Appendix: Supporting Information for "Whistleblowing and Compliance in the Judicial Hierarchy"

Deborah Beim
Alexander V. Hirsch
Jonathan P. Kastellec

American Journal of Political Science

SI-1 Proofs

We first state and prove the following Lemma employed in many of the subsequent proofs.

Lemma 9 *The function $G(\phi_d(c_L, c_W)) - G(\phi_{nd}(c_L, c_W))$ is decreasing in c_W when $F(\cdot)$ and $G(\cdot)$ are uniform.*

Proof: With $G(\cdot)$ uniform it sufficies to show $\phi_d(c_L, c_W) - \phi_{nd}(c_L, c_W)$ is decreasing in c_W. The function $\phi_d(c_L, c_W)$ may be rewritten as

$$\frac{\phi_{nd}(c_L, c_W)}{P(x \in [c_W, H] \mid x > c_W)} + (E[x \mid x \in [c_W, H]] - E[x \mid x \in [c_L, c_W]]),$$

and with substitution and algebra $\phi_d(c_L, c_W) - \phi_{nd}(c_L, c_W)$ may be rewritten as

$$P(x > H \mid x > c_W) \cdot (H - E[x \mid x \in [c_W, H]]) + (E[x \mid x \in [c_W, H]] - E[x \mid x \in [c_L, c_W]]).$$

Now for $F(\cdot)$ uniform $H - E[x \mid x \in [c_W, H]] = \left(\frac{\bar{x}}{2}\right) \cdot P(x \in [c_W, H])$. Substituting back in and simplifying we then have that $\phi_d(c_L, c_W) - \phi_{nd}(c_L, c_W) =$

$\left(\frac{\bar{x}}{2}\right) \cdot P(x > H)(1 - P(x > H \mid x > c_W)) + (E[x \mid x \in [c_W, H]] - E[x \mid x \in [c_L, c_W]])$ which $\rightarrow$

$$\begin{aligned}\frac{\partial}{\partial c_W}(\phi_d(c_L, c_W) - \phi_{nd}(c_L, c_W)) &= -\left(\frac{\bar{x}}{2}\right) \cdot P(x > H) \frac{f(c_W)}{P(x > c_W)} P(x > H \mid x > c_W) \\ &\quad + \frac{\partial}{\partial c_W}(E[x \mid x \in [c_W, H]] - E[x \mid x \in [c_L, c_W]]) \\ &= -\left(\frac{\bar{x}}{2}\right) \cdot f(c_W) \cdot [P(x > H \mid x > c_W)]^2 < 0\end{aligned}$$

since $\frac{\partial}{\partial c_W}(E[x \mid x \in [c_W, H]]) = \frac{\partial}{\partial c_W}(E[x \mid x \in [c_L, c_W]])$ for the uniform. ■

Proof of Lemma 1 As described in the beginning of Section 3 and footnote 4, in the main text we restrict attention to equilibria where dissent increases the probability of review – in other words, where $\phi_d > \phi_{nd}$. However, there sometimes exists a second more fragile class of equilibria with a different structure, in which "dissent" decreases the probability of review (i.e. $\phi_{nd} > \phi_d$) because it signals that noncompliance occured but that it was relatively minor. Such "dissents" are more easily interpeted as "concurrences." Below we prove a more general statement about the form of all equilibria; Lemma 1 is a straightforward corollary of this more general statement.

Lemma 10 *All equilibria are in* ***cutpoint strategies*** $(c_L^*, c_W^*, \phi_s^*, \phi_{ns}^*)$ *with* $c_L^* \in [L, H]$ *and* $c_W^* < H$. *There are two types of equilibria.*

- *In a* ***dissent*** *equilibrium,*
  - *The lower court* L *rules liberally for* $x \geq c_L^*$ *and conservatively otherwise.*
  - *The potential whistleblower* W *never dissents following a conservative ruling, and issues a costly dissent following a liberal ruling whenever the facts are sufficiently conservative* ($x \leq c_W^*$).
  - *The higher court* H *never reviews conservative rulings, and sometimes reviews liberal rulings. Specifically, he reviews a liberal ruling i.f.f.* $k \leq \phi_s^*$ *when* W *dissents, and i.f.f.* $k \leq \phi_{ns}^* < \phi_s^*$ *when* W *does not dissent.*
- *In a* ***concurrence*** *equilibrium,*
  - *The lower court* L *rules liberally for* $x \geq c_L^*$ *and conservatively otherwise.*
  - *The potential whistleblower* W *never concurs with a conservative ruling, and issues a costly concurrence following a liberal ruling when the facts are in* $x \in [c_W^*, H]$.
  - *The higher court* H *never reviews conservative rulings, and sometimes reviews liberal rulings. Specifically, he reviews a liberal ruling i.f.f.* $k \leq \phi_{ns}^*$ *when* W *fails to concur, and i.f.f.* $k \leq \phi_s^* < \phi_{ns}^*$ *when* W *concurs.* ■

Proof: The interpretation of the costly signal is dependent on equilibrium. Specifically, it can either signal that the case facts are more conservative – in which case it is interpreted as a *dissent* – or it can signal that the case facts are more liberal – in which case it is interpreted as a *concurrence*. In the former instance it raises the probability of review, while in the latter instance it lowers it. Thus, we denote the whistleblower's actions using the agnostic label $j \in \{s, ns\}$ – that is, she either issued the costly signal or she did not.

When H is called to play he is the final mover, and the history $h \in \{lib, con\} \times \{s, ns\}$ can take four possible values. For each history he calculates a *net gain* from review that is derived using Bayes' rule and the other players' strategies – denote this net gain $\phi_{i,j}$ where i denotes the ruling and j denotes the signal value. Because H is the last mover, his best-response takes the form of a cutpoint for each history – he reviews i.f.f $k < \phi_{i,j}$, where k is the cost of review.

Now consider the whistleblower W. If faced with a compliant ruling of either *lib* or *con* she will never send the costly signal, since the ruling will stand whether or not H reviews. Now suppose she is faced with a noncompliant ruling of *lib*. If she issues the costly signal then she will pay an up front cost of d. If the signal raises the chance of review ($\phi_{l,s} > \phi_{l,ns}$) then her net gain will be

$$(G(\phi_{l,s}) - G(\phi_{l,ns})) \cdot ((W - x) - \alpha\varepsilon),$$

since whenever $k \in (\phi_{l,s}, \phi_{l,ns})$ the signal results in a review and a reversal that otherwise would not have occurred. If the signal lowers the chance of review ($\phi_{l,s} < \phi_{l,ns}$) then her net gain will be

$$(G(\phi_{l,ns}) - G(\phi_{l,s})) \cdot ((x - W) + \alpha\varepsilon),$$

since whenever $k \in (\phi_{l,ns}, \phi_{l,s})$ the signal *prevents* a review and reversal that would otherwise have occured. Recalling that she will never dissent on a compliant liberal ruling, her best response is either to signal when

$$x < \min\left\{(W - \alpha\varepsilon) - \frac{d}{G(\phi_{l,s}) - G(\phi_{l,ns})}, H\right\}$$

if the signal increases review, or to signal when

$$x \in \left[(W - \alpha\varepsilon) + \frac{d}{G(\phi_{l,s}) - G(\phi_{l,ns})}, H \right]$$

if the signal decreases review. Thus, her strategy takes the desired forms.

