A Theory of Claim Resolution

Scott Baker
Washington University in St. Louis

Lewis Kornhauser
New York University

Preliminary and Incomplete

September 10, 2019

Abstract

We build a model of claim resolution. A claim consists of two facts, only one of which is observable to the principal. The agent might disagree with the principal about (1) the importance each fact should play in the claim's resolution (formal bias) or (2) the overall proportion of claims which should be found valid (substantive bias) or both. We first characterize the principal's optimal review policy. Our second result shows the benefit of "diversifying" agency bias. Specifically, assuming the inevitability of disagreement, the principal prefers an agent who exhibits a mixture of formal and substantive bias. In calibrating the sources of disagreement, the principal can ensure that conflict arises most for claims where the agent's information is least valuable. The paper thus provides a theory for why a principal might be concerned with both the substance and form of an agent's

decision-making. We also derive results concerning which is worse for the principal: an over-confident or meek agent.

1 Introduction

The executive search firm, Spencer Stuart, pitches itself as "provid[ing] insight[s] into how candidates are likely to behave, make decisions and operate in the [executive] role." In looking for a new member of the C-suite, the CEO tries to decipher (with the help of a search firm) how the candidate would make decisions.

No one wants a CFO or chief information officer who makes decisions that benefit himself at the expense of the corporation. Yet top candidates must display more than a limited appetite for self-dealing; they must also demonstrate good judgment.

Good judgment is a complex quality. A candidate might rely too much on gut instinct rather than objective metrics in making hiring and firing decisions or deciding which product to bring to market. On the other hand, a candidate might be too timid. She might always go by the numbers rather than taking risks based on instinct or feel – the latter of which is necessary to move the company forward.

In this paper, we investigate two different ways an agent can disagree with her principal. The agent might disagree substantively; she might prefer more claims be decided as valid than the principal does. The agent might also disagree about the formal features of decision makng, specifically how to weight different criteria that bear on the decision.

The vast principal/agent literature in economics and political science has focused on substantive disagreement. The sender and receiver in a cheap talk game, for example, differ over the substantive decision that will be taken in each state (Crawford and Sobel, 1982).

The cheap talk model does not admit errors of judgment by the agent. Yet these concerns may dominate a principal's choice of agent, particularly when the set of decisions that an agent will confront is unknown. We address these issues in a simple model.

In our model, the agent must resolve a claim. The validity of the claim depends on two facts, of which the principal observes only one, the "global" fact. The agent observes both the global fact and a "local" fact, which is her private information. In addition to the global fact, the principal observes the agent's resolution of the claim. Based on these two pieces of information, the principal must decide whether to reverse or affirm the agent's resolution.

Consider formal bias first. We define a timid agent as one who places less weight on her own private information than the principal desires. In short, a timid agent relies too much on the global fact; she doesn't trust her instincts. By contrast, an overconfident agent places too much weight on what she alone observes.

As noted above, the agent and the principal may also disagree on purely substantive grounds. In this pure case, the agent accords the same weight to global and local facts as the principal but she nonetheless believes, in some cases, that the appropriate resolution of the claim differs from the principal's.

We generalize substantive bias to instances in which the agent and the principal disagree formally as well by aggregating the substantive differences between the two. A substantively biased agent believes that a different proportion of possible claims are valid than the principal does. A substantively unbiased agent believes that the same proportion of possible claims are valid as the principal. When the agent is formally biased, the principal and agent would thus disagree on the appropriate resolution of some claims though they would resolve the same percentage of claims as valid.

We ask three questions. First, given an agent infused with both substantive and formal bias, what is the optimal review practice for the principal? Which resolutions should he reverse and why? Second, which is worse for the principal: a timid agent or an overconfident agent? Third, assuming some bias is inevitable, should the principal seek an agent who is more prone to disagree about substance, form, or perhaps a little of both?

We first find that the principal's review strategy depends on the type of agency bias. Suppose first that the agent disagrees with the principal purely about substance: she, say, prefers to find more claims valid than the principal. In that setting, the principal affirms any decision that goes against the agent's natural inclination (i.e., an invalid resolution). She also affirms valid resolution unless there is sufficient information in the global fact to raise a red flag about that resolution. Reacting to a red flag, the principal reverses the valid resolution with positive probability. The fear of reversal, then, induces the agent to partially comply with the wishes of the principal. The principal thus creates a one-sided bound to the agent's discretion.

Next consider an agent who disagrees about process. Suppose, say, the agent is over-confident, placing more weight on her gut instinct than the principal prefers. In that case, the optimal review strategy involves two bounds on the agent's authority with an allocation of complete discretion in between. If the global fact points sufficiently toward a finding of invalidity, the principal reverses valid resolutions – the one that raises a red flag – with positive probability. Likewise, if the global fact points sufficiently toward a finding of validity, the principal affirms any valid resolutions and reverses invalid resolutions with positive probability. In between – where the global fact doesn't speak one way or the other – the principal affirms all resolutions. In this situation, the agent faces a two-sided bound on his discretion.

Having studied the interaction between the principal and agent, we move on to the principal's selection criteria. We find that the principal prefers a brash to a meek agent.

The reason resonates. The timid agent is reluctant to rely on his own private information. That means the agent's resolution provides little additional information above and beyond what is contained in the global fact. By contrast, the overconfident agent relies too much on her gut. While timidity only brings costs, over-confidence brings costs and benefits for the principal. Overconfidence implies that any decision will reflect a hefty dose of private information; private information is valuable when the global fact is uninformative. But overconfidence also implies that the agent has a tendency to disregard, or at least underweight, the information available to everybody else, including the principal. This disregard imposes costs on the principal in those cases in which she finds the global fact highly informative. But the principal can partially temper the agent's overconfidence in these contexts by credibly threatening, conditionally on the realized global fact, to reverse at least some decisions. And this credible threat, in turn, induces some – albeit imperfect – compliance by the over-confident agent.

Finally, we show that the principal can be better off appointing a more biased rather than a less biased agent. As noted, the overall level of bias is the proportion of cases where the principal and agent disagree. The model shows that the principal doesn't only care about overall disagreement, but instead the nature of that disagreement. If the agent and principal disagree the most in cases where the principal can make a pretty good decision based on the global fact alone, the disagreement doesn't hurt the principal.

How can the principal make that happen? She can do so by selecting an over-confident agent, specifically an agent whose interests are aligned with the principal's in the case where the global fact is uninformative. Further, heated disagreement occurs in cases where the principal does not need access to the agent's private information to make the proper decision.

While abstract, our model has many real world applications, including:

1. **The Judicial Hierarchy**: In any judicial system, there are trial courts and appellate courts. The trial court makes a decision:

liable or not liable. That decision can be appealed to an appellate court. The appellate court has access to some facts (e.g., the text of the contract under dispute) but not all the facts (the appellate court does not observe the demeanor of the witnesses). The appellate court must decide whether to affirm or reverse the trial court's disposition.

2. **CEO and CIO**. The CEO of a firm oversees its operation. The Chief Information Officer is charged with the development of an information security system. The CIO makes a decision, which can be reversed by the CEO. The local fact or private information would be, say, the CIO's insights into the expertise of the IT staff at the firm. The global fact would be what other companies in the industry are using for information security.

3. **The Office of Information and Regulatory Affairs (OIRA)** OIRA is part of the executive branch of government. It reviews "significant regulatory actions" by federal agencies. OIRA is the principal in this setting, the individual agencies are the agents. It has access to global facts about, say, the president's policy positions. It does not have information about the pros and cons of a regulatory review, which is available to the agencies. The agency proposes a rule and OIRA must review it. It can affirm, reverse, or revise the rule.

The paper unfolds as follows: Section 2 sets up the model. Section 3 characterizes the equilibrium when the principal faces (1) a zealot and (2) a zealot who is also over-confident. Section 4 shows that, given a fixed amount of disagreement, the principal prefers an agent who is over-confident but not a zealot (meaning she places too much weight on her private information but is equally likely to disagree when it comes to finding a claim valid as to finding a claim invalid). Section B. demonstrates that the principal prefers an over-confident to a timid agent assuming both agents disagree in the same fraction of cases. Section 5 compares and contrasts our results to the previous work on this topic. Section 6 provides a short conclusion. All proofs not found in the text can be found in the appendix.

