# Why Stare Decisis?*

Luca Anderlini
*(Georgetown University)*

Leonardo Felli
*(London School of Economics)*

Alessandro Riboni
*(Université de Montréal)*

June 2012

**Abstract.** All Courts rule ex-post, after most economic decisions are sunk. This might generate a time-inconsistency problem. From an ex-ante perspective, Courts will have the (ex-post) temptation to be excessively lenient. This observation is at the root of the principle of *stare decisis*.

*Stare decisis* forces Courts to weigh the benefits of leniency towards the current parties against the beneficial effects that tougher decisions have on future ones.

We study these dynamics and find that *stare decisis* guarantees that precedents evolve towards ex-ante efficient decisions, thus alleviating the Courts' time-inconsistency problem. However, the dynamics do not converge to full efficiency.

JEL Classification: C79, D74, D89, K40, L14.
Keywords: Stare Decisis, Precedents, Time-Inconsistency, Case Law.
Address for correspondence: Leonardo Felli, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. lfelli@econ.lse.ac.uk

*Part of the research presented here was carried out while Luca Anderlini was visiting EIEF in Rome and the LSE, and Leonardo Felli was visiting the Anderson School at UCLA and while all authors were at the Studienzentrum, Gerzensee. Their generous hospitality is gratefully acknowledged. We greatly benefited from comments by Sandeep Baliga, Margaret Bray, Hugh Collins, Ross Cranston, Carola Frege, Nicola Gennaioli, Gillian Hadfield, Andrea Mattozzi, Jean-Laurent Rosenthal, Alan Schwartz, David Scoones and Jean Tirole.

# 1. Introduction

## *1.1. Motivation*

*None the less, in a system so highly developed as our own, precedents have so covered the ground that they fix the point of departure from which the labor of the judge begins. Almost invariably his first step is to examine and compare them. If they are plain and to the point there may be need of nothing more.* Stare decisis *is at least the everyday working rule of our law.* [...]

[...] *It is when the colors do not match, when the references in the index fail, when there is no decisive precedent, that the serious business of the judge begins. He must then fashion law for the litigants before him. In fashioning it for them he will be fashioning it for others.* — Benjamin Cardozo (1921, p. 20–21).

Justice Cardozo paints a clear picture. Stare decisis is the workhorse of the judicial process.[1] When it works, it does so almost mechanically. However, when "the colors do not match" the "serious business" of considering *both* the parties *currently* before the Court and *future* ones begins in earnest.

Henry Campbell Black (1886, p. 745–746), citing James Kent (1896, Part III, Lect. 21, p. 476) is very clear about *why* stare decisis is needed in the first place.

[...] *It would, therefore, be extremely inconvenient to the public, if precedents were not duly regarded and implicitly followed. It is by the notoriety and stability of such rules that professional men can give safe advice to those who consult them; and people in general can venture with confidence to buy and trust, and to deal with each other.* [...]

This paper pursues a complementary answer to the "stability" rationale for stare decisis.[2] After all, citing Cardozo (1921, p. 28) once again: *"Nothing is stable. Nothing absolute. All is fluid and changeable."* And this is one of the strengths of Case Law which easily adapts and, in his view and that of many others (Posner, 2003, 2004), evolves towards efficiency as a result.

[1] Stare decisis translates literally from the Latin as "to stand by things decided." According to the Oxford English Dictionary stare decisis is "The legal principle of determining points in litigation according to precedent."

[2] See also Peters (1996).

A rationale for stare decisis that does not revolve on the predictability of Court behavior is also comforting in the world of modern economic models, populated by rational decision makers. Their predictive powers are so strong that the predictability of Courts only becomes an issue when asymmetric information is present, perhaps coupled with judicial bias.

In short, we argue that whenever Courts exercise *discretion*,[3] they are afflicted by a time-inconsistency problem that potentially generates a present-bias in their behavior.[4] Stare decisis is then a device that works in two ways. The first is mechanical, much as Cardozo (1921) has it. Precedents, if they have evolved in the "right direction" will often bind the Courts to avoid the temptation to be present-biased. The second is that welfare-maximizing Courts, when they are not bound by precedents, will have an incentive *not* to succumb to their present-bias temptation because of the effect of their decision on *future* litigants via stare decisis. This trade-off between the effects of current and future decisions is a fundamental part of the "serious business" that Cardozo (1921) refers to above.

So, what is the source of the Courts' present-bias temptation? It is hard to disagree with the observation that since Courts rule *ex-post*, they choose after most or all economic decisions have been taken; when the relevant economic choices are strategically *sunk*.

In a variety of situations, ex-post the interests of the current litigants may be best served by a degree of leniency which, when stacked against the optimal *ex-ante* incentives may well be inefficiently high. Our point of departure in this paper is precisely that Courts that have discretion will be tempted to be excessively lenient when they consider the welfare of the parties currently before them.

We return at some length to some examples of economic situations that fit well our claim in Section 2 below. For the time being we note that both ex-post debt-restructuring versus ex-ante investor protection, and ex-post patent infringement versus ex-ante incentives for

[3]We use the word discretion in the standard sense that it has acquired in Economics. Legal scholars are often uneasy about the term. Another way to express the same concept would be to say that Courts exercise "flexibility." Given that Courts in our model are always welfare-maximizers, it would be appropriate to say that, under Case Law, Courts exercise "flexibility with a view to commercial interest." We are grateful to Ross Cranston for making us aware of this terminological issue.

[4]The term "time-inconsistency" is a standard piece of modern economic jargon that goes back to at least Strotz (1956) and subsequently to the classic contributions of Phelps and Pollak (1968) and Kydland and Prescott (1977). It can be used whenever an ex-ante decision is potentially reversed ex-post. The term "present-bias" describes well the type of time-inconsistency that afflicts the Courts in our set-up. We use the two terms in an interchangeable way.

R&D are situations of first order economic importance to which our analysis applies.

### 1.2. *Preview*

Our model comprises a "pool" of cases; a draw from this pool materializes each period. In each period a Court of Law can, in principle, either take a forward looking, *tough*, or a myopic, *weak* (lenient) decision.

Each Court may be either constrained by precedents (which evolve according to a dynamic process specified below) or unconstrained.[5] In the latter case the Court has complete discretion to either take the tough or the weak decision.[6]

Whenever a Court of Law exercises discretion it does so necessarily *ex-post*. This biases the Court's decision away from ex-ante efficiency (in our stylized model always towards weak decisions). If the Court just maximizes the (ex-post) welfare of the parties in Court, the weak decision will always be taken.

The Courts' bias towards weak decisions is the dominant force determining their rulings in the absence of stare decisis. In such a hypothetical legal regime, a Court's ruling does not directly affect the decisions of future Courts. Therefore it is optimal for a Court to maximize the welfare of the current litigants and choose the weak decision. In the absence of stare decisis all Courts will take the weak decision and precedents will never evolve.

This bias is, instead, mitigated, although not entirely resolved, by the *dynamics of precedents* generated by the principle of stare decisis. Taking the tough decision, through precedent-setting, increases the probability that *future* Courts will be *constrained* to do the same, thus raising ex-ante welfare. The choice of each welfare-maximizing Court between a weak or a tough decision is determined by the trade-off between an instantaneous gain from a weak decision, and a future gain from a tough one, via the dynamics of precedents.

One of our key findings is that it is stare decisis that guarantees the evolution of precedents towards the ex-ante efficient, tough, decision. However, the time-inconsistency problem

[5]In reality, of course, it is seldom the case that a Court is either completely constrained or completely unconstrained by precedents. Each case has many dimensions, and precedents can have more or less impact according to how "fitting" they are to the current case. We model this complex interaction in a simple way. With a certain probability existing precedents "apply," and with the complementary probability existing precedents simply "do not apply." We do not believe that the main flavor of our results would change in a richer model capturing more closely this complex interaction, although the latter obviously remains an important target for future research.

[6]See footnote 3 above.

prevents Case Law from reaching full efficiency. Eventually the Courts *must* succumb to the present-bias. This is because they trade off a present increase in *(ex-post)* welfare, which does not shrink as time goes by, against a *marginal* effect on the decisions of future Courts. The latter eventually shrinks to be arbitrarily small.[7]

### *1.3. Related Literature*

The hypothesis that Case Law evolves towards efficiency has been widely investigated by the literature on Law and Economics. According to Posner (2004), judge-made laws are more efficient than statutes mainly because Courts, unlike legislators, have personal incentives to maximize efficiency.[8] Evolutionary models of the Common Law have called attention to explanations other than judical preferences. For instance, it has been argued that Case Law moves towards efficiency because inefficient rules are more often (Priest, 1977, Rubin, 1977) or more intensively (Goodman, 1978) challenged in Courts than efficient ones.[9]

More recently, a few papers have studied the dynamics of Case Law under the assumption that judges have biased preferences.[10] Gennaioli and Shleifer (2007b) consider a model of "distinguishing" where judges are able to limit the applicability of previous precedents by introducing new material dimensions to adjudication. In particular, their goal is to investigate the claim by Cardozo (Cardozo, 1921, p. 177), among others, that Case law converges to efficient rules even in the presence of judicial bias. Their results partially support this hypothesis. On the one hand, sequential decision making improves efficiency on average by making Case Law more precise. But on the other, Gennaioli and Shleifer (2007b) show that Case Law reaches full efficiency only under very stringent conditions.[11] Interestingly, Gennaioli and Shleifer (2007b) find that some judicial bias may be welfare improving; the

[7]At least since Cardozo (1921), the economic efficiency properties of the process of evolution of precedents have been the subject of intense scrutiny. We return to this point in the following Subsection 1.3. when we review some related literature.

[8]In Hadfield (1992), however, efficiency-oriented Courts may fail to make efficient rules because of the bias in the sample of cases observed by Courts.

[9]In a recent paper, Bustos (2008) studies the evolution of Common Law with forward-looking (and efficiency-oriented) judges by explicitly modeling the decision of the disputing parties to bring the case to Court.

[10]Judicial bias is interpreted in a broad sense that ranges from "idiosyncracies" in the judges' preferences (Bond, 2009, Gennaioli and Shleifer, 2007b, among others) to plain "corruption" of the Courts (Ayres, 1997, Bond, 2008, Legros and Newman, 2002, among others).

[11]When considering a model of overruling, as opposed to one of "distinguishing," Gennaioli and Shleifer (2007a) show that the case for efficiency in the Common Law is even harder to make.

intuition for this result is that polarized judges have stronger incentives to distinguish the existing precedent in order to correct the bias of the previous Court.[12]

In a model where judges can overrule previous precedents by incurring an adjustment cost, Ponzetto and Fernandez (2008) show that Case Law converges to an asymptotic distribution with mean equal to the efficient rule: in the long run judges' heterogeneous biases balance one another and Courts make better and more predictable decisions. It bears mentioning that in their model the rule of precedent has ambiguous welfare predictions: strong adherence to previous decisions slows down the convergence to the efficient rule, but it implies less variability in the long run. However, when judges are assumed to be forward looking (as it is always the case in our paper), they show that stare decisis is welfare decreasing since judges make more extreme decisions in order to have a longer-lasting impact on the law. In Gennaioli and Shleifer (2007a), for a given level of judicial polarization, welfare in Case Law is independent of the strength of stare decisis, as measured by the fixed cost of overruling the precedent.

We abstract completely from "judicial bias." This is not because we do not subscribe to the "pragmatist" view of the judicial process that can be traced back to at least Cardozo (1921) and subsequently Posner (2003). It is mainly to make sure that our results can be clearly attributed to the particular source of inefficiency we choose to focus on (present-bias temptation). Introducing judicial bias may well have ambiguous effects on welfare when Courts have more discretion; more specifically, it may strengthen the incentives of the current Court to take the tough decision in order to constrain future Courts via precedents.[13]

We also ignore the distinction between "lower" and "appellate" Courts.[14] As with judicial bias, we prefer to maximize the transparency of our results and leave the distinction out of the model. In our model *all* Courts have, in principle, the same ability to create precedents that affect future Courts. Clearly, in reality, appellate Courts differ from lower Courts in this respect. For instance Gennaioli and Shleifer (2007b) insist, realistically, that the Court that

[12]Similarly, Ponzetto and Fernandez (2011) find that the evolution of Case Law towards efficiency is more likely when judges are sufficiently polarized.

[13]See again Gennaioli and Shleifer (2007b) and Ponzetto and Fernandez (2011), where judicial polarization may actually improve the efficiency of the Common Law.