Now consider the lower court L. It is strictly dominant to rule compliantly when it and the higher court agree ($x < L$ or $x > H$) since this ensures its desired outcome and there is no chance of being reversed (even upon review). For cases within $x \in [L, H]$ it is always the case that ruling liberally elicits a higher probability of review when $x < c_W^*$ than when $x > c_W^*$. If $\phi_s > \phi_{ns}$ then W signals when $x < c_W^*$, (thereby raising the probabiliy of review) and if $\phi_{ns} > \phi_s$ then W fails to signal when $x < c_W^*$ (also raising the probability of review). Consequently, L must also play a cutpoint strategy c_L^*; if it is unwilling to comply on some x then it is also unwiling to comply on some $x' > x$ where the benefits of the liberal outcome are greater and the probability of review is (weakly) lower.

Finally, because L always rules *lib* when $x > H$, any conservative ruling must be compliant. Since W never signals on a compliant ruling, PBE requires that $\phi_{c,ns} = 0$; that is, H evaluates the net gain of reviewing a conservative ruling without a signal to be 0 and never reviews it. A conservative ruling accompanied by a costly signal is off-path – in PBE $\phi_{c,s}$ is unrestricted, and the value will generate some off-path best response behavior for the whistleblower when she observes a noncompliant conservative ruling ($x > H$). However, choosing these pairs arbitrarily does not perturb equilibrium because ruling *lib* is stricty dominant for L whenever $x > H$ and a conservative ruling would be noncompliant; thus we leave these values unspecified. This completes the proof. ■

Proof of Lemmas 2 – 5 Lemmas 2 – 4 follow immediately from the in-text analysis. The necessary and sufficient condition for cutpoints $(c_L^*, c_W^*, \phi_{nd}^*, \phi_d^*)$ to be an equilibrium in Lemma 5 is a straightforward assembly of the best-response characterizations in the preceding Lemmas; that is, strategies are an equilibrium if and only if every player is best-responding down every path of play given the strategies of the other players. A more explicit statement of the assembled necessary and sufficient conditions is included below for clarity.

1. $\phi_d^* = \phi_d(c_L, \max\{c_W, c_L\})$ and $\phi_{nd}^* = \phi_{nd}(c_L, \max\{c_W, c_L\})$

   (higher court best-response)

2. $c_W^* = \min\{c_W(\phi_d^*, \phi_{nd}^*), H\}$

   (whistleblower best-response)

3. $c_L^* = c_L(c_W^*, \phi_d^*, \phi_{nd}^*)$

   (lower court best-response). ■

Proof of Proposition 1 The proof proceeds in two parts. First, we show that the benchmarks $\underline{c}_L$ and $\overline{c}_L$ exist, are unique, and satisfy $L < \underline{c}_L < \overline{c}_L < H$. Second, we prove the main body of the statement.

Part 1

The cutpoint $\underline{c}_L$ solves $c_L = x^*(G(\phi_{nd}(c_L, c_L)))$ and the cutpoint $\overline{c}_L$ solves $c_L = x^*(G(\phi_d(c_L, c_L)))$. Using the definitions of $\phi_d(\cdot)$, $\phi_{nd}(\cdot)$, $x^*(\cdot)$, and that $G(0) = 0$, it is easily verified that the right hand sides of both equalities are (1) greater than L when $c_L = L$, (2) equal to L when $c_L = H$, and (3) strictly decreasing in c_L. Thus, both have a unique solution interior to (L, H). Finally, to see that $\overline{c}_L > \underline{c}_L$ it suffices to observe that $\phi_d(c_L, c_L) > \phi_{nd}(c_L, c_L)$ $\forall c_L$ and $x^*(G(\phi))$ is decreasing in ϕ.

Although the following is inessential to the proof, we now also briefly explain why $\underline{c}_L$ and $\overline{c}_L$ are the unique equilibrium levels of compliance in the no information and complete information 2 player games, respectively.

In the no information game absent the whistleblower, the higher court will use a single threshold ϕ for reviewing a liberal disposition, and this threshold must equal the expected benefit of review given his beliefs about the lower court's behavior. Applying the analysis in Section 3.3, the lower court must use a compliance cutpoint $c_L = x^*(G(\phi))$ in a best response. Given a cutpoint strategy by the lower court, the higher court must believe upon observing a liberal disposition that the case facts are $x > c_L$; it is easily verified that the net benefit of review under these circumstances is equal to $\phi_{nd}(c_L, c_L)$. Combining these two best response conditions yields the equilibrium condition $c_L = x^* (G(\phi_{nd}(c_L, c_L)))$.

In the complete information game the higher court observes the case facts to be x. Her expected benefit of reviewing and reversing the case, should the lower court rule noncompliantly, is equal to $H - x = \phi_d(x, x)$. Thus, she will review and reverse a noncompliant liberal ruling with probability $G(\phi_d(x, x))$. The lower court's net gain from noncompliance is $(1 - G(\phi_d(x, x)))(x - L)$ and the cost is $G(\phi_d(x, x)) \cdot \varepsilon$. The net gain is increasing and equal to 0 at $x = L$, and the cost is decreasing and equal to 0 at $x = H$; hence there is a unique interior cutpoint $\hat{c}_L$ below which the lower court will comply and above which she will not, which satisfies

$$(1 - G(\phi_d(\hat{c}_L, \hat{c}_L)))(\hat{c}_L - L) = G(\phi_d(\hat{c}_L, \hat{c}_L)) \cdot \varepsilon \iff \hat{c}_L = x^* (G(\phi_d(\hat{c}_L, \hat{c}_L))),$$

which is the definition of $\bar{c}_L$

Part 2

We seek to characterize a *partial equilibrium* $(c_L^*, \phi_{nd}^*, \phi_d^*)$ where the lower and higher courts are best responding to each other and the whistleblower, and the whistleblower's strategy is to use a dissent cutpoint of c_W. In other words, we seek values of $(c_L^*, \phi_{nd}^*, \phi_d^*)$ that jointly satisfy Lemmas 2 and 4 given c_W.

