



## Why stare decisis? ☆



Luca Anderlini<sup>a</sup>, Leonardo Felli<sup>b</sup>, Alessandro Riboni<sup>c,\*</sup>

<sup>a</sup> Georgetown University, United States

<sup>b</sup> London School of Economics, United Kingdom

<sup>c</sup> Department of Economics, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France

### ARTICLE INFO

#### Article history:

Received 8 June 2013

Received in revised form 12 December 2013

Available online 20 February 2014

#### JEL classification:

C79

D74

D89

K40

L14

#### Keywords:

Stare decisis

Dynamics of precedents

Time-inconsistency

Case Law

### ABSTRACT

All Courts rule *ex-post*, after most economic decisions are sunk. This can 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 rule of precedent, known as *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.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

### 1.1. Motivation

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.

In many circumstances, the *ex-post* nature of Court decisions will generate a *time-inconsistency* problem stemming from the fact that the optimal *ex-ante* decision is different from the optimal *ex-post* one. The optimal decision *ex-ante* typically will take into account incentives for actions that are already taken once a Court is called upon to rule.<sup>1</sup>

☆ This paper partially replaces a previous one entitled “Statute Law or Case Law?” (Anderlini et al., 2008). Part of the research presented here was carried out while Luca Anderlini was visiting EIEF in Rome and the LSE, Leonardo Felli was visiting the Anderson School at UCLA and Humboldt University in Berlin and while all authors were at the Studienzentrums, Gerzensee. Their generous hospitality is gratefully acknowledged. We are grateful to the Editor of the journal and two anonymous referees for their insightful comments. We greatly benefited from the suggestions of 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.

\* Corresponding author.

E-mail address: [alessandro.riboni@polytechnique.edu](mailto:alessandro.riboni@polytechnique.edu) (A. Riboni).

<sup>1</sup> 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*).

Consider for instance 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 to invest in R&D, and the social cost of monopoly power exercised by the patent owner.<sup>2</sup> 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. The optimal decision differs according to whether we look at the problem ex-ante or ex-post.

A Court that is called upon to rule in a case concerning debt-restructuring for a firm in financial distress may also face a time-inconsistency problem. Ex-ante, affording maximum protection to the firm's lenders facilitates the supply of funds, fosters high quality investment projects, and deters low-quality ones. Ex-post, allowing debt-restructuring avoids the social costs of liquidation. Once again, the ex-ante and ex-post optimal decisions differ.<sup>3</sup> Again, the "myopic" or "weak" decision is optimal ex-post, while ex-ante the "tough" or "forward-looking" decision dominates.

A key ingredient of a system of Common Law (for instance the UK or the USA) is the role of *precedents* in shaping the decisions of Courts. While Courts often have considerable latitude in the interpretation of the Law, they are in general bound by an evolving body of precedents. The principle by which "previous rulings apply" is known as *Stare Decisis* — literally Latin for "to stand by things decided."<sup>4</sup>

Our model centers on the role of stare decisis as a mechanism to alleviate the time inconsistency problem that Courts face in a Common Law environment.

Each Court may be bound by stare decisis to take the tough decision, depending on the details of the case (which we model as a random draw). When it is *not* bound by stare decisis in its *current ruling* even a forward-looking Court would take the weak decision were it not for the effect that its current ruling has on *future courts*, via stare decisis again.

Thus, a forward-looking Court faces a trade off between the welfare gains from taking a weak decision on the case before it today (which it examines ex-post) against the future gains that a tough decision today may have on future Courts. Stare decisis implies that today's tough decision increases the likelihood that *future Courts* are bound by precedents to take the ex-ante optimal tough decision.

We study the open-ended dynamics generated by a sequence of forward-looking Courts facing the trade off we have just described. Two main insights emerge from the analysis. The first is that identifying the time-inconsistency problem facing the Courts provides a new rationale for the principle of stare decisis. The second is that our model indicates that while precedents evolve towards efficiency, they are necessarily bounded away from their optimal form, even in the long run.

The reason that the evolution of precedents stops short of full efficiency is that the trade-off facing the Courts that we described above changes through time. As precedents evolve towards efficiency the gains from *further constraining* future Courts shrink since stare decisis already imposes the tough decision on future Courts with high probability. On the other hand, the gains from taking the weak decision in the case before the current Court stay constant through time. The latter effect eventually dominates and precedents stop evolving.

Throughout the paper, we focus on Markov-Perfect Equilibria of the dynamic game among successive Courts. We characterize the unique equilibrium, which involves an evolving sequence of randomized Court decisions.

The rest of the paper is organized as follows. In the next section we review some related literature. In Section 3 we set up the model. In Section 4 we proceed to characterize the unique Markov-Perfect Equilibrium of our model, and to provide a numerical illustration. Section 5 concludes the paper. For ease of exposition, all proofs are in [Appendix A](#).

## 2. Related literature

The literature on the role of stare decisis goes back at least to the end of the nineteenth century. The most prominent argument is that the rule of precedents provides predictability to the rulings of Courts. 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.* [...]