[14]The efficiency rationale for the existence of an appeal system has also receive vigorous scrutiny in recent years (Daughety and Reinganum, 1999, 2000, Levy, 2005, Shavell, 1995, Spitzer and Talley, 2000, among others)

changes the relevant body of precedents is an appellate Court. Their appellate Courts are immune from the potential time-inconsistency problem we identify here because, by assumption, judges' utility does not depend on the resolution of the current case. Provided that appellate Courts suffer at least to some extent from the same potential time-inconsistency problem as our Courts, the general flavor of our results would be unaffected by an explicit distinction between these two levels of judgement.

Case Law dynamics have been studied mainly from a theoretical perspective. One exception is Niblett, Posner, and Shleifer (2010), who analyze the evolution of a doctrine, known as the economic loss rule (ELR hereafter), over a period of 35 years.[15] Their contribution is directly related to our work because the application of the ELR in Courts is likely to be subject to credibility problems.[16] Niblett, Posner, and Shleifer (2010) show that while convergence to what they regard as the ex-ante efficient rule (ELR) was quite apparent (at least in some States) for about 20 years starting from 1970, in the early 1990's things changed and appellate Courts started accepting more and more exceptions to the ELR. The conclusion they draw is that Case Law not only did not converge to the efficient rule, but did *not converge at all* over the span of time they analyze. Their findings are obviously consistent with our theoretical indication that Case Law is unlikely to converge to the efficient rule.[17]

Finally, this paper is obviously related to the vast literature on time consistency problems. Since the classic contributions of Phelps and Pollak (1968) and Kydland and Prescott (1977), the literature has explored mechanisms that substitute for commitment and make credibility problems less severe. The institutional mechanism adopted by Case Law (the rule of prece-

[15]This rule broadly states that one cannot sue in tort for a loss that is not accompanied by personal injury or property damage. In the words of Judge Posner in *Miller v United States Steel Corp*: "Tort law is a superfluous and inapt tool for resolving purely commercial disputes. We have a body of law designed for such disputes. It is called contract law." (902 F.2d 573, 574, 7th Cir. 1990). In other words, the ELR encourages parties to solve their potential problems through contracts.

[16]It is conceivable that, at an ex-post stage, a judge may have sympathy for a wronged plaintiff—for example because the warranty specified in the contract has just expired—and be tempted to accept an exception to the ELR.

[17]In our model, Case Law matures and settles into a regime in which the Courts that have discretion succumb to the time-inconsistency problem that afflicts them and issue narrow rulings (idiosyncratic exceptions in their terminology) whenever they are not bound by precedents. The gap between their findings and the predictions of our model is that they observe the fraction of exceptions first decreasing and then *raising* rather than settling down as our model would predict. To reconcile the two, one would have to consider a version of our model where narrow rulings had a small effect on the body of precedents (instead of no impact as in our model); Case Law would likely not settle in this case. These issues are obviously ripe for future research, but clearly remain beyond the scope of the present paper.

dent) constitutes a distinct and ingenious solution to time consistency problems which, to our knowledge, has not yet been studied. The peculiarity of stare decisis is that the constraints to discretion (the precedents) are not imposed by an external mechanism designer but arise *endogenously* as a result of Courts' decisions.

Two papers from the literature on time consistency problems that are closer to ours are Phelan (2006), and Hassler and Rodríguez Mora (2007). Their models analyze credibility problems in a capital taxation model. Similarly to us, they focus attention on Markov-Perfect Equilibria. The mechanism through which policy makers in their models can (partly) overcome time consistency problems is different from ours, however. Hassler and Rodríguez Mora (2007), in a model where agents are loss-averse, show that the current government may keep capital taxes low in order to raise the households' reference level for consumption in the next period, so as to make it more costly for future governments to confiscate capital. In Phelan (2006), an opportunistic policy maker (whose type cannot be observed by households) may choose low taxes in order to increase his reputation.

Similarly to our characterization, the Markov perfect equilibria in their models may involve a randomization between "myopic" (confiscation) and "strategic (low taxes) behavior. The underlying intuition of why the equilibrium involves mixing is that the expectation of myopic behavior with certainty in the future generates an incentive to refrain from confiscation in the current period. Conversely, the expectation that future governments will refrain from confiscation induces the current government to raise taxes. In our model, as we discuss in Section 5 below, the incentive to procrastinate tough decisions is the reason behind the Courts' randomization. However, the same incentives to procrastinate do not apply if the decision is myopic, thanks to the Court's ability to control the "breadth" of its ruling.

### *1.4. Overview*

In Section 2 we briefly describe three leading examples of how time-inconsistency problems of the kind we consider here may arise. In Section 3 we set up the static model. We then proceed to lay down the model of precedents and hence our dynamic model. In Section 4 we report our first result that highlights the effect of time-inconsistency on the evolution of precedents (Proposition 1) as well as role of stare decisis in guaranteeing that precedents do evolve (Proposition 2). In Section 5 we impose some further restrictions on the precedents technology that allow us to characterize the equilibrium behavior of our model (Proposition

3). Section 6 concludes the paper. For ease of exposition, all proofs are in the Appendix.

## 2. Time-Inconsistency: Three Leading Examples

As we mentioned above, our point of departure is the observation that Courts examine the disputes brought before them at an *ex-post* stage. Many decisions will have been taken and much uncertainty will have been realized by the time a Court is asked to rule.

It is key to our results that the optimal decision for our benevolent Court may be different when evaluated ex-ante, or at the actual ex-post stage.[18] There are many examples of spheres in which the potential time-inconsistency we work with occurs. Here, we briefly describe three of these that we think are both important and fit well our setup.

Our first example is related to the "topsy-turvy" principle in corporate finance (see Tirole, 2005, Ch.16). Projects requiring finance can be of, say, high or low quality (ex-ante) and can be affected or not by a liquidity crisis (ex-post). Socially, it is optimal to let only high quality projects be financed ex-ante. Lenders cannot observe project quality, nor can they distinguish at an ex-post stage whether the borrower's state of distress is due to a low quality project or a liquidity problem.

Providing maximum protection to the lenders so that all projects in distress are liquidated achieves ex-ante selection in the sense that only borrowers that know to have high quality projects apply for funds. On the other hand, for projects of high quality, the social cost of re-deploying resources in a new activity after liquidation is high. Hence if only high quality projects are financed in the first place, ex-post it is optimal to lower lenders' protection and allow debt-restructuring. This avoids the social loss from redeploying resources away from high quality projects. The ex-ante and ex-post optimal Court decisions differ.

Our second example concerns patents. As in the first case, the specifics could take a variety of different forms, of which we only mention one. Consider a Court that examines a patent infringement case. From an ex-ante perspective, as is standard, the optimal breadth of the patent will be determined taking into account the trade-off between the incentives

[18]The distinction between "forward looking" decisions (that maximize ex-ante welfare) and ones that focus on the parties currently before the Court can be found in some of the extant literature. Kaplow and Shavell (2002) distinguish between "welfare" (ex-ante) and "fairness" (ex-post). Summers (1992) distinguishes between "goal reasons" (ex-ante) and "rightness reasons" (ex-post).

to invest in R&D, and the social cost of monopoly power exercised by the patent owner.[19] Ex-post, however, since the R&D investments are sunk, it is always socially optimal to rule in favor of the infringer and thus open the market to competition. So, once again, the optimal decision for the Court may differ according to whether we look at the problem ex-ante or ex-post.

The model in Anderlini, Felli, and Postlewaite (2011) involves a buyer and a seller in a model with relationship-specific investment, asymmetric information and incomplete contracting. In this world, it may be optimal for a Court to actively intervene in the parties' relationship and void some of the contracts that they may want to write. This is because without Court intervention inefficient pooling would obtain in equilibrium, and the ex-ante expectation of Court intervention will destroy the pooling equilibrium and hence raise ex-ante welfare. On the other hand, once a contract has been written and the parties have agreed which widget to trade, the optimal Court decision at an ex-post stage is to let the contract stand so that the parties can in fact trade. While intervention and voiding the contract is optimal ex-ante, the opposite is true ex-post.

## 3. The Model

### *3.1. The Static Environment*

The Court can take one of two possible decisions denoted $\mathcal{W}$ for "weak," or myopic, and $\mathcal{T}$ for "tough," or forward-looking. The Court's "ruling" is denoted by $\mathcal{R}$, with $\mathcal{R} \in \{\mathcal{W}, \mathcal{T}\}$.

Since our Courts are benevolent, their payoffs coincide with the parties' welfare, and we will use the two terms interchangeably.[20] The Court's payoffs are determined by the ruling it chooses and, critically, they may be different viewed from ex-ante and ex-post. Let $\Pi^{\mathcal{A}}(\mathcal{R})$ and $\Pi^{\mathcal{P}}(\mathcal{R})$, with $\mathcal{R} \in \{\mathcal{W}, \mathcal{T}\}$, denote the ex-ante and ex-post payoffs respectively. We assume that the optimal ruling is different from an ex-ante and an ex-post point of view. Ex-ante the optimal decision is the tough one, but ex-post the optimal ruling is instead the weak one. In other words, the time-inconsistency problem arises. Formally, we have

$$\Pi^{\mathcal{A}}(\mathcal{W}) < \Pi^{\mathcal{A}}(\mathcal{T}) \tag{1}$$

[19] See for instance the classic references of Nordhaus (1969) and Scherer (1972). For a discussion of the recent literature on (ex-ante) optimal patent length and breadth see Scotchmer (2006).

[20] As discussed above we assume benevolent Courts for simplicity.

and

$$\Pi^{\mathcal{P}}(\mathcal{W}) > \Pi^{\mathcal{P}}(\mathcal{T}) \tag{2}$$

### *3.2. Time-Inconsistent Courts*

As we mentioned before, our model relies on the key observation that Courts will be asked to rule on contractual disputes at an *ex-post* stage. Consider a (benevolent) Court that is unconstrained (by precedents) and that only considers the *present* case, without looking at any effect that its ruling might have on future Courts. Then, (2) tells us that its ruling will be $\mathcal{W}$. According to (1) from an ex-ante point of view the correct choice is instead the tough ruling $\mathcal{T}$. This is the source of the time-inconsistency problem, or present-bias, that afflicts the Courts.

### *3.3. The Nature of Precedents*

Each case comes equipped with *its own* specific *legal characteristics*, which determine, as we will explain shortly, whether the current *body of precedents* applies.

We model the legal characteristics of the case as a random variable $\ell$ uniformly distributed over $[0, 1]$.[21] This allows us to specify the body of precedents in a particularly simple way.

The body of precedents $\mathcal{J}$ is represented by a pair of numbers in $[0, 1]$ so that $\mathcal{J} = (\tau, \omega)$ with the restriction that $\tau \leq \omega$. Once a case comes to the attention of a Court its legal characteristics are determined: $\ell$ is realized.

The interpretation of $\mathcal{J} = (\tau, \omega)$ is straightforward. The body of precedents is seen to either apply or not apply and in which direction. If $\ell \leq \tau$ the body of precedents constrains the Court to a *tough* decision, if $\ell \geq \omega$ the body of precedents constrains the Court to a *weak* decision, while if $\tau < \ell < \omega$ the Court has *discretion* over the case.

Whenever the precedents bind the Court towards one decision or the other, we are in a situation in which the Court's ruling is determined by stare decisis. Whenever the precedents

[21] The fact that we take the legal characteristics of a case to be represented by a single-dimensional variable is obviously simplistic. While a richer model of this particular feature of a case would be desirable, it is completely beyond the scope of our analysis here. The modeling route we follow is just the simplest one that will do the job in our set-up.

do not bind, the case at hand is sufficiently idiosyncratic to escape the doctrine of stare decisis.[22]

Finally, note that in each period the contracting parties *observe* the body of precedents, and the legal characteristics of the case. Therefore, they know whether the Court will be constrained by precedents or not and in which direction if so. They will also correctly forecast the Court's decision if it has discretion. In other words, the parties anticipate correctly whether the Court will take a tough or weak decision.[23]

### *3.4. The Dynamics of Precedents*

The present-bias or time-inconsistency problem that afflicts the Courts is mitigated in two distinct ways. One possibility is that stare decisis applies and the Court's decision is predetermined by the past. Another is that the ruling of the current Court will affect the body of precedents that future Courts will face. A forward looking Court will clearly take into account the effect of its ruling on the ruling of future Courts. In doing so, it will evaluate the payoffs of future Courts from an ex-ante point of view.

