

Lemma A13. Assume A does not remand when indifferent. A remands if and only if $\pi \leq \tilde{\eta} < \eta$ and $\eta \geq \hat{\eta}$.

Proof. Denote $U_A(\sigma, \pi)$ as A's ex ante utility from the baseline model with procedure π. Since A does not remand when indifferent, it does not remand if $\pi = \eta$. Then, A remands if $\pi < \eta$ if

$$U_A(\sigma, \eta) \geq U_A(\sigma, \pi)$$

Lemma A12 indicates that if π does not alter whether T receives deference, then $\pi = \eta$. Accordingly, A does not remand. In the following cases, π does not alter whether T receives deference:

- If $\tilde{\eta} < 0$, then T never receives deference for all $\pi \leq \eta$.
- If $0 \leq \eta \leq \tilde{\eta}$, then T receives deference for all $\pi \leq \eta$.

Now, we consider the remaining case:

- Suppose $\eta > \tilde{\eta}$. If $\pi \in (\tilde{\eta}, \eta)$, then T receives non-deference under either procedure and A does not remand. If $\pi \leq \tilde{\eta}$, A remands if

$$U_A(\sigma^n, \eta, e^n) > U_A(\sigma^0, \pi) \iff e^n\eta > \frac{\varepsilon(1-\delta_T) - (1-2\delta_A)}{\varepsilon\delta_A} = \tilde{\eta}$$

If $e^n = 0$, then this does not hold and A does not remand. If $e^n = 1$ this condition holds and A remands.

Therefore, we have shown that A remands if and only if $\pi \leq \tilde{\eta} < \eta$ and $e^n = 1$. Finally, $e^n = 1$ if $\eta \geq \hat{\eta}$ by Lemma A10. Then, A remands if $\pi \leq \tilde{\eta} < \eta$ and $\eta \geq \hat{\eta}$ □

Let $\tilde{\delta} \equiv 1 - \eta\delta_A - \frac{1}{\varepsilon}(1-2\delta_A)$ so that $\delta_T \leq \tilde{\delta} \iff \eta \leq \tilde{\eta}$. Moreover, note that $\tilde{\eta} < 0 \iff \delta_T > \overline{\delta}$.

Proposition A5 (Procedural Discretion)**.** In the equilibrium of the extended game, T exercises procedural discretion if $\delta_T \in (\tilde{\delta}, \overline{\delta}]$ and $\eta < \hat{\eta}$ and sets $\pi = \tilde{\eta} < \eta$.

Proof. T exercises procedural discretion if it is a best response, given A's sequentially rational remands, characterized by Lemma A13. Moreover, Lemma A12 indicates that when $\tilde{\eta} < 0$ or $0 \leq \eta \leq \tilde{\eta}$, $\pi = \eta$. Equivalently, if $\delta_T > \overline{\delta}$ or $\delta_T \leq \tilde{\delta}$, then $\pi = \eta$.

Now suppose $0 \leq \tilde{\eta} < \eta < 1$, or equivalently $\delta_T \in (\tilde{\delta}, \overline{\delta}]$. **Case 1.** If $\eta < \hat{\eta}$, then setting $\pi \in (0, \tilde{\eta}]$ gets T deference instead of non-deference and A does not remand. Since T is better off under deference than non-deference (Lemma A11), it is optimal to set $\pi \leq \tilde{\eta}$. Moreover, since T's utility is increasing in π holding fixed A's review strategy (Lemma A12), then T optimally sets $\pi = \tilde{\eta} < \eta$. **Case 2.** If $\eta \geq \hat{\eta}$, then A remands after any $\pi < \eta$. Accordingly, T chooses $\pi = \eta$. □

3 Formal Results for De Facto Law

In this section, assume that adjudication happens according to the extended model with procedural discretion. Then, conditional on δ_T, the proportion of pro-defendant outcomes is

$$\Delta_T = \begin{cases} 1 - \varepsilon(1-\delta_T) & \text{if } (\delta_T \leq \overline{\delta} \text{ and } \eta < \hat{\eta}) \text{ or } (\delta_T \leq \tilde{\delta}_T \text{ and } \eta \geq \hat{\eta}) \\ \delta_A e^n \varepsilon\eta & \text{otherwise} \end{cases}$$

Definition A4. T is a **moderately biased** judge if

$$\delta_T \leq \begin{cases} \overline{\delta} & \text{if } \eta < \hat{\eta} \\ \tilde{\delta}_T & \text{if } \eta \geq \hat{\eta} \end{cases}$$

and an **extremely biased** judge otherwise.

Proposition A6 (De Facto Law)**.** If T is moderately biased, then $1 - \varepsilon(1 - \delta_T) > \delta_A$. If T is extremely biased, then $e\delta_A\varepsilon\eta < \delta_A$.

Proof. We prove this directly considering two mutually exclusive and exhaustive cases.

Case 1. ($\delta_T \leq \overline{\delta}$ and $\eta < \hat{\eta}$) or ($\delta_T \leq \tilde{\delta}_T$ and $\eta \geq \hat{\eta}$). Then, by Proposition A2, T receives deference and $\Delta_T = 1 - \varepsilon(1 - \delta_T)$. Then, note that for all $\varepsilon \in [0, 1]$ and $0 < \delta_A < \frac{1}{2} < \delta_T < 1$:

$$1 - \varepsilon(1 - \delta_T) > \delta_A$$

We have shown directly that $\Delta_T > \delta_A$.

Case 2. ($\delta_T > \tilde{\delta}_T$ and $\eta \geq \hat{\eta}$) or $\delta_T > \overline{\delta}$. Then, by Proposition A1, T does not receive deference and $\Delta_T = \delta_A e^n \varepsilon\eta$. Then, then for all $\varepsilon, \eta \in (0, 1)$:

$$\delta_A e^n \varepsilon\eta < \delta_A$$

We have shown directly that $\Delta_T < \delta_A$. □