Appellate Lawmaking in a Judicial Hierarchy*

Adam B. Badawi † and Scott Baker ‡

† School of Law, Washington University in St. Louis

‡ School of Law, Washington University in St. Louis

January 29, 2014

Abstract

We examine an infinite horizon model of appellate court lawmaking that incorporates review of trial court dispositions. We consider two types of trial courts: realist and legalist. Realist trial courts want to get their preferred disposition subject to the threat of reversal. Legalist trial courts, by contrast, exert effort analogizing their case to one of two existing appellate court precedents (a "liable" precedent and a "not liable" precedent). Distant analogies are more expensive for the legalist trial court to make than close analogies. Each period, the appellate court audits trial court dispositions. A successful audit provides an opportunity to create precedent. Precedent changes future trial court dispositions by providing new cases from which the legalist trial court can draw analogies. This, in turn, alters the appellate court's scrutiny of these dispositions – its audit strategy – going forward. We use the model to provide an account for affirmances with opinions, the practice of dicta, and to explore how appellate scrutiny will differ depending on the appellate court's hunch of the likely merits of the lower court disposition. We also demonstrate how settlement of cases before appeal can improve the performance of appellate review. Throughout, we relate the findings to existing evidence and derive testable predictions.

*For valuable feedback, we thank David Law, Claudio Mezzetti, Judge Richard Posner and participants at the workshop at Washington University School of Law.

1 Introduction

A central function of appellate courts is to develop law. In a common law system, appellate courts fulfill this role by providing guidance to lower courts through case-by-case adjudication. The relationship between lower courts and appellate court is, however, symbiotic. Appellate courts rely on lower courts to supply the cases that they will use to clarify and guide the resolution of future cases. When appellate courts write opinions that clarify areas of law that were previously murky, they help their own cause by enabling lower courts to resolve novel cases more easily. Put another way, making new precedent allows lower courts to analogize future cases to that precedent with less effort. Over time, this process should allow appellate courts to define the law more completely. Despite the centrality of this interaction to appellate court functioning this dynamic has received little attention in the literature.

We develop a formal model that shows how developing the law can serve the goals of a resource constrained, but policy-motivated appellate court and use that model to explain several durable features of common law systems. First, this approach provides one account of why appellate judges invest their limited resources in writing opinions. By doing so, they make it easier for lower courts to resolve cases in accordance with the appellate court's preferred disposition. This account sheds light on common judicial behaviors that are hard to explain through models of judicial hierarchy that focus on policy differences or error correction. For example, it is relatively common for an appellate court to affirm a lower court and write an opinion that explains why it is doing so. If appellate courts care only about policy outcomes or correcting errors, they would not invest the effort to explain a case that is consistent with their policy preferences or is not erroneous.[1] Our approach suggests a reason why appellate courts might do so: it makes it easier for lower courts to analogize future cases that are similar to the case decided by the appellate court.

Second, the model shows that an appellate court will review trial dispositions that it disagrees with more intensely than those that it favors. One reason for this result is obvious: the appellate court suffers a loss from a disposition with which it might disagree; a grant of a habeas petition perhaps. But our model suggests a less obvious reason as well. Dispositions that the appellate

[1] Of course, an appellate court might publish a decision to influence sister courts or to facilitate certiorari by a higher court. Our account provides a reason for writing opinions for application in the trial courts and having trial courts refer to those opinions.

court might disagree with provide a better opportunity to make new law than dispositions that the appellate court views favorably. Imagine, as an example, that the appellate court must determine the amount of precaution that is necessary to escape liability for a tort. Further suppose that the appellate court wants to limit a prior case that found the defendant not liable to its facts. The best case to do so is an analogous case located right next to that precedent. If the appellate court issues an opinion in such a case finding the defendant liable, that liable precedent will impact a large number of subsequent cases in the lower courts. Given the prior not liable precedent on the books (the one whose reach the appellate court wishes to limit), the trial court is apt to find the defendant not liable in any closely analogous case. A finding of no liability thus signals to the appellate court that the case is closely analogous and, as a result, a good candidate for reversing and limiting the prior holding. If instead the trial court has found the defendant liable, the appellate court can infer that the case is unlikely to lie close to the not liable precedent and thus be a less effective vehicle for making law.

Third, the model contributes to the longstanding debate among legal scholars about the degree to which legalist and realist concerns motivate judges. Legalist values capture those that are internal to the legal system such as the importance of following precedent and the method of reasoning by analogy. Realist concerns, on the other hand, involve preferences for external policy outcomes. We dynamically model how appellate courts are likely to scrutinize lower courts that are solely motivated by legalist values and those that act on policy alone. Appellate scrutiny of these two categories of trial courts differs.

In the limit (after the law converges), the appellate court continues to scrutinize decisions by trial courts with a realist bent, but defers to dispositions by legalist trial courts. The appellate court continues to suffer losses: it doesn't get its preferred outcome in every case. That happens as a byproduct of delegating initial decisionmaking to an agent (the realist trial court) with conflicting preferences. Importantly, the appellate court losses from this delegation are smaller when it makes law than when it doesn't. Lawmaking, in other words, helps to reduce the agency costs associated with the delegation of initial decisionmaking to lower courts.

Fourth, our account provides an explanation of how appellate judges might go about writing opinions. One persistent, but puzzling, judicial practice is to include dicta in opinions. Appellate courts are traditionally reluctant to include any information about how they would resolve cases that are not before

them. But by including dicta, appellate courts often do just that. We explain this strategy as a way to lower the decision costs of lower courts. When an appellate court has an actual case in front of it, that court may be able to say something useful about closely related hypothetical cases. But the usefulness of those statements to lower courts decreases as their distance from the actual case increases. Our model is consistent with the familiar belief among lawyers and legal scholars that statements that are unconnected from a concrete controversy do little to help the resolution of future cases. But a hypothetical that slightly modifies the facts of an actual case can be illuminating for future trial courts that consider similar cases. We thus characterize dicta as a maximization problem; the appellate court wants to set dicta at a level that optimizes the lower court's use of those statements. We show that the use of dicta makes the legalist trial court and the long-lived appellate court strictly better off.[2]

Finally, the model provides yet another rationale for making appellate court opinions "public." Appellate courts could limit the availability of their opinions to other judges or to participants in litigation. They do not do so. Rather, opinions are public, available to anyone who wants to read them. Jurisprudence scholars suggest that "public" opinions enhance the legitimacy of the courts (Llwellyn (1960); White (1995)) or allow the courts to alter primary behavior. Our model suggests another benefit. By making opinions public, the appellate court encourages settlement of cases where the law is clear. That means that any case going to trial is apt to involve new issues. And new issues are what the appellate court needs to make new law. In this way, settlement increases the "yield" of the appellate court's audit of trial court dispositions – i.e., it increases the percentage of trial court dispositions that present the opportunity to create new law. This feedback effect, then, improves the performance of appellate review.

The next subsection reviews the related literature. Part 2 is a numerical example, which serves to illustrate many of the themes of the formal model. Section 3 lays out a one-period model to benchmark the dynamic model that comes in Section 4. That section derives the results regarding how the appellate court will make law and review trial court dispositions over time. Section 5 studies dicta. Section 6 extends the model to consider settlement. A short conclusion follows and an appendix contains all proofs. Throughout the exercise,

[2]By grounding our model on beliefs about legal reasoning, we differentiate our approach from those that view the limited nature of judicial opinions as a way to preserve room for future judges to make law (Rasmusen, 1994).

we identify testable predictions of the model and suggest consistencies between our model and the available evidence.

1.1 Related Literature

Scholars across a range of disciplines are interested in the relationship between upper and lower courts. Political scientists view the relationship as one of principal and agent. The principal is the upper court and it has a policy preference. The agent is the lower court and it has a conflicting policy preference. The upper court's threat of reversal keeps the lower court agent in check. Subject to limited resources, the upper court seeks to reverse lower court decisions it disagrees with (Songer et al. (1995); McNollGast (1995); Spitzer and Talley (2000); Daughety and Reinganum (2000)). Gennaioli and Shleifer (2008) treat the determination of fact as a choice in the trial court. The trial court manipulates facts to achieve its preferred outcome and avoid reversal. These models consider a one-period interaction between the upper and lower courts. Thus, they are not well-suited to explore how appellate lawmaking and appellate review of trial courts changes over time – two issues tackled here.

Law professors offer a different rationale for the behavior of trial courts. They assume that lower court judges intrinsically care about following precedent and about policy outcomes (Cross (2005); Kim (2007)). The issue, then, is which preference dominates in a particular case. Unsurprisingly, assuming lower courts care about following precedent generates a predication that they will tend to follow precedent. But why do lower court judges care about appellate court precedent? What function does it serve? Our model suggests that appellate court precedent lowers the cost of making analogies for the trial courts, thereby making it cheaper for them to resolve cases.

Other scholars have looked at law creation in the appellate courts. For example, appellate judges in Gennoili and Schliefer (2007) search for distinctions between their case and the prior precedent. The quest to distinguish improves the overall performance of the law. Baker & Mezzetti (2012) show that, given limited resources, a long-lived, imperfectly informed appellate court will always follow precedent. Precedent reveals information about the way case closes to the precedent should be decided. Rather than spend resources taking a fresh look, the appellate court summarily decides those cases by reference to precedent. In Stephenson and Bueno De Mesquita (2002), the appellate court judge often follows precedent with which he disagrees. He does so because adding new

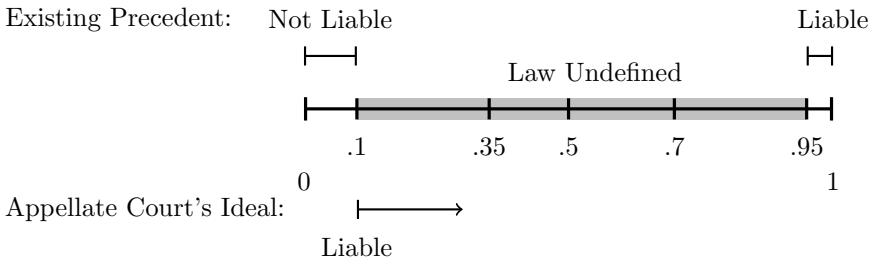
precedent to the stock of existing precedent improves communication between himself and the trial judge. All these papers take the flow of cases as given and consider how appellate judges react to their own precedent (horizontal stare decisis).

We study how trial courts are influenced by appellate court precedent (vertical stare decisis). The model then investigates the interaction between the two tasks of the appellate courts: review of trial courts (auditing) and law creation. We find that auditing trial courts and making new law are interdependent decisions by the appellate court, both of which evolve over time.

2 Numerical Example

This section works through a numerical example that we base on the model. The context for this example is a socially beneficial activity that can be harmful to others if a party does not take sufficient precaution – a tort. When a lawsuit results, courts must decide whether the amount of precaution taken by the defendant should result in liability or no liability. Imagine that we line up the cases from the least reckless (i.e. the defendant took extreme precaution) to the most reckless (i.e. the defendant took no precaution). So, case 0 represents the safest possible activity and Case 1 represents the most dangerous activity. Case .5 is halfway in between. In the example there are seven possible cases and each has 1/7 chance of occurring. Figure 1 shows each of the cases.