Our next step is to describe the mechanics of the dynamics of precedents: *the precedents technology.* This is literally the mechanism by which the current body of precedents, paired with the current ruling will determine the body of precedents in the next period.

Consider a body of precedents for date $t$, $\mathcal{J}^t = (\tau^t, \omega^t)$. Let $d^t = \omega^t - \tau^t$ be the probability that the $t$-th Court has discretion, so that the $t$-th Court is constrained by precedents with probability $1 - d^t$. To streamline the analysis, we assume that if the $t$-th Court is constrained by precedents then the body of precedents simply does not change between period $t$ and period $t + 1$ so that $\mathcal{J}^{t+1} = \mathcal{J}^t$.

When a Court is *not* constrained by precedents (with probability $d^t$), it can choose the tough or weak decision at its discretion.

[22] As we mentioned above, the assumption that precedents either do or do not apply, without intermediate possibilities is obviously an extreme one. It seems a plausible first cut in the modeling of stare decisis and the role of precedents that we wish to pursue here.

[23] It should be emphasized that, despite their correct expectations, we assume that our parties always go to Court. The Court then rules, and thus affects the body of precedents. This is an unappealing assumption. We nevertheless proceed in this way as virtually all the extant literature does. The question of why, in equilibrium (and therefore with "correct" expectations), contracting parties go to Court is a key question that is ripe for rigorous scrutiny. However, it clearly remains well beyond the scope of this paper.

A key feature of our model is that a Court that exercises discretion can also choose the *breadth* of its ruling. For simplicity, we take this to be a binary decision $b^t \in \{0,1\}$, with $b^t = 0$ interpreted as a narrow ruling, and $b^t = 1$ as a broad one. Broad rulings have more impact on the body of precedents than narrow ones do. We return to this critical point at length in Subsection 4.1 below.

The discretionary ruling $\mathcal{R}^t \in \{\mathcal{T}, \mathcal{W}\}$ of the $t$-th Court, together with the breadth of its ruling determine how the body of precedents $\mathcal{J}^t$ is modified to yield $\mathcal{J}^{t+1}$, on the basis of which the $t+1$-th Court will operate. Therefore, the precedents technology can be viewed as a map $\boldsymbol{\mathcal{J}} : [0,1]^2 \times \{\mathcal{T}, \mathcal{W}\} \times \{0,1\} \to [0,1]^2$, so that $\mathcal{J}^{t+1} = \boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, b^t)$.[24] We will later use the notation $\boldsymbol{\omega}(\mathcal{J}^t, \mathcal{R}^t, b^t)$ and $\boldsymbol{\tau}(\mathcal{J}^t, \mathcal{R}^t, b^t)$ to denote the first and second element of $\boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, b^t)$.

Typically, the map $\boldsymbol{\mathcal{J}}$ will embody the workings of a complex set of legal mechanisms and constitutional arrangements, which may entail complex interaction effects among its arguments. Some of our results hold under surprisingly general conditions on the precedents technology, while a more stringent characterization of the equilibrium behavior of our model requires more hypotheses. We return to these at length below.

### *3.5. Dynamic Equilibrium*

We assume that all Courts are forward looking in the sense that they assign weight $1-\delta$ to the current payoff, weight $(1-\delta)\,\delta$ to the per-period Court payoff in the next period, weight $(1-\delta)\,\delta^2$ to the per-period Court payoff in the period after, and so on.[25] Critically, when the current Court takes into account the payoffs of future Courts it does so using the *ex-ante* payoffs satisfying (1) above.

The $t$-th Court inherits $\mathcal{J}^t$ from the past. Given $\mathcal{J}^t$, it observes the outcome of the draw that determines the legal characteristics of the case ($\ell$). Together with $\mathcal{J}^t$, this determines whether the $t$-th period Court has discretion or not. If it has discretion, the $t$-th Court then

[24] Note that a $\mathcal{J}^t$ with $\omega^t < \tau^t$ is just *not feasible.* Hence a typical $\mathcal{J}^t$ is not an element of $[0,1]^2$ but of the subset of the unite square on or above the $45^o$ line. Throughout the paper we abuse notation referring to the space in which a typical $\mathcal{J}^t$ lives as $[0,1]^2$ with out further qualification.

[25] We interpret $\delta$ as the common discount factor shared by the Court and the parties. Notice, however, that $\delta$ could also be interpreted as the probability that the same type of case will occur again in the next period. This probability would then be taken to be independent across periods. Clearly in this case $\delta$ should be part of the legal characteristics of the case. This reinterpretation would yield no changes to the role that $\delta$ plays in the equilibrium characterization (see Sections 4 and 5 below).

chooses $\mathcal{R}^t$ and $b^t$ — the ruling and its breadth. Together with $\mathcal{J}^t$ this determines $\mathcal{J}^{t+1}$, and hence the decision problem faced by the $t+1$-th Court. If the Court does not have discretion then the precedents fully determine the Court's decision, and $\mathcal{J}^{t+1} = \mathcal{J}^t$.

Some new notation is necessary at this point to describe the strategy of the Courts when they are not constrained by precedents. The choice of ruling $\mathcal{R}^t$ depends on $\mathcal{J}^t$. We let $\mathcal{R}^t = \boldsymbol{\mathcal{R}}(\mathcal{J}^t)$ denote this part of the Court's strategy. Similarly, we let the Court's (contingent) choice of breadth be denoted by $b^t = \boldsymbol{b}(\mathcal{J}^t)$. Notice that, in principle, the choices of the $t$-th Court could depend on the entire history of past rulings, breadths, legal characteristics (including the ones at time $t$) and parties' behavior. We restrict attention to behavior that depends only on the body of precedents $\mathcal{J}^t$. These are clearly the only *payoff relevant* state variables for the $t$-th Court. In this sense our restriction is equivalent to saying that we are restricting attention to the set of *Markov-Perfect Equilibria*.[26] We will do so throughout the rest of the paper.

With this restriction, we can simply refer to *the* strategy of the Court, regardless of the time period $t$. This will sometimes be written concisely as $\boldsymbol{\sigma} = (\boldsymbol{\mathcal{R}}, \boldsymbol{b})$. Given $\mathcal{J}^t$ and $\boldsymbol{\sigma}$, using our new notation and the one in (1) and (2), the expected (as of the beginning of period $t$) payoff accruing in period $t$ to the $t$-th Court, can be written as follows.

$$\boldsymbol{\Pi}^{\mathcal{A}}(\mathcal{J}^t, \boldsymbol{\sigma}) = \tau^t \, \Pi^{\mathcal{A}}(\mathcal{T}) + (1 - \omega^t)\Pi^{\mathcal{A}}(\mathcal{W}) + d^t \, \Pi^{\mathcal{A}}(\boldsymbol{\mathcal{R}}(\mathcal{J}^t)) \tag{3}$$

The interpretation of (3) is straightforward. The first two terms refer to the cases in which the Court is constrained (to a tough and weak decision respectively). The third term is the Court's payoff given its discretionary ruling $\boldsymbol{\mathcal{R}}(\mathcal{J}^t)$.

Given the (stationary) preferences we have postulated, the overall payoffs to each Court can be expressed in a familiar recursive form. Let a $\boldsymbol{\sigma}$ be given. Let $\mathbf{Z}(\mathcal{J}^t, \boldsymbol{\sigma})$ be the expected (as of the beginning of the period) overall payoff to the $t$-th Court, given $\mathcal{J}^t$ and $\boldsymbol{\sigma}$.[27] We can then write this payoff as follows.

$$\mathbf{Z}(\mathcal{J}^t, \boldsymbol{\sigma}) = (1-\delta)\boldsymbol{\Pi}^{\mathcal{A}}(\mathcal{J}^t, \sigma^t) + \delta[(1-d^t)\mathbf{Z}(\mathcal{J}^t, \boldsymbol{\sigma}) + d^t\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}^t, \boldsymbol{\mathcal{R}}(\mathcal{J}^t), \boldsymbol{b}(\mathcal{J}^t)), \boldsymbol{\sigma})] \tag{4}$$

---

[26] See Maskin and Tirole (2001) or Ch. 13 in Fudenberg and Tirole (1991).

[27] The function $\mathbf{Z}(\cdot, \cdot)$ is independent of $t$ because we are restricting attention to stationary Markov-Perfect Equilibria.

The interpretation of (4) is also straightforward. The first term on the right-hand side is the Court's period-$t$ payoff. The first term that multiplies $\delta$ is the Court's continuation payoff if its ruling turns out to be constrained by precedents so that $\mathcal{J}^{t+1} = \mathcal{J}^t$. The second term that multiplies $\delta$ is the Court's continuation payoff if the Court's choices at $t$ are $[\boldsymbol{\mathcal{R}}(\mathcal{J}^t), \boldsymbol{b}(\mathcal{J}^t)]$.

Now recall that the $t$-th Court decides wether to take a tough or a weak decision (if it is given discretion) and chooses the breadth of its ruling ex-post, after the parties' actions are *sunk*. Hence the $t$-th Court *continuation* payoffs viewed from the time it is called upon to rule will have two components. One that embodies the period-$t$ payoff, which will be made up of ex-post payoffs as in (2) reflecting the Court's present-bias. And one that embodies the Court's payoffs from period $t+1$ onwards, which on the other hand will be made up of *ex-ante* payoffs as in (4) since all the relevant decisions lie ahead of when the $t$-th Court makes its choices.

It follows that, given $\mathcal{J}^t$ and $\boldsymbol{\sigma}$, the decisions of the $t$-th Court can be characterized as follows. Suppose that the $t$-th Court is *not* constrained by precedents to either a tough or a weak decision.[28] Then, the values of $\mathcal{R}^t = \boldsymbol{\mathcal{R}}(\mathcal{J}^t) \in \{\mathcal{T}, \mathcal{W}\}$ and $b^t = \boldsymbol{b}(\mathcal{J}^t) \in \{0,1\}$ must solve

$$\max_{\mathcal{R}^t \in \{\mathcal{T},\mathcal{W}\}, b^t \in \{0,1\}} (1-\delta)\, \Pi^{\mathcal{P}}(\mathcal{R}^t) \;+\; \delta \left\{ \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, b^t), \boldsymbol{\sigma}) \right\} \tag{5}$$

It is now straightforward to define what constitutes an equilibrium.

**Definition 1.** *<u>Case Law Equilibrium Behavior</u>: An equilibrium is a $\boldsymbol{\sigma}^* = [\boldsymbol{\mathcal{R}}^*, \boldsymbol{b}^*]$ such that, for every $t = 0, 1, 2\ldots$ and for every possible $\mathcal{J}^t$, the pair $[\boldsymbol{\mathcal{R}}^*(\mathcal{J}^t), \boldsymbol{b}^*(\mathcal{J}^t)]$ is a solution to the following problem.*[29]

$$\max_{\mathcal{R}^t \in \{\mathcal{T},\mathcal{W}\}, b^t \in \{0,1\}} (1-\delta)\, \Pi^{\mathcal{P}}(\mathcal{R}^t) \;+\; \delta \left\{ \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, b^t), \boldsymbol{\sigma}^*) \right\} \tag{6}$$

For any given equilibrium behavior as in Definition 1 we can compute the value of the expected payoff to the Court of period $t = 0$, as a function of the initial value $\mathcal{J}^0$. Using the

[28] Recall that if the ruling turns out to be constrained by precedents, the $t$-th Court does not make any choice and the body of precedents remains the same so that $\mathcal{J}^{t+1} = \mathcal{J}^t$.

[29] It should be noted that in equilibrium the decision of the $t$-th Case Law Court is required to be optimal given *every possible* $\mathcal{J}^t$, and not just those that have positive probability given $\boldsymbol{\sigma}^*$ and $\mathcal{J}^0$. This is a standard "perfection" requirement.

notation we already established, this is denoted by $\mathbf{Z}(\mathcal{J}^0, \boldsymbol{\sigma}^*)$.

## 4. The Evolution of Case Law

### *4.1. Residual Discretion and Zero Breadth*

We are able to derive our first two results imposing a surprisingly weak structure on $\boldsymbol{\mathcal{J}}$. This is embodied in the following two assumptions.

**Assumption 1.** *Residual Discretion: Assume that $\mathcal{J}^t$ is such that $d^t > 0$. Then $\mathcal{J}^{t+1} = \boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, b^t)$ is such that $d^{t+1} > 0$ whatever the values of $b^t$ and $\mathcal{R}^t$.*

Assumption 1 simply asserts that the influence of precedents is never able to take discretion completely away from future Courts. This seems a compelling element of the very essence of Case Law.