2 The Model

The model involves a principal and an agent. At the start of the game, the agent is presented with a "case." Denoted by s, a case has two facts: a global fact x and a local fact y. The global fact is observable to both the principal and the agent. The local fact, is private information, observable only by the agent.

To flesh out this abstraction, consider the application where the principal is an appellate court and the agent is a trial court. There, the global fact might be the text of the contract under dispute, whereas the local fact might be the demeanor of the witness testifying about the intentions of the parties.

The global fact and the local fact are each randomly drawn from independent uniform distributions on $[0, 1]$. The space of the possible cases is thus the unit square denoted by S.

In terms of timing, once presented with a case, the agent must declare it valid ($d(s) = 1$) or invalid ($d(s) = 0$). The principal observes the agent's resolution and the global fact. Based on these two pieces of information, the principal must decide whether to reverse ($\gamma = 1$), affirm ($\gamma = 0$) or mix between the two ($\gamma \in (0, 1)$).

A. Preferences

The principal and the agent have preferences over case outcomes. Cutlines through the unit square define the preferences. These cutlines partition the square into cases where the agent or principal prefer to find the claim valid and cases where they prefer to find the claim invalid. The conflict in preference is modeled as the players having different cutlines or preferred partitions of the claim space.

The principal prefers that the case or claim be resolved as valid if

$$y > f(x) = 1 - x.$$

By contrast, the agent prefers that the case or claim be found valid if

$$y > g(x) = b - ax$$

Let S^{n_v} denote the set of cases where the principal prefers to find the claim valid (that is, $S^v = \{s \in S | y > 1 - x\}$) and T^v denote the set of cases where the agent prefers to find the claim valid (that is, $T^v = \{s \in S | y > b - ax\}$). Denote as S^{nv} and T^{nv} as the sets of cases where principal and the agent, respectively, prefer to find the claim invalid.

Figure 1 illustrates the preferences when $a = \frac{1}{2}$ and $b = \frac{3}{4}$. The blue line is the principal's cutline. The orange line is the agent's cutline. As noted in the introduction, the principal and the agent might disagree in one of two ways. First, the principal and agent might disagree about whether more overall claims should be found valid or not. We denote this as substantive bias or zealotry.

More interestingly, the principal and the agent might disagree about balance or relative importance of the global and local facts. We denote this formal bias.[1]

Formal bias measures the degree of over-confidence or meekness of the agent. Specifically, the over-confident agent weighs his gut instinct (e.g., the private information) more than the principal wants him to. The meek agent weighs it less.

The parameter b (holding a constant) measures the amount of zealotry. This corresponds to the traditional agency conflict parameter found in the prior literature (**?**Dessein, 2002).

[1]Gennaioli and Shleifer (2007) present a model that assumes yet another kind of judicial disagreement in a two-dimensional case space model. There, the judges agree about what cases represent an error, but disagree as to the weight to place on type I versus type II error.

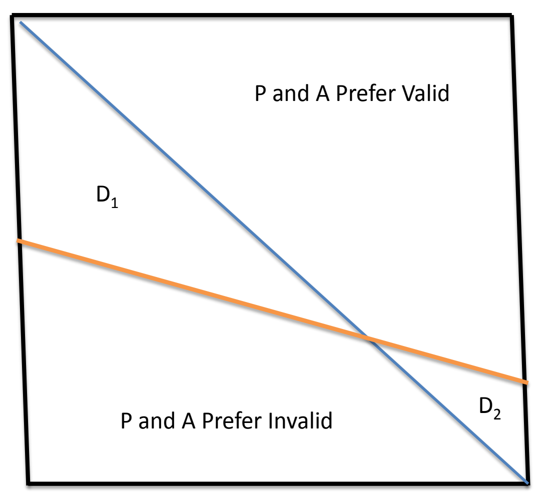

Figure 1: Preference Conflict

The parameter a (holding b constant) represents the degree of "formal bias."

B. Payoffs and Equilibrium Concept

The principal and the agent suffer losses if the case is resolved incorrectly from their point of view.

To motivate the assumption that incorrect dispositions generate losses, take the judicial hierarchy application. If the court declares a defendant liable when it shouldn't, the court will induce too many expenditures on precautions in the future. On the other hand, if the judge declares a defendant not liable when it should be, the court will induce too few expenditures on precautions.

Further, we assume that the loss associated with an incorrect disposition increases in the vertical distance between the case and the principal or agent's cutline. In words, bigger mistakes result in larger losses.

Denote the final resolution of the case as r_f. Since the principal has the power to reverse, the final resolution depends on the resolution by the agent and whether the principal affirms or reverses the agent.[2]

Given this formulation, the the principal's payoff is

$$U_p(r_f, s) = \begin{cases} -(y - f(x)) & \text{if} \quad s \in S^v \text{ and } r_f = 0 \\ -(f(x) - y) & \text{if} \quad s \in S^{nv} \text{ and } r_f = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The agent's payoff has two parts. First, mirroring the principal's payoff, the agent has preferences over the final resolution of the case.

[2] If the agent finds the claim valid ($d(s) = 1$) and the principal affirms ($\gamma = 0$) then $r_f = 1$; if the agent finds the claim valid and the principal reverses then $r_f = 0$; if the agent finds the claim invalid and the principal affirms then $r_f = 0$; finally, if the agent finds the claim invalid and the principal reverses then $r_f = 1$

Those are:

$$u_a(r_f, s) = \begin{cases} -(y - g(x)) & \text{if} \quad s \in T^v \text{ and } r_f = 0 \\ -(g(x) - y) & \text{if} \quad s \in T^{iv} \text{ and } r_f = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Second, the agent suffers a reputation cost from reversal. That is,

$$v_a(\gamma) = \begin{cases} -k & \text{if} \quad \gamma = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The agent's total payoff is thus $U_a = u_a + v_a$.

As noted, the principal observes the agent's resolution, and x, the global fact. From this (and the agent's equilibrium strategy), the principal updates her beliefs about the location of y – the local fact.

Let $f(y|r_a, x)$ be the principal's posterior belief given the agent's initial resolution (r_a) and the global fact, x.

A perfect Bayesian equilibrium of the game consists of a set of strategies, $(d^\star, \gamma^\star)$, and system of posterior beliefs, $f^\star$, such that

- $$r_a = d^\star(s) \in \underset{r_a}{\operatorname{argmax}}\, U_a(\gamma^\star, r_a, s).$$
- $$\gamma^\star(r_a^\star, x) \in \underset{\gamma}{\operatorname{argmax}} \int_0^1 U_p({r^\star}_a, \gamma, s) f(y|r, x) dy.$$
- On the equilibrium path, and, to the extent possible, off the equilibrium path,the prinicipal's beliefs $f^\star$ satisfy Bayes' Rule

Intuitively, in a Perfect Bayesian equilibrium, the agent resolves the case optimally given the equilibrium reversal strategy of the principal. The principal reverses only when doing so maximizes her expected utility, given her posterior beliefs about y induced by the agent's equilibrium action.

Following actual judicial practice, we view this game as one where the agent resolves the case and the principal must decide whether to reverse or affirm.

Alternatively, the game could be seen as a cheap talk game, where the agent is restricted to sending one of two messages (valid or not valid). The agent, in effect, makes a recommendation to the principal as to whether the claim should be valid or invalid. The agent's recommendation is costless (all that matters for payoffs is the ultimate resolution imposed by the principal), but the agent suffers a reputation consequence when its advice is ignored.

3 Equilibrium

A. The Zealous Agent

We start our equilibrium analysis behavior by considering a "purely zealous" agent. Such an agent shares the principal's preferences that the global and local facts should be weighted equally in making a decision; that is to say, the pure zealot's cutline also has a slope of -1. The pure zealot however prefers to hold more claims valid that the principal does. To capture this fact, let $b < 1$.

Figure 2 depicts the principal and pure zealot cutlines. Area D is the area of disagreement. In these cases, the principal prefers the claim be decided as invalid whereas the zealot prefers to find the claim valid. Close inspection of the figure above reveals three points, which will then provide the intuition for the equilibrium.