Figure 1: Numerical Example



The appellate court inherits some existing precedent, either from the Supreme Court or from a prior decision of the appellate court. This inherited precedent says that if a defendant takes a level of precaution at or below .1, that defendant should not be found liable. It also says that the activities above .95 should

result in liability. There is no guidance on the cases in between .1 and .95. It is in this range that the appellate court will fulfill its traditional role of developing the law (Frisch, 2003).

Following much of the literature on judicial decision making, we assume that appellate judges have an ideal point in the case space (Kornhauser (1992a); Kornhauser (1992b); Lax (2007)). In this example, we imagine an appellate court that wishes to restrict the reach of the no liability precedent only to cases .1 and below. It wants defendants to be liable for all activities that are more dangerous than .1. We make the standard assumption that the appellate court suffers a loss if a case is resolved against its preferences (McNollgast, 1995). The question is how the appellate court will minimizes its losses, given that it has scarce resources and there is a cost to reviewing trial court dispositions.

To answer that question we need to understand how the trial court will behave. For the purpose of the example, we look only at a "legalist" trial court.[3] This type of court has no preferences over outcomes. Instead, this trial court cares about effort it must devoted to analogical or legal reasoning. This approach follows the traditional principles of legal reasoning; it is generally considered more persuasive to make a close analogy than a stretched one (Sunstein, 1993). As lawyers sometimes put it, it is better to have a case that is "on point." We capture this effect through the effort that a trial court must make to make an analogy. It takes less effort to associate a new set of facts with a close case than it does to associate the new case with a distant case. Applying this approach, the legalist trial court resolves cases 0,.1, .35 and .5 in favor of the defendant (no liability). It resolves cases .7, .95 and 1 in favor of the plaintiff (liability).

This legalist approach is consistent with descriptions that judges give of their behavior. As Judge Posner explains: "in all cases of reasoning by analogy, sound analysis requires attending to the policy considerations that align the case at hand with one or another line of precedents." (Posner, 2006, p.766). Our model of the legalist trial court attempts to capture this process. Often, the trial court will be confronted with a case with precedent on both sides, but no precedent directly on point. To reach one result or the other requires that the trial court reason by analogy from the precedent. All the model says is that the closer the

[3]Some might construe the term legalist to imply that cases have a "right" answer or that judges should avoid any reference to ideology. We do not use the term in this way. Rather, we suggest that reasoning to the closest available precedent is a way for judges to economize on the effort required to resolve a case.

case is to the precedent, the easier (and cheaper) it is to make the analogy.

The appellate court observes whether the trial court held the defendant liable or not liable. It does not, however, observe the location of the case. This assumption reflects the fact that it is the trial court, and not the appellate court, that hears the evidence (James et al., 1992). It takes effort for the appellate court to discover where along the spectrum the case falls. If, say, the trial court finds the defendant not liable, the appellate court doesn't know whether the seriousness of the case was 0, .1, .35, or .5. Only a close examination of the trial court record will reveal this information. The appellate court must decide when it will expend its scarce resources to learn where the case lies.

We first consider a one-period interaction between the trial court and appellate court. What happens when the appellate court observes a trial court disposition that favors the plaintiff (liable)? In this situation, the appellate court suffers no loss because that outcome is consistent with its preferences. Accordingly, it spends the minimum possible effort affirming these decisions. For example, the appellate court might issue a one-line order or a cursory, unpublished decision. What about decisions where the trial court determines that the defendant is not liable? Assume that the appellate court suffers a loss of 100 if the case goes against its preference, which occurs when it is either .35 or .5. Given the no liability disposition, the appellate court can infer that each of these outcomes occurs 1/4 of the time. Thus, the appellate court will spend up to 50 (((1/2)*(100)) reviewing trial court outcomes that find no liability. If the appellate audit costs more than 50, the appellate court will simply affirm and suffer the loss. If it costs less, she will spend resources reviewing the record closely. If, through its investigation, the appellate court uncovers cases .35 or .5, it reverses. If it uncovers case 0 or .1, the appellate court affirms, as those cases are decided in a way the appellate court prefers.

This one-period example replicates the error correction models in the literature (Kornhauser (1992b); Shavell (1995)). The appellate court audits the trial court. If the appellate court finds a case in the undefined area that goes against its preferences, it reverses. Notably, there is no reason for the appellate court to develop law in this one-period interaction. The appellate court does not write an opinion because doing so would be costly (De Mesquita and Stephenson, 2002) and it would provide no future benefit given that the interaction lasts for one period.

Moving to a two-period interaction changes matters. This longer term perspective allows us to understand how developing the law might serve the long

term interests of the appellate court. In period 1, suppose that the appellate court reviews a case where the trial court held the defendant liable. In so doing, the appellate court uncovers case .7. The appellate court affirms and declares that, in all future cases, the plaintiff prevails in all cases .7 and higher. What happens in period 2?

Let's say that the trial court draws case .5. Given the new precedent promulgated in period 1, the closest analogous precedent is the .7 one. That precedent says the the plaintiff wins: the defendant is liable. The establishment of new precedent alters the way the trial court resolves case .5. Unlike the one-period model, the trial court now resolves cases 0, .1, and .35 in favor of the defendant. It resolves cases .5, .7, .95, and 1 in favor of the plaintiff. Recall the appellate court wants the defendant to be liable in cases .35, .5, .7, .95 and 1. By establishing new precedent in the first period, the appellate court gets closer to its desired resolution of the cases in the second period. Now, the appellate court is only willing to spend 33 (1/3*100) auditing not liable dispositions. If the cost of auditing exceeds 33, the appellate court will let the disposition with no review. If the cost of auditing is less than 33, she will review. In the one period model, the appellate court suffers a loss of, at most, of 50. In the two period model, the appellate court suffers a loss of, at most, 33.

Establishing precedent provides a new case from which the trial court can draw analogies and, in so doing, alters the way trial courts decide future cases. This ability to establish precedent, then, feeds back into the appellate court's initial review of trial court dispositions. Unlike the one-period model, the appellate court spends resources auditing dispositions with which it agrees. It might affirm and write an opinion about what counts as too little precaution. The appellate court isn't just trying to correct trial court errors; it wants to develop the law in a way that aids the trial courts in the future.

Consider two examples of this practice. In Nightingale Home HealthCare, Inc. v. Anodyne Therapy, LLC[4], the trial court awarded the defendant attorney's fees in a trademark case, which the Lanham Act permits in exceptional cases. In an opinion by Judge Posner, the appellate court affirmed. An express goal of the opinion was to "clarify" when attorneys' fees should be awarded in trademark cases. Judge Posner rooted the test in abuse of process. He then applied the test to the facts, saying that the plaintiff had abused the process because it filed the claim in an attempt to "coerce a price reduction" from the

[4]626 F.3d 958 (2010).

defendant.

Miller v. Othello Packers, Inc.[5] involved a contract between a crop grower and a crop processor. The parties left the price open, to be determined by "tonnage and grading as the beans went through the processor's plant." The trial court awarded the grower a reasonable value of the crop. The Supreme Court of Washington affirmed. It explained that the harvester failed to act in good faith because it had "left three truckloads of bean vines in the grower's fields." In so doing, the harvester reduced the value of the harvested crop and, with it, the price he would have to pay on the contract.

In both cases, the appellate court took time to write an opinion affirming the trial court. The opinions establish markers for future trial courts to follow. Before Nightingale a trial court may have found itself at sea when it tried to determine whether a set of facts qualified as exceptional circumstances. With little precedent it would take significant effort to justify an award of attorneys' fees. In this situation, the trial court may believe that the proper course is to analogize to the more common cases where fees are not shifted. But after Nightingale, trial courts understand that using trademark litigation to extract price concessions counts as abuse of process and, as a result, merits shifting fees. The next case might present actions close – but not identical to – a price concession. In light of Nightingale, the trial court will feel comfortable shifting fees in that future case; she can cite the Nightingale precedent and reason by analogy. Prior to Nightingale, this option was not available. Similarly, Miller tells courts that allowing a product to rot amounts to bad faith in a contract whose price based on the value of the processed seeds. Even though a subsequent case might involve a substantially different product–say delayed delivery of components for a new model of smartphone–the ability to analogize to the time-sensitive nature of the bean vines in Miller makes it easier for a future trial court to resolve the case.

These cases, we submit, are not atypical. Indeed, studies show that a significant number of published opinions affirm the result below (Mead, 2001). Yet the principal-agent model of judicial hierarchy struggles to explain this observed behavior. In that model, the appellate court sits in judgment of the trial court. It corrects any decision with which they disagree subject to resource constraints. Neither Nightingale nor Miller involve error correction of a deviant trial court. In our model, the reason for these cases is readily apparent: by making law, the

[5] 410 P.2d 33 (1966)

appellate court makes it easier (cheaper) for trial courts to decide future cases.

Three other insights follow from the observation that there is a benefit to developing precedent. First, the appellate court will spend more resources reviewing decisions it expects to reverse than decisions it expects to affirm. In this example, in period 1, the appellate court expects to reverse half the cases where the trial court finds the defendant not liable. Reversal carries a large potential for precedent building. To see why go back to period 1, where the trial court found the defendant not liable in cases 0, .1, .35, and .5.

Suppose that the appellate court audits a not liable disposition and uncovers case .5. The appellate court reverses and says that defendants should be found liable in this case. Now, in period 2, the trial court will "flip" the disposition of case .35 and case .5 from not liable to liable. The precedent is directly on point for any subsequent .5 case. Case .35 is closer to the liable precedent (resting at .5) than to the not liable precedent (resting at .1). If instead the appellate court audits a first period disposition finding liability, it knows that the case lies at .7. .95 or 1. The best the appellate court can hope for is to establish the .7 precedent. In so doing, the appellate court changes the trial court's disposition of the .5 case in the next period. Notice that the appellate court can never use a liable disposition to alter how trial court resolves case .35. The closest liable disposition, .7, is too far away from the .35 case to provide useful precedent to the trial court.

Put yourself in the mind of the appellate court. It knows that reversing a not liable disposition has the potential to change the trial court disposition in two kinds of cases in the future (.5 and .35). It also knows that affirming a liable disposition has the potential to change the trial court disposition in one kind of case in the future (the .5 case). Because two cases going its way are better than one, the appellate court is willing to spend more resources auditing decisions where it expects to reverse. The desire to create effective precedent leads to asymmetric appellate scrutiny: the appellate court looks harder at cases that have a chance of being reversed.

Turning to our next insight, consider dicta. Most commentary on the practice of dicta focuses on definitional issues (Abramowicz and Stearns, 2004). We take the hornbook definition as given – dicta are statements unnecessary for the resolution of the case. We ask why a court might issue dicta in the first place.

As scholars emphasize, reasoning by analogy relies on a comparison of the details in concrete controversies (Holmes (1870); Sunstein (1993)). With this guidance in mind, imagine a .7 case where the appellate court announces that

all cases .1 and above should result in liability for the defendant. Because the statements about a hypothetical .1 case are far and unconnected from the facts of the actual .7 case, the dicta might not be that helpful to the trial court. Put another way, the trial court cannot easily analogize an actual .1 case to what the appellate court has said about a hypothetical .1 case. To provide effective guidance, dicta needs to be related to the actual case before the appellate court. The statements need to be close, but not too close to the holding. Statements to close to the holding fail to capitalize on the potential benefits of dicta for aiding the lower court's analogical reasoning. Statements too far away are ignored. The appellate court must balance these two concerns.