Our next assumption concerns the effect of the Court's choice of breadth of its ruling.

**Assumption 2.** *Zero Breadth: For any ruling $\mathcal{R}^t$, we have that $\mathcal{J}^t = \boldsymbol{\mathcal{J}}(\mathcal{J}^t, \mathcal{R}^t, 0)$ (so that in this case $\mathcal{J}^{t+1} = \mathcal{J}^t$).*

Assumption 2 states that, regardless of the ruling it issues, any Court can ensure (setting $b^t = 0$) that its ruling is sufficiently narrow so as to have *no effect* on future Courts. This clearly merits some further comments.

First of all, Assumption 2 greatly simplifies the technical side of our analysis. In particular it implies certain monotonicity properties of the dynamics that are used in our arguments below, including our characterization of equilibrium. However, it should also be noted that the basic trade-off between the present-bias temptation and the effect of precedents on future Courts does not depend on the availability of zero breadth rulings.

Finally, the possibility that a Court might decide to narrow down on purpose the effect that its ruling has, through precedents, on future cases does correspond to reality. For instance, in the US, a commonly used formula is for a Court to declare that they wish to "restrict the holding to the facts of the case." In some other instances the Court may choose not to publish the opinion in an official Reporter. Unpublished opinions are collected by various services and so are available to lawyers. However, the decision not to publish in an official Reporter, is regarded by future Courts as a signal that the Court does not want its decision to have value as a precedent.[30]

### *4.2. Mature Case Law*

Given $\boldsymbol{\sigma}^*$ and an initial body of precedents $\mathcal{J}^0$, as the randomness in each period is realized (whether the precedents bind or not, and how) a sequence of Court rulings will also be realized.

We first show that the realized number of times that the Courts have discretion ("discretionary Courts") and will take the tough decision ($\mathcal{T}$) has an *upper bound.* Case Law eventually *matures*, and, after it does, all discretionary Courts succumb to the (time-inconsistent) temptation to rule $\mathcal{W}$ instead of $\mathcal{T}$.

**Proposition 1.** *Evolving and Mature Case Law: Let any equilibrium $\boldsymbol{\sigma}^*$ be given.*

*Then, there exists an integer $q$, which depends on $\delta$ but not on $\mathcal{J}^0$ or on the particular equilibrium $\boldsymbol{\sigma}^*$, with the following property.*

*Along any realized path of uncertainty, the number of times that the Courts have discretion and rule $\mathcal{T}$ does not exceed $q$.*

Intuitively, each time a Court rules $\mathcal{T}$, it must be that the future gains from constraining future Courts via precedents exceed the instantaneous gain the Court can get ex-post giving in to the temptation to rule $\mathcal{W}$. While this gain remains constant through time, the effect on future Courts must eventually become small. This is a consequence of Assumptions 1 and 2. In particular, because the Courts can always choose to select a breadth of zero for their ruling, it is not hard to see that along any realized path of uncertainty the (long-run) equilibrium payoff of the Courts cannot decrease through time (see Lemma A.1 in the Appendix). The

[30] We are indebted to Alan Schwartz for useful guidance on these points. A particularly stark example of a formula that tries to limit (for a variety of possible reasons) the effect that the Court's decision will have on future cases can be found in the ruling of the US Supreme Court in the Bush v. Gore case: "[...] *Our consideration is limited to the present circumstances, for the problem of equal protection in election processes generally presents many complexities."* (Bush v. Gore (00-949). US Supreme Court Per Curiam).

It should be noticed that we also rule out the more extreme possibility (which could be though of as "negative breadth") that Courts are able to increase the probability that future Courts will be constrained to take a given decision while applying a different one to the current case. On this point, the US Supreme Court recently prohibited federal courts from applying a decision only prospectively (see Harper v. Virginia Dept. of Taxation, 91-794, 509 U.S. 86, 1993) As Justice Scalia, concurring in that judgment, notes: *"Prospective decision-making is* [...] *the born enemy of stare decisis."*

future gain from taking the tough decision $\mathcal{T}$ today consists in raising the probability that a future Court will be forced by precedents to rule $\mathcal{T}$. In other words, future gains stem from the upwards effect on some future $\tau^t$ (the probability that the Court is constrained to choose $\mathcal{T}$ at some future date $t$) that a tough decision today may have. It is then apparent that, since $\tau^t$ cannot exceed 1,[31] eventually we must have "decreasing returns" in the future gains stemming from a tough decision today. Eventually, Case Law becomes mature in the sense that, in the eyes of today's Court, future Courts are already sufficiently constrained to rule $\mathcal{T}$ so that any future gains from choosing $\mathcal{T}$ today are washed out by the current temptation to choose the weak decision $\mathcal{W}$.

In general, in our model, Case Law undergoes two phases: a transition, which lasts a finite number of periods, and a mature (or steady) state. Along the transition, precedents evolve and become more binding following a (finite) sequence of tough decisions (with positive breadth) taken by discretionary Courts. In the steady state, only the Courts that are bound by precedents to choose $\mathcal{T}$ take the efficient decision. The ones that are unconstrained (recall that by Assumption 1 precedents cannot end up being completely binding) take instead the weak decision with zero breadth in order to keep the body of precedents intact.

Note that stare decisis plays the key disciplinary role in our model. In the absence of the rule of precedent, it is straightforward to show that Courts always take the present-bias decision whenever they have discretion so that precedents do not evolve.

**Proposition 2.** *Absence of Stare Decisis: If $\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{R}, 1)$ does not depend on $\mathcal{R}$ (i.e., if future precedents do not depend on the current ruling even when $b = 1$), in any equilibrium the number of times Case Law courts have discretion and rule $\mathcal{T}$ is exactly zero.*

Intuitively, if legal rules are not affected by judicial rulings, in our model all Courts succumb to the time inconsistent temptation to be lenient towards the parties presently before them. Their decisions have no effect on the future and hence the welfare of the current parties entirely dictates the behavior of Courts.

[31] From Assumption 1, unless the very first Court has no discretion at all, it is clear that $\tau^t$ will in fact always be strictly below 1.

## 5. Equilibrium Characterization

### 5.1. Mixed Strategies

The characterization of the equilibrium behavior is somewhat intricate. To appreciate some of the difficulties involved, recall that from Proposition 1 we know that along any path of resolved uncertainty the Courts can only take the tough decision $\mathcal{T}$ a finite number of times $q$.

Suppose now that we are in a configuration of parameters (a $\delta$ not "too low" is, for instance, necessary) such that in equilibrium the Courts initially rule $\mathcal{T}$ with $b = 1$ to constrain future Courts to do the same with higher probability. Now consider "the last" Court to rule $\mathcal{T}$ with $b = 1$ exercising its discretion to do so. In other words, suppose that the (Markov perfect) equilibrium prescribes that some Court that has discretion rules $\mathcal{T}$ with breadth 1, knowing that from that point on all future Courts will rule $\mathcal{W}$ (with $b = 0$) and they are not bound by precedents. In other words, suppose that the equilibrium involves a state of precedents that generates the "last tough Court," with all subsequent Courts succumbing to the time-inconsistency problem.

This "natural conjecture" as to how a typical Markov perfect equilibrium might play out is in fact contradictory in some cases. To see this, consider the possibility that the last tough Court, say that this occurs at time $t$, deviates and takes instead the weak decision $\mathcal{W}$, but with breadth 0 so that its decision has *no effect* on the future. If it does so, the next Court that has discretion will face the same body of precedents, and (by stationarity) *it* will be the last tough Court. To make the argument more straightforward, suppose that $\omega^t = 1$ so that the current precedent does not constrain at all the current Court to choose $\mathcal{W}$. Note that the $t$-th period Court gains in two distinct ways from the deviation. First of all, it has an instantaneous gain at time $t$ since it rules ex-post and $\Pi^{\mathcal{P}}(\mathcal{W}) > \Pi^{\mathcal{P}}(\mathcal{T})$. Second, it puts one of the future Courts (the first one to have discretion) in the position of being the *last* tough Court, and hence to rule $\mathcal{T}$ while without the deviation the ruling would have been $\mathcal{W}$. Since the $t$-th Court evaluates these payoff from an ex-ante point of view, this is also a gain because $\Pi^{\mathcal{A}}(\mathcal{T}) > \Pi^{\mathcal{A}}(\mathcal{W})$.[32]

[32]If instead our expository assumption that $\omega^t = 1$ does not hold (and $\omega^t$ is appropriately "low") and the precedents technology is such that a $\mathcal{T}$ decision with breadth 1 decreases the probability that future Courts are constrained to choose $\mathcal{W}$ the deviation we are describing may not be profitable. In this case, besides the

The solution to the puzzle we have just outlined is that a typical Markov Perfect Equilibrium of our model may require *mixed strategies*. Before Case Law matures, Courts randomize between the $\mathcal{T}$ decision (with positive breadth) and the $\mathcal{W}$ decision (with zero breadth). This in turn allows Case Law to begin with tough discretionary decisions with $b = 1$ without violating Proposition 1 (Case Law eventually must mature) and without running into the difficulty we have outlined. No Court is certain to be the last to have discretion and take a $\mathcal{T}$ decision. The mixing probabilities used before Case Law matures depend on many details of the equilibrium. However, it is not too hard to see that that each Court that acts before Case Law matures can be kept indifferent between the two decisions by an appropriate choice of the mixing probabilities employed by future Courts.

Before moving on, we remark on the juxtaposition of the mixed strategy equilibria we find here with the rationale for stare decisis based on the predictability and stability of Court decisions commonly found in the literature.[33] Since it generates mixed strategy equilibria, at face value, the rationale for stare decisis that we pursue here — an instrument available for each Court to alleviate the time inconsistency of future Courts — *decreases* the predictability of Court behavior. While this is an interesting observation, we believe it should be treated with some caution to avoid misleading interpretations.

First of all, everyone in our model is rational in the standard all-encompassing sense of modern Game Theory, with access to unlimited computational resources. Presumably one of the underlying reasons for valuing stare decisis as something that affords predictability of Court behavior is a belief that not all participants are infinitely adept at forecasting complex behavior. Second, and very much following from our last point, in a mixed strategy equilibrium agents *do* predict correctly the behavior of others in the sense that they can exactly *compute the probabilities* (the mixed strategies) that guide everyone else's behavior.

### *5.2. Well-Behaved Precedents*

Our task in this Section is to characterize a Markov perfect equilibrium of our model. Given the delicate nature of the construction that stems from our considerations above, we proceed

current gains described above, procrastination may have a cost since future Courts may be more likely overall to choose the inefficient ruling as a result of the deviation. When this implies that the deviation described above is not profitable overall, a pure strategy equilibrium as in the "natural conjecture" above will in fact exist.

[33]See Subsections 1.1 and 1.3 in the Introduction for a discussion.

to impose further structure on the precedents technology. This keeps the problem tractable, while it still allows us to bring out the main features of the equilibrium behavior of the model.

The regularity conditions on the precedents technology that we work with are summarized next. We comment on each condition immediately after their statement. An example of precedent technology that satisfies Assumptions 1 to 3 will be discussed at length in Subsection 5.4 below.

**Assumption 3.** *Well-Behaved Precedents Technology: The map $\boldsymbol{\mathcal{J}}$ satisfies:*

(i) Continuity and Monotonicity: *For any ruling $\mathcal{R}$, $\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{R}, 1)$ is continuous in $\mathcal{J}$. Moreover, $\boldsymbol{\tau}(\tau, \omega, \mathcal{T}, 1) > \tau$ and $\boldsymbol{\omega}(\tau, \omega, \mathcal{T}, 1) \geq \omega$ whenever $\tau \in (0, 1)$.*[34] *Finally, $\boldsymbol{\tau}(\tau, \omega, \mathcal{W}, 1) \leq \tau$ and $\boldsymbol{\omega}(\tau, \omega, \mathcal{W}, 1) < \omega$ whenever $\omega \in (0, 1)$.*[35]

(ii) Everywhere Decreasing Returns from $\mathcal{T}$ Decisions: *The function $\boldsymbol{\tau}(\tau, \omega, \mathcal{T}, 1)$ is concave in $\tau$, for every $\tau \in [0, 1]$.*

(iii) Independent Impact of $\mathcal{T}$ Decisions: *Consider $\mathcal{J} = (\tau, \omega)$ and $\widetilde{\mathcal{J}} = (\tau, \widetilde{\omega})$ with $\omega \neq \widetilde{\omega}$. Then $\boldsymbol{\tau}(\mathcal{J}, \mathcal{T}, 1) = \boldsymbol{\tau}(\widetilde{\mathcal{J}}, \mathcal{T}, 1)$.*

(iv) Reversibility: *For every $\mathcal{J} \in [0, 1]^2$ we have that $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{T}, 1), \mathcal{W}, 1) = \mathcal{J}$, and symmetrically $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{W}, 1), \mathcal{T}, 1) = \mathcal{J}$.*

Besides continuity (which simplifies the analysis in the obvious way), (i) of Assumption 3 rules out "perverse" shapes of the precedents technology. For instance, it rules out that a tough decision with breadth equal to one might lead to a decrease in the probability that future Courts will be constrained to take a tough decision in the same environment.