1. Any claim that the agent prefers to find invalid, the principal also prefers to find invalid. As a result, in equilibrium, such a resolution will not raise a red flag and the principal will be apt to affirm.

2. For a case with a global fact located at $\frac{1}{2}$, the principal cannot make an informed decision based on the global fact alone.

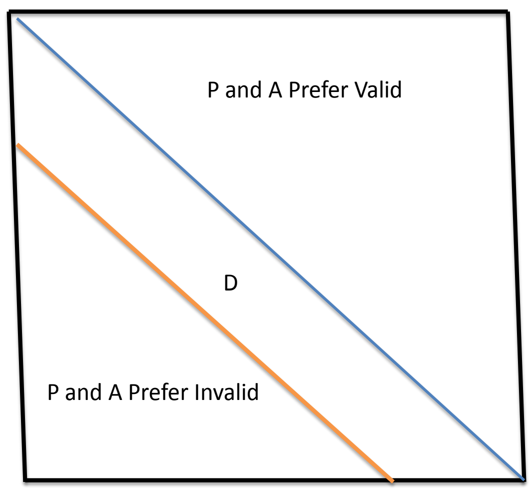

Figure 2: Zealot Preference Disagreement

Absent any other information, both an invalid and a valid resolutions lead to the same loss for the principal. It is in these cases where the principal will have the strongest desire to affirm a valid resolution by the agent.

3. For a case with a global fact located at 0, the principal knows that, no matter the realization of the local fact, the claim should be found invalid. As a result, the principal will have an incentive to reverse any valid resolution.

Points (2) and (3) suggests that there exists some value of a global fact between 0 and $\frac{1}{2}$ where the agent decides according to their preferred cutline (that is, they find claims valid if $y > b - x$) and the principal affirms that resolution, but just barely has the incentive to do so.

To find the location of that global fact, suppose that the agent resolved all cases according to her preferred disposition. The principal's would then infer from the valid resolution that the value of y is uniform with support $[b-x, 1]$.

Given these beliefs, the principal's loss from affirming a valid resolution is:

$$\begin{aligned}\text{Affirming Loss} &= \frac{\int_{b-x}^{1-x}(1-x-y)dy}{pr(valid)} \\ &= \frac{(1-b)^2}{2pr(valid)}\end{aligned}$$

The principal's loss from reversing the valid resolution is

$$\begin{aligned}\text{Reversing Loss} &= \frac{\int_{1-x}^{1}(y-(1-x))dy}{pr(valid)} \\ &= \frac{x^2}{2pr(valid)}\end{aligned}$$

Setting the two losses equal provides a "lower bound" on the global fact where the principal will affirm the disposition of a completely defiant agent; specifically, if the global fact lies above $\underline{x} = 1-b$, the principal affirms all valid and invalid resolutions.

The agent, then, decides all cases in her preferred way.

The intuition appears in figure 3. Take a case with global fact x. Suppose that the agent decides all cases above $b-x$ as valid.

If the principal reverses, she suffers a loss if $y > 1-x$ – the green dotted line. If the principal affirms, she suffers a loss if $y \in [b-x, 1-x]$ – the red dotted line. Since the green dotted line is longer than the red dotted line, the principal always affirms.

At $\underline{x}$, the green and red lines are equal, and this defines a lower bound of unbridled discretion.

Next, consider what happens when the global fact lies below the lower bound.

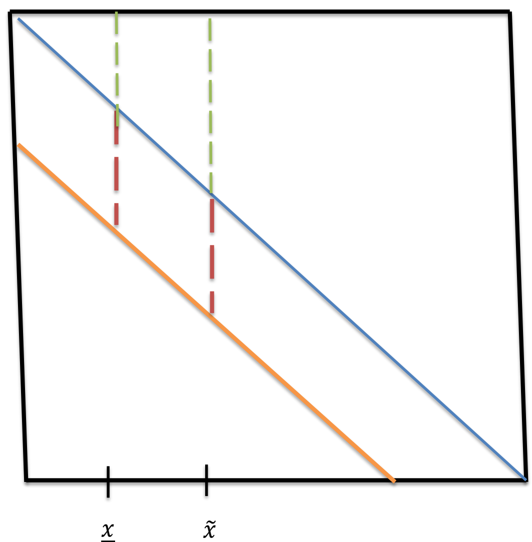

Figure 3: Zealot Equilibrium

Here, if the agent decided each claim her desired way, the principal would reverse the valid resolution. Then, the agent would not only suffer a loss because the case, after reversal, didn't go her way. She would also suffer the reputation penalty from being reversed. Hence, the agent deciding each case as she prefers cannot be an equilibrium.

Suppose instead that agent fully complied with the principal's wishes and only found cases valid where the local information was such that $y > 1 - x$. Assuming the agent followed that strategy, the principal would affirm every valid resolution.

This cannot be an equilibrium either. Knowing that affirmance is in the offing, the agent would prefer to deviate and find more cases valid.

Having rejected these two candidates, the remaining option has the principal mixing between reversing and affirming a valid resolution.

Specifically, for each global fact below $\underline{x}$, the agent picks a $y^*(x)$ such that the principal is indifferent between affirming and reversing the valid resolution (remember the invalid resolution is always affirmed).

Next, the principal picks a reversal probability such that the agent is indifferent between finding the claim valid and invalid at that exact same $y^\star(x)$.

We are now in the position to formally state our first proposition.

Proposition 1. *Suppose that the principal faces a zealot whose cutline is defined by $b - x$, where $b < 1$. The equilibrium can be characterized by two regions.*

1. *If the global fact lies below $\underline{x}$, the principal affirms all invalid resolutions. The agent finds a claim valid if the local fact exceeds $y^*(x) = 1 - 2x$; and the principal reverses valid resolutions with probability $\gamma(x) = \frac{1-x-b}{1-x-b+k}$. The principal's belief following a valid resolution is that the y is uniform with support $[y^*(x), 1]$.*
2. *If the global fact exceeds $\underline{x}$, the principal affirms all resolutions by the agent. The principal's beliefs following a valid resolution is that y is uniform with support $[b - x, 1]$. The agent resolves claims according his cutline $y = b - x$*

Proof. Consider first values of x less than $1 - b$. According to the candidate equilibrium strategy of the agent, she resolves claims according to the strategy $y^\star = 1 - 2x$, which lies strictly below the principal's cutline of $1 - x$. As a result, for any invalid resolution, the principal's payoff from affirming is 0, while he suffers a loss from reversal; thus affirming invalid resolutions is his best response.

For valid resolutions, the principal's loss from affirming is

$$\frac{\int_{1-2x}^{1-x}(1-x-y)dy}{Prvalid}.$$

or $\frac{x^2}{2Pr(Valid)}$ The principal's loss from reversing the valid resolution is

$$\frac{\int_{1-x}^{1}(y-(1-x)dy}{Pr(valid)} = \frac{x^2}{2Pr(valid)}$$

Since the losses from each action are the same, the principal is willing to mix between reversing and affirming.

To pin down the reversal probability, it must be the case that the agent is indifferent between finding the claim valid and invalid at $y^\star$. And so we have that

$$y-(b-x) = \gamma(y-(b-x)+k)$$

The LHS is the agent's loss from finding the claim invalid (and thus avoiding reversal). The RHS is the agent's loss from finding the claim valid (and risking reversal). The mixing probability that solves this expression is given by

$$\gamma = \frac{y-(b-x)}{y-(b-x)+k}$$

Plugging in for $y^\star = 1-2x$ yields the expression in the proposition (Note that the LHS increases faster in y than the RHS, meaning there is just one point of indifference).

Next consider claims with global facts greater than $1-b$. By the argument in the text, the principal's payoff to affirming a valid resolution exceeds his payoff from reversing when the agent decides according to his cutline. Thus, the agent's best reply is to do so and the principal affirms both valid and invalid resolutions. □

The equilibrium involves "one-sided" delegation to the agent. A finding of invalidity is always affirmed. The reason is that the principal knows that the agent's bias involves a bias for valid claims. Since the finding of invalidity goes against this bias, it doesn't raise any red flags.

Second, the principal only reverses valid resolutions if the global fact – the text of the contract, say – provides enough evidence that invalidity is indeed the principal's preferred result.