Our model predicts that appellate courts will issue dicta and lower courts will make use of this guidance. The dicta statements won't involve an assertion of what the appellate judge prefers in all cases on a topic. Instead, they will be statements about what the appellate judge prefers in some closely related cases. Consider Myers v. Loudon County Public Schools[6], a constitutional challenge to the Pledge of Allegiance in the Fourth Circuit. Even though the Supreme Court had recently declined to decide the issue in Elk Grove v. Newdow,[7] the Myers majority upheld the Pledge on the basis of dicta in previous Supreme Court cases. In another instance of this pattern, an Oregon Court of Appeals had to evaluate the suspension of a driver's license through a questionable procedure.[8] The court expressly relied on dicta from the Oregon Supreme Court to reverse the suspension. More systematic studies of dicta suggest that these cases are not anomalies. Klein and Devins (2013) code a sample of appellate and trial court opinions for the use of dicta. They find that lower courts overwhelmingly follow dicta when they are available, which is consistent with the predictions of our model.

Finally, imagine that the parties settle cases where the law is certain. In this simple example, that means that cases 0, .1, .95, and 1 are never brought. Since these case settle, the appellate court knows that scrutinizing any remaining case will allow it to make new law. Appellate audits never "misfire" and this makes the appellate court better off. Settlement improves the performance of appellate review. We now are ready for the formal model. It establishes these points more precisely in an infinite-horizon setting. The model also allows us to consider what happens if some trial courts have policy preferences that conflict

[6]418 F.3d 395 (4th Cir. 2005).
[7]542 U.S. 1 (2004).
[8]Hays v. DMV, 216 P.3d 902 (Or. Ct. App. 2009).

with the appellate court.

3 The Model

A long-lived appellate court faces a series of short-lived trial courts. The idea that trial courts are short lived is consistent with a judicial system involving many trial courts and a single appellate court. With many trial courts, free riding will be prevalent, making coordination over actions which impact future payoffs to other trial courts hard to do. Each trial court thus defaults to maximizing its own one-period payoff. Time is discrete, indexed by $t = 0, 1, 2$. Each period, a trial court draws a case $x \in [0, 1]$ from the uniform distribution A case is a set of facts that describes an activity. The larger x is, the higher the social costs and the lower the social benefits of the activity. At time 0, an initial precedent is exogenously established. This precedent instructs that (1) no liability should arise in cases below $\underline{x} < \frac{1}{2}$ and (2) liability should result for cases above $\overline{x} = 1 - \underline{x}$. The precedent might come from an (unmodeled) Supreme Court or from the appellate court itself. At time 0, the interval of uncertain law is $[\underline{x}, \overline{x}]$. The appellate court does not wish to expand the no liability precedent It wants to restrict the holding from the Supreme Court or from its own prior case law to its facts. One can thus think of $\underline{x}$ as the appellate court's preferred cutoff, slicing the case space into activities that result in no liability and activities that result in liability.[9] The appellate court's per period loss from having a case decided in a way it disfavors is λ.

At time 0, the appellate court has not spoken about which activities beyond the initial precedent should result in liability. Over time, the appellate court will set law about when a defendant should be found liable. Define the lowest such activity as of time t as $\widetilde{x}_t$.

Trial courts fall into two categories: $i \in \{1, 2\}$. The category 1 trial court judge is legalist. She wants to minimize the effort she must expend to analogize a new case to appellate precedent. This legalist trial court judge has no preferences over who prevails in the litigation. The category 2 trial court is a realist; she is the attitudinalist from the political science literature that cares only about policy outcomes. Specifically, the realist trial court judge wants to

[9] We study the case where the preferred cutoff can be found in the inherited prior case law. The issue is that the law is undefined in the area above the cutoff. That is to say, the trial courts do not know whether to limit that precedent or face large analogical reasoning costs in doing so. The appellate court provides guidance by issuing new decisions. Baker and Mezzetti (2012) study the case where the preferred cutoff is unknown to the appellate court.

expand the applicability of the initial "no liability" precedent beyond its present position. The realist trial court suffers a loss of l if the defendant is found liable anywhere along the unit interval.[10]

After receiving a case, the trial court must render a decision and justify it.[11] Justification requires effort. The effort required to find the defendant not liable depends on how analogous the case is to the not liable precedent case. The closer the case is, the less effort will be required to justify ruling for the defendant. As an example, suppose a trial court wishes to write a summary judgment decision in favor of the defendant. Judicial norms demand that the trial court provide reasons for deciding the case as a matter of law. In this model, reasons take the form of saying that the case is analogous to the no liability precedent. Given a short distance between the case and the precedent, the trial court can easily write the summary judgment decision. The further away the case is from the precedent border, the harder it is for the trial court to write this decision.[12] Formally, we model the effort cost incurred in finding the defendant not liable as $x - \underline{x}$. Likewise – and for similar reasons– the effort cost associated with finding the defendant liable is $\widetilde{x}_t - x$. Figure 2 depicts these costs.

A skeptic might respond that a trial court can always exert no effort whatsoever. Perhaps the trial court could flip a coin to see determine which side prevails.[13] Doing so would, however, violate the Federal Rules of Civil Procedure. In a bench trial, Rule 52 requires that "the court must find the facts specially and state its conclusions of law separately." In awarding summary judgment, the trial court must "state on the record the reasons for granting or denying the motion.for summary judgment." We assume that trial courts look to appellate court precedent to justify dispositions.

After drawing a case, the trial court either finds the defendant liable or not liable and the case is appealed. We assume that the appellate court doesn't

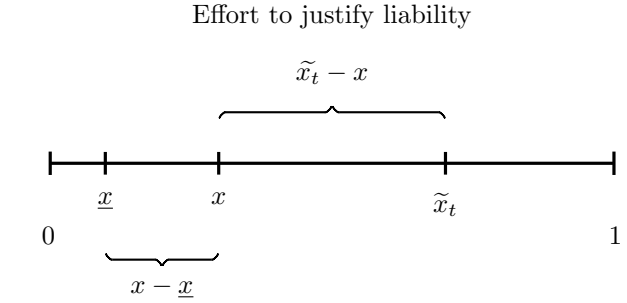
[10] More generally, we might say that the trial court's utility function is $U(e, l)$, where e is the cost of effort and l is the loss from a disfavored case resolution. To make matters easy, we assume the legalist trial court only cares about e and the realist trial court only cares about l.

[11] Kim (2007) notes that "[W]hen deciding cases, judges do at least two things. They determine the outcome of the dispute before them, and they offer reasons for their decision that connect the facts to applicable legal doctrine." The latter step requires effort and is what we model here.

[12] We have bench decisions in mind in the model. But we might also think that the trial court in instructing the jury might emphasize one set of precedents over another.

[13] For an analysis of where the legal system tolerates, and does not tolerate, random outcomes see (Samaha, 2009)

Figure 2: Analogizing Cases



immediately know the true location of x. It can only observe the the type of the trial court, legalist or realist, and the outcome, liable or not liable. The appellate court cannot tell how much effort it took the trial court to justify the decision. The appellate court can, however, spend effort auditing the trial court's decision. It can examine the trial record in detail, for example. We model this decision as the appellate court selecting audit probabilities. Denote the probability of a successful audit of a no liability disposition rendered by trial court i as $p_i \in [0,1]$. The probability of a successful audit of a liable finding is $q_i \in [0,1]$.[14] Upon a successful audit, the appellate court learns the true location of x.

Since the appellate court has limited resources, there is a cost to auditing. The cost is the same for liable and not liable dispositions and is represented by $c(\cdot)$. The cost function satisfies the usual conditions, namely $c(0) = 0, c' > 0, c'' > 0; c'(0) = 0; c'(1) = \infty$. If the audit fails, the appellate court doesn't learn enough about the location of x to reverse the trial court (it cannot find reversible error). Thus, it lets the trial court decision stand.[15]

[14] We will alternatively refer to the audit probabilities as probabilities or appellate court effort per case. They are interchangeable, given that any level of effort corresponds to a level of audit success. For a similar approach to modeling the successful discovery of information, see Che and Karthnik (2009).

[15] The assumption that the appellate court must locate the case before reversing rules out the following course of events: The realist trial court finds the defendant not liable. The appellate court – knowing a realist trial court made the ruling – reverses summarily, without expending any effort. Anticipating immediate reversal if his type is found out, the realist trial court spends effort trying to convince the appellate court it was a legalist trial court instead. Rather than investigate this asymmetric information, we assume that the trial court's type is observable, but the case location is not. Then, in line with actual practice, we assume

We next examine how the trial court will decide cases. Take the legalist trial court hearing a case $x \in [\underline{x}, \widetilde{x}_t]$. Suppose the trial court finds the defendant not liable. For this disposition, the reasoning or effort cost is $x - \underline{x}$. If instead the trial court finds the defendant liable, rationalizing the finding costs $\widetilde{x}_t - x$. Given these costs, the decision rule of the legalist trial court is simple:

$$\begin{aligned} x &< \frac{\underline{x} + \widetilde{x}_t}{2} = x_1^* \qquad \text{Not Liable} \\ x &\geq \frac{\underline{x} + \widetilde{x}_t}{2} = x_1^* \qquad \text{Liable} \end{aligned} \tag{1}$$

Distance from the precedent border determines the trial court's decision. If the case lies closer to the not liable precedent than the liable precedent, the trial court finds the defendant not liable. Otherwise, he finds the defendant liable. This, we submit, is one way to model a effort-motivated or legalist judge.[16] She just cares about which appellate precedent is closer to the case at hand because close analogies are cheaper in terms of legal reasoning costs to do than distant analogy. The cutoff is the average value of the precedent in the region of uncertain law. The legalist trial court doesn't care whether the plaintiff or the defendant wins. Thus, we assume they are indifferent to reversal.[17]

Next take the realist trial court. If she finds the defendant not liable, she will be reversed with probability p_2 and suffer a policy loss of l. If instead, she finds the defendant liable she suffers a loss for sure. Thus, the realist trial court always finds the defendant not liable, hoping to avoid reversal if the appellate court cannot verify the location of x and thus defers to its disposition.

A few remarks are in order here. First, the realist trial court finds the

that, before reversing a trial court, the appellate court must be comfortable saying there was reversible error. Put another way, the appellate court must expend the effort and locate x in order to reverse.

[16] The same model can accommodate a trial court whose objective is to follow the appellate court precedent. Suppose the trial court suffers a loss if he decides the case incorrectly from the appellate court's point of view. The trial court does not know the appellate court's cutoff point. It must infer the cutoff from the precedent interval $[\underline{x}, \widetilde{x}_t]$. Given this interval, the trial court treats the appellate court's cutoff as a random variable – call it x^C – distributed uniformly on $[\underline{x}, \widetilde{x}_t]$. After drawing a case x, suppose the trial court finds the defendant not liable. The probability he believes this decision is mistaken is $prob\{x^C < x\} = \frac{x - \underline{x}}{\widetilde{x}_t - \underline{x}}$. Suppose instead he finds the defendant liable. The probability he believes this decision is mistaken is $prob\{x^C > x\} = \frac{\widetilde{x}_t - x}{\widetilde{x}_t - \underline{x}}$. Setting equal the two expected losses gives the same decision rule, x_1^*, as in the text.