As we discussed above, in a neighborhood of $\tau = 1$ the mapping $\boldsymbol{\tau}$ necessarily satisfies a form of decreasing returns in $\tau$. Condition (ii) extends this property to the whole interval $[0, 1]$, which turns out to be analytically very convenient.

Condition (iii) guarantees that the effect of a tough decision on the probability that future Courts will be constrained to take a tough decision does not depend on the probability that the current Court is constrained to take a weak decision instead.

Finally, condition (iv) is technically extremely convenient. It guarantees that opposite consecutive decisions by Courts (with positive breadth) cancel each other out. In effect, this

[34] If $\tau = 1$, and hence $\omega = 1$, then $\boldsymbol{\tau}(\tau, \omega, \mathcal{T}, 1) = \boldsymbol{\omega}(\tau, \omega, \mathcal{T}, 1) = 1$.
[35] If $\omega = 0$, and hence $\tau = 0$, then $\boldsymbol{\tau}(\tau, \omega, \mathcal{W}, 1) = \boldsymbol{\omega}(\tau, \omega, \mathcal{W}, 1) = 0$.

allows us to narrow down dramatically the cardinality of the set of possible pairs $(\tau, \omega)$ that need to be considered overall along any particular equilibrium path.[36]

### *5.3. A Markov Perfect Equilibrium*

Using Assumption 3 as well as Assumptions 1 and 2 we can now proceed with a detailed characterization of the equilibrium behavior of our model.

Some extra pieces of notation will prove useful. Because of (iii) of Assumption 3 we know that $\boldsymbol{\tau}(\mathcal{J}, \mathcal{T}, 1)$ with $\mathcal{J} = (\tau, \omega)$ does not in fact depend on $\omega$, but only on $\tau$ itself. We then let $\boldsymbol{\kappa}(\tau) = \boldsymbol{\tau}(\mathcal{J}, \mathcal{T}, 1) - \tau$, so that $\boldsymbol{\kappa}(\tau)$ is the *increment* in $\tau$, stemming from a tough decision today with $b = 1$.

Suppose that there exist some value(s) of $\tau \in [0, 1]$ such that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] < \delta\, \boldsymbol{\kappa}(\tau)\, [\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \tag{7}$$

Notice that from (i) of Assumption 3, we know that $\boldsymbol{\kappa}(\tau)$ is non-negative, and equal to zero if $\tau = 1$. It can also be shown that (iv) of Assumption 3 implies that $\boldsymbol{\kappa}(\tau) = 0$ if $\tau$ is equal to zero. By (i) and (ii) of Assumption 3 it is then immediate that if (7) holds, there exist precisely two values of $\tau \in (0, 1)$ such that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] = \delta\, \boldsymbol{\kappa}(\tau^*)\, [\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \tag{8}$$

From now on, we denote the lower of these two values by $\tau_*$ and the higher one by $\tau^*$, if they exist. Otherwise, we leave them to be undefined.

**Proposition 3.** *A Markov Perfect Equilibrium: Suppose that Assumptions 1, 2 and 3 hold. Suppose also that there exist values of $\tau \in (0, 1)$ such that (7) holds, so that $\tau_*$ and $\tau^*$ are well defined.*

*Then there is an equilibrium (see Definition 1) of our model $\boldsymbol{\sigma}^* = (\boldsymbol{\mathcal{R}}^*, \boldsymbol{b}^*)$ with the following properties.*

(i) *There exists a threshold value $\underline{\tau} \in (0, \tau_*]$ as follows. Whenever $\mathcal{J} = (\tau, \omega)$ is such that either $\tau \leq \underline{\tau}$ or $\tau \geq \tau^*$, then $\boldsymbol{\mathcal{R}}^* = \mathcal{W}$ and $\boldsymbol{b}^*(\mathcal{J}) = 0$. In other words, given $\omega$, if $\tau \in [0, \underline{\tau}] \cup [\tau^*, 1]$, then, if it has discretion, the ruling of the Court is $\mathcal{W}$ with breadth 0.*

[36]See our discussion of the bound $\underline{\tau}$ of Proposition 3 at the end of Subsection 5.3 below.

(ii) *If instead* $\mathcal{J} = (\tau, \omega)$ *is such that* $\tau_* < \tau < \tau^*$, *if it has discretion the Court randomizes between a ruling of* $\mathcal{T}$ *with breadth* 1 *and a ruling of* $\mathcal{W}$ *with breadth* 0. *We denote the probability of a* $\mathcal{T}$ *ruling with breadth* 1 *by* $p(\tau, \omega) \in (0, 1]$, *so that the probability of a* $\mathcal{W}$ *ruling with breadth* 0 *is* $1 - p(\tau, \omega) \in [0, 1)$.

The equilibrium behavior captured by Proposition 3 is rich because of the temptation of time-inconsistent behavior. This can be seen focusing on the case in which the initial $\mathcal{J}^0$ has $\tau_* < \tau^0 < \tau^*$. In this case, the initial body of precedents and the other parameters of the model are such that the instantaneous gain from taking the $\mathcal{W}$ decision (appropriately weighted by $1 - \delta$) is smaller than the future gains (appropriately weighted by $\delta$) from the increase in $\tau$ stemming from a $\mathcal{T}$ decision with $b = 1$ — inequality (7) holds.

However, when inequality (7) holds, for the reasons we described in Subsection 5.1 above, a *pure strategy* equilibrium in which a finite sequence of $\mathcal{T}$ decisions with $b = 1$ are taken may not be viable. The equilibrium then involves the Courts who have discretion mixing between a $\mathcal{T}$ ruling with $b = 1$ and a $\mathcal{W}$ ruling with $b = 0$. Each Court which randomizes in this way is kept indifferent between the two choices by the randomization with appropriate probabilities of future Courts.

While the randomizations take place, the value of $\tau$ increases stochastically through time, as the tough ruling with breadth 1 is chosen. Eventually, this process puts the value of $\tau$ over the threshold $\tau^*$. At this point Case Law is mature. All Courts from this point on, if they have discretion, issue ruling $\mathcal{W}$ with breadth 0.

We conclude noting that the region ($\tau^t \leq \underline{\tau}$) near zero in which the equilibrium prescribes that Case Law will not evolve is essentially an artifact of the regularity conditions (namely (iv) of Assumption 3) we have imposed on the precedents technology. More complex equilibria that do not display this region can be constructed when these conditions are dropped. The amount of technicalities involved makes the material intricate and we omit a treatment for reasons of space. Finally, it is not hard to show that the region near zero of non-evolving Case Law can be made arbitrarily small as discounting decreases — $\underline{\tau}$ approaches zero as $\delta$ approaches one.

### *5.4. An Example*

The class of equilibria characterized in Proposition 3 is probably best understood via an example in which the equilibrium behavior of the Courts is computed explicitly. This is our next task.

To proceed we need to specify an actual precedents technology that complies with requirements of Assumption 3. Consider then the following specification for $\mathcal{J}$.

Since precedents do not change when the Court does not have discretion, nor when it chooses $b^t = 0$, we only need to specify what happens to the thresholds $\tau$ and $\omega$ when the Court chooses $b^t = 1$ and chooses $\mathcal{T}$ and $\mathcal{W}$ respectively. These are as follows. Fix a value of $\alpha \in (0,1)$, then let

$$\tau^{t+1} = \begin{cases} (\tau^t)^\alpha & \text{if} \quad \mathcal{R}^t = \mathcal{T} \\ (\tau^t)^{\frac{1}{\alpha}} & \text{if} \quad \mathcal{R}^t = \mathcal{W} \end{cases} \tag{9}$$

and

$$\omega^{t+1} = \begin{cases} (\omega^t)^\alpha & \text{if} \quad \mathcal{R}^t = \mathcal{T} \\ (\omega^t)^{\frac{1}{\alpha}} & \text{if} \quad \mathcal{R}^t = \mathcal{W} \end{cases} \tag{10}$$

It is a matter of straightforward algebra to check that the both Assumption 1 and Assumption 3 are met by the precedents technology specified in (9) and (10); we omit the details.

To continue with the characterization of explicit values for an equilibrium, we now fix the following levels of all the quantities involved in the computation.

$$\alpha = \frac{1}{2}, \quad \delta = 4/5, \quad \Pi^{\mathcal{A}}(\mathcal{T}) = 15, \quad \Pi^{\mathcal{A}}(\mathcal{W}) = 5, \quad \Pi^{\mathcal{P}}(\mathcal{T}) = 15, \quad \Pi^{\mathcal{P}}(\mathcal{W}) = 20 \tag{11}$$

The initial values of the thresholds defining $\mathcal{J}^0$ that we will use are

$$\tau^0 = \left(\frac{3}{4}\right)^4, \qquad \omega^0 = 1 \tag{12}$$

We note immediately that setting $\omega^0 = 1$ as in (12) together with (10) implies that $\omega^t = 1$ for every value of $t$, which simplifies our computations considerably.

We then use equation (8) to see that the two thresholds $\tau_*$ and $\tau^*$ are well defined and given by

$$\tau^* = \frac{3+2\sqrt{2}}{8} \qquad \text{and} \qquad \tau_* = \frac{3-2\sqrt{2}}{8} \tag{13}$$

Using the values in (11), (12) and (13) we can now proceed to compute the specifics of the equilibrium in Proposition 3 in this case. Figure 1 below provides a graphical representation designed to aid the exposition.

Since $\tau_* < \tau^0 < \tau^*$, to begin with, when it has discretion, the Court will randomize between a ruling of $\mathcal{T}$ with breadth 1 and a ruling of $\mathcal{W}$ with breadth 0. The numerical values we have picked imply that Case Law will mature in two steps in this case. That is after the Court has had discretion and the outcome of its randomization has been to rule $\mathcal{T}$ with breadth 1 twice, the equilibrium prescribes that no further randomization on the part of the Court will occur. The court will always take the weak decision with breadth zero when it has discretion.

In this steady state, the value (denoted $\tau''$) of $\tau$ is 3/4 as is evident from the value of $\tau^0$ in (12), the fact that two tough decisions with breadth 1 have been taken, and finally that $\alpha$ = 1/2 in (9). Similarly, the intermediate value of $\tau$ after one tough decision with breadth 1 has been taken can be seen to be $\tau' = (3/4)^2$.

The probabilities with which the Court randomizes while $\tau$ is moving from its initial value $\tau^0$ to $\tau'$ and then from $\tau'$ to $\tau''$ are computed so as to keep the Court appropriately indifferent between the tough and the weak decision with breadth 1 and 0 respectively. The first value (probability of tough decision) is $p(\tau^0, 1) \simeq 0.18$ while the second is $p(\tau', 1) \simeq 0.10$.[37]

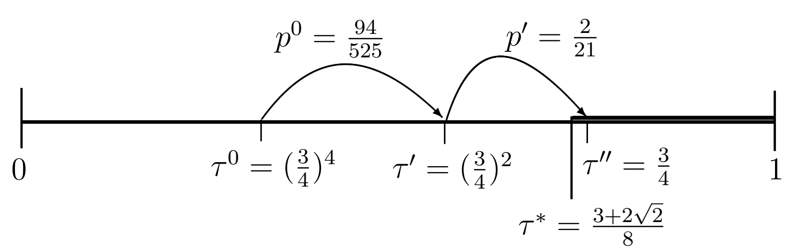


Figure 1: Dynamics of Precedents.

To see how these probabilities are computed, note that we can proceed backwards from the steady state.

[37]The exact values are $p(\tau^0, 1) = \frac{94}{525}$ and $p(\tau', 1) = \frac{2}{21}$.