The upshot is that an agent who disagrees about "substance" will have a review that leans against the bias. Yet that review is only triggered when the valid resolution raises a sufficiently bright red flag (where the "brightness" of the red depends on what can be learned from the global fact). Fear of this review, then, will generate some partial compliance by the agent. By way of contrast, the next section considers an agent who is over-confident as well as substantively biased.

B. The Over-Confident Zealot

We say that an agent is over-confident if he relies more on his private information – his gut instinct – than the principal prefers. On the other hand, we say that the agent is timid or meek if he relies too little on his private information – a meek agent is inclined instead to decide according to the global fact.

In this model, over-confidence and meekness relate to the slope of the agent's cutline, whereas zealotry relates to the intercept. An over-confident agent's cutline has a slope greater than -1. The meek agent's cutline has slope less than -1.

The extreme over-confidence agent would rely on her private information only, and her cutline would be horizontal.

By contrast, the extreme timid agent would rely only on the global fact, and her cutline would be vertical.

This section considers an over-confident zealot. She has a cutline of $g(x) = b - ax$, where $b < 1$ and $a < 1$. We also restrict attention to the parameter space where $b > a$.

Figure 4 (a reproduction of figure 1) depicts the preference conflict between the over-confident zealot and the principal. Note that the agent prefers to find more claims valid than the principal (i.e., she

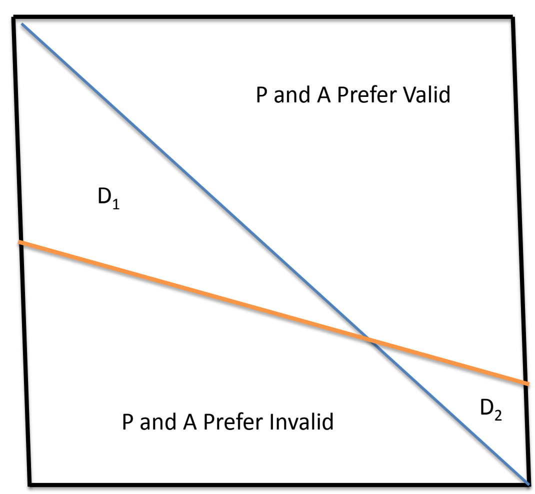

Figure 4: The Over-Confident Zealot

prefers more than $\frac{1}{2}$ of all claims be decided as valid.) At the same time, the over-confident zealot weights her private information more heavily than the principal prefers.

It is fruitful, then, to compare figure 3 and figure 4.

Note that adding over-confidence to zealotry creates an easy case for delegation. Because the cutline from the over-confident zealot crosses the principal's cutline, there will be a global fact where the two preferences are fully aligned. That doesn't occur with a zealot, as represented in figure 4.

Second, the over-confident zealot prefers to hold invalid some claims that the principal prefers to find valid (the region of D-2 in figure 3). The reason: the over-confident agent weights his private information too much and, as a result, there will be realizations of the private information that point in favor of invalidity when the global fact does not. There, the agent prefers invalidity and the principal does not. That course of events happens when the global fact lies close to 1.

Turning now to the equilibrium much of the analysis from the previous section goes through. There will be cases around the "easy case" where the principal affirms all resolutions: valid and invalid. That is, there is will be band of complete discretion.

Where the global fact points overwhelmingly towards valid or invalid (where the global fact is close to 0 or close to 1), the principal will reverse with positive probability resolutions that cut against the global fact. The agent, then, will respond with partial compliance.

This intuition is represented formally in the next proposition.

Proposition 2. *Facing an over-confident zealot, the equilibrium can be characterized by three intervals:*

1. *If $x < \frac{1-b}{2-a}$, the principal affirms an invalid disposition and reverses a valid disposition with probability $\gamma(x) = \frac{y-(b-ax))}{y-(b-ax)+k}$. The agent finds the claim valid if $y > y^\star(x) = 1-2x$ and invalid otherwise. The principal believes following a valid resolution that y is uniform with support $[1-2x, 1]$.*

2. *If* $x \in [\frac{1-b}{2-a}, \frac{2-b}{2-a}]$*, the principal affirms all dispositions. The agent resolves cases according her own preferences. The principal's beliefs about* y *are uniform with support* $[b - ax, 1]$ *after a valid resolution.*

3. *If* $x > \frac{2-b}{2-a}$ *the principal affirms any valid resolution. The agent finds claims valid if* $y > y^\star = 2(1-x)$ *and invalid otherwise. The principal reverses invalid resolutions with probability* $\gamma(x) = \frac{b-2+(2-a)x}{b-2+(2-a)x+k}$*. Following a valid disposition, the principal believes* y *is uniform with support* $[2(1-x), 1]$*.*

Proof. See appendix. □

The model with an over-confident zealot sheds light on a number of aspects of claim resolution. First, it provides a theory of red flags. A red flag is a resolution that goes against the principal's "instincts," her prior beliefs, given x, about the appropriate resolution of the claim. What are the chances that the star witness testimony was so convincing that his testimony alone overwhelmed the clear contrary text of contract? The answer to that question depends on just how clear the contrary text is. In other words, both whether a red flag is raised and its "brightness" depend on the informativeness of the global fact.

Second, in many circumstances principals uphold decisions even if they think the decision is wrong. For example, consider the standard of review for an appellate court. These standards guide the appellate court's review of a decision by a trial court or agency. These review standards instruct the appellate court to look for a "manifest" or "clear" error, to ask whether the agent made a reasonable (not necessarily a correct) decision. The model explains the difference between a grave and a minor error, a reasonable and an unreasonable decision. Specifically, the deference interval $[\underline{x}, \overline{x}]$ determines the range of reasonableness. For claims in this interval, the principal has a conjecture based on the global fact x about the right answer. The principal nonetheless defers and affirms all resolutions. In other words, she concludes that the agent has acted reasonably, no matter the outcome reached. A manifest error, by contrast, occurs when the resolution is in sufficient tension with the location of the global fact.

Third, unlike many models of delegation (Athey et al., 2005), the trial court faces two bounds on its permissible behavior, not one. The trial court doesn't just face, say, an inflation cap or target. Instead, the trial court's decisions below and above certain thresholds are potentially reversed. The dual bounds exist because there is "balance-disagreement" as well as "litigant-bias."

This section closes with a legal application. Consider a tort claim. The law states that the appellate court reviews finding of fact for clear error. The appellate court reviews findings of law de novo. The appellate court's review of the application step is more muddy. The appellate court reviews mixed questions of law and fact on a sliding scale. But what makes something "more factual" or "more legal."? Who knows? Indeed, the Supreme Court has stated that the difference between law and fact cannot be gleaned from the case law. In Pullman-Standard v. Swint, 456 U.S. 273 (1982), for example, the Supreme Court stated:

> This Court has previously noted the vexing nature of the distinction between questions of law and questions of fact. . . Rule 52(a) does not furnish particular guidance Nor do we yet know of any other rule or principle that will unerringly distinguish a factual finding from a legal conclusion.

The model clarifies why the standard of review for mixed questions of law and fact is so murky and inconsistent in practice. The appellate court doesn't know the true value of y. The appellate court knows that the trial court might overstate a finding of fact to justify the imposition of liability. Deference to the trial court's resolution (the application step) turns on how important the principal thinks that y finding is. If critical, the appellate court defers, even if the resolution seems suspect based on the global fact. If not critical, the appellate court reverses with some positive probability. The lack of commitment (because appellate courts move after the trial court resolution) leads to some range of complete delegation with blurred boundaries of authority, a finding consistent with the Supreme Court's statement about the lack of clarity of the law/fact distinction.

4 The Choice of Agent in a Second Best Setting

In this section, we investigate how the principal evaluates agents that exhibit different types of bias. Obviously, fixing the type of bias, the principal always prefers an agent with less bias rather than more. As the next subsection shows, however, conditional on the same level of disagreement, the principal an agent who exhibits formal bias over an agent who weighs the global and local facts in the same way she does.

A. Formally biased vs purely substantively biased Agents

Recall that a purely substantively biased agent has a cutline with the same slope as the principal's cutline. The presence of formal bias implies that the slopes are different.