[17] One might suspect that trial courts who don't care about outcomes might still care about reversal. Reversal signals they are bad judges. Reversal costs are sometimes modeled as resulting from a failure of a lower court judge to get his preferred outcome (McNollgast 1995). We follow this approach here.

defendant not liable in more cases than the legalist trial court. Second, appellate precedent matters. It influences the effort the legalist trial court must expend in reaching certain outcomes. By setting new law, the appellate court makes it cheaper to render certain dispositions (by providing a close analogous case for the legalist trial court to refer to as precedent).

4 One-Period Benchmark

Start with a one-period benchmark. The benchmark allows for an easy comparison for the dynamic model to come. The timing of the game is as follows:

1. The appellate court picks: (a) the probability of successfully auditing a not liable decision by each category of trial court $\{p_1, p_2\}$; and (b) the probability of successfully auditing a liable decision by the legalist trial court q_1.

2. Nature draws a case, x, and a category of trial court to hear the case.

3. The legalist trial court decides the case according to its cutoff strategy, x_1^*. The realist trial court finds the defendant not liable.

4. The appellate court applies the audit probability suited for that trial court's resolution.

5. The appellate court reverses if it discovers that the trial court found the defendant not liable in a case above $\underline{x}$. The appellate court affirms if: (1) it discovers that the trial court found the defendant not liable in a case below $\underline{x}$; (2) it discovers that trial court found the defendant liable in a case above $\underline{x}$; or (3) the audit fails and it doesn't learn the location of x.

6. Payoffs are realized.

The appellate court anticipates the choices of the trial courts. Specifically, understanding the locations of x_1^* the appellate court decides its review strategy to maximize

$$\max_{0\leq p_1\leq 1;0\leq p_2\leq 1;0\leq q_2\leq 1} \beta_1\{-\lambda(1-p_1)[x_1^*-\underline{x}]-\Pr[NL]c(p_1)-\Pr[L]c(q_1)\} \\ +\beta_2\{-\lambda(1-p_2)[1-\underline{x}]-c(p_2)\}$$

The first term is the expected loss when legalist trial court hears the case. It consists of (1) the probability the case lies above the appellate court's preferred cutoff times the loss from a failed audit of a not liable disposition; (2) the cost of auditing the not liable disposition times the probability of a not liable disposition; and (3) the cost of auditing a liable disposition times the probability of a liable disposition.[18] The second term is the expected loss when the realist trial court hears the case. This term has two components: (1) the expected loss from a failed audit; and (2) the cost of auditing not liable dispositions. Since the realist trial court finds all defendants not liable, there is no cost of auditing liable dispositions in this part of the payoff.

The results of this maximization program are intuitive. The appellate court never audits dispositions where the defendant is found liable. The appellate court suffers no loss from this outcome. Further, it gains nothing from using a review of liable disposition to set precedent. On the other hand, the appellate court always audits when the defendant is found not liable. It then reverses upon successful detection of a case above its preferred cutoff.[19]

Formally, we have

Proposition 1 *In a one-period model, the appellate court audits trial court decisions finding no liability with positive probability, $p_1^{**} > 0$; $p_2^{**} > 0$. It never audits liability dispositions by the legalist trial court $q_1^{**} = 0$.*

In the one-period model, the appellate court does not issue opinions. The practice of defining the law for application in the trial courts is absent. The

[18]To derive the payoff with respect to the legalist trial court another way, define $f(x|NL)$ as the density conditional on a no liability finding and $f(x|L)$ as the density conditional on a finding of liability. Given uniformity, we have

$$f(x|NL) = \frac{1}{\Pr\{NL\}} \quad \text{if } x \in [0, x_1^*]; \text{ 0 otherwise}$$
$$f(x|L) = \frac{1}{\Pr\{L\}} \quad \text{if } x \in [x_1^*, 1]; \text{ 0 otherwise}$$

There are two states of the world: liable and not liable. The expected loss can be written as

$$\Pr[NL](-\lambda(1-p_1)\int_{\underline{x}}^{x_1^*} f(x|NL)dx - c(p_1))$$
$$+ \Pr[L](-c(q_1)$$

Multiplication gives the expression in the text.

[19]We use a double-star superscript to denote the values of the appellate court audit strategy that solves the one-period program. That way, we retain the use of the single star for the dynamic model to come.

one-period model amounts to error correction by the appellate court, with the inclusion of a cost of analogical reasoning in the utility function of the legalist trial court.

5 The Dynamic Optimization Problem

In the dynamic model, we ask whether the appellate court makes law and, if so, how. The legalist trial court's decision cutoff depends on the average value of cases where precedent provides no conclusive guidance. The bigger this average, the higher the cutoff is: the more cases where the trial court finds the defendant not liable instead of liable. If the appellate court has little precedent defining activities that result in liability, the trial court must work hard to justify holding the defendant liable. The analogous appellate precedent doesn't exist. Rather than expend effort reasoning from distant precedent, the trial court finds the defendant not liable instead. As noted in the numerical example, by making new law, the appellate court creates precedent that justifies a holding of liability that is closer to any subsequent case arising in the trial courts. Given this new precedent, the trial court will be more apt to find the defendant liable because it is cheaper to do so. This outcome, of course, is what the appellate court prefers. The appellate court audit of trial courts will be done with an eye toward these future benefits

In terms of the optimization, the state variable is $\widetilde{x}_t$ – the lowest activity where the appellate court has previously spoken and said the case should result in liability. The control variables are threefold. First we have p_{1t} and p_{2t} – the probability of successfully auditing a not liable finding of category 1 and 2 trial courts respectively. Second, there is q_{1t} – the probability of successfully auditing a liability finding by the legalist trial court. The choice of these variables depends on how much of the law the appellate court has defined previously – the location of $\widetilde{x}_t$.

Assuming a discount factor δ, the appellate court selects the control variables to maximize its discounted stream of payoffs. Denote the value function $V(\widetilde{x}_t)$. It can be expressed as

$$V(\widetilde{x}_t) = \max_{0\leq p_{1t}\leq 1;0\leq q_{1t}\leq 1;0\leq p_{2t}\leq 1} \beta_1\{-\lambda(1-p_1)[x_1^*-\underline{x}]-\Pr[NL]c(p_1)-\Pr[L]c(q_1)\} + \beta_2\{-\lambda(1-p_2)[1-\underline{x}]-c(p_2)\}+\delta E_t V$$

In this expression, E_tV is the appellate court's expected value function in the next period. This expected value function consists of a number of terms and equals

$$E_tV = V(\widetilde{x}_t)\left([1-\widetilde{x}_t+\underline{x}]+\beta_1\{(1-p_{1t})\int_{\underline{x}}^{x_1^*}dx+(1-q_{1t})\int_{x_1^*}^{\widetilde{x}_t}dx\}+\beta_2(1-p_{2t})\int_{\underline{x}}^{\widetilde{x}_t}dx\}\right)$$
$$+\beta_1\left(p_{1t}\int_{\underline{x}}^{x_1^*}V(x)dx+q_{1t}\int_{x_1^*}^{\widetilde{x}_t}V(x)dx\right)+\beta_2p_{2t}\int_{\underline{x}}^{\widetilde{x}_t}V(x)dx$$

The two regions of certain law are $[0,\underline{x}]$.and $[\widetilde{x}_t,1]$. In our numerical example, these two sets were {0,.1} and {.95,1}. If existing precedent is on point with the appealed case, the appellate court cannot use the case to make new law. In the model (as opposed to numerical example), the appellate court also cannot make new law if the audit fails to locate x. The first term adds these two probabilities together and multiplies them by $V(\widetilde{x}_t)$. Why? The environment doesn't change if the appellate court cannot make new law. It thus faces the same choices in period t+1 that it did in period t.

The second term is the expected payoff from successfully auditing the legalist trial court. The interval of uncertain law is separated into cases where the defendant is found not liable and cases where the defendant is found liable. If the audit is successful, the state of the appellate court case law shifts in next period. The amount of the change depends on the location of the case on the unit interval. That is why the value function is integrated over the case space in the second term. The third term is expected payoff from successfully auditing the realist trial court.

Lemma 1 in the appendix shows that, if the appellate court's loss from erroneous decisions is large, a unique and differentiable value function, $V(\widetilde{x}_t)$, exists. It decreases in $\widetilde{x}_t$ and reaches it minimum at $\widetilde{x}_t=\underline{x}$ – when the appellate court has defined the law for all possible cases.

The next proposition discusses how the appellate court makes law in this dynamic setting.

Proposition 2 *(a) In the dynamic model, the appellate court audits trial court decisions whose resolution do not result in a loss to the appellate court (i.e., $q_{1t}^*>0$). (b) The audit of the legalist trial court is asymmetric ($p_{1t}^*>q_{1t}^*$).*

Proposition 2 explains a number of institutional features of the judiciary. Depending on the location of the case and the success of the audit, the appellate court might (1) affirm and make law; (2) affirm in an unpublished decision; or (3) reverse and make new law. Specifically, if the case lies in the range of settled law, the appellate court issues an unpublished decision. If the audit fails–meaning that the appellate court cannot ascertain the location of x with sufficient confidence–the appellate court issues an unpublished decision that affirms the trial court's disposition. If the audit of a not liable disposition succeeds and the case lies in the region of uncertainty, the appellate court reverses and makes new law.[20] If the audit of a liable finding succeeds and the case lies in the region of uncertainty, the appellate court affirms and makes new law.

Part (a) of proposition 2 establishes that the appellate court audits cases whose disposition it knows it will agree with. The appellate court wants to make law. Liable findings offer a potential opportunity to do so. Indeed, the appellate court might hold oral argument or ask for additional briefing, all in a case where it knows an affirmance is in the offing.

Part (b) of proposition 2 predicts that the appellate court will scrutinize not liable dispositions more intensely than liable dispositions. In other words, the appellate court's auditing resources will tilt toward review of cases where the appellate court might disagree with the trial court's disposition. The obvious reason comes first: not liable findings can impose a loss on the appellate court, a loss avoided by a successful review and reversal. Liable findings never impose a loss on the appellate court. There is a less obvious reason too, one that accounts for the benefit of law creation. Not liable findings arise in cases closer to the not liable precedent border than liable findings. Reversal in such cases is especially useful for shifting down the liability precedent. The best case for setting law from the appellate court's perspective rests a smidge above $\underline{x}$. The appellate court wants to take this case, reverse and instruct that all cases above $\underline{x}$ result in liability. That way, the appellate court can make the most law in a single decision, thereby minimizing the extent – and cost – of future audits.

Part (b) of this proposition yields the first testable prediction from the model.

Prediction 1: Appellate courts should be more likely to make new law

[20] There is some evidence that judges act strategically when it comes to the decision whether to publish a case (Law, 2004). Our model of a single long-lived appellate court means that these strategic elements do not come into play. We leave the modeling of publication decisions on multi-member courts to future research.

through reversals of trial courts than through affirmances

Next consider how appellate scrutiny of trial court dispositions changes as the law becomes more defined.

Proposition 3 *(a) In the dynamic model, the appellate court devotes more resources to auditing the realist trial court than in the single period model; (b) If* $\lambda > c'(p_{1t}^*)$*, appellate scrutiny of not liable dispositions decreases as the law becomes more defined; that is,* $\frac{\partial p_{2t}^*}{\partial \widetilde{x}_t} > 0$ *and* $\frac{\partial p_{1t}^*}{\partial \widetilde{x}_t} > 0$.