By substituting the best-response conditions for the higher court into the best response condition for the lower court, we derive the following necessary and sufficient condition for existence of a partial equilibrium with compliance level c_L:

$$c_L = \min\{x^* (G(\phi_d(c_L, \max\{c_W, c_L\}))), \max\{x^* (G(\phi_{nd}(c_L, \max\{c_W, c_L\}))), c_W\}\} \quad (7)$$

The right hand side of the above is a function of c_L and c_W, and we henceforth denote it $\hat{c}_L(c_L; c_W)$. Intuitively, $\hat{c}_L(c_L; c_W)$ is the lower court's best response cutpoint when the higher court *believes it* to be using cutpoint c_L, and everybody believes the whistleblower to be using cutpoint c_W.[14] A partial equilibrium level of compliance is a fixed point of this function.

We now show that for every c_W there exists a *unique* value of c_L satisfying the equality in (7), and that value is equal to the cutpoint $c_L^*(c_L)$ described in the Proposition. First, using the definitions of $\phi_d(\cdot)$, $\phi_{nd}(\cdot)$, $x^*(\cdot)$, and that $G(0) = 0$, the following facts about the r.h.s. of the equality are easily verified: (1) it is weakly decreasing in c_L (since it is the middle value of three weakly decreasing functions), 2) $c_L^*(L, c_W) > L$, and 3) $c_L^*(H, c_W) = L$. This establishes that there is a unique solution interior to (L, H).

For the next steps also recall that $\underline{c}_L = x^*(G(\phi_{nd}(\underline{c}_L, \underline{c}_L)))$ and $\bar{c}_L = x^*(G(\phi_d(\bar{c}_L, \bar{c}_L)))$.

Region 1: To see that $\underline{c}_L = \hat{c}_L(\underline{c}_L; c_W)$ when $c_W < \underline{c}_L$, note that the latter implies

$$\begin{aligned} \hat{c}_L(\underline{c}_L, c_W) &= \min\{x^*(G(\phi_d(\underline{c}_L, \underline{c}_L))), \max\{x^*(G(\phi_{nd}(\underline{c}_L, \underline{c}_L))), c_W\}\} \\ &= \min\{x^*(G(\phi_d(\underline{c}_L, \underline{c}_L))), \max\{\underline{c}_L, c_W\}\} = \underline{c}_L \end{aligned}$$

Intuitively, when c_W is less than the cutpoint $\underline{c}_L$ that the lower court would use absent the whistleblower, then the lower court's partial equilibrium compliance cutpoint is the same as absent W – the whistleblower never dissents on path, absent dissents the higher court draws the same inference as he would absent the whistleblower, and so the lower court complies to the same degree. In this case, the degree of compliance is constant in c_W and equal to $\underline{c}_L$, the probability of review after dissent is $G(\phi_d(\underline{c}_L, \underline{c}_L))$ and also constant, and the probability of dissent is 0 (since $c_W < \underline{c}_L$).

Region 2: To see that $c_L^*(c_W) = c_W \iff c_W = \hat{c}_L(c_W, c_W)$ when $c_W \in [\underline{c}_L, \bar{c}_L]$, note that the latter implies (from the definitions of $\underline{c}_L$ and $\bar{c}_L$) that $c_W > x^*(G(\phi_{nd}(c_W, c_W)))$ and that $c_W < x^*(G(\phi_d(c_W, c_W)))$. Hence,

$$\hat{c}_L(c_W; c_W) = \min\{x^*(G(\phi_d(c_W, c_W))), \max\{x^*(G(\phi_{nd}(c_W, c_W))), c_W\}\} = c_W.$$

Intuitively, suppose $c_W \in [\underline{c}_L, \bar{c}_L]$ and the lower court were to comply exactly up to c_W. A

[14] Note that c_W is a complete contingent description of how W would behave after a liberal ruling on any case $x \in X$. L cannot "change c_W" off-equilibrium path. Rather, L's ruling, combined with the case facts, determine whether or not W dissents *based on* c_W. If, for example, W and L's strategies are described by cutpoints $L < c_L < c_W < H$, then W's strategy specifies precisely what would happen if L were to go "off-path" by ruling liberally on a case $x \in [L, c_L]$ – it would trigger a dissent.

dissent would perfectly signal that the case facts were at c_W, and thus L would want to comply for all $x < c_W$ since c_W is less than the cutpoint $\bar{c}_L$ it would use if H were perfectly informed about the case facts. Conversely, the absence of dissent would signal that $x > c_W$; since L would not comply on such cases if H drew the inference that $x \in [\underline{c}_L, H]$, she also would not comply when H draws the weaker inference that $x \in [c_W, H]$. Consequently, L's best response cutpoint is exactly at $\hat{c}_L(c_W) = c_W$ and we have a partial equilibrium. In this region, compliance is clearly increasing since it is equal to c_W, and the probability of review after dissent is $G(\phi_d(c_W, c_W)) = G(H - c_W)$ which is also decreasing in c_W.

Region 3: Suppose that $c_W > \bar{c}_L$ and denote $c_L^d(c_W)$ as the value of c_L that solves $c_L = x^*(G(\phi_d(c_L, c_W)))$. It is easily verified using steps identical to those for $\underline{c}_L$ and $\bar{c}_L$ that $c_L^d(c_W)$ is unique, well defined, and in (L, H). We now wish to prove that $c_W > \bar{c}_L \rightarrow c_L^d(c_W) = \hat{c}_L\left(c_L^d(c_W), c_W\right)$, meaning that the partial equilibrium compliance cutpoint is exactly $c_L^d(c_W)$. First, note that $\bar{c}_L = x^*(G(\phi_d(\bar{c}_L, \bar{c}_L))) \rightarrow \bar{c}_L > x^*(G(\phi_d(\bar{c}_L, c_W)))$ for $c_W > \bar{c}_L$ (since $x^*(G(\phi_d(c_L, c_W)))$ is decreasing in c_W). This then implies that the solution to $c_L^d(c_W) = x^*\left(G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)\right)$ is $< \bar{c}_L < c_W$. Now from the definition of $\hat{c}_L(c_{L;}c_W)$, the lower court's best response cutpoint for any (c_L, c_W) such that $c_W > x^*(G(\phi_d(c_L, \max\{c_W, c_L\})))$ is equal to $x^*(G(\phi_d(c_L, \max\{c_W, c_L\})))$; thus, the best response cutpoint to $\left(c_L^d(c_W), c_W\right)$ is $c_L^d(c_W)$ and we have a partial equilibrium.