Consider a principal with a choice between two agents that are equally biased. One of them exhibits only pure substantive bias; the other is formally biased. Somewhat surprisingly, the principal will prefer the agent that exhibits the formal bias.

To see this, contrast figures 3 and 4. The agents in Figures 3 and 4 do not exhibit the same amount of substantive bias. Equal substantive bias implies that the area D between the two cutlines in figure 3 equals the sum of the areas of D_1 and D_2

Superficially, it may seem that the principal would prefer the agent with the cutline as in figure 3. But there is always an agent with only formal bias but substantive bias equal to D whom the principal prefers to the pure zealot. That agent has a cutline that intersects the principal's cutline at $\frac{1}{2}, \frac{1}{2}$ and has a slope greater than -1. Figure 5 illustrates this agent. Notice that this over-confident agent is

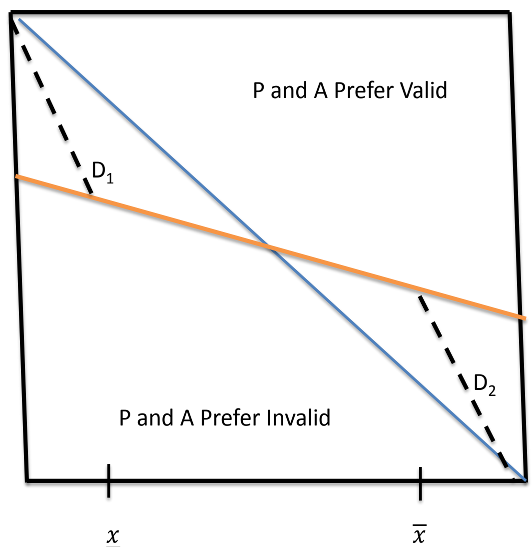

Figure 5: Optimal Second Best Agent

equally likely to make a mistake (from the principal's perspective) as to finding of validity or a finding of invalidity (the areas of D_1 and D_2 are the same). Further, this agent is induced by a fear of reversal to moderate his own preferences and do closer to what the principal wants towards the end points of the unit interval. This "compliance" effect is illustrated by the dark dotted lines.

Why is this sort of agent such a good find for the principal? The intuition is straightforward. The agent's signal at $x = \frac{1}{2}, \frac{1}{2}$ is maximally informative. At that global fact, the principal willingly defers to the agent who will decide as she would. For points close to this x, the agent's resolution of the claim is also highly informative. With the zealous agent of figure 3, by contrast, there is no global fact at which the agent's disposition of the claim is so informative.

Proposition 3. *1. Given overall disagreement between the principal and agent of $A < \frac{1}{4}$, the principal maximizes his expected utility by selecting an over-confident agent who does not display any zealotry, whose cutline has an intercept of $b^\star = 1 - 2A$ and a slope of $a^\star = 1 - 4A$; such a cutline generates the easy case (the case of fully aligned preferences) at $x = \frac{1}{2}$.*

2. The principal prefers an over-confident agent whose cutline is defined by $a^\star = 1 - 4A$ and $b^\star = 1 - 2A$ to a zealot presenting the same level of disagreement.

Proof. See Appendix □

Consider the pure zealot with substantive disagreement of A. Proposition 3 identifies an overconfident agent, also with substantive disagreement of A that the principal strictly prefers to the pure zealot. Consequently, given continuity, there will be an overconfident agent with substantive disagrement slightly greater than A that the principal prefers to the pure zealot.

B. Timid versus Over-Confident Agents

Conjecture 4. *Facing a choice between a timid agent and on over-confident zealot who share the same overall bias and the same easy case (i.e., the global fact is the same where the principal and agent's cutlines intersect), the principal prefers the over-confident zealot.*

Figures 6 and 7 provide the intuition behind the conjecture. Figure 6 is an extremely timid agent. She doesn't ever make a decision based on the local fact.

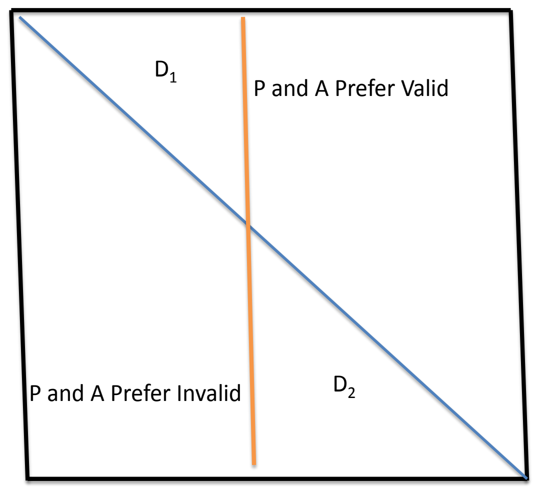

Figure 6: Extremely Timid Agent

Figure 7 is an extremely over-confident agent, who never makes a decision based on the global fact. Both agents disagree with the principal in $\frac{1}{4}$ of the cases.

Notice that, in figure 7, the over-confident agent, in equilibrium exhibits some partial compliance to avoid reversal (reflected in the dotted black lines outside the interval of delegation, i.e., $[\underline{x}, \overline{x}]$. That effect mitigates the loss from hiring the over-confident agent. Such mitigation does not occur with a timid agent.

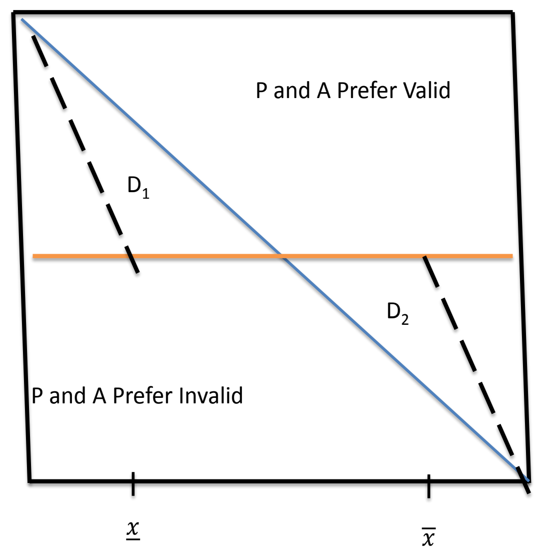

Figure 7: Extremely Over-Confident Agent

5 Relationship to Other Models

This section identifies key features of our model, discusses the implications they have for our results, and relates them to the most relevant literature.

Our model differs from the standard model of the principal-agent relation in two ways.[3] First, we assume that the principal cannot commit to a review strategy. Typically, the inability to commit transforms delegation games to cheap talk games. In our framework, however, the principal effectively delegates to the agent in the interval $[\underline{x}, \overline{x}]$ in which she affirms the agent's decision with probability one, even though the

[3]The principal agent literature is vast, Mas-Colell et al. (1995) presents the standard model in chapter 14.

principal moves after observing the resolution. The absence of commitment leads to the red flag property of our equilibrium: outside the interval, the principal reverses unexpected resolutions with some probability less than one.

Second, we model the preference conflict with the agent differently. The standard principal-agent models either assume that the agent faces a cost to effort or he is "zealous" in the sense that his payoffs differ from the principal's by a fixed amount. The latter assumption apparently derives from the canonical discussion of cheap talk in Crawford and Sobel (1982). In these models, the principal wants to match her action to the state of the world while the agent wants to match his action to the state of world plus some constant.

Our model permits three forms of bias: ex ante, interim, and ex post. (In the standard model with constant bias, these three measures collapse into one but in our model they are distinct.) Ex ante bias is measured by the proportion of the claim space on which principal and agent disagree over the correct resolution of the claim.

Clearly, the extent of delegation – with the overconfident zealot or the simple zealot the length of the discretion interval where all resolutions are affirmed increases as the overall level of disagreement falls. Though the extent of ex ante bias determines the extent of the delegation, the variation in interim bias drives our results. Interim bias refers to the expected bias of the agent once the global fact x is realized. We might measure it by the length of the interval on the vertical line on which principal and agent disagree about the resolution of the claim. At the easy case with the over-confident zealot, principal and agent perfectly agree on how to resolve all claims. When the principal observes any other claim, however, she disagrees with the agent and the extent of the disagreement increases with the distance from the easy case, i.e. as bias increases. Again delegation occurs only when interim bias is sufficiently small.