Part (a) reveals a link between appellate review of the realist trial court and the creation of law for legalist trial courts. By assumption, the law has no impact on the decisions of the realist trial court. That said, the appellate court can use a successful review of a realist court's disposition to make law. That new law, then, changes the behavior of the legalist trial court going forward, saving the appellate court audit costs and losses from dispositions it disfavors. Because of the benefit of law creation, the appellate court spends more resources auditing decisions of the realist trial court in the dynamic model.

As the law gets more defined, it becomes less likely that an audit of a not liable finding will uncover a case in the range of uncertain law. Instead, a successful audit is likely to reveal a case overlapping with an existing appellate precedent. Given the overlap, the appellate court can't use the case to create new law. Since the benefit of the audit falls as the interval of uncertain law shrinks, the appellate court spends less auditing these dispositions of the trial courts.[21]

Taken together, proposition 3 suggests that appellate courts will invest more heavily in reviewing trial court dispositions when the law is its infancy, perhaps shortly after a statute has passed or just after the Supreme Court issues a decision. In each scenerio, the appellate court will be searching for cases decided in the trial courts to make law. Given the lack of precedent, this search will often be fruitful. This leads to the next prediction from the model.

Prediction 2: Over time, the appellate court will publish fewer and fewer

[21] The condition that $\lambda > c'(p_{1t}^*)$ ensures that $\frac{\partial p_{1t}^*}{\partial \widetilde{x}_t} > 0$. Intuitively, the condition means that, as the law becomes less well-defined, the marginal benefit of an audit of a not liable increases faster than the marginal cost. As to the audit of liable findings, the comparative statics are ambiguous; they can't be signed without knowing more about the aspects of the function V. We can say that, as the law converges, the audit of a liable finding by a legalist court goes to zero.

decisions on a specific topic; that is, the ratio unpublished to published decisions will rise.

The available evidence is consistent with this prediction. One prominent example is habeas petitions. The law in this area is fairly well defined and, consequently, it is unlikely to be an effective use of appellate court resources to publish many decisions in this area. Though systematic studies of the degree to which court fill in habeas law do not appear to exist, empirical evidence shows that the ratio of unpublished to published opinions for habeas appeals is substantially higher relative to other types of appeals (Mead, 2001). There are other areas of law where auditing will not be as intense. For example, one judge has suggested appeals of immigration cases and Social Security cases tend not to implicate the "law-declaring function" of appellate courts because most issues are determined by precedent (Jones, 1995).

This prediction suggests another likely pattern. When there is an exogenous shock to precedent–such as a Supreme Court opinion that reverses longstanding law–appellate courts should publish more opinions. The exogenous change can create a new area of undefined law and it will be in the interest of appellate courts to provide guidance in that interval. Examples of this phenomenon include the need to define what a substantive reasonable sentence is in the wake of Gall v. United States[22] and the development of a standard of review for Second Amendment claims after District of Columbia v. Heller[23] and McDonald v. City of Chicago[24] (Levy, 2013).

To close this section, we emphasize that the appellate court does strictly better as it defines the law. Its overall losses – given the optimal review strategy – shrink as the interval of uncertain law collapses. The appellate court does better by articulating new precedent because precedent changes how legalist trial courts decide future cases.[25]

[22] 552 U.S. 38 (2007).
[23] 554 U.S. 570 (2008).
[24] 130 S.Ct. 3020(2010).
[25] The predictions of the dynamic model differ from, say, McNollgast (1995). In that model, the Supreme Court sets law anticipating a single round of auditing of lower court decisons. In this model, the appellate court audits today in order to review a case where it can make law, which then enables it to save on auditing costs and losses from dispositions it disfavors in the future.

5.1 Convergence of Law

The appellate court can put in little effort and hope it successfully uncovers a case in the interval of uncertainty (this follows since $c'(0) = 0$). Given this assumption, the appellate court audits until the interval of uncertain law vanishes ($\underline{x} = \overline{x}_t$). If the law spans the entire spectrum, the legalist trial court will always find an appellate court case on point and follow it. The legalist trial court need not devote any effort to reasoning from an analogous case. The appellate court understands as much. It thus affirms the legalist court's disposition via unpublished decision.

The realist trial court is a different story. Given this trial court's preferences over outcomes, this court will continue to find the defendant not liable. After the law converges, then, the steady state involves (1) the realist trial court finding the defendant not liable in the face of a liable precedent directly on point and (2) the appellate court spending resources auditing not liable dispositions from the realist trial court.

We predict, then, asymmetric treatment of trial courts in the limit. The realist trial court faces appellate scrutiny. The legalist trial court does not.

Proposition 4 *(A) Eventually, the law converges – a precedent is defined for every possible case. (B) After convergence, the appellate court no longer audits dispositions by legalist trial courts; it continues to audit not liable dispositions by realist trial courts.*

After the law converges, the appellate court's long run loss is given by

$$V(\underline{x}) = \frac{-\beta_2\{(1-\underline{x})\lambda(1-p_2^{**}) - c(p_2^{**}))\}}{1-\delta}$$

Realist trial courts do not respond to precedent and appellate review is not perfect. As a result, the appellate court suffers an expected loss anytime a realist trial court hears a case. Some cases where the appellate court would have preferred liability go undetected and are not reversed.

In the long run, the appellate court suffers losses even though it has fully defined the law. This outcome arises because (1) the initial decision is delegated to an agent who holds different, conflicting preferences and (2) appellate review is costly. The long run losses are the agency costs of having a judicial hierachy. The appellate court can control the realist trial court if it devotes enough resources to detect each time they deviate from established precedent.

Yet devoting so many resources to detection is not cost-justifed for the appellate court.

The appellate court's per period loss is less in the long run than in the one period model. By making law, the appellate court enables the legalist trial court to costlessly follow its precedent. It provides cases on point, eliminating the effort cost associated with analogical reasoning. In so doing, the appellate court eliminates all its losses from delegating decisionmaking to the legalist trial court.

One testable prediction flows from this proposition.

Prediction 3: Appellate courts will target the amount of resources for review based on the type of trial court rendering the disposition. The appellate court will more heavily audit trial courts who hold conflicting policy preferences.

Measuring the resources that judges devote to the review of lower court opinions is, of course, a difficult task. But a number of empirical studies show that the behavior of lower courts appears to differ in a way that depends on the policy preferences of reviewing courts. These findings suggest that lower courts respond to the threat of more searching review by appellate courts that have conflicting policy preferences. Schanzenback and Tiller(2008) show this result in the sentencing context. They find that federal district court judges are more likely to depart from the federal sentencing guidelines when the reviewing court is politically aligned with the reviewing court. From this evidence, it appears that appellate courts are likely to expend more effort reviewing a sentence from a lower court with conflicting policy preferences and, moreover, lower courts appear to be responding to that preference. Another study shows that the degree of political conflict with reviewing courts appears to affect whether federal district courts choose to publish their decisions (Choi et al., 2012). This finding suggests that, when the stakes are high, policy conflicts will lead appellate judges to scrutinize cases more closely. Lower courts respond to this threat by choosing not to publish cases that the appellate court is likely to reverse. While these studies do not provide a direct measure of audit intensity, they permit inferences that are consistent with our model.

6 Holding and Dicta

Thus far, we have assumed that the promulgation of law took a specific form. If the appellate court uncovered a case located at, say, .7, it announced that all cases above .7 should be decided against the defendant. Using this case as a vehicle, the appellate court could not make law about cases below .7. The appellate court's opinion was limited by the facts as presented the trial court. This section relaxes this assumption. If the appellate court reviews a case at .7, it can say in the opinion that all cases above, say, .5 should result in liability. The opinion, then, consists of two parts: (1) a holding – all cases .7 and above should result in liability and (2) dicta – cases in the interval [.5,.7] should result in liability. Black's Law Dictionary (2009) defines "obiter dictum" as "[a] judicial comment made while delivering a judicial opinion, but one that is unnecessary to the decision in the case and therefore not precedential (although it may be considered persuasive)."[26]

In our example, any statements about case .5 are unnecessary for holding the defendant liable who has a .7 level of activity.

To study the impact of dicta, suppose that, if the the appellate court uncovers a case in the interval of uncertain law, it can issue dicta, $\Delta \in [\underline{x}, \widetilde{x}_t]$. In finding the defendant liable, the legalist trial court can then justify its decision by reference to the prior holding or dicta, or a combination of the two Suppose that this trial court's effort cost of holding the defendant liable is given by

$$[\alpha\Delta + (1-\alpha)\widetilde{x}_t] - x$$

In this expression, α is the weight the trial court allocates to dicta and $(1-\alpha)$ is the weight allocated to the holding. According to legal scholars, the further away the dicta is from the holding, the less likely it is to be persuasive (Dorf, 1994). In other words, the distance between dicta and the holding determines its influence. To capture this effect in the simplist way, suppose that

$$\alpha = \frac{\Delta - \underline{x}}{\widetilde{x} - \underline{x}}$$

If the dicta spans the entire interval of uncertain law, the trial court places no weight on it. The closer the dicta is to the holding, the more weight the trial court places on the statements. Suppose the appellate court aggressively issues

[26] On further definitions of dicta, see (Abramowicz and Stearns, 2004).

dicta in a case involving liability – setting Δ close to the no liability precedent border. With this move, the appellate court maximizes the chance that the trial court will ignore the dicta and rely solely on the holding. In other words, the trial court won't be able to put the dicta into context and thus it won't help reduce its cost in justifying a finding of liability. As the dicta gets closer and closer to the holding, it becomes more likely that the trial court will rely on the dicta. At the same time, the dicta doesn't span much of the interval – it doesn't say much beyond the holding As a result, the dicta doesn't lower the costs of finding the defendant liable by all that much. As we will see, the appellate court extends the dicta until the increase in the chance it fails to persuade the trial court just offsets the increase in liability findings if, in fact, it does persuade.

With this understanding of dicta in place, the trial court's dispositional cutoff becomes

$$x < \frac{\underline{x} + [\alpha\Delta + (1-\alpha)\widetilde{x}]}{2} = x_1^* \qquad \text{Not Liable}$$
$$x \geq \frac{\underline{x} + [\alpha\Delta + (1-\alpha)\widetilde{x}]}{2} = x_1^* \qquad \text{Liable}$$

The appellate court can (and will) issue a different amount of dicta for each case in the interval of uncertainty. Technically, the ability to issue dicta changes the state variable in the optimization problem. Before the state variable was the lowest activity where the appellate court had previously found the defendant liable. Now the state variable is the dicta/holding combination associated with that activity, $\widetilde{x}_t$. Define this new state variable as

$$\widehat{x}_t = \alpha\Delta(\widetilde{x}) + (1-\alpha)\widetilde{x}$$

The timing of the interaction each period runs like this:

1. The appellate court selects an audit review strategy $\{p_1, q_1, p_2\}$ and its plan for dicta, $\Delta(x)$, for each $x \in [\underline{x}, \widetilde{x}_t]$

2. A case and trial court is drawn. The trial court renders a decision: not liable or liable.

3. The appellate court applies the audit strategy associated with that disposition and category of trial court. If it discovers a case in the interval of uncertainty, the appellate court issues new dicta ($\Delta(x)$) and a new holding $\widetilde{x}_t$. In that case, the state variable changes accordingly.