On the other hand it has also long been understood that precedents evolve since genuinely new issues are bound to arise. When judges stray from precedents, they take their influence on future rulings into account. In the words of Benjamin Cardozo (1921, p. 20–21).<sup>5</sup>

[...] *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.*

<sup>2</sup> 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).

<sup>3</sup> See for instance Tirole (2005, Ch. 16) on the "topsy-turvy" principle in Corporate Finance.

<sup>4</sup> According to the Oxford English Dictionary (2013) stare decisis is "The legal principle of determining points in litigation according to precedent."

<sup>5</sup> For a modern perspective on these issues see also Calabresi (1982).

This paper pursues a complementary answer to the “stability” rationale for stare decisis.<sup>6</sup> As we noted above, the role of stare decisis in our model is to mitigate the Courts’ time-inconsistency problem.

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.<sup>7</sup> Evolutionary models of the Common Law have called attention to explanations other than judicial 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.

The first to address the evolution of precedents in the Economics literature are Gennaioli and Shleifer (2007b). They consider a model of “distinguishing” where judges are able to limit the applicability of precedents by introducing new material dimensions to adjudication. In particular, they investigate the claim by Cardozo (1921, p. 177) that Case law evolves towards efficient rules even in the presence of judicial bias. Their results support this hypothesis. Sequential decision making improves efficiency on average by making Common Law more precise. The 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. “Judicial bias” is the source of innovation.<sup>8</sup>

In Gennaioli and Shleifer (2007b) the evolution of precedents is driven by judicial heterogeneity and new information is added as precedents evolve. Judicial heterogeneity is key to the process since the difference in judges’ preferences is what drives their desire to introduce new information and change previous decisions. In our set-up all Courts are identical, both ex-ante and ex-post. But because they act sequentially, and they have time-inconsistent preferences the period- $t$  Court “disagrees” with the period- $t + 1$  Court on the optimal ruling at time  $t + 1$ . In our model, this tension is what drives the evolution of precedents, via stare decisis. To reiterate the point if the Courts in our model could meet at the same time, they would agree on everything. Time-inconsistency generates a rich dynamical system that is substantially different from Gennaioli and Shleifer (2007b). In particular the desire to postpone tough decisions forces the use of mixed strategies in equilibrium.<sup>9</sup>

The long-run fate of precedents has been addressed by several other recent contributions. 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.<sup>10</sup> By contrast, in a fully dynamic model where Courts have to spend time and resources investigating a case, Backer and Mezzetti (2012) find that precedents might converge to an inefficient set of legal rules.<sup>11</sup>

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). As is the case in Gennaioli and Shleifer (2007b), strictly speaking all Courts in our model should be viewed as appellate Courts since. With positive probability, they all have the ability to change the state of precedents.

Case Law dynamics have been studied mainly from a theoretical perspective. One exception is Niblett et al. (2010), who analyze the evolution of a doctrine, known as the economic loss rule (ELR hereafter), over a period of 35 years.<sup>12</sup> Their contribution is directly related to our work because the application of the ELR in Courts is likely to be subject to credibility problems.<sup>13</sup> Niblett et al. (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 results of Gennaioli and Shleifer (2007b) are consistent with the decrease in exceptions to the ELR observed by Niblett et al. (2010). By contrast, our model generates a frequency of weak decisions (exceptions to the ELR) that increases thorough time, consistent with the later trend in Niblett et al. (2010). This is because in our model initially Courts take the tough decisions with higher probability to shift the body of precedents towards efficiency, while later the frequency of tough decisions decreases since stare decisis forces the tough decision with increasing probability.

<sup>6</sup> See also Peters (1996).

<sup>7</sup> 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.

<sup>8</sup> Judicial bias is interpreted in a broad sense that ranges from “idiosyncracies” in the judges’ preferences (Bond, 2009; Gennaioli and Shleifer, 2007a) to plain “corruption” of the Courts (Ayres, 1997; Bond, 2008; Legros and Newman, 2002, among others).

<sup>9</sup> We analyze this point in detail in Section 4.4 below.

<sup>10</sup> In another recent paper, Ponzetto and Fernandez (2012) find that the evolution of Case Law towards efficiency is more likely when judges are sufficiently polarized.

<sup>11</sup> See also Bustos (2008) who models 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.

<sup>12</sup> 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.

<sup>13</sup> 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.

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 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. 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. In our model 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.

### 3. The model

Consider an infinite horizon model and an infinite sequence of *one-period-lived* Courts. A *forward-looking* Court rules in each period.

#### 3.1. The static environment

A 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. 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  $\Pi^A(\mathcal{R})$  and  $\Pi^P(\mathcal{R})$ , 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

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

and

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

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.

#### 3.2. 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  $[0, 1]$ . This allows us to specify the body of precedents in a particularly simple way.

The body of precedents is represented by a number in  $\tau \in [0, 1]$ . Once a case comes to the attention of a Court its legal characteristics are determined:  $\ell$  is realized.

The interpretation of  $\tau$  is straightforward. The body of precedents is seen to either apply or not apply. If  $\ell \leq \tau$  the body of precedents constrains the Court to a *tough* decision, if instead  $\tau < \ell \leq 1$  the Court has *discretion* over the case.

Whenever the precedents bind the Court