To see the comparative statics, first note $c_L^d(c_W) = x^*\left(G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)\right)$ is decreasing in c_W since $\phi_d(\cdot)$ is decreasing in c_W, $G(\phi)$ is increasing in ϕ, and $x^*(q)$ is increasing in q. The probability of dissent $F(c_W) - F\left(c_L^d(c_W)\right)$ is then increasing in c_W since $c_L^d(c_W)$ is decreasing in c_W. Finally, to see that the probability of review given dissent $G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)$ is decreasing in c_W, implicitly differentiate the definition to get,

$$\frac{\partial}{\partial c_W}\left(c_L^d(c_W)\right) = \frac{\partial x^*\left(G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)\right)}{\partial q} \cdot \frac{\partial}{\partial c_W}\left(G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)\right)$$

Since $\frac{\partial x^*(\cdot)}{\partial q} > 0$, $\frac{\partial}{\partial c_W}\left(G\left(\phi_d\left(c_L^d(c_W), c_W\right)\right)\right)$ inherits the sign of $\frac{\partial}{\partial c_W}\left(c_L^d(c_W)\right)$ which as previously shown is negative. $\blacksquare$

Proof of Proposition 2

Part 1 - Equilibrium Characterization and Existence

A necessary and sufficient condition for a profile of cutpoints $(c_L^*, c_W^*, \phi_{nd}^*, \phi_d^*)$ to be an equilibrium of the complete model is that they jointly satisfy Lemmas 2 – 4. By Proposition 1, if the whistleblower uses cutpoint c_W then equilibrium requires that $c_L^* = c_L^*(c_W)$ (which is uniquely defined) and thus that $\phi_d^* = \phi_d(c_L^*(c_W), \max\{c_L^*(c_W), c_W\})$ and $\phi_{nd}^* = \phi_{nd}(c_L^*(c_W), \max\{c_L^*(c_W), c_W\})$, which are also uniquely defined. Because the necessary values of the other players strategies in an equilibrium are uniquely pinned down for every c_W, we can substitute these values into the whistleblower's best response characterization in Lemma 3 to yield a necessary and sufficient condition for existence of an equilibrium with whistleblowing cutpoint c_W:

$$c_W = \min\{c_W(\phi_d(c_L^*(c_W), \max\{c_L^*(c_W), c_W\}), \phi_{nd}(c_L^*(c_W), \max\{c_L^*(c_W), c_W\})), H\}.$$

Observe that the right hand side is a function of c_W alone, and we henceforth denote it $\hat{c}_W(c_W)$. Equilibrium values of c_W are fixed points of this function; the equilibrium condition in the main text is identical except with the definition of $c_W(\phi_d, \phi_{nd})$ substituted in. Intuitively, $\hat{c}_W(c_W)$ is the whistleblower's best response cutpoint when the lower and higher court *believe* her to be using cutpoint c_W, and play their corresponding partial equilibrium strategies.

Existence of an equilibrium whistleblowing cutpoint satisfying $c_W = \hat{c}_W(c_W)$ that is $\leq H$ (and hence an equilibrium of the complete model) then follows immediately from the fact that $\hat{c}_W(c_W) \leq H\ \forall c_W$.

Part 2 - Necessary and Sufficient Condition for Whistlebower Effects

It is helpful to more-explicitly write the definition of $\hat{c}_W(c_W)$ by substituting in the partial equilibrium values of the lower court's compliance $c_L^*(c_W)$. We have that

$$\hat{c}_W(c_W) = \begin{cases} \min\left\{(W-\alpha\varepsilon) - \frac{d}{G(\phi_d(\underline{c}_L,\underline{c}_L)) - G(\phi_{nd}(\underline{c}_L,\underline{c}_L))}, H\right\} & \text{for } c_W \leq \underline{c}_L. \\ \min\left\{(W-\alpha\varepsilon) - \frac{d}{G(\phi_d(c_W,c_W)) - G(\phi_{nd}(c_W,c_W))}, H\right\} & \text{for } c_W \in [\underline{c}_L, \overline{c}_L] \\ \min\left\{(W-\alpha\varepsilon) - \frac{d}{G\left(\phi_d\left(c_L^*(c_W),c_W\right)\right) - G\left(\phi_{nd}\left(c_L^*(c_W),c_W\right)\right)}, H\right\} & \text{for } c_W \geq \overline{c}_L \end{cases}$$

In the main proof we employ the following useful properties of $\hat{c}_W(c_W)$. First, it is constant for $c_W \leq \underline{c}_L$. Second, it is weakly decreasing for $c_W \in [\underline{c}_L, \overline{c}_L]$, and strictly decreasing if $\hat{c}_W(c_W) < H$. Third, $\hat{c}_W(\underline{c}_L) \geq \hat{c}_W(c_W)$ for any $c_W > \underline{c}_L$.

To show these properties, first notice that c_W only affects $\hat{c}_W(c_W)$ through the denominator of the fraction in the first term – this is the probability that dissent is *pivotal* for review given cutpoint c_W. Now, the first property is immediate from the definition. To prove the second and third properties, we argue as an intermediate step that $G(\phi_d(c_L, c_W)) - G(\phi_{nd}(c_L, c_W)) > G(\phi_d(c_L', c_W')) - G(\phi_{nd}(c_L', c_W'))$ for $c_L' > c_L$, $c_W' > c_W$, $c_L \leq c_W$ and $c_L' \leq c_W'$. This is because

$$\begin{aligned} G(\phi_d(c_L, c_W)) - G(\phi_{nd}(c_L, c_W)) &> G(\phi_d(c_L, c_W')) - G(\phi_{nd}(c_L, c_W')) \text{ (by Lemma 9)} \\ &> G(\phi_d(c_L', c_W')) - G(\phi_{nd}(c_L', c_W')) \text{ (by definitions)} \end{aligned}$$

This then implies both that $G(\phi_d(c_W, c_W)) - G(\phi_{nd}(c_W, c_W))$ is strictly decreasing for $c_W \in [\underline{c}_L, \overline{c}_L]$ (implying the second property) and that $G(\phi_d(\underline{c}_L, \underline{c}_L)) - G(\phi_{nd}(\underline{c}_L, \underline{c}_L)) > G(\phi_d(c_L^*(c_W), c_W)) - G(\phi_{nd}(c_L^*(c_W), c_W))$ for $c_W > \underline{c}_L$ (implying the third property).

We now proceed to the main proof. By the definition of $c_L^*(c_W)$, whistleblower effects – that is, greater compliance than $\underline{c}_L$ – occur in an equilibrium if and only if the equilibrium whistleblowing cutpoint c_W^* is $> \underline{c}_L$. To show that the desired condition is necessary and sufficient, it then suffices to show that (1) if it fails all equilibrium whistleblowing cutpoints are $\leq \underline{c}_L$, and (2) if it holds there exists an equilibrium whistleblowing cutpoint $> \underline{c}_L$.

(Necessity). If the condition fails, then $H > \underline{c}_L > (W-\alpha\varepsilon) - \frac{d}{G(\phi_d(\underline{c}_L,\underline{c}_L)) - G(\phi_{nd}(\underline{c}_L,\underline{c}_L))}$ which implies that $\underline{c}_L > \hat{c}_W(\underline{c}_L)$. Since $\hat{c}_W(\underline{c}_L) \geq \hat{c}_W(c_W)\ \forall c_W > \hat{c}_W(\underline{c}_L)$, there are no equilibrium whistleblowing cutpoints greater than $\underline{c}_L$ and hence no equilibria with whistleblower effects.