Given a state variable $\widehat{x}_t$, the appellate court selects $\{p_{1t}\ q_{1t}, p_{2t}\}$ and $\Delta_t(\cdot)$ to maximize

$$V^{\Delta}(\widehat{x}_t) = \max_{0\leq p_{1t}\leq 1; 0\leq p_{2t}\leq 1; 0\leq p_{1t}\leq 1;\ \{\Delta_t(\cdot)\in[\underline{x},\widetilde{x}_t]\}} \beta_1\{-\lambda(1-p_1)[x_1^*-\underline{x}]-x_1^* c(p_1)-(1-x_1^* c(q_1)\} + \beta_2\{-\lambda(1-p_2)[1-\underline{x}]-c(p_2)\} + \delta E_t V^{\Delta}$$

In light of the possibility of dicta, the value function expected in the next period can be written as

$$E_t V^{\Delta} = V(\widehat{x}_t)\left([1-\widetilde{x}+\underline{x}] + \beta_1\{(1-p_1)\int_{\underline{x}}^{x_1^*} dx + (1-q_1)\int_{x_1^*}^{\widetilde{x}_t} dx\} + \beta_2(1-p_2)\int_{\underline{x}}^{\widetilde{x}_t} dx\}\right)$$
$$+\beta_1\left(p_1\int_{\underline{x}}^{x_1^*} V^{\Delta}(\widehat{x}(x,\Delta)dx + q_1\int_{x_1^*}^{\widetilde{x}_t} V^{\Delta}(\widehat{x}(x,\Delta)dx\right) + \beta_2 p_2\int_{\underline{x}}^{\widetilde{x}_t} V^{\Delta}(\widehat{x}(x,\Delta))dx$$

There are two differences from the baseline model. First, the cutoff for the legalist trial court is lower. Second, if the appellate court successfully audits a case in the interval of uncertain law, it can issue a holding and dicta. Thus, in the second line the value function (the expected value from a successful audit) depends both on the case draw (x) and the amount of dicta $\Delta(x)$.

As noted, the benefit of dicta is that it lowers the trial court's decision costs associated with finding the defendant liable. The trial court, however, ignores appellate court statements too far removed from the holding. Each period, the appellate court balances these two concerns and issues dicta optimally. Formally, we have

Proposition 5 *(A) The appellate court always issues dicta (i.e.,*$\Delta^*(x) = \frac{x+\underline{x}}{2}$*); (B) The appellate court is strictly better off when it can issue dicta than when it cannot; (C) The legalist trial court is strictly better off when the appellate court issues dicta.*

The literature on dicta provides support for the predictions of our model. While scholars and judges debate the extent to which judicial opinions should go beyond the specific facts at issue, there is no question that they regularly do so. Schauer (1995) goes so far as to argue that the presence of dicta is inevitable if judges have to give reasons to support their rulings. As he puts it, "every time a court gives a reason it is, in effect, giving an advisory opinion." If this

proposition is correct, that means that, to some degree, an appellate court issues dicta whenever it writes an opinion. This near-universal use of the practice is what the model predicts.

The commentary also suggests that dicta is useful to lower courts. As Judge Pierre Leval has explained, dicta "can assist future courts to reach sensible, well-reasoned results" (2006).[27] Some scholars emphasize that dicta can have the beneficial effect of providing guidance in future cases (Katyal, 1998) and others argue that broad statements are consistent with the need for judges to settle disputes in an authoritative manner (Alexander and Schauer, 1997). And, as we noted earlier, empirical studies of dicta confirm that lower courts often use the guidance that dicta provides (Klein and Devins, 2013). These findings are consistent with the model's predictions.

7 Settlement and Effective Appellate Review

This final section considers settlement. There are many models of settlement (Bebchuk (1984); Reinganum and Wilde (1986); Spier (1992); Shavell (1996)). Typically, they address when and whether parties will settle. The answer depends on which party holds the private information and who makes the offer. Scholars have also recognized that settlement can also influence the development of law (Galanter 1974). If the court only sees certain kinds of cases, it might skew the law in one direction or another, say, in favor of repeat player litigants.

Here, we take a blackbox approach to settlement. We assume that parties settle cases before the legalist trial court where the law is clear (that is, $x \notin [\underline{x},\widetilde{x}]$). The question is what effect, if any, will settlement have on the review strategy of the appellate court of cases that don't settle. Settlement, as we will show, improves the yield rate on the audit of the dispositions by the legalist trial court. Given settlement, the appellate court knows that case that go to trial and are appealed must involve new issues, issues where the law is uncertain. These cases present the opportunity to make law and, as a result, alter the future decisions in the legalist trial courts. Unlike in the baseline dynamic model, the appellate court will never audit successfully and be forced to write an unpublished decision. To analogize, settlement guarantees that the pool of cases appealed from the legalist trial court is stocked with fish. The higher stock makes the appellate court strictly better off. It never wastes its audit resources.

[27] While Judge Leval does believe that dicta has this benefit, he argues that courts engage in the practice too broadly and in a manner that exceeds their appropriate authority.

This, in turn, makes the appellate review more effective. We summarize the effect of settlement as follows.

Proposition 6 *The appellate court is strictly better off if it can encourage settlement before the legalist trial courts.*

This proposition is consistent with – and provides a rationale for – some appellate court practices. Opinions are public. The appellate court doesn't just issue opinions to fellow judges. They make the opinions available to parties outside the judicial system. Scholars of jurisprudence claim this practice provides legitimacy. We suggest something different: appellate courts want parties to settle to make for more effective appellate review of the cases that remain. Second, trial courts often say that a case presents a novel issue. They might say so in the disposition itself. The trial court thus signals that the case is one upon which the appellate court can make law. Such signaling helps reduce the review cost in the appellate court.

Prediction 4: The appellate court will publish more opinions if litigants have strong incentives to settle given clear law.

Settlement can create benefits at all levels of adjudication. Resolving a dispute before it gets before a court can save substantial resources for the parties and for the court system. Our model predicts that appellate courts will share this interest in settlement because it allows them to maximize the use of their own scarce resources. By filtering out cases clearly decided under current law, judges can spend more time writing opinions about previously unresolved areas of law. As we have suggested, making opinions public accomplishes this goal by alerting parties that a dispute may be resolved by existing precedent. But there are other ways that appellate courts encourage parties to resolve cases with little use of court resources. For example, many federal circuit courts have mediation programs that encourage parties to resolve their disputes before the court devotes much attention to the case (Ganzfried, 1996). This makes sense as a way to conserve a scarce resource–and it is a practice that fits with the predictions of our model.

8 Conclusion

Appellate courts typically do two tasks. First, they review and correct dispositions in the trial courts. Second, they make new law to be applied by trial courts. The model shows that the two tasks are interrelated. Appellate scrutiny of trial court cases is necessary to find cases upon which new law can be made. Making new precedent changes the way trial courts resolve cases, enabling the appellate court to devote fewer resources to auditing those dispositions. Models in the literature typically consider either one task or the other. Such models have trouble explaining affirmances with opinions, the practice of dicta and how appellate scrutiny changes over time. By considering both tasks in a dynamic model, the model explains these common features of the judiciary as consistent with a dynamic optimization problem. We can also explain why trial courts focus so much on finding a case on point and why appellate court want to provide more of those cases. To reiterate one last time, it makes it easier for trial courts to resolve cases by analogical reasoning.

The model leaves out important aspects of the legal system. We assume that all cases are appealed. There will be selection into appeal, which the appellate court audit strategy should account for and respond to. Second, the trial courts are short-lived. They do not care how today's decision impacts the payoff to future trial courts. If they did, one might suppose they would flag cases likely to supply good precedent for analogical reasoning and thus reduce their future decision costs. Finally, the model treats the appellate court as a unitary actor. Issues of bargaining among judges over the contours of an opinion or the decision to publish were suppressed. We leave these questions for future work.

References

Abramowicz, M. and Stearns, M. (2004). Defining dicta. *Stan. L. Rev.*, 57:953.

Alexander, L. and Schauer, F. (1997). On extrajudicial constitutional interpretation. *Harv. L. Rev.*, 110:1359.

Baker, S. and Mezzetti, C. (2012). A theory of rational jurisprudence. *Journal of Political Economy*, 120(3):513–551.

Bebchuk, L. A. (1984). Litigation and settlement under imperfect information. *Rand Journal of Economics*, 15(3):404–415.

Blackwell, D. (1965). Discounted dynamic programming. *The Annals of Mathematical Statistics*, 36(1):226–235.

Che, Y.-K. and Kartik, N. (2009). Opinions as incentives. *Journal of Political Economy*, 117(5):815–860.

Choi, S. J., Gulati, M., and Posner, E. A. (2012). What do federal district judges want? an analysis of publications, citations, and reversals. *Journal of Law, Economics, and Organization*, 28(3):518–549.

Cross, F. (2005). Appellate court adherence to precedent. *Journal of Empirical Legal Studies*, 2(2):369–405.

Daughety, A. F. and Reinganum, J. F. (2000). Appealing judgments. *The Rand Journal of Economics*, pages 502–525.

De Mesquita, E. B. and Stephenson, M. (2002). Informative precedent and intrajudicial communication. *American Political Science Review*, 96(04):755–766.

Dorf, M. C. (1994). Dicta and article iii. *University of Pennsylvania Law Review*, 142(6):1997–2069.

Frisch, D. (2003). Contractual choice of law and the prudential foundations of appellate review. *Vand. L. Rev.*, 56:57.

Ganzfried, J. J. (1996). Bringing business judgment to business litigation: Mediation and settlement in the federal courts of appeals. *Geo. Wash. L. Rev.*, 65:531.

Gennaioli, N. and Shleifer, A. (2007). The evolution of common law. *Journal of Political Economy*, 115(1):43–68.

Gennaioli, N. and Shleifer, A. (2008). Judicial fact discretion. *J. Legal Stud.*, 37:1–577.

Holmes, O. W. (1870). Codes, and the arrangement of the law. *Am. L. Rev.*, 5:1.

James, F., Hazard, G., and Leubsdorf, J. (1992). *Civil procedure*. Little, Brown.

Jones, E. H. (1995). Back to the future for federal appeals courts: Rationing federal justice by recovering limited jurisdiction. *Tex. L. Rev.*, 73:1485–1821.

Katyal, N. K. (1998). Judges as advice givers. *Stan. L. Rev.*, 50:1709.

Kim, P. (2007). Lower court discretion. *New York University Law Review*, 82(2):383–442.

Klein, D. and Devins, N. (2013). Dicta, schmicta: Theory versus practice in lower court decision making. *William & Mary Law Review*, 54(6).

Kornhauser, L. A. (1992a). Modeling collegial courts i: Path-dependence. *International Review of Law and Economics*, 12(2):169–185.

Kornhauser, L. A. (1992b). Modeling collegial courts. ii. legal doctrine. *Journal of Law, Economics, & Organization*, 8(3):441–470.

Law, D. S. (2004). Strategic judicial lawmaking: Ideology, publication, and asylum law in the ninth circuit. *U. Cin. L. Rev.*, 73:817.

Lax, J. R. (2007). Constructing legal rules on appellate courts. *American Political Science Review*, 101(3):591.