The result that the degree of delegation decreases as the agent's bias increases parallels the comparative statics result in Dessein (2002), who assumes constant bias.[4]

In our model, the variation in the degree of interim bias sustains a no-commitment equilibrium.

Specifically, whether an agent is biased depends on the realization of not just the global fact, but the local fact too. For some realizations, the agent shares the principal's preferences over the resolution of the claim and consequently resolves the claim in the interests of both. In models of constant bias, by contrast, the preferences of agent and principal never perfectly coincide so ex post bias always exists.

At least in the context of claim resolution, our assumptions on bias are more plausible than the standard assumption of constant bias. In part, the plausibility of our assumption results from the dichotomous nature of claim resolution. It is implausible to believe that principal and agent disagree on the resolution of every claim; it may be plausible in the standard model which typically assumes a continuous action and state space that the agent always wants just a little more or a little less of some outcome. But claim resolution, as done by courts, administrators of social security and veteran's affairs, and parole boards, does not seem to fall into this setting.[5] At a more applied level, our motivating institutional example is the judicial hierarchy. As such the paper closely relates to that literature (Kastellec, 2017). For example, Beim (2017) investigates how the Supreme Court might learn from decisions by rival court of appeals. She finds that Supreme Court benefits from observing multiple circuit court decisions on a single issue, even if resource constraints permit the Supreme Court to review only one case at a time. Related, Clark and Kastellec (2013) model the Supreme Court as facing an optimal stopping problem when

[4]Dessein (2002) primary result shows the conditions under which the principal prefers delegation to communication.

[5]Recall that, in our setting, the agent (and the principal) are resolving claims, not announcing policy. So while it might be, for instance, that an agent always prefers a more claimant-favorable policy than the principal, they may still agree about the resolution of specific claims.

establishing a precedent. In that model, percolation of decisions in the lower courts generates useful information for the Supreme Court, but also more legal uncertainty. In both these models, the higher court uses the lower courts to learn. In our model, the question differs: we ask first, how does the appellate court commit to follow a rule and second, how do lower courts and agencies make or break their reputation with the appellate court. Other models of judicial hierarchy, like Cameron et al. (2000), focus on auditing by an relatively uninformed appellate court. These models typically assume that agents are (purely) overzealous rather than overconfident or timid. As noted above, the assumption of pure zealotry means that the audit decision is always asymmetric: the appellate court only audits either the grant of the relief or the denial of relief, never both types of decisions.[6]

Finally, in a more recent paper, Guimaraes and Salama (2017) study a static model in which a court uses the statute as a signal about the true state of the world and hence of the appropriate standard to apply. The model is superficially similar to ours but in fact is radically different. The court's signal is noisy because there are two types of lawmakers. "Bad" lawmakers choose the standard randomly; "good" lawmakers announce a statutory standard identical to the optimal standard. As each type is mechanical, the court faces a decision problem. There is no strategic interaction between the upper and lower court. The optimal decision of the court is to adopt the standard if it falls within an interval defined by the signal and the court's prior on states of the world. In our model, of course, the biased agent and the principal are both strategic actors.

[6] In their model, player preferences exhibit constant differences in dispositional value. In this model, both parties' preferences exhibit increasing differences in dispositional value.

6 Conclusion

We have studied a simple principal-agent model in which the principal must make a dichotomous decision that depends on the state $s(x, y)$ of the world. Both the principal and the agent know x but only the agent knows y. This simple two-dimensional structure complicates the ways in which principal and agent may disagree: the parties may disagree both substantively and formally.

In this framework, the principal generally delegates to the agent but the nature of the delegation depends on the nature of the disagreement between principal and agent. When the disagreement is purely substantive, the delegation is bounded, as in the standard cheap talk model, on one side only. When the disagreement is also formal, the delegation is bounded on both sides.

Moreover, in our model, the informativeness of the local fact y varies systematically with the global fact x. Consequently, the delegation occurs even though the principal cannot commit to a delegation strategy. The bounds of delegation are supported by the self-interest of the principal.

Finally, in our model, the principal has preferences over the bias that infects her agent. Conditional on a fixed level of substantive disagreement, the principal prefers an overconfident to a timid agent.

Our model has several applications. Delegation is widespread in both private and public bureaucracies. Our first result, however, has special leverage on delegation in the public sphere. Generally, in the public sector, the principal cannot commit to her delegation through an enforceable contract. In public bureacuracies, this inability to commit derives from the limitations on the employment contract. In the judiciary, though a hierarchy of courts exists, there are no mechanisms of control other than affirmance and reversal of the decisions of the lower court. An appellate court, thus, cannot commit to defer to the decisions of a lower court or an administrative agency. Our analysis shows that, nonetheless, when the agent has private information that is valuable to her, the principal will rationally delegate many decisions to the agent. Specifically, she will defer to the decision of the agent over a suitable region.

Further our results reveal how variation in the nature of the disagreement between the principal and agent may explain the variation in the structure of delegation. In the adjudicatory context, the delegation to trial courts or administrative agencies typically has two bounds. On our account, the presence of two bounds rather than one suggests that the parties have formal disagreements. In the presence of pure substantive disagreement, we expect to one-sided bounds only.

Our second result shows the value of over-confidence in a agent. Such an agent will want to rely on her private information too much. The principal, then, can use the threat of reversal to leverage this desire to be most informative in cases where the principal is most uninformed. The timid agent, by contrast, doesn't bring much to the table; the principal might as well just make the decision on her own. Finally, we showed that the principal second best option (assuming some bias is inevitable) is to select an over-confident zealot, where the easy case – the case of fully aligned preferences occurs when the global fact is least informative.

7 Appendix

A. Proof of Proposition 2

First define the point at which the cutlines cross as x_c – this is the location of the easy case.

The lowest global fact, x where the principal always affirms a finding of validity solves

$$\frac{\int_{b-ax}^{1-x}(1-x-y)dy}{Pr(Liable)} - \frac{\int_{1-x}^{1}(y-(1-x))dy}{Pr(liable)} = 0$$

The first integral is thus

$$\frac{(1-b-(1-a)x)^2}{2}$$

The second integral can be written as

$$\int_0^x udu = \frac{x^2}{2}$$

Thus, the lower bound of the global fact must solve

$$\frac{(1-b-(1-a)x)^2}{2Pr(Liable)} - \frac{x^2}{2Pr(Liable)} = 0$$

The positive solution of this equation is $\underline{x} = \frac{1-b}{2-a}$ Next consider the maximum value of the global fact where the principal will affirm an agent's invalid disposition. That upper bound solves this expression:

$$\frac{\int_{1-x}^{b-ax}(y-(1-x))dy}{Pr(NotLiable} - \frac{\int_0^{1-x}(1-x-y)dy}{Pr(NotLiable} = 0$$

The first integral is

$$\int_0^{b-ax-1+x)} udu = \frac{(b-ax-1+x)^2}{2}$$

The second integral is becomes

$$-\int_{1-x}^0 udu = \int_0^{1-x} udu = \frac{(1-x)^2}{2}.$$

Therefore, the upper bound on the interval of delegation must solve

$$\frac{(b-ax-1+x))^2}{2Pr(NotLiable)} - \frac{(1-x)^2}{2Pr(NotLiable)} = 0$$

The positive solution to this expression is $\overline{x} = \frac{2-b}{2-a}$.

Now consider the review strategy of the principal for global facts outside the interval of delegation. Here, he must be indifferent to mix between affirming and reversing the disposition that raises a red flag. Take first low values x, values below $\underline{x}$. To mix, the principal must be indifferent given the optimal strategy of the agent (denoted by $y^\star(x)$). Thus, the equilibrium strategy of the trial court solves

$$\frac{\int_y^\star(x)^(1-x)(1-x-y)dy}{Pr(liable} - \frac{\int_{1-x}^1(y-(1-x))dy}{Pr(Liable)} = 0 \tag{1}$$

After a change of variables, we see that this expression becomes:

$$\frac{(1-x-y^\star(x))^2}{2Pr(Liable)} - \frac{x^2}{2Pr(Liable)} = 0$$

The solution yields a different value of $y^\star$ for each x. Specifically, we have $y^\star(x) = 1 - 2x$.