Plugging in the correct numerical values, using (A.15) we can compute the value of the ex-ante expected payoff to the Court in the steady state. The ex-ante expected payoff to the Court when $\tau$ is equal to $\tau'$ can then be written with the randomization probability as a free variable as in equation (A.16). Finally, the value computed using (A.16) as a function of the probability can be plugged into equation (A.17) that stipulates that the Court must be indifferent between the tough and the weak decision with breadth 1 and 0 respectively. Solving (A.17) for $p(\tau', 1)$ now gives its value as above.

To compute the value of $p(\tau^0, 1)$ we can proceed backwards one more step in essentially the same way.

## 6. Conclusions

Courts intervene in economic relationships at the *ex-post* stage (if at all). Because of sunk strategic decisions this might generate a time-inconsistency — a present-bias in the decisions of Courts that exercise discretion. We argue that this is one of the roots of the principle of stare decisis that disciplines Case Law.

As well as in many others, in some situations of first order economic magnitude such as debt restructuring and patent infringement cases, the optimal ex-ante decision is typically "tougher" than the ex-post optimal decision. Ex-ante, the parties need incentives (for appropriate risk-taking and R&D investment respectively) that are of no use ex-post. Hence, ex-post a more lenient "weaker" decision is optimal when only the parties currently before the Court are considered.

When a benign forward looking Court chooses between the tough and the weak decision it trades off the temptation to favor the parties currently before it, versus the effects of a tough decision today on future Courts via the evolution of precedents and stare decisis.

In our simple framework, without stare decisis the Courts never have an incentive to take the tough decision since they have no effect on the decisions of future Courts — a tough decision today does not increase the probability of a tough decision by future Courts. Hence without stare decisis, in our model precedents do not evolve at all and all Courts succumb to the time inconsistency problem that afflicts them.

The evolution of precedents under stare decisis generates a dynamic process that does not converge to full efficiency. Eventually, the effect of tough decisions via precedents and stare decisis must become small since it is a marginal one. The temptation to take the ex-post optimal decision on the other hand does not shrink through time. Hence, at some point Case Law "matures" in the sense that precedents are already sufficiently likely to constrain future Courts to take the (ex-ante) efficient tough decision. This undoes the incentives to set the "right" precedents whenever the present Court has the chance to do so. Bounded away from full efficiency, Case Law stops evolving and settles into, narrow, lenient decisions whenever precedents do not bind.

Finally, we characterize a class of equilibria under additional assumptions. This indicates that, in a robust set of cases, the Courts use a mixed strategy along the stochastic path that characterizes the dynamics of precedents. We argue that this is worthy of notice when juxtaposed with the common justification for stare decisis based on the predictability of Court behavior, but should not be taken as an argument to invalidate it.

## Appendix

**Lemma A.1**: *Let $\boldsymbol{\sigma}^* = (\boldsymbol{\mathcal{R}}^*, \boldsymbol{b}^*)$ be any equilibrium. Then expected welfare is weakly monotonically increasing in the sense that for any $\mathcal{J} \in [0,1]^2$ we have that*

$$\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), \boldsymbol{b}^*(\mathcal{J}), \boldsymbol{\sigma}^*) \geq \mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma}^*) \tag{A.1}$$

**Proof:** By Definition 1 for every $\mathcal{J} \in [0,1]^2$ the values $\mathcal{R} = \boldsymbol{\mathcal{R}}^*(\mathcal{J})$ and $b = \boldsymbol{b}^*(\mathcal{J})$ must solve

$$\max_{\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}, b \in \{0,1\}} (1-\delta)\, \Pi^{\mathcal{P}}(\mathcal{R}) \,+\, \delta\, \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{R}, b), \boldsymbol{\sigma}^*) \tag{A.2}$$

Suppose now that for some $\mathcal{J}$ inequality (A.1) were violated. Then, using Assumption 2, setting $b = 0$ yields

$$\mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma}^*) = \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), 0), \boldsymbol{\sigma}^*) > \mathbf{Z}^*(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), \boldsymbol{b}^*(\mathcal{J})), \boldsymbol{\sigma}^*) \tag{A.3}$$

and hence

$$(1-\delta)\, \Pi^{\mathcal{P}}(\boldsymbol{\mathcal{R}}^*(\mathcal{J})) \,+\, \delta\, \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), 0), \boldsymbol{\sigma}^*) > (1-\delta)\, \Pi^{\mathcal{P}}(\boldsymbol{\mathcal{R}}^*(\mathcal{J})) \,+\, \delta\, \mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), \boldsymbol{b}^*(\mathcal{J})), \boldsymbol{\sigma}^*) \tag{A.4}$$

which contradicts the fact that $\boldsymbol{b}^*(\mathcal{J})$ and $\boldsymbol{\mathcal{R}}^*(\mathcal{J})$ must solve (A.2). ■

**Lemma A.2**: *Let $\boldsymbol{\sigma}^* = (\boldsymbol{\mathcal{R}}^*, \boldsymbol{b}^*)$ be any equilibrium. Suppose that for some $\mathcal{J} \in [0,1]^2$ we have that*

$$\boldsymbol{\mathcal{R}}^*(\mathcal{J}) = \mathcal{T} \tag{A.5}$$

*then it must be that*

$$\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \boldsymbol{\mathcal{R}}^*(\mathcal{J}), \boldsymbol{b}^*(\mathcal{J})), \boldsymbol{\sigma}^*) - \mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma}^*) \;\geq\; \frac{1-\delta}{\delta}\left[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})\right] \tag{A.6}$$

**Proof:** From (6) of Definition 1, we know that for every $\mathcal{J} \in [0,1]^2$ the values $b = \boldsymbol{b}^*(\mathcal{J})$ and $\mathcal{R} = \boldsymbol{\mathcal{R}}^*(\mathcal{J})$ must solve

$$\max_{\mathcal{R}\in\{\mathcal{T},\mathcal{W}\}, b\in\{0,1\}} (1-\delta)\,\Pi^{\mathcal{P}}(\mathcal{R}) \,+\, \delta\,\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\mathcal{R},b),\boldsymbol{\sigma}^*) \tag{A.7}$$

Since (A.5) must hold it must then be that

$$(1-\delta)\,\Pi^{\mathcal{P}}(\mathcal{T}) \,+\, \delta\,\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\boldsymbol{\mathcal{R}}^*(\mathcal{J}),\boldsymbol{b}^*(\mathcal{J})),\boldsymbol{\sigma}^*) \geq (1-\delta)\,\Pi^{\mathcal{P}}(\boldsymbol{\mathcal{R}}^*(\mathcal{J})) \,+\, \delta\,\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\boldsymbol{\mathcal{R}}^*(\mathcal{J}),0),\boldsymbol{\sigma}^*) \tag{A.8}$$

Using Assumption 2 we know that $\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\boldsymbol{\mathcal{R}}^*(\mathcal{J}),0),\boldsymbol{\sigma}^*) = \mathbf{Z}^*(\mathcal{J},\boldsymbol{\sigma}^*)$. Hence (A.8) directly implies (A.6). ■

**Proof of Proposition 1:** Let $q$ be the smallest integer that satisfies

$$q \,\geq\, \left(\frac{\delta}{1-\delta}\right)\frac{\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})}{\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})} + 1 \tag{A.9}$$

Notice next that $\mathbf{Z}(\mathcal{J},\boldsymbol{\sigma}^*)$ is obviously bounded above by $\Pi^{\mathcal{A}}(\mathcal{T})$ and below by $\Pi^{\mathcal{A}}(\mathcal{W})$.

Suppose now that the proposition were false and therefore that along some realized history $h^t = (\mathcal{J}^0, \ldots, \mathcal{J}^{t-1})$ the Court were given discretion and ruled $\mathcal{T}$ for $q$ or more times. Then using Lemmas A.1 and A.2 we must have that

$$\mathbf{Z}(\mathcal{J}^{t-1},\boldsymbol{\sigma}^*) \;\geq\; q\left(\frac{1-\delta}{\delta}\right)\left[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})\right] + \Pi^{\mathcal{A}}(\mathcal{W}) \tag{A.10}$$

Using (A.9), it is immediate that the right-hand side of (A.10) is greater than $\Pi^{\mathcal{A}}(\mathcal{T})$. Since the latter is an upper bound for $\mathbf{Z}(\mathcal{J},\boldsymbol{\sigma}^*)$, this is a contradiction and hence it is enough to establish the claim. ■

**Proof of Proposition 2:** Let $\boldsymbol{\sigma}^* = (\boldsymbol{\mathcal{R}}^*,\boldsymbol{b}^*)$ be any equilibrium. Suppose that for some $\mathcal{J} \in [0,1]^2$ we have that

$$\boldsymbol{\mathcal{R}}^*(\mathcal{J}) = \mathcal{T} \tag{A.11}$$

then it must be that

$$(1-\delta)\,\Pi^{\mathcal{P}}(\mathcal{T}) \,+\, \delta\,\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\mathcal{T},\boldsymbol{b}^*(\mathcal{J})),\boldsymbol{\sigma}^*) \geq (1-\delta)\,\Pi^{\mathcal{P}}(\mathcal{W}) \,+\, \delta\,\mathbf{Z}(\boldsymbol{\mathcal{J}}(\mathcal{J},\mathcal{W},\boldsymbol{b}^*(\mathcal{J})),\boldsymbol{\sigma}^*) \tag{A.12}$$

When $\boldsymbol{\mathcal{J}}(\mathcal{J},\mathcal{R},b)$ does not depend on $\mathcal{R}$, (A.12) becomes

$$\Pi^{\mathcal{P}}(\mathcal{T}) \;\geq\; \Pi^{\mathcal{P}}(\mathcal{W}) \tag{A.13}$$

which is obviously impossible because of (2). ■

**Proof of Proposition 3:** We proceed in 9 steps to verify that, under the hypotheses of the proposition, we can find a $\boldsymbol{\sigma^*} = (\boldsymbol{\mathcal{R}^*}, \boldsymbol{b^*})$ that satisfies (i) and (ii).

**Step 1**: *There is no profitable deviation from $\boldsymbol{\sigma^*}$ whenever $\tau \geq \tau^*$.*

**Proof:** Recall that $\boldsymbol{\sigma^*}$ prescribes that whenever $\tau \geq \tau^*$ then $\boldsymbol{\mathcal{R}^*}(\mathcal{J}) = \mathcal{W}$ and $\boldsymbol{b^*}(\mathcal{J}) = 0$. Fix any value of $\omega$ together with the given $\tau \geq \tau^*$, and let $\mathcal{J} = (\tau, \omega)$.

Begin by considering a deviation to setting $\mathcal{R} = \mathcal{W}$ and $b = 1$, keeping the continuation equilibrium fixed as given by $\boldsymbol{\sigma^*}$, as set out in the statement of Proposition 3.

We need to distinguish two cases: $\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1) \geq \tau^*$ and $\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1) < \tau^*$. Consider first $\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1) \geq \tau^*$, so that the continuation equilibrium involves a choice of $\mathcal{W}$ with breadth zero in every period, whenever the Court is not constrained by precedents to rule $\mathcal{T}$. The deviation to $\mathcal{R} = \mathcal{W}$, $b = 1$ is not profitable given inequality (1) and (2), Assumption 2 and Conditions (i) and (iii) of Assumption 3.

Next, consider the case in which $\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1) < \tau^*$. Recall that by Assumption 3 (iv) $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{T}, 1), \mathcal{W}, 1) = \mathcal{J}$. Hence it must be that the equilibrium strategy $\boldsymbol{\sigma^*}$ at $\mathcal{J}' = (\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1), \boldsymbol{\omega}(\mathcal{J}, \mathcal{W}, 1))$ prescribes a decision of $\mathcal{T}$ (with breadth one) with positive probability. By (A.6) this immediately implies that $\mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma^*}) > \mathbf{Z}(\mathcal{J}', \boldsymbol{\sigma^*})$ and hence that the hypothesized deviation to $\mathcal{R} = \mathcal{W}$ and $b = 1$ is not profitable in this case (in fact it entails a positive loss).

Next, consider the deviation to setting $\mathcal{R} = \mathcal{T}$ and $b = 1$, again of course keeping the continuation equilibrium fixed as given by $\boldsymbol{\sigma^*}$. By definition of $\tau^*$, whenever $\tau \geq \tau^*$ we must have that

$$(1 - \delta) \left[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})\right] \geq \delta\, \boldsymbol{\kappa}(\tau) \left[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})\right] \tag{A.14}$$

with the equality holding if and only if $\tau = \tau^*$. Since, whenever the Court is not constrained by precedents to rule $\mathcal{T}$, the continuation equilibrium after the deviation to $\mathcal{R} = \mathcal{T}$ and $b = 1$ involves a choice of $\mathcal{W}$ with breadth zero in every period, (A.14) directly implies that the proposed deviation cannot be profitable.