Leval, P. N. (2006). Judging under the constitution: Dicta about dicta. *NYUL Rev.*, 81:1249.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. MIT Press.

Llwellyn, K. N. (1960). The common law tradition.

McNollGast (1995). Politics and the courts: A positive theory of judicial doctrine and the rule of law. *Southern California Law Review*, 68:1631–1684.

Mead, R. A. (2001). "unpublished" opinions as the bulk of the iceberg: Publication patterns in the eighth and tenth circuits of the united states courts of appeals. *Law Library Journal*, 94(4):589–607.

Posner, R. A. (2006). Reasoning by analogy. *Cornell L. Rev.*, 91:761–1383.

Rasmusen, E. B. (1994). Judicial legitimacy: An interpretation as a repeated game. *Journal of Law, Economics, and Organization*, 10:63–83.

Reinganum, J. F. and Wilde, L. L. (1986). Settlement, litigation, and the allocation of litigation costs. *The RAND Journal of Economics*, pages 557–566.

Samaha, A. M. (2009). Randomization in adjudication. *Wm. & Mary L. Rev.*, 51:1.

Schanzenbach, M. M. and Tiller, E. H. (2008). Reviewing the sentencing guidelines: Judicial politics, empirical evidence, and reform. *The University of Chicago Law Review*, pages 715–760.

Schauer, F. (1995). Giving reasons. *Stan. L. Rev.*, 47:633.

Shavell, S. (1995). The appeals process as a means of error correction. *Journal of Legal Studies*, 24.

Shavell, S. (1996). Any frequency of plaintiff victory at trial is possible. *J. Legal Stud.*, 25:493.

Songer, D. R., Cameron, C. M., and Segal, J. A. (1995). An empirical test of the rational-actor theory of litigation. *Journal of Politics*, 57:1119–1129.

Spier, K. E. (1992). The dynamics of pretrial negotiation. *The Review of Economic Studies*, 59(1):93–108.

Spitzer, M. and Talley, E. (2000). Judicial auditing. *The Journal of Legal Studies*, 29(2):649–683.

Stokey, N. L. and Lucas Jr, R. E. (2012). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Sunstein, C. R. (1993). On analogical reasoning. *Harvard Law Review*, 106(3):741–791.

White, J. B. (1995). What's an opinion for? *The University of Chicago Law Review*, pages 1363–1369.

Appendix

Proof of Proposition 1

(A) Let μ_{p_1}, μ_{p_2}, and μ_{q_1} be the non-negative Lagrangian multipliers on the constraints $p_1 \geq 0$,$p_2 \geq 0$, $q_1 \geq 0$. Since $c'(1) = \infty$, it is obvious that $p_1^* \neq 1$, $p_2^* \neq 1$and $q_1^* \neq 1$; thus those constaints don't bind. The Lagrangian can be written as

$$L = \sum_{i=1}^{2} \beta_i \pi_i + \mu_{p_1} p_1 + \mu_{q_1} q_1 + \mu_{p_2} p_2$$

where

$$\pi_1 = -(x_1^* - \underline{x})\lambda(1 - p_1) - x_1^* c(p_1) - (1 - x_1^*)c(q_1)$$
$$\pi_2 = -(1 - \underline{x})\lambda(1 - p_2) - c(p_2)$$

The first order condition with respect q_1 is

$$\frac{\mu_{q_1}}{\beta_1(1 - x_1^*)} = c'(q_1^*)$$

Suppose that q_1^* is in the interior (i.e., $q_i^* \in (0, 1)$). In that case, complementary slackness implies that $\mu_{q_1} = 0$. The first order condition must be

$$0 = c'(q_1^*)$$

a contradiction since $c'(q_1^*)$ is positive when $q_i^* \in (0, 1)$. Thus, $q_1^* = 0$. The first order conditions respect to p_1 and p_2 are

$$\lambda \frac{(x_1^* - \underline{x})}{x_1^*} + \frac{\mu_{p_1}}{\beta_1 x_1^*} = c'(p_1^*) \tag{2}$$
$$(1 - \underline{x})\lambda + \frac{\mu_{p_2}}{\beta_2} = c'(p_2^*) \tag{3}$$

Suppose that $p_1 = p_2 = 0$ (so $c' = 0$). Rearranging results in

$$\frac{\mu_{p_1}}{\beta_1 x_1^*} = -\lambda[x_1^* - \underline{x}]$$
$$\frac{\mu_{p_2}}{\beta_1 x_1^*} = -\lambda(1 - \underline{x})$$

a contradiction since μ_{p_i} must be non-negative. Thus, p_1^* and p_2^*. are at the

interior. The second order condition holds since $c'' > 0$

Lemma 1. A value function $V(\widetilde{x}_t)$ exists. The value function is decreasing and uniquely defined by

$$V(\widetilde{x}_t) = \max_{\substack{0 \leq p_{1t} \leq 1, 0 \leq q_{1t} \leq 1 \\ 0 \leq p_{2t} \leq 1,}} \sum_{i=1}^{2} \beta_i \pi_i$$

$$+\delta V(\widetilde{x}_t) \left([1 - \widetilde{x} + \underline{x}] + \beta_1 \{(1 - p_{1t}) \int_{\underline{x}}^{x_1^*} dx + (1 - q_{1t}) \int_{x_1^*}^{\widetilde{x}} dx\} + \beta_2 (1 - p_{2t}) \int_{\underline{x}}^{\widetilde{x}} dx \} \right)$$

$$+ \delta \beta_1 \left(p_{1t} \int_{\underline{x}}^{x_1^*} V(x) dx + q_{1t} \int_{x_1^*}^{\widetilde{x}} V(x) dx \right) + \delta \beta_2 p_{2t} \int_{\underline{x}}^{\widetilde{x}} V(x) dx$$

Proof

To prove the lemma, we follow the approach in Baker & Mezzetti (2012), which applies the sufficiency conditions from Blackwell's Theorem (Blackwell 1965) to a similar problem (on this approach, see also Ljungqvist & Sargent (2012)).

Define S as the metric space of continuous functions mapping $[\underline{x}, \overline{x}]$ into the real line. Consider the complete metric $d(v, w) = \sup_{x \in [\underline{x}, \overline{x}]} |v(x) - w(x)|$. Let T be an operator mapping continuous function, v, into a new function, Tv. Specifically let

$$Tv = \max_{\substack{0 \leq p_{1t} \leq 1, 0 \leq q_{1t} \leq 1 \\ 0 \leq p_{2t} \leq 1,}} \sum_{i=1}^{2} \beta_i \pi_i$$

$$+ \delta v(\widetilde{x}_t) \left([1 - \widetilde{x}_t + \underline{x}] + \beta_1 \{(1 - p_{1t}) \int_{\underline{x}}^{x_1^*} dx + (1 - q_{1t}) \int_{x_1^*}^{\widetilde{x}_t} dx\} + \beta_2 (1 - p_{2t}) \int_{\underline{x}}^{\widetilde{x}_t} dx \} \right)$$

$$+ \delta \beta_1 \left(p_{1t} \int_{\underline{x}}^{x_1^*} v(x) dx + q_{1t} \int_{x_1^*}^{\widetilde{x}_t} v(x) dx \right) + \delta \beta_2 p_{2t} \int_{\underline{x}}^{\widetilde{x}_t} v(x) d$$

Given v, Tv is the guess at the value function. We must prove that Tv is a contraction mapping and, as a result, a unique and continuous value function

exists. To do so, we must show monotonicity and discounting (Blackwell, 1965). Suppose that $w(x) < v(x)$ for all $x \in [\underline{x}, \overline{x}]$. It is immediate that $Tw < Tv$ (the integration of the $w(x)$ terms is always smaller than the intergration of the $v(x)$ terms if $w(x) < v(x)$).

To show discounting, take a constant c, and the function $v + c$. We need to show that $T(v + c) = Tv + \gamma c$, where $\gamma \in (0, 1)$. We have

$$
\begin{aligned}
T(v+c) = & \max_{\substack{0 \leq p_{1t} \leq 1, 0 \leq q_{1t} \leq 1 \\ 0 \leq p_{2t} \leq 1,}} \sum_{i=1}^{2} \beta_i \pi_i \\
& + \delta[v(\widetilde{x}_t) + c] \left([1 - \widetilde{x}_t + \underline{x}] + \beta_1 \{(1 - p_{1t}) \int_{\underline{x}}^{x_1^*} dx + (1 - q_{1t}) \int_{x_1^*}^{\widetilde{x}_t} dx\} + \beta_2 (1 - p_{2t}) \int_{\underline{x}}^{\widetilde{x}_t} dx \} \right) \\
& + \beta_1 \left(p_{1t} \int_{\underline{x}}^{x_1^*} (v(x) + c) dx + q_{1t} \int_{x_1^*}^{\widetilde{x}_t} (v(x) + c) dx \right) + \beta_2 p_{2t} \int_{\underline{x}}^{\widetilde{x}_t} (v(x) + c) d
\end{aligned}
$$

which equals

$$T(v + c) = Tv + c\delta$$

Discounting holds.

To show that the value function is negative, notice that if v is negative, then, Tv is negative Finally, we show that Tv maps decreasing functions into decreasing function (see Stokey and Lucas (2012)) and, as a result, the value function increases in the amount of defined law. Suppose v is decreasing (i.e., $v' < 0$). Evaluated at the optimal values, we have

$$
\begin{aligned}
\frac{\partial Tv(\widetilde{x}_t)}{\partial \widetilde{x}_t} = & \beta_1 \frac{\partial x_1^*}{\partial \widetilde{x}_t} [-\lambda(1 - p_{1t}^*) - c(p_{1t}^*) + c(q_{1t}^*) + \delta(p_{1t}^* - q_{1t}^*)[v(x_1^*) - v(\widetilde{x}_t)] \\
& + \{v'(\widetilde{x}_t)[1 - \beta_1(p_{1t}^*(x_1^* - \underline{x}) - q_{1t}^*(\widetilde{x}_t - x_1^*)] - \beta_2 p_{2t}^*(\widetilde{x}_t - \underline{x})]\} < 0
\end{aligned}
$$

We know that $\frac{\partial x_i^*}{\partial \widetilde{x}_t} = \frac{1}{2}$, $v < 0$ and $v' < 0$. The term on the second line in curly brackets is always negative (the coefficient on v' is positive for all values of the parameters and choice variables). The first term is negative if λ is sufficiently large. Note that the term is not necessarily negative for all values of λ. Because v is decreasing, the last term in that part can be positive (it will be if $p_{1t}^* > q_{1t}^*$). The condition amounts to saying that the appellate court prefers that the case

on the margin between liable and not liable, flip from not liable to liable. If it does, the appellate court gains some utility in the current period. At the same time, it loses with some probability the opportunity to learn about the marginal case (this happens whenever $p^*_{1t} > q^*_{1t}$). The condition says that this lost opportunity to learn is outweighed by the gains in fewer losses in the current period. That will always happen if we set λ large enough.

Proof of Proposition 2

To prove part (a), take the first order condition with respect to q_{1t}

$$\frac{\partial Obj}{\partial q_{1t}} = \frac{\delta \int_{x_1^*}^{\widetilde{x}_t} (V(x) - V(\widetilde{x}_t))dx}{1 - x_1^*} = c'(q_1^*)$$

Since V is negative and decreasing, $\int_{x_1^*}^{\widetilde{x}_t} (V(x) - V(\widetilde{x}_t))dx$ must be positive. The audit probability q^*_{1t} is positive so long as there is space in the interval $[x_1^*, \widetilde{x}]$. Since $x_1^* = \frac{\widetilde{x}+\underline{x}}{2}$, this space exists until the law completely converges (i.e., $\widetilde{x}_t = \underline{x}$).