(Sufficiency) If the condition holds then $(W-\alpha\varepsilon) - \frac{d}{G(\phi_d(\underline{c}_L,\underline{c}_L)) - G(\phi_{nd}(\underline{c}_L,\underline{c}_L))} < \underline{c}_L < H$

which implies that $\underline{c}_L < \hat{c}_W(\underline{c}_L)$. Since $\hat{c}_W(\underline{c}_L)$ is constant for $c_W < \underline{c}_L$, this implies that all equilibrium whistleblowing cutpoints c_W^* are $> \underline{c}_L$. Consequently, all equilibria exhibit whistleblower effects (which is in fact stronger than the desired property). ■

Proof of Lemma 6 For the purposes of this proof it is helpful to explicitly express the dependence of the best response mapping on the whistleblower's parameters, i.e. $\hat{c}_W(c_W; W, d, \alpha)$. Several substantively unimportant subtleties are now worth noting. First, the mapping from the parameter space to maximum equilibrium compliance $\tilde{c}_L(W, d, \alpha)$ is not necessarily continuous. Second, in the comparative statics for each parameter (W, d, α) there need not always be a region with partial and *strictly diminishing* whistleblower effects. Instead, the region with full whistleblower effects may jump to one with partial whistleblower effects where compliance is constant. Third, the regions may be truncated for the cost of dissent d because it cannot fall below 0;[15] for example, if $W = L$ and $\alpha = 0$, then there would be no whistleblower effects in a dissent equilibrium for any feasible value of $d \geq 0$.

In this proof we formally describe the steps for the parameter W; steps for d, α are the same except the order of the regions is reversed (and thus identical for $-d$ and $-\alpha$) with the understanding that the parameter space for d is truncated. The proof proceeds in three parts. First, we show that when there are multiple equilibria the compliance maximizing equilibrium is the one with the lowest c_W; we denote this whistleblowing cutpoint $\tilde{c}_W(W, d, \alpha) = \min\{c_W : c_W = \hat{c}_W(c_W; W, d, \alpha)\}$. Hence, maximum equilibrium compliance $\tilde{c}_L(W, d, \alpha)$ is equal to the composite mapping $c_L^*(\tilde{c}_W(W, d, \alpha))$. Second, we prove several properties of $\tilde{c}_W(W, d, \alpha)$. Third, we apply parts 1 and 2 to show the desired result.

Part 1

We argue that when there are multiple equilibria, the compliance maximizing equilibrium is the one with the lowest c_W^*. Intuitively, this holds because to sustain a higher equilibrium whistleblowing cutpoint, the whistleblower's probability of being pivotal must be higher,

[15] This is not an assumption but an observation; if dissent in the literal real-world sense were beneficial rather than costly, then choosing not to dissent would be the costly signal of "dissent" in the model.

and since more whistleblowing reduces the probability of being pivotal *ceteris paribus*, this necessarily requires less compliance.

Recall from the proof of Proposition 2 that $\hat{c}_W(c_W; W, d, \alpha)$ is constant over $c_W \leq \underline{c}_L$ and larger at $\underline{c}_L$ than at any $c_W > \underline{c}_L$. As a result, there cannot be multiple equilibria where one exhibits no whistleblower effects and others do—either there is a unique equilibrium with no whistleblower effects ($c_W^* \leq \underline{c}_L$) or there are one or more equilibria all of which exhibit whistleblower effects ($c_W^* > \underline{c}_L$).

Now, if there are two equilibria with whistleblower effects $\hat{c}_W^* > c_W^* > \underline{c}_L$, then by definition,

$$\begin{aligned} c_W^* &= \min\left\{(W - \alpha\varepsilon) - \frac{d}{G\left(\phi_d\left(c_L^*\left(c_W^*\right), c_W^*\right)\right) - G\left(\phi_{nd}\left(c_L^*\left(c_W^*\right), c_W^*\right)\right)}, H\right\} \quad \text{and} \\ \hat{c}_W^* &= \min\left\{(W - \alpha\varepsilon) - \frac{d}{G\left(\phi_d\left(c_L^*\left(\hat{c}_W^*\right), \hat{c}_W^*\right)\right) - G\left(\phi_{nd}\left(c_L^*\left(\hat{c}_W^*\right), \hat{c}_W^*\right)\right)}, H\right\} \end{aligned}$$

This implies that,

$$G\left(\phi_d\left(c_L^*\left(\hat{c}_W^*\right), \hat{c}_W^*\right)\right) - G\left(\phi_{nd}\left(c_L^*\left(\hat{c}_W^*\right), \hat{c}_W^*\right)\right) > G\left(\phi_d\left(c_L^*\left(c_W^*\right), c_W^*\right)\right) - G\left(\phi_{nd}\left(c_L^*\left(c_W^*\right), c_W^*\right)\right),$$

which in turn could only be true if $c_L^*\left(\hat{c}_W^*\right) < c_L^*\left(c_W^*\right)$, since by Lemma 9 the difference $G\left(\phi_d\left(c_L, c_W\right)\right) - G\left(\phi_{nd}\left(c_L, c_W\right)\right)$ is decreasing in c_W. This shows the desired property.

Part 2

We now show that $\tilde{c}_W(W, d, \alpha)$ satisfies the following three properties:

1. it is weakly increasing in W

2. for any value of $c_W^* \in [-\infty, \bar{c}_L]$, there $\exists$ a unique W^* s.t. $\tilde{c}_W(W^*, d, \alpha) = c_W^*$

3. there $\exists W$ s.t. $\tilde{c}_W(W, d, \alpha) = H$

To see (1) consider two values of the whistleblower $W' > W$. By definition of $\tilde{c}_W(W, d, \alpha)$, any value of $c_W < \tilde{c}_W(W, d, \alpha)$ is also less than $\hat{c}_W(c_W; W, d, \alpha)$ (because it is less than the lowest fixed point). Since $\hat{c}_W(c_W; W, d, \alpha)$ is increasing in W, this furthermore implies that any value of $c_W < \tilde{c}_W(W, d, \alpha)$ is also less than $\hat{c}_W(c_W; W', d, \alpha)$. Thus, the lowest fixed point $\tilde{c}_W(W', d, \alpha)$ for W' must be $\geq$ the lowest fixed point $\tilde{c}_W(W, d, \alpha)$ for W.

To see (2), it is easy to verify from the equilibrium definition in Proposition 2 that for any value of $c_W^* < H$ there is a *unique* W^* s.t. c_W^* is *an* equilibrium whisteblowing cutpoint. However, this is not enough to show that $\tilde{c}_W(W^*, d, \alpha) = c_W^*$, i.e., that c_W^* is the *lowest* equilibrium whistleblowing cutpoint for W^*. We now show this property must also hold for any $c_W^* \leq \bar{c}_L$. To do so, it suffices to recall from the proof of Proposition 2 that the best-response mapping $\hat{c}_W(c_W; W, d, \alpha)$ is weakly decreasing for $c_W \leq \bar{c}_L$. Consequently, there is at most one fixed point $\leq \bar{c}_L$, so if such an equilibrium exists it must be the lowest one.

To see (3), observe that the probability dissent is pivotal $G(\phi_d(c_L^*(c_W), c_W)) - G(\phi_{nd}(c_L^*(c_W), c_W))$ is $> G(\phi_d(\bar{c}_L, H)) - G(\phi_{nd}(\bar{c}_L, H)) > 0$ by $c_L^*(c_W) \leq \bar{c}_L$ and Lemma 9. Thus, for any whistleblower W such that,

$$(W - \alpha\varepsilon) - \frac{d}{G(\phi_d(\bar{c}_L, H)) - G(\phi_{nd}(\bar{c}_L, H))} > H$$

the best response mapping $\hat{c}_W(c_W; W, d, \alpha)$ is also $> H$ for any c_W, and consequently the unique equilibrium involves whistleblowing cutpoint $c_W^* = H$.

Part 3

The properties proved in Part 2 jointly imply that (1) $\tilde{c}_W(W, d, \alpha)$ is first strictly increasing, continuous in W, and onto $[-\infty, \bar{c}_L]$ (2) beyond $\bar{c}_L$ the function continues to increase (but potentially discontinuously) until it reaches H, and (3) it is constant thereafter. Consequently, maximum equilibrium compliance $c_L^*(\tilde{c}_W(W, d, \alpha))$ exhibits the regions as described mirroring the regions of c_W – first with no whistleblower effects, followed by continuously increasing whistleblower effects up to $\bar{c}_L$, followed by (potentially discontinuously decreasing) partial whistleblower effects, and finally constant and partial whistleblower effects when W is fully reporting all instances of noncompliance. Properties and analysis are identical for $(-\alpha, -d)$ except that for d the regions may be truncated from the top.

Proof of Lemma 7 The proof proceeds in two parts. First, we characterize H's expected utility as a function of the whistleblower's cutpoint c_W in a partial equilibrium, as well as the derivative of that utility. Second, we use this analysis to prove the main results.

Part 1

In Region I of Proposition 1, H's expected utility as a function of c_W is constant since $c_L^*(c_W) = \underline{c}_L$ and dissent is off path. In Regions II and III H's complete expected utility, taking into account his review costs, is the expression:

$$\int_{-\infty}^{c_L^*(c_W)} \left(\frac{H-x}{2}\right) f(x)\,dx + \int_H^{\infty} \int_0^{\phi_{nd}\left(c_L^*(c_W),c_W\right)} -kg(k)\,dkf(x)\,dx$$
$$+\int_{c_L^*(c_W)}^{c_W} \left(\int_0^{\phi_d\left(c_L^*(c_W),c_W\right)} \left(\left(\frac{H-x}{2}\right)-k\right) g(k)\,dk + \int_{\phi_d\left(c_L^*(c_W),c_W\right)}^{\infty} \left(\frac{x-H}{2}\right) g(k)\,dk\right) f(x)\,dx$$
$$+\int_{c_W}^{H} \left(\int_0^{\phi_{nd}\left(c_L^*(c_W),c_W\right)} \left(\left(\frac{H-x}{2}\right)-k\right) g(k)\,dk + \int_{\phi_{nd}\left(c_L^*(c_W),c_W\right)}^{\infty} \left(\frac{x-H}{2}\right) g(k)\,dk\right) f(x)\,dx$$

This expression has a simple and easily interpretable derivative in the whistleblower's cutpoint c_W which is derived using Leibniz rule and canceling:

$$\left(\frac{\partial c_L^*(c_W)}{c_W}\right) f\left(c_L^*(c_W)\right) \left(\int_0^{\phi_d\left(c_L^*(c_W),c_W\right)} kg(k) + \left(1-G\left(\phi_d(\cdot)\right)\right)\left(H-c_L^*(c_W)\right)\right)$$
$$+f(c_W)\left(\int_{\phi_{nd}\left(c_L^*(c_W),c_W\right)}^{\phi_d\left(c_L^*(c_W),c_W\right)} \left((H-c_W)-k\right) g(k)\,dk\right). \tag{8}$$

The first line is the net gain resulting from the change in L's compliance behavior $c_L^*(c_W)$. It is the product of three subterms: 1) the density $f\left(c_L^*(c_W)\right)$ of cases at the compliance cutpoint, 2) the marginal change $\frac{\partial c_L^*(c_W)}{c_W}$ in the compliance cutpoint, and 3) the marginal benefit $\int_0^{\phi_d\left(c_L^*(c_W),c_W\right)} kg(k) + \left(1-G\left(\phi_d(\cdot)\right)\right)\left(H-c_L^*(c_W)\right)$ of switching from noncompliance to compliance at case $x = c_L^*(c_W)$. (This is because when $k < \phi_d\left(c_L^*(c_W), c_W\right)$ the outcome doesn't change but H saves the review cost, while when $k > \phi_d\left(c_L^*(c_W), c_W\right)$ H would not have reviewed either way but now gets a compliant outcome for free.)

The second line is the net gain or loss from the whistleblower sending the costly rather than free signal at case $x = c_W$, which results in H inferring that x is in $[c_L, c_W]$ rather than $[c_W, \bar{x}]$. This net gain is comprised of the density of cases $f(c_W)$ at the whistleblowing cutpoint, times the net benefit of obtaining the conservative outcome through a review when $k \in [\phi_{nd}(\cdot), \phi_d(\cdot)]$.

Part 2

We now show that the utility-maximizing whistleblowing cutpoint for H is strictly less than H and weakly greater than L; we do so by showing that the derivative is < 0 at H and > 0 at $c_W \in (\underline{c}_L, \bar{c}_L)$.

At $c_W = H$ we have $\left.\frac{\partial c_L^*(c_W)}{c_W}\right|_{c_W=H} < 0$ (so the first term is negative) and the second term reduces to $-f(c_W)\int_{\phi_{nd}\left(c_L^*(H),H\right)}^{\phi_d\left(c_L^*(H),H\right)} kg(k)\,dk < 0$. Intuitively, complete reporting of noncompliance is both costly in terms of compliance, and more reporting than H wants even absent the compliance effect.

For $c_W \in (\underline{c}_L, \bar{c}_L)$ we have $c_L^*(c_W) = c_W$ and the derivative simplifies to,

$$f(c_W)\left(\int_0^{\phi_{nd}(c_W,c_W)} kg(k)\,dk + \int_{\phi_{nd}(c_W,c_W)}^{\infty}(H - c_W)\right) > 0.$$

Intuitively, more whistleblowing is all gain since it converts the marginal case from one where the lower court is compliant only when reviewed, to one on which the lower court complies for sure. Thus, the utility maximizing cutpoint is $\geq \bar{c}_L$. ■

Proof of Lemma 8 Recall from the analysis in the proof of Lemma 7 that H's preferences for changes in the whistleblower's cutpoint c_W involves a trade off between the *equilibrium compliance cost* of more whistleblowing against the *marginal informational benefit*.

From equation (8), this marginal informational benefit (henceforth "MIB") is equal to,

$$f(c_W)\left(\int_{\phi_{nd}\left(c_L^*(c_W),c_W\right)}^{\phi_d\left(c_L^*(c_W),c_W\right)}((H - c_W) - k)\,g(k)\,dk\right)$$

The proof requires two substeps. First, we show that the MIB satisfies a single crossing property and is equal to 0 at some unique $c_W^{**} \in (\bar{c}_L, H)$. Second, we show that the equilibrium of the game where the whistleblower is a perfect agent who internalizes H's review costs involves that whistleblower using cutpoint c_W^{**}, and the higher and lower courts jointly best responding exactly as in the baseline model. These two properties then jointly imply that at the unique equilibrium with a perfect agent, the whistleblower's cutpoint is strictly greater than the cutpoint maximizing H's utility. The reason is that a necessary condition for some $\hat{c}_W$ to maximize H's utility (from eqn. 8) is for the MIB be equal to the marginal compliance cost. Since the marginal compliance cost is always strictly positive in Region III, at the utility maximizing $\hat{c}_W$ the MIB must also be strictly positive, so $\hat{c}_W$ must be strictly less than c_W^{**}.

Intuitively, a "clone" of H as whistleblower dissents too much because – lacking commitment power and responding to her interim incentives – she only takes into account the MIB of more whistleblowing and not the marginal compliance cost. (A clone who could commit ex-ante to her whistleblowing behavior would indeed induce the optimum for H).

Part 1

From the proof in Lemma 7, recall that the MIB is positive in Region II and negative at $c_W = H$. In Region III it can be rewritten as,

$$\left(\frac{\phi_d(\cdot) - \phi_{nd}(\cdot)}{\bar{x}\bar{k}}\right) \cdot \left((H - c_W) - \frac{\phi_d\left(c_L^*(c_W), c_W\right) + \phi_{nd}\left(c_L^*(c_W), c_W\right)}{2}\right)$$

To show this satisfies a single crossing property it suffices to show that the second term is decreasing in c_W, which can be written as,

$$\frac{1}{2}\left(\left((H - c_W) - \phi_d\left(c_L^*(c_W), c_W\right)\right) + \left((H - c_W) - \phi_{nd}\left(c_L^*(c_W), c_W\right)\right)\right)$$

It is simple to verify that $(H - c_W) - \phi_d\left(c_L^*(c_W), c_W\right)$ is decreasing in c_W. Hence for the desired property it suffices to show that $(H - c_W) - \phi_{nd}\left(c_L^*(c_W), c_W\right)$ is also decreasing in c_W, which in turn holds if $\frac{\partial(\phi_{nd}(\cdot))}{\partial c_W} > -1$. Taking this derivative we have

$$\frac{\partial}{\partial c_W}\left(\frac{(H - c_W)^2}{2(\bar{x} - c_W)}\right) = -\frac{H - c_W}{\bar{x} - c_W} + \frac{(H - c_W)^2}{2(\bar{x} - c_W)},$$

which immediately shows the desired property because $\frac{H - c_W}{\bar{x} - c_W} < 1$.

Part 2

In the slightly modified game where W internalizes H's review costs, H still uses cutpoint strategies as in the baseline model. When $\phi_d > \phi_{nd}$, W's net benefit of dissenting on a liberal ruling on x – conditional on that dissent being pivotal for review (i.e. $k \in [\phi_{nd}, \phi_d]$ – is now modified to be equal to $((W - x) - \alpha\varepsilon) - E\left[k \mid k \in [\phi_{nd}, \phi_d]\right]$, because she internalizes k. The net cost of dissent is again d. Thus as in the baseline model W uses a cutpoint strategy of dissenting whenever x is less than the minimum of H and

$$(W - \alpha\varepsilon - E\left[k \,|\, k \in [\phi_{nd}, \phi_d]\right]) - \frac{d}{G(\phi_d) - G(\phi_{nd})}$$

This in turn implies that L uses a cutpoint strategy, and that the form of the equilibrium and the partial equilibrium conditions from Proposition 1 are unchanged.

Now when W is a perfect agent in the sense of preferences, her ideal cutpoint is $= H$, $\alpha = 0$, and $d = 0$. Thus, her best response cutpoint is $H - E\left[k \,|\, k \in [\phi_{nd}, \phi_d]\right]$, which must be equal to the other players' beliefs about it in equilibrium. Substituting in the partial equilibrium conditions implies that a necessary and sufficient condition for an equilibrium with whistleblower cutpoint c_W is then

$$\begin{aligned} & H - \frac{\phi_d\left(c_L^*\left(c_W\right), c_W\right) + \phi_{nd}\left(c_L^*\left(c_W\right), c_W\right)}{2} = c_W \\ \iff \quad & \left(H - c_W\right) - \frac{\phi_d\left(c_L^*\left(c_W\right), c_W\right) + \phi_{nd}\left(c_L^*\left(c_W\right), c_W\right)}{2} = 0 \end{aligned}$$

which is equivalent to the condition for the MIB to be $= 0$. ■

SI-2 Institutional Design Analysis

In this supplemental analysis to the main text, we consider how an institutional designer would select the parameters in the whistleblower's utility function—her conservatism W, the cost of dissent d, and her share of the sanction α—to maximize compliance. Compliance is maximized when the whistleblower's payoffs are calibrated so that dissent is attractive, but not too attractive. Specifically, she must be willing to dissent *exactly up* to the intermediate "limit to compliance" ($\bar{c}_L$) derived in Proposition 1 and no further. Any less dissenting and compliance gains are foregone because the threat of dissent can induce more compliance. Any more dissenting is counterproductive due to the negative equilibrium effect of reducing the impact of dissent. The condition for dissenting precisely up to the limit $\bar{c}_L$ to constitute an equilibrium is stated in the following proposition.

Proposition 3 *Holding other parameters of the model fixed, compliance by the lower court is maximized when W, d, and α are jointly chosen so that the following equality holds:*

$$W - \alpha\varepsilon = \bar{c}_L + \frac{d}{G\left(\phi_d\left(\bar{c}_L, \bar{c}_L\right)\right) - G\left(\phi_{nd}\left(\bar{c}_L, \bar{c}_L\right)\right)}. \tag{9}$$

Using Proposition 3, we can extract a number of substantively interesting results about compliance-maximizing institutional design. We first consider how an institutional designer with power to choose only one of the cost of dissent d, the whistleblower's share of the sanction α, or the whistleblower's indifference point W, would change the parameter of interest in response to changes in one of the other two.

Corollary 2 *A compliance-maximizing institutional designer choosing d, α, or W would:*

- *lower the cost of dissent d if the whistleblower became more liberal or if her share of the sanction α increased;*[16]
- *choose a more conservative whistleblower if the cost of dissent d increased or the whistleblower's share of the sanction α increased;*

[16] Unless d were already 0, in which case she would leave it there.