Finally, consider the deviation to setting $\mathcal{R} = \mathcal{T}$ and $b = 0$, as before keeping the continuation equilibrium fixed as given by $\boldsymbol{\sigma^*}$. Since deviating to $\mathcal{R} = \mathcal{T}$ and $b = 0$ does not change the continuation equilibrium path (because of Assumption 2), this deviation cannot be profitable given inequality (2).

**Step 2**: *Consider the sequence of numbers in $[0, \tau*]$ obtained repeatedly applying a $\mathcal{W}$ decision with breadth one starting with $\tau^*$. Formally, let $\tau^* = \hat{\tau}^0$ and then define recursively $\hat{\tau}^n = \boldsymbol{\tau}(\hat{\tau}^{n-1}, \omega, \mathcal{W}, 1)$ for $n = 1, \ldots, \infty$, and note that because of Assumption 3 (iii) this defines a unique sequence $\{\hat{\tau}^n\}_{n=0}^{\infty}$, regardless of the corresponding values of $\omega$.*

*Denote by $\mathcal{I}^n$ each of the half open intervals $[\hat{\tau}^n, \hat{\tau}^{n-1})$, and by convention set $\mathcal{I}^0 = [\tau^*, \boldsymbol{\tau}(\tau^*, \omega, \mathcal{T}, 1))$. Then,*

(i) *The sequence $\{\hat{\tau}^n\}_{n=0}^{\infty}$ is strictly decreasing and $\lim_{n \to \infty} \hat{\tau}^n = 0$.*

(ii) *For every $n = 1, \ldots, \infty$, if $\tau^n \in \mathcal{I}^n$ then $\boldsymbol{\tau}(\tau^n, \omega, \mathcal{W}, 1) \in \mathcal{I}^{n+1}$ and $\boldsymbol{\tau}(\tau^n, \omega, \mathcal{T}, 1) \in \mathcal{I}^{n-1}$*

**Proof:** Claim (i) is a direct consequence of (i) (monotonicity) and (ii) (concavity) of Assumption 3, and of the fact that (as a consequence of Assumption 3) we know that $\boldsymbol{\tau}(0, \omega, \mathcal{W}, 1) = \boldsymbol{\tau}(0, \omega, \mathcal{T}, 1) = 0$. Claim (ii) is a direct consequence of (i) (continuity and monotonicity) of Assumption 3.

**Step 3**: *Let a Markov Perfect Equilibrium $\boldsymbol{\sigma}^*$ as in Proposition 3 be given, and for any given $\tau < \tau^*$ and $\omega > \tau$, let $p(\tau, \omega) \in [0, 1]$ be the probability that, according to $\boldsymbol{\sigma}^*$, the Court rules $\mathcal{T}$ with $b = 1$ and $1 - p(\tau, \omega) \in [0, 1]$ be the probability that the Court rules $\mathcal{W}$ with $b = 0$.*

*Note that we are allowing $p(\tau, \omega) = 0$ since (for the moment) we are not restricting $\tau$ to be strictly greater than the $\underline{\tau}$ of the statement of Proposition 3.*

*Then, the value function $\mathbf{Z}(\tau, \omega, \boldsymbol{\sigma}^*)$ can be computed as follows.*

*Let $m$ be such that the arbitrarily given $\tau$ is in $\mathcal{I}^m$, one of the intervals defined in Step 2. Using (i) and (ii) of Step 2, we can construct a decreasing sequence $\{\tau^n\}_{n=0}^m$ with $\tau^m = \tau$, $\tau^0 \geq \tau^*$, and $\tau^{n+1} = \boldsymbol{\tau}(\tau^n, \omega, \mathcal{W}, 1)$ for every $n = 0, \ldots, m-1$ so that $\tau^n \in \mathcal{I}^n$ for every $n = 0, \ldots, m$. Start also by letting $\omega^m = \omega$, and then construct another sequence $\{\omega^n\}_{n=0}^m$ by setting recursively $\omega^{n-1} = \boldsymbol{\omega}(\omega^n, \omega, \mathcal{T}, 1)$ for every $n = m, \ldots, 1$. Let $\mathcal{J}^n = (\tau^n, \omega^n)$ and $d^n = (\omega^n - \tau^n)$ for every $n = 0, \ldots, m$.*

*Since $\tau^0 \geq \tau^*$ it is immediate that*

$$\mathbf{Z}(\mathcal{J}^0, \boldsymbol{\sigma}^*) = \tau^0\, \Pi^{\mathcal{A}}(\mathcal{T}) + (1 - \tau^0)\, \Pi^{\mathcal{A}}(\mathcal{W}) \tag{A.15}$$

*Proceeding recursively backwards from $\mathcal{J}^0 = (\tau^0, \omega^0)$ (that is, increasing the index $n$), directly from the properties of $\boldsymbol{\sigma}^*$ in Proposition 3, we get that for every $n = 0, \ldots, m-1$*

$$\begin{aligned}\mathbf{Z}(\mathcal{J}^{n+1}, \boldsymbol{\sigma}^*) = \tau^{n+1} &\left[(1-\delta)\, \Pi^{\mathcal{A}}(\mathcal{T}) + \delta\, \mathbf{Z}(\mathcal{J}^{n+1}, \boldsymbol{\sigma}^*)\right] + \\ &\left[(1 - \omega^{n+1}) + d^{n+1}(1 - p(\tau^{n+1}, \omega^{n+1}))\right] \left[(1-\delta)\, \Pi^{\mathcal{A}}(\mathcal{W}) + \delta\, \mathbf{Z}(\mathcal{J}^{n+1}, \boldsymbol{\sigma}^*)\right] + \\ &d^{n+1}\, p(\tau^{n+1}, \omega^{n+1}) \left[(1-\delta)\, \Pi^{\mathcal{A}}(\mathcal{T}) + \delta\, \mathbf{Z}(\mathcal{J}^{n}, \boldsymbol{\sigma}^*)\right]\end{aligned} \tag{A.16}$$

*For future reference, we also note that using (A.1) of Lemma A.1 and inequality (1), it is immediate that the right-hand side of (A.16) is increasing in $p(\tau^{n+1}, \omega^{n+1})$.*

**Step 4**: *Let an $\mathcal{J} = (\tau, \omega)$ with $\tau < \tau^*$ be given. Let $m \geq 1$ be such that $\tau \in \mathcal{I}^m$. Construct the associated sequences $\{\tau^n\}_{n=0}^m$ and $\{\omega^n\}_{n=0}^m$ as in Step 3. (Recall that, by construction $\tau^m = \tau$ and $\omega^m = \omega$.)*

*We can now construct the probability $p(\tau, \omega) \in [0, 1]$ with which, according to $\boldsymbol{\sigma}^*$, the Court rules $\mathcal{T}$ with $b = 1$, with $1 - p(\tau, \omega) \in [0, 1]$ be the probability that the Court rules $\mathcal{W}$ with $b = 0$.*

*Note that once again we are allowing $p(\tau, \omega) = 0$ since (for the moment) we are not restricting $\tau$ to be strictly greater than the $\underline{\tau}$ of the statement of Proposition 3.*

*Construct the value function backwards as in Step 3, beginning with (A.15). Beginning with $n = 0$ consider the equality*

$$(1-\delta)\, \Pi^{\mathcal{P}}(\mathcal{W}) + \delta\, \mathbf{Z}(\mathcal{J}^{n+1}, \boldsymbol{\sigma}^*) = (1-\delta)\, \Pi^{\mathcal{P}}(\mathcal{T}) + \delta\, \mathbf{Z}(\mathcal{J}^{n}, \boldsymbol{\sigma}^*) \tag{A.17}$$

*where of course the left-hand side depends on $p(\tau^{n+1}, \omega^{n+1})$ as determined by (A.16), while the right-hand side is given, either because $n = 0$ (and hence $\mathbf{Z}(\mathcal{J}^0, \boldsymbol{\sigma}^*)$ is given by (A.15)) or because the values of $p(\tau^h, \omega^h)$ with $h = n, \ldots, 1$ have been set in previous rounds of the recursive procedure we are describing here.*

*The values of $p(\tau^n, \omega^n)$ for $n = 1, \ldots, m$ are then set proceeding recursively backwards (increasing $n$). Since (A.16) is increasing as we noted in Step 3 above, by Assumption 3 (i) (continuity), the following three cases are exhaustive of all possibilities.*

(i) *If (A.17) cannot be satisfied because for every value of* $p(\tau^n, \omega^n)$ *its left-hand side is strictly greater than the right-hand side we set* $p(\tau^n, \omega^n) = 0$.

(ii) *If (A.17) cannot be satisfied because for every value of* $p(\tau^n, \omega^n)$ *its left-hand side is strictly lower than the right-hand side we set* $p(\tau^n, \omega^n) = 1$.

(iii) *If (A.17) can be satisfied, then we set the value of* $p(\tau^n, \omega^n)$ *so that it in fact holds.*

**Step 5***: Let a* $\mathcal{J} = (\tau, \omega)$ *with* $\tau < \tau^*$ *and consider the probabilities constructed in Step 4 and the overall* $\boldsymbol{\sigma^*}$ *of Proposition 3.*

*Taking the continuation equilibrium as given, the Court has no profitable deviation available from choosing* $\mathcal{T}$ *with* $b = 1$ *with probability* $p(\tau, \omega)$ *and choosing* $\mathcal{W}$ *with* $b = 0$ *with probability* $1 - p(\tau, \omega)$ *as prescribed by* $\boldsymbol{\sigma^*}$.

**Proof:** We start by showing that the Court never has a profitable deviation to playing $\mathcal{T}$ with $b = 1$ and $\mathcal{W}$ with $b = 0$ with probabilities other than the ones constructed in Step 4. We will then show that deviating to playing $\mathcal{T}$ with $b = 0$ or $\mathcal{W}$ with $b = 1$ is not profitable either.

Given $\tau < \tau^*$ and $\omega$, let $m$ and the two sequences $\{\tau^n\}_{n=0}^m$, and $\{\omega^n\}_{n=0}^m$ and $\{p(\tau^n, \omega^n)\}_{n=0}^{m-1}$ be as in Step 4. Recall that by construction $\tau = \tau^m$ and $\omega = \omega^m$. Let $\mathcal{J} = (\tau, \omega)$

Clearly, the left-hand side of (A.17) is the Court's (ex-post) continuation payoff after a choosing $\mathcal{W}$ and $b = 0$ with probability one. Similarly, the right-hand side of (A.17) is the Court's (ex-post) continuation payoff after a choosing $\mathcal{T}$ and $b = 1$ with probability one.

Hence if we are in case (i) of Step 4 it is optimal to set $p(\tau^m, \omega^m) = 0$ and if we are in (ii) it is optimal to set set $p(\tau^m, \omega^m) = 1$. If we are in case (iii) of Step 4, then the Court is indifferent between choosing $\mathcal{W}$ and $b = 0$ and choosing $\mathcal{T}$ and $b = 1$, hence setting any value of $p(\tau, \omega) \in [0, 1]$ is optimal in this case.

To see that it is not profitable for the Court to deviate and choose $\mathcal{T}$ and $b = 0$, it is now sufficient to notice that the Court's (ex-post) continuation payoff in the case is $(1 - \delta)\,\Pi^{\mathcal{P}}(\mathcal{T}) + \delta\,\mathbf{Z}(\mathcal{J}^{n+1}, \boldsymbol{\sigma^*})$ which, using (2), is dominated by the left-hand side of (A.17).

Therefore, it only remains to show that it is not profitable for the Court to deviate to choosing $\mathcal{W}$ and $b = 1$. Continuing along the sequence, a $\mathcal{W}$ and $b = 1$ decision will yield a movement from $\tau^m$ to $\tau^{m+1}$ and from $\omega^m$ to $\omega^{m+1}$. Here we need to distinguish two further cases. Using the method we spelled out in Step 4, either $p(\tau^m, \omega^m) = 0$, or $p(\tau^m, \omega^m) > 0$. Suppose first that $p(\tau^m, \omega^m) = 0$, then the continuation payoff from a $\mathcal{W}$ and $b = 1$ decision is

$$(1 - \delta)\Pi^{\mathcal{P}}(\mathcal{W}) + \delta[\tau^{m+1}\Pi^{\mathcal{A}}(\mathcal{T}) + (1 - \tau^{m+1})\Pi^{\mathcal{A}}(\mathcal{W})] \tag{A.18}$$

which, since $\tau^{m+1} < \tau^m$, is less than the left-hand side of (A.17) when we set $n + 1 = m$. Hence the deviation to $\mathcal{W}$ and $b = 1$ cannot be profitable in this case.