To prove part (b), at the interior solution, the first order conditions with respect to p_{1t} and q_{1t} are

$$\frac{\partial Obj}{\partial p_{1t}} = \frac{1}{x_1^*}\{\lambda(x_1^* - \underline{x}) + \delta \int_{\underline{x}}^{x_1^*} [V(x) - V(\widetilde{x}_t)dx\} = c'(p^*_{1t})$$

$$\frac{\partial Obj}{\partial q_{1t}} = \frac{\delta \int_{x_1^*}^{\overline{x}} (V(x) - V(\widetilde{x}))dx}{1 - x_1^*} = c'(q^*_{1t})$$

We made the following assumption about the initial liability precedent, $\overline{x} = 1 - \underline{x}$. Any new liability precedent will always be set less than $\overline{x}$ (i.e., $\widetilde{x}_t < \overline{x}$). Since $x_1^* = \frac{\widetilde{x}+\underline{x}}{2}$ this means that $x_1^* < 1 - x_1^*$. Because the denominator on the marginal benefit side (the LHS) is smaller in first equation, it follows that $p^*_{1t} > q^*_{1t}$ if $\int_{\underline{x}}^{x_1^*} V(x)dx > \int_{x_1^*}^{\widetilde{x}_t} V(x)dx$, which must be true if $V(x)$ is negative and decreasing. In other words, the expression on the left-hand side of the inequality

is larger because: (1) the expression inside the integral is always smaller; and (2) the integration is over the same range since x_1^* lies at the average of $\underline{x}$ and $\widetilde{x}_t$.

Proof of Proposition 3

Part (a)

In the dynamic model, the first order condition with respect to p_{2t} is

$$\frac{\partial Obj}{\partial p_{2t}} = (1-\underline{x})\lambda + \delta \int_{\overline{x}}^{\widetilde{x}_t} (V(x) - V(\widetilde{x}_t))dx = c'(p_{2t}^*)$$

In the one period model, the optimal p_2 was defined by

$$(1-\underline{x})\lambda = c'(p_2^{**})$$

Since $\delta \int_{\overline{x}}^{\widetilde{x}_t} (V(x) - V(\widetilde{x}_t))dx > 0$, it follows that $p_{2t}^* > p_2^{**}$.

Part (b)

Notice that each first order condition is independent of the other choice variables. Thus, we can totally differentiate each with respect to $\widetilde{x}_t$ to obtain the comparative statics. Doing so yields

$$\frac{\partial p_{2t}^*}{\partial \widetilde{x}_t} = \frac{\delta \int_{\underline{x}}^{\widetilde{x}_t} dx V'(\widetilde{x}_t)}{-c''} > 0$$

$$\frac{\partial p_{1t}^*}{\partial \widetilde{x}_t} = \frac{\{c'(p_{1t}^*) - \lambda - \delta\{V(x_1^*) - V(\widetilde{x}_t)\}\frac{\partial x_1^*}{\partial \widetilde{x}_t} + \delta \int_{\underline{x}}^{x_1^*} dx V'(\widetilde{x}_t)}{-x_1^* c''} > 0$$

The first inequality follows since V' is less than zero. In the second expression, the numerator is negative since $\lambda > c'(p_{1t}^*)$ by assumption; V is decreasing (thus, $V(x_1^*) > V(\widetilde{x}_t)$); $\frac{\partial x_1^*}{\partial \widetilde{x}_t} > 0$; and $V'(\widetilde{x}_t) < 0$.

Proof of Proposition 4

The appellate court loses the opportunity to make new law if it stops auditing ($p_{1t}^* = p_{2t}^* = q_{1t}^* = 0$). To demonstrate that the appellate court makes law for all possible cases (the liability precedent, $\widetilde{x}_t$ eventually equals $\underline{x}$), we show that the audit probabilities will not equal zero unless $\widetilde{x}_t = \underline{x}$. Denote the non-negative

Lagrangian multipliers on these constraints as $\mu_{p_{1t}}, \mu_{p_{2t}}$, and $\mu_{q_{1t}}$. Accounting for the constraints, the first order conditions of the Lagrangian are

$$\frac{\partial Obj}{\partial p_{1t}} = \frac{1}{x_1^*}\{\lambda(x_1^* - \underline{x}) + \delta \int_{\underline{x}}^{x_1^*} [V(x) - V(\widetilde{x})dx\} - c'(p_{1t}^*) + \frac{\mu_{p_{1t}}}{\beta_1 x_1^*} = 0$$

$$\frac{\partial Obj}{\partial q_{1t}} = \frac{\delta \int_{x_1^*}^{\overline{x}} (V(x) - V(\widetilde{x}_t))dx}{1 - x_1^*} - c'(q_{1t}^*) + \frac{\mu_{q_{1t}}}{\beta_1(1 - x_1^*)} = 0$$

$$\frac{\partial Obj}{\partial p_{2t}} = \delta \int_{x_1^*}^{\overline{x}} (V(x) - V(\widetilde{x}))dx - c'(q_{1t}^*) + \frac{\mu_{p_{2t}}}{\beta_2} = 0$$

Suppose that $p_{1t}^* = 0$. In that case, we have

$$\frac{\mu_{p_{1t}}}{\beta_1 x_1^*} = -\{\frac{1}{x_1^*}\{\lambda(x_1^* - \underline{x}) + \delta \int_{\underline{x}}^{x_1^*} [V(x) - V(\widetilde{x})dx\}$$

which is a contradiction if the RHS is less than zero (recall thar $\mu_{p_{1t}}$ is non-negative). The RHS is strictly negative unless $x_1^* = \underline{x}$. Recall that $x_1^* = \frac{\underline{x} + \widetilde{x}_t}{2}$. Thus, the only time that $x_1^* = \underline{x}$ is if the law has converged completely (i.e., $\widetilde{x}_t = \underline{x}$). Proofs for p_{2t}^* and q_{1t}^* follow similar logic.

To prove part (B), note that, after convergence, $x_1^* = \underline{x}$. Thus, the disposition cutoff for the legalist trial court matches the appellate court's cutoff. Thus it doesn't make sense to even spend a small amount auditing the legalist trial court.

Proof of Proposition 5

At the interior solution, the first order conditions with respect to $\Delta(\cdot)$ for each $x \in [\underline{x}, x_1^*]$ and $x \in [x_1^*, \widetilde{x}_t]$ are respectively

$$\frac{\partial Obj}{\partial \Delta(x)} = \frac{\partial\{\alpha\Delta(x) + (1-\alpha)x\}}{\partial \Delta}\{\delta V^{\Delta\prime}(\cdot)\{\beta_1 p_1^* + \beta_2 p_2^*\} = 0$$

$$\frac{\partial Obj}{\partial \Delta(x)} = \frac{\partial\{\alpha\Delta(x) + (1-\alpha)x\}}{\partial \Delta}\{\delta V^{\Delta\prime}(\cdot)\{\beta_1 q_1^* + \beta_2 p_2^*\} = 0$$

Note that all the audit probabilities are positive. If λ is sufficiently large, then, $V^{\Delta\prime} < 0$ in this new optimization program with dicta available. Thus, these

first order conditions only hold if

$$\frac{\partial\{\alpha\Delta(x)+(1-\alpha)x\}}{\partial\Delta}=0$$

Recall that $\alpha=\frac{\Delta-\underline{x}}{\widetilde{x}-\underline{x}}$. Thus, we can write this expression as

$$\frac{\partial\{\frac{\Delta^2-\Delta\underline{x}+x-\Delta x}{\widetilde{x}-\underline{x}}\}}{\partial\Delta}=2\Delta^*-\underline{x}-x\Delta^*=0$$

which implies that $\Delta^*=\frac{\underline{x}+x}{2}$. Since $\underline{x}<x$, the dicta is always set less than the holding. A restriction saying no dicta forces the appellate court to set $\Delta=x$. Such a constraint must make the appellate court worse off since it prefers $\Delta=\Delta^*$ and $\Delta^*=\frac{\underline{x}+x}{2}<x$.

To prove part (C), note that the legalist trial court suffers effort losses if the case lies in the interval, $[\underline{x},\widehat{x}_t]$. Otherwise there is a case plus dicta sufficiently close to be on point. By using dicta, the appellate court always reduces this interval. The restriction of dicta makes the interval where the trial court suffers a loss equal to $[\underline{x},\widetilde{x}_t]$, which is strictly larger.

Proof of Proposition 6

Let I_t be the probability the appellate court encourages settlement prior to appeal of cases $x\notin[\underline{x},\widetilde{x}_t]$ before the legalist trial court. We treat I as a choice variable for the appellate court. It of course can depend on $\widetilde{x}_t$. The appellate court can make its decisions public or make the ruling more or less clear, thereby effectively communicating the rules to parties outside the judiciary. The maximization problem for this new program

$$V^S(\widetilde{x}_t)=\max_{\substack{0\le p_{1t}\le 1,\,0\le q_{1t}\le 1\\ 0\le p_{2t}\le 1}}\beta_1\pi_1'+\beta_2\pi_2$$

$$+\delta V^S(\widetilde{x}_t)\left(\beta_1(1-I)[1-\widetilde{x}_t+\underline{x}]+\beta_2[1-\widetilde{x}_t+\underline{x}]+\beta_1\{(1-p_{1t})\int_{\underline{x}}^{x_1^*}dx+(1-q_{1t})\int_{x_1^*}^{\widetilde{x}_t}dx\}+\beta_2(1-p_{2t})\int_{\underline{x}}^{\widetilde{x}_t}dx\}\right)$$

$$+\delta\beta_1\left(p_{1t}\int_{\underline{x}}^{x_1^*}V^S(x)dx+q_{1t}\int_{x_1^*}^{\widetilde{x}_t}V^S(x)dx\right)+\delta\beta_2p_{2t}\int_{\underline{x}}^{\widetilde{x}_t}V^S(x)dx$$

where

$$\pi_1' = (x_1^* - \underline{x})[-\lambda(1 - p_{1t}) - c(p_{1t})] - \underline{x}(1 - I_t)c(p_{1t})$$
$$- (\widetilde{x}_t - x_{1t}^*)c(q_{1t}) - (1 - \widetilde{x}_t)(1 - I)c(q_{1t})$$

and

$$\pi_2 = -(1 - \underline{x})\lambda(1 - p_{2t}) - c(p_{2t})$$

One can show that monotoncity and discounting hold for this new program; so the $V^S(\widetilde{x})$ exists and is unique. The first order condition with respect to I is

$$\frac{\partial Obj}{\partial I_t} = \underline{x}c(p_{1t}) + (1 - \widetilde{x}_t)c(q_{1t}) - \delta V^S(\widetilde{x}_t)\beta_1[1 - \widetilde{x}_t + \underline{x}] > 0$$

So, we are at a corner. It is a optimal to set $I_t^* = 1$ and encourage settlement. The constrained program where settlement is unavailable forces the appellate court to set $I = 0$, which must be strictly worse.