Next consider high values of the global fact x and do the same analysis. That agent's equilibrium strategy thus must solve:

$$\frac{\int_{1-x}^{y^\star}(y-(1-x))dy}{Pr(NotLiable} - \frac{\int_0^{1-x}(1-x-y)dy}{Pr(NotLiable} = 0$$

This expression reduces to

$$\frac{(y^\star-(1-x))^2}{2Pr(NotLiable)} - \frac{(1-x)^2}{2Pr(NotLiable)} = 0.$$

Solving, the agent's equilibrium cutline in this region is given by $y^\star(x) = 2(1-x)$.

To find the mixing probability for the principal one just plugs to ensure that the probability of reversal induces the agent to be indifferent between finding the claim valid and invalid at $y^\star$.

B. Proof of Proposition 3, Part 1

We first show that, given a fixed amount of bias, among the overconfident zealots, the principal prefers to select an agent who is equally likely to make an error (from the principal's perspective) in a finding of validity as she is to make a error in a finding of invalidity. In other words, her cutline slopes through the point $(\frac{1}{2}, \frac{1}{2}$. We start by deriving the expected welfare of the principal given an agent whose cutline has an intercept of b and a slope of a. It is fruitful to divide the calculation into four regions of global facts: (1) facts below $\underline{x}$, where the welfare total is denoted W_1; (2) facts between $\underline{x}$ and x_c (the value of x where the cutlines cross). Denote this welfare as W_2 (3) facts between x_c and $\overline{x}$, where this welfare value is denoted by W_3; and (4) facts between $\overline{x}$ and 1, where the welfare value is denoted W_4. The total expected welfare this thus $W = W_1 + W_2 + W_3 + W_4$

i. global facts below $\underline{\mathbf{x}}$

The expected welfare is

$$W_1 = -\{\int_0^{\underline{x}} \int_{1-x}^{1} \gamma(x)(y-(1-x))dydx + \int_0^{\underline{x}} \int_{y^\star(x)}^{1-x} (1-\gamma(x)(1-x-y)dydx\}$$

In this region, findings of invalidity are always affirmed and result in the correct disposition from the principal's point of view. Findings of validity, by contrast, are reversed with probability $\gamma(x)$. The first term is the probability of reversal times the expected error associated with that reversal. The second term is the probability of affirmance times the expected error from affirming that disposition.

We can rewrite the first term above as

$$\begin{aligned} integral &= \int_0^{\underline{x}} \Big(\int_{1-x}^{1} \gamma(x)(y-(1-x))dy \Big) dx \\ &= \int_0^{\underline{x}} \Big[\gamma(x)\frac{u^2}{2} \big]_0^x dx \\ &= \int_0^{\underline{x}} \gamma(x)\frac{x^2}{2} dx \end{aligned}$$

$$
\begin{aligned}
insideintegral &= \Big[\gamma(x)(\frac{y^2}{2} - (1-x)y)\Big]_{1-x}^{1} \\
&= \gamma(x)\Big((\frac{1}{2} - (1-x)) - (\frac{(1-x)^2}{2} - (1-x)^2)\Big) \\
&= \gamma(x)\Big((x - \frac{1}{2}) - (-\frac{(1-x)^2}{2})\Big) \\
&= \gamma(x)\Big((x - \frac{1}{2}) + \frac{(1-2x+x^2)}{2})\Big) \\
&= \gamma(x)\Big((x - \frac{1}{2}) + \frac{1}{2} - x + \frac{x^2}{2})\Big) \\
&= \gamma(x)\frac{x}{2}
\end{aligned}
$$

Note that $\gamma(x)$ is treated as a constant when we integrate with respect to y. That is to say, the reversal probability can only depend on the global fact, not the local fact.

We can rewrite the second term above as

$$
\begin{aligned}
integral &= \int_0^{\underline{x}} \int_{y^\star(x)}^{1-x} (1 - \gamma(x)(1-x-y)dydx \\
&= -\int_0^{\underline{x}} \Big[((1-\gamma(x)(u)^2\Big]_{y^\star(y*(x))}^{0} dx \\
&= \int_0^{\underline{x}} (1-\gamma(x))\frac{(1-x-y^\star(x))^2}{2}dx \\
&= \int_0^{\underline{x}} (1-\gamma(x))\frac{x^2}{2}dx \qquad \text{since the equilibrium value } \ y^\star(x) = 1-2x.
\end{aligned}
$$

As a result, we can write this part of the welfare expression as

$$
\begin{aligned}
W_1 &= -\{\int_0^{\underline{x}} \gamma(x)\frac{x^2}{2}dx + \int_0^{\underline{x}} (1-\gamma(x))\frac{x^2}{2}dx \\
&= -\{\int_0^{\underline{x}} (\gamma(x)\frac{x^2}{2} + (1-\gamma(x))\frac{x^2}{2})dx\} \quad (\text{by } \int_a^b f(x)dx + \int_a^b g(x)dx = \int_a^b (f(x)+g(x))dx) \\
&= -\int_0^{\underline{x}} \frac{x^2}{2}dx \\
&= -\frac{\underline{x}^3}{6} \\
&= -\frac{(1-b)^3}{6(2-a)^3}
\end{aligned}
$$

where the last equality follows because $\underline{x} = \frac{1-b}{2-a}$.

C. Global Facts Between $\underline{\mathbf{x}}$ and x_c

In this range (1) the principal's cutline lies above the agent's cutline and (2) the principal always affirms. The ex ante expected welfare is

$$
\begin{aligned}
W_2 &= -\{\int_{\underline{x}}^{x_c} \int_{b-ax}^{1-x} (1-x-y)dydx\} \\
&= +\{\int_{\underline{x}}^{x_c} \left[\frac{(1-x-y)^2}{2}\right]_{b-ax}^{1-x-(b-ax)} dx \\
&= -\{\int_{\underline{x}}^{x_c} \frac{(1-x-(b-ax))^2}{2}\} \\
&= -\{\int_{\underline{x}}^{x_c} \frac{(1-b-(1-a)x))^2}{2}\}
\end{aligned}
$$

or, after doing the integration.

$$
\begin{aligned}
W_2 &= \frac{1}{1-a}\int_{(1-b)-(1-a)\underline{x}}^{0} \frac{u^2}{2} du \\
&= -\frac{1}{1-a}\int_{0}^{(1-b)-(1-a)\underline{x}} \frac{u^2}{2} du \\
&= -\frac{(1-b)-(1-a)\underline{x})^3}{6(1-a)} \\
&= -\frac{(1-b)^3}{6(1-a)(2-a)^3}
\end{aligned}
$$

where the last equality follows because $\underline{x} = \frac{1-b}{2-a}$.

D. Global Facts Between x_c and $\overline{x}$

In this range, the agent's cutline lies above the principal's cutline and the principal affirms all dispositions. The principal's expected welfare is

$$
\begin{aligned}
W_3 &= -\{\int_{x_c}^{\overline{x}}\int_{1-x}^{b-ax}(y-(1-x))dydx\} \\
&= -\{\int_{x_c}^{\overline{x}}\Big(\int_{1-x}^{b-ax}(y-(1-x))dy\Big)dx\} \\
&= -\}\int_{x_c}^{\overline{x}}\Big[\frac{(y-(1-x))^2}{2}\Big]_{1-x}^{b-ax} dx \\
&= -\{\int_{x_c}^{\overline{x}}\frac{\Big(b-ax)-(1-x)\Big)^2}{2} dx.
\end{aligned}
$$

Doing some substitution and integration (recalling that $\overline{x} = \frac{2-b}{2-a}$), we see that

$$
W_3 = -\frac{(b-a)^3}{6(1-a)(2-a)^3}
$$

E. Global Facts Between $\overline{x}$ and 1

In this region findings of validity are always affirmed and result in the correct disposition from the principal's point of view. Findings of invalidity, by contrast, are reversed with probability γx The expected welfare is

$$W_4 = -\{\int_{\overline{x}}^{1}\int_{0}^{1-x}\gamma(x)(1-x-y)dydx + \int_{\overline{x}}^{1}\int_{1-x}^{y^\star(x)}(1-\gamma(x))(y-(1-x))dydx$$

which can be written as

$$\begin{aligned} W_4 &= -\{\int_{\overline{x}}^{1}\gamma(x)\frac{(1-x)^2}{2}dx + \int_{\overline{x}}^{1}(1-\gamma(x))\frac{(1-x)^2}{2}dx\} \\ W_4 &= -\int_{\overline{x}}^{1}\frac{(1-x)^2}{2}dx \end{aligned}$$

In this last integral, let $u = 1-x$; thus, $dx = -du$. When $x = 1$, then $u = 0$. When $x = \overline{x}$, then $u = 1-\overline{x}$. After this change of variables, the integral becomes:

$$\begin{aligned} W_4 &= \int_{\overline{x}}^{0}\frac{u^2}{2}du \\ &= -\int_{0}^{\overline{x}}\frac{u^2}{2}du \\ &= -\frac{(1-\overline{x})^3}{6} \\ &= -\frac{(b-a)^3}{6(2-a)^3} \end{aligned}$$

where the last line follows because $\overline{x} = \frac{2-b}{2-a}$.