Finally, consider that case $p(\tau^m, \omega^m) > 0$. Recall that by Assumption 3 (iv) $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{J}}(\mathcal{J}, \mathcal{T}, 1), \mathcal{W}, 1) = \mathcal{J}$. Since $p(\tau^m, \omega^m) > 0$, the equilibrium strategy $\boldsymbol{\sigma^*}$ at $\mathcal{J}' = (\boldsymbol{\tau}(\mathcal{J}, \mathcal{W}, 1), \boldsymbol{\omega}(\mathcal{J}, \mathcal{W}, 1))$ prescribes a decision of $\mathcal{T}$ (with breadth one) with positive probability. By (A.6) this immediately implies that $\mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma^*}) > \mathbf{Z}(\mathcal{J}', \boldsymbol{\sigma^*})$

and hence that the hypothesized deviation to $\mathcal{R} = \mathcal{W}$ and $b = 1$ is not profitable in this case (in fact it entails a positive loss).

**Step 6***: Consider the probabilities associated with* $\boldsymbol{\sigma}^*$ *computed in Step 4 above. Fix any* $\omega$*, and assume* $\tau \in (\tau_*, \tau^*)$*. Then* $p(\tau, \omega) > 0$.

**Proof:** We deal first with the claim pertaining to $\tau \in (\tau_*, \tau^*)$. By definition of $\tau^*$ and $\tau_*$ it must be that (7) holds. Rearranging terms, this implies directly that

$$\begin{aligned}(1-\delta)\Pi^{\mathcal{P}}(\mathcal{T}) + \delta[\boldsymbol{\tau}(\tau,\omega,\mathcal{T},1)\Pi^{\mathcal{A}}(\mathcal{T}) + (1-\boldsymbol{\tau}(\tau,\omega,\mathcal{T},1))\Pi^{\mathcal{A}}(\mathcal{W})] > \\ (1-\delta)\Pi^{\mathcal{P}}(\mathcal{W}) + \delta[\tau\Pi^{\mathcal{A}}(\mathcal{T}) + (1-\tau)\Pi^{\mathcal{A}}(\mathcal{W})]\end{aligned} \tag{A.19}$$

This, using (A.16), implies that setting $p(\tau^{n+1}, \omega^{n+1}) = 0$ makes the right-hand side of (A.17) strictly greater than the left-hand side. Hence, we cannot be in case (i) of Step 4 and hence it cannot be that $p(\tau, \omega) = 0$.

**Step 7***: Consider a* $\mathcal{J} = (\tau, \omega)$ *with* $\tau < \tau_*$*. Let* $\{\tau^n\}_{n=0}^m$*, and* $\{\omega^n\}_{n=0}^m$ *and* $\{p(\tau^n, \omega^n)\}_{n=0}^{m-1}$ *be as in Step 4 associated with* $\boldsymbol{\sigma}^*$*. Then, if* $p(\tau^m, \omega^m) > 0$*, it must be the case that* $p(\tau^n, \omega^n) > 0$ *for every* $n = m-1, \ldots, 1$.

**Proof:** To see that this must be the case observe that if the claim were false then, using Step 6, for some $m - 1 \leq n \leq 1$ we should have that $p(\tau^{n+1}, \omega^{n+1}) > 0$ and $p(\tau^n, \omega^n) = 0$, with $\tau^{n+1} < \tau_*$ and $\tau^n \leq \tau_*$. Following the prescription of $\boldsymbol{\sigma}^*$ which entails choosing $\mathcal{T}$ and $b = 1$ with positive probability, the Court's (ex-post) continuation payoff is

$$(1-\delta)\Pi^{\mathcal{P}}(\mathcal{T}) + \delta[\tau^n\Pi^{\mathcal{A}}(\mathcal{T}) + (1-\tau^n)\Pi^{\mathcal{A}}(\mathcal{W})] \tag{A.20}$$

since after the equilibrium transition from $\tau^{n+1}$ to $\tau^n$, all Courts will choose $\mathcal{W}$ and $b = 0$ whenever they are not constrained by precedents to choose $\mathcal{T}$.

If instead the current Court deviates to choosing $\mathcal{W}$ and $b = 0$ with probability one, its (ex-post) continuation payoff can be written as

$$\begin{aligned}&(1-\delta)\Pi^{\mathcal{P}}(\mathcal{W}) + \\ &\quad\delta\{\tau^{n+1}\Pi^{\mathcal{A}}(\mathcal{T}) + (1-\omega^{n+1})\Pi^{\mathcal{A}}(\mathcal{W}) + \\ &\qquad(1-\tau^{n+1}-\omega^{n+1})[(1-\delta)\Pi^{\mathcal{A}}(\mathcal{T}) + \delta[\tau^n\Pi^{\mathcal{A}}(\mathcal{T}) + (1-\tau^n)\Pi^{\mathcal{A}}(\mathcal{W})]\}\end{aligned} \tag{A.21}$$

where the term that multiplies $(1 - \tau^{n+1} - \omega^{n+1})$ embodies the fact that the Court following the current one will follow the equilibrium prescription and pick $\mathcal{T}$ and $b = 1$ with positive probability (possibly equal to one), but after the precedents transition from $\tau^{n+1}$ to $\tau^n$, all Courts will choose $\mathcal{W}$ and $b = 0$ whenever they are not constrained by precedents to choose $\mathcal{T}$. Using (1) and (2), and the fact that since $\tau^m \leq \tau_*$ we know that (7) holds as a weak inequality, it is immediate that the quantity in (A.21) is larger than the quantity in (A.20). Hence the deviation to $\mathcal{W}$ and $b = 0$ with probability one is profitable, and this establishes our claim.

**Step 8***: Consider the probabilities associated with* $\boldsymbol{\sigma^*}$ *computed as in Step 4 above. Then there exists a* $\underline{\tau} \in (0, \tau_*]$ *such that* $\mathcal{J} = (\tau, \omega)$ *with* $\tau \leq \underline{\tau}$ *implies that* $p(\tau, \omega) = 0$.

Consider a $\mathcal{J} = (\tau, \omega)$ with $\tau < \tau_*$. Let $\{\tau^n\}_{n=0}^m$, and $\{\omega^n\}_{n=0}^m$ and $\{p(\tau^n, \omega^n)\}_{n=0}^{m-1}$ be as in Step 4 associated with $\boldsymbol{\sigma^*}$.

Let $q$ be the bound on the number of realized $\mathcal{T}$ and $b = 1$ decisions along any equilibrium path identified in Proposition 1. Let the $\mathcal{I}^m$ be the half-open intervals defined in Step 2 above. Suppose that for some $m > q$ we have that $\tau \in \mathcal{I}^m$ and $p(\tau, \omega) > 0$. Then using Step 7, we know that $p(\tau^n, \omega^n) > 0$ for every $n = m, \ldots, 1$. Hence the equilibrium path generated by $\boldsymbol{\sigma^*}$ would have to exceed the bound $q$ with positive probability. Since this contradicts Proposition 1 we conclude that $\tau \in \mathcal{I}^m$ and $m > q$ imply that $p(\tau, \omega) = 0$, which clearly suffices to prove the claim.

**Step 9***: To conclude the proof of Proposition 3 we simply notice that the construction in Step 4 yields an equilibrium which, as a consequence of Steps 6 and 8, satisfies properties (i) and (ii) as required by the statement of the proposition.*

## References


Anderlini, L., L. Felli, and A. Postlewaite (2011): "Should Courts Always Enforce What Contracting Parties Write?," *Review of Law and Economics*, Vol. 7, Iss. 1, Article 2.

Anderlini, L., L. Felli, and A. Riboni (2008): "Statute Law or Case Law?," CESifo Working Paper No. 2358.

Ayres, I. (1997): "The Twin Faces of Judicial Corruption: Extortion and Bribery," *Denver University Law Review*, 74, 1231–1253.

Black, H. C. (1886): "The Principle of Stare Decisis," *The American Law Register (1852-1891)*, 34, 745–757.

Bond, P. (2008): "Persistent Court Corruption," *Economic Journal*, 85, 58–80.

——— (2009): "Contracting in the Presence of Judicial Agency," *The B.E. Journal of Theoretical Economics*, Volume 9, Issue 1 (Advances), Article 36.

Bustos, A. (2008): "A Dynamic Theory of Common Law Courts," Northwestern Law & Econ Research Paper No. 08-12.

Cardozo, B. N. (1921): *The Nature of the Judicial Process*. New Haven: Yale University Press. Available via Google Books.

DAUGHETY, A. F., AND J. F. REINGANUM (1999): "Stampede to Judgement: Persuasive Influence and Herding Behavior by Courts," *American Law and Economics Review*, 1, 158–189.

——— (2000): "Appealing Judgements," *RAND Journal of Economics*, 31, 502–525.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, Massachusetts: The MIT Press.

GENNAIOLI, N., AND A. SHLEIFER (2007a): "Overruling and the Instability of Law," *Journal of Comparative Economics*, 35, 309–328.

——— (2007b): "The Evolution of Common Law," *Journal of Political Economy*, 115, 43–68.

GOODMAN, J. C. (1978): "An Economic Theory of the Evolution of the Common Law," *Journal of Legal Studies*, 7, 393–406.

HADFIELD, G. K. (1992): "Bias in the Evolution of Legal Rules," *Georgetown Law Journal*, 80, 583–616.

HASSLER, J., AND J. V. RODRÍGUEZ MORA (2007): "Political Commitment and Loss Aversion," University of Edinburgh, *mimeo*.

KAPLOW, L., AND S. SHAVELL (2002): *Fairness Versus Welfare*. Cambridge Massachusetts: Harvard University Press.

KENT, J. (1896): *Commentaries on American Law*, vol. I. XIV edn. Boston MA: Little, Brown and Company. XIV edn. Available via Google Books. First edn. 1826.

KYDLAND, F. E., AND E. C. PRESCOTT (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 87, 473–492.

LEGROS, P., AND A. F. NEWMAN (2002): "Courts, Contracts, and Interference," *European Economic Review*, 46, 734–744.

LEVY, G. (2005): "Carrierist Judges and the Appeals Process," *Rand Journal of Economics*, 36, 275–297.

Maskin, E., and J. Tirole (2001): "Markov Perfect Equilibrium I. Observable Actions," *Journal of Economic Theory*, 100(2), 191–219.

Niblett, A., R. A. Posner, and A. Shleifer (2010): "The Evolution of a Legal Rule," *Journal of Legal Studies*, 39(2), 325–58.

Nordhaus, W. (1969): *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, MA: The MIT Press.

Peters, C. J. (1996): "Foolish Consistency: On Equality, Integrity, and Justice in Stare Decisis," *Yale Law Journal*, 105, 2031–2115.

Phelan, C. (2006): "Public Trust and Government Betrayal," *Journal of Economc Theory*, 127(1), 27–43.

Phelps, E. S., and R. A. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35, 185–199.

Ponzetto, G. A. M., and P. A. Fernandez (2008): "Case Law vs. Statute Law: An Evolutionarty Comparison," *Journal od Legal Studies*, 37, 379–430.

——— (2011): "Stare Decisis: Rethoric ad Substance," *Journal of Law Economics & Organization*, forthcoming.

Posner, R. A. (2003): *Economic Analysis of Law*. $6^{th}$ edn. New York: Aspen.

——— (2004): "The Law and Economics of Contract Interpretation," *Texas Law Review*, 83, 1581–1614.

Priest, G. (1977): "The Common Law Process and the Selection of Efficient Rules," *Journal of Legal Studies*, 6, 65–82.

Rubin, P. (1977): "Why is the Common Law Efficient?," *Journal of Legal Studies*, 6, 51–63.

Scherer, F. M. (1972): "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation," *The American Economic Review*, 62(3), 422–427.

Scotchmer, S. (2006): *Innovation and Incentives*. Cambridge, MA: The MIT Press.

Shavell, S. (1995): "The Appeals Process as a Means of Error Correction," *Journal of Legal Studies*, 24, 379–426.

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

Strotz, R. H. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165–180.

Summers, R. (1992): *Essays on the Nature of Law and Legal Reasoning*. Berlin: Duncker and Humblot.

Tirole, J. (2005): *The Theory of Corporate Finance*. Princeton, NJ: Princeton University Press.