- *decrease the whistleblower's share of the sanction α if the cost of dissent d increased or the whistleblower became more liberal.*

Intuitively, these comparative statics arise from the fact that decreasing the cost of dissent d, decreasing the sanction share α, and increasing whistleblower's conservatism W, are *substitutable* ways of increasing the whistleblower's willingness to dissent. Since compliance maximization requires intermediate whistleblowing, a compliance-maximizing institutional designer should respond to an increase in the whistleblower's intrinsic willingness to dissent by tamping down on the incentive to dissent—either by increasing the whistleblower's sanction share, increasing the cost of dissent, or by choosing a more liberal whistleblower.

Next, we consider how an institutional designer choosing one of d, α, or W would change that parameter in response to changes in the lower court's willingness to comply, either through a change in the cost of sanction ε or the conservatism of the lower court L. The effect of changing these latter parameters is to shift the limit to compliance ($\bar{c}_L$) in the equality in equation (9).

Corollary 3 *In response to an increase in the lower court's willingness to comply—either through an increase in the cost of sanction ε or its conservatism L—a compliance-maximizing institutional designer should* ***increase*** *the incentive to dissent by:*

- *lowering the cost of dissent d;*[17]
- *decreasing the whistleblower's share of the sanction α;*
- *choosing a more conservative whistleblower.*

The corollary states that when the lower court becomes *more* willing to comply, a compliance-maximizing institutional designer should adjust the whistleblower's parameters to *further encourage dissent*. This is counterintuitive: one might expect that as the lower court's propensity to comply increases, the need for dissent to inform the higher court of noncompliance would *decrease*.

[17] Again, unless it were already 0 in which case she would leave it there.

The reason for this surprising result is that the institution of dissent plays two interrelated, but distinct, roles as a tool to increase compliance. First, it *informs* the higher court that noncompliance has occurred. In this informational role, dissent is a substitute for direct mechanisms of control like increasing sanctions or ideological alignment with the higher court, or both. Second, because the higher court may take a costly action following a dissent, dissent is a *threat*. And, as a threat, dissent can increase compliance even if it is not carried out. But there is a limit to its ability to do so—specifically, the limit to compliance $\bar{c}_L$ derived in Proposition 1. The effect of increasing direct mechanisms of control such as reversal sanctions ε and the lower court's conservatism L is *both* to increase what compliance would be absent a whistleblower (i.e. $\underline{c}_L$) and to increase the limit to dissent with a whistleblower (i.e. $\bar{c}_L$). In other words, increasing direct mechanisms of control increases both compliance *and* the effectiveness of dissent as a threat for inducing even more compliance. The compliance-maximizing institutional response is therefore to have more dissent.

Combined Proof of Proposition 3 and Corollaries 2 and 3 Maximum feasible compliance in equilibrium is $\bar{c}_L$, and which occurs i.f.f. $c_W^* = \bar{c}_L$. It is straightforward to verify from Proposition 2 that $c_L^* = c_W^* = \bar{c}_L$ is an equilibrium if and only if the equality in Proposition 3 holds.

To prove the institutional design comparative statics in Corollaries 2 and 3, note first that $\bar{c}_L$ is a function of L and ε that is implicitly defined as $\bar{c}_L(\varepsilon, L) = L + \frac{\varepsilon(H-\bar{c}_L(L,\varepsilon))}{\bar{k}-(H-\bar{c}_L(L,\varepsilon))}$. It is easy to verify that $\bar{c}_L(\varepsilon, L)$ is strictly increasing in L and ε. We now prove comparative statics on the compliance-maximizing choice of W, which we denote $\bar{W}(d, \alpha, \varepsilon, L)$. From the proposition, this quantity is defined as

$$\bar{W}(d, \alpha, \varepsilon, L) - \alpha\varepsilon = \bar{c}_L(\varepsilon, L) + \frac{d\bar{k}}{\phi_d(\bar{c}_L(\varepsilon, L), \bar{c}_L(\varepsilon, L)) - \phi_{nd}(\bar{c}_L(\varepsilon, L), \bar{c}_L(\varepsilon, L))}.$$

The properties in Corollary 2 clearly follow from the fact that l.h.s. is increasing in W and decreasing in α, and the r.h.s. is increasing in d. The properties in Corollary 3 follow from the fact that the r.h.s. is increasing in $\bar{c}_L$ (which in turn follows from Lemma 9) and that

$\bar{c}_L(\varepsilon, L)$ is strictly increasing in L and ε. Nearly identical steps prove the properties for $\bar{\alpha}(W, d, \varepsilon, L)$, the compliance-maximizing share of the sanction.

A slight wrinkle arises for the compliance-maximizing cost of dissent, since d cannot go below 0. First we define $\bar{d}(W, \alpha, \varepsilon, L)$ to be the dissent cost satisfying

$$W - \alpha\varepsilon = \bar{c}_L(\varepsilon, L) + \frac{\bar{d}(W, \alpha, \varepsilon, L) \cdot \bar{k}}{\phi_d(\bar{c}_L(\varepsilon, L), \bar{c}_L(\varepsilon, L)) - \phi_{nd}(\bar{c}_L(\varepsilon, L), \bar{c}_L(\varepsilon, L))}.$$

$\bar{d}(W, \alpha, \varepsilon, L)$ has straightforward comparative statics like the previous implicit characterizations. Moreover, when $\bar{d}(W, \alpha, \varepsilon, L) > 0$ it is the compliance-maximizing dissent cost.

Next we argue that when $\bar{d}(W, \alpha, \varepsilon, L) < 0$, the compliance-maximizing dissent cost is 0. To see this, recall from the proof of Lemma 6 that the compliance-maximizing equilibrium whistleblowing cutpoint $\tilde{c}_W(W, d, \alpha)$ is decreasing in d, and the lower court's best-response $c_L^*(c_W)$ is increasing in c_W when $c_W < \bar{c}_L$. Hence, to show that increasing d above 0 decreases maximum equilibrium compliance only requires showing that $\tilde{c}_W(W, 0, \alpha) < \bar{c}_L$. This follows from the observations that 1) $\tilde{c}_W(W, 0, \alpha) = W - \alpha\varepsilon$, and 2) $\bar{d}(W, \alpha, \varepsilon, L) < 0$ implies that $W - \alpha\varepsilon < \bar{c}_L(\varepsilon, L)$.

Finally, since the compliance-maximizing dissent cost is $\max\left\{0, \bar{d}(W, \alpha, \varepsilon, L)\right\}$, and the function $\bar{d}(W, \alpha, \varepsilon, L)$ satisfies the desired monotone comparative statics, $\max\left\{0, \bar{d}(W, \alpha, \varepsilon, L)\right\}$ must also (weakly) satisfy them. $\blacksquare$