F. Total Welfare

Combining the ex ante welfare in each of these four regions, we have

$$
\begin{aligned}
W &= W_1 + W_2 + W_3 + W_4 \\
&= -\Big(\frac{(1-b)^3}{6(2-a)^3} + \frac{(1-b)^3}{6(1-a)(2-a)^3} + \frac{(b-a)^3}{6(1-a)(2-a)^3} + \frac{(b-a)^3}{6(2-a)^3}\Big) \\
&= -\Big(\frac{(1-b)^3(1-a) + (1-b)^3 + (b-a)^3 + (b-a)^3(1-a)}{6(1-a)(2-a)^3}\Big) \\
&= -\Big(\frac{(1-b)^3(2-a) + (b-a)^3(2-a)}{6(1-a)(2-a)^3}\Big) \\
&= -\Big(\frac{(2-a)((1-b)^3 + (b-a)^3)}{6(1-a)(2-a)^3}\Big) \\
&= -\frac{(1-b)^3 + (b-a)^3}{6(1-a)(2-a)^2} \\
&= -\frac{(1-b)^3 + (b-a)^3}{6(1-a)(1+1-a)^2}
\end{aligned}
$$

The principal's program is to maximize

$$W = -\frac{(1-b)^3 + (b-a)^3}{6(1-a)(1+1-a)^2}$$

subject to the following constraints

1. $$a \geq 0$$

2. $$a \leq 1$$

3. $$b \geq 0$$

4. $$b \leq 1$$

5.

$$b - a \geq 0$$

6.

$$\frac{(1-b)^2}{2(1-a)} + \frac{(b-a)^2}{2(1-a)} - A \geq 0$$

It is convienent to express the objective function as a function of the y-intercept of the agent's cutline, β (the amount of party bias) and the amount of slope, σ (which captures the amount of balance-bias):

$$\beta = 1 - b$$
$$\sigma = b - a$$

After this transformation, the objective function becomes:

$$W = -\frac{\beta^3 + \sigma^3}{6(\sigma+\beta)(1+\sigma+\beta)^2}$$
$$W = -\frac{(\beta+\sigma)(\beta^2 - \beta\sigma + \sigma^2)}{6(\sigma+\beta)(1+\sigma+\beta)^2} \text{factor with} x^3 + y^3 = (x+y)(x^2 - xy + y^2)$$
$$W = -\frac{\beta^2 + \sigma^2 - \beta\sigma}{6(1+\beta+\sigma)^2}$$

The constraints are:

1. $\beta \geq 0$.
2. $\beta \leq 1$.
3. $\sigma \geq 0$.
4. $1 - \beta - \sigma \geq 0$.
5. $1 - \beta - \sigma \leq 1$.
6. $\frac{\beta^2+\sigma^2}{2(\beta+\sigma)} - A \geq 0$.

Assume that only constraint on the amount of disagreement is binding (conjecture: this is true if $A < \frac{1}{4}$. Note that

$$\frac{\partial W}{\partial \beta} = \frac{-3\beta\sigma - 2\beta + 3\sigma^2 + \sigma}{6(1+\beta+\sigma)^3},$$

and

$$\frac{\partial W}{\partial \sigma} = \frac{-3\beta\sigma - 2\sigma + 3\beta^2 + \beta}{6(1+\sigma+\beta)^3}$$

Further, notice that

$$\frac{\partial C}{\partial \beta} = \frac{\beta^2 + 2\beta\sigma - \sigma^2}{(\sigma+\beta)^2},$$

and

$$\frac{\partial C}{\partial \sigma} = \frac{\sigma^2 + 2\beta\sigma - \beta^2}{(\sigma+\beta)^2}$$

Denote by λ the multiplier on the constraint (6), the constraint reflecting the area of disagreement. The first order conditions for the principal's program with respect to β and σ are:

$$\frac{\partial W}{\partial \beta} + \lambda \frac{\partial C}{\partial \beta} = 0, \tag{2}$$

$$\frac{\partial W}{\partial \sigma} + \lambda \frac{\partial C}{\partial \sigma} = 0 \tag{3}$$

Solving each expression for λ gives the following equality at the solution,

$$\frac{\frac{\partial W}{\partial \beta}}{\frac{\partial W}{\partial \sigma}} = \frac{\frac{\partial C}{\partial \beta}}{\frac{\partial C}{\partial \sigma}}$$

which holds when $\beta^\star = \sigma^\star$. Plugging those values into the constraint and solving yields:

$$\sigma^\star = \beta^\star = 2A$$

At the solution, then, we can transform the optimal values of β and σ into values of a and b. Doing so yields:

$$b^\star = 1 - 2A,$$
$$a^\star = 1 - 4A$$

And the solution always goes has the agent's cutline slicing through $(\frac{1}{2}, \frac{1}{2})$.

G. Proof of Proposition 3, Part 2

Given a loss of A, the principal's payoff from deploying the optimal second best over-confident agent is

$$W_{over} = -\frac{2A^2}{3(1+4A)^2}$$

The area of disagreement with a pure zealot is

$$\frac{1-b^2}{2} = A$$

The ex ante expected welfare is

$$W_z = -\Big(\frac{(1-b)^3}{3} + \frac{(2b-1)(1-b)^2}{2}\Big).$$

Solving the constraint for b yields $b = \sqrt{1-2A}$.

Plug this value into the welfare expression yielding

$$W_z(A) = -\Big(\frac{(1-\sqrt{1-2A})^3}{3} + \frac{(2(\sqrt{1-2A})-1)(1-\sqrt{1-2A})^2}{2}\Big).$$

Define a new function as follows:

$$H(A) = |W_z| - |W_{over}|$$

Note that $H(0) = 0$ and $H(\frac{1}{4}) > 0$ and $H'(A) > 0$. Thus, the loss from the pure zealot is always higher than the loss from the optimal second-best over-confident agent, proving the result (note, here, that we have restricted attention to cases where $A < \frac{1}{4}$.

References

Athey, S., Atkeson, A., and Kehoe, P. J. (2005). The optimal degree of discretion in monetary policy. *Econometrica*, 73:1431–1475.

Beim, D. (2017). Learning in a judicial hierarchy. *The Journal of Politics*, 79:591–604.

Cameron, C. M., Segal, J. A., and Songer, D. (2000). Strategic auditing in a political hierarchy: An informational model of the supreme court's certiorari decisions. *American Political Science Review*, 94:101–116.

Clark, T. S. and Kastellec, J. P. (2013). The supreme court and percolation in the lower courts: An optimal stopping model. *The Journal of Politics*, 75:150–168.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50:1431–1451.

Dessein, W. (2002). Authority and communication in organizations. *Review of Economic Studies*, 69:811–838.

Gennaioli, N. and Shleifer, A. (2007). The evolution of common law. *The Journal of Political Economy*, 115:43–68.

Guimaraes, B. and Salama, B. M. (2017). Contingent judicial deference: theory and application to usury law. *mimeo*.

Kastellec, J. P. (2017). The judicial hierarchy: A review essay. In *Oxford Research Encyclopedia of Politics*.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York, New York, 1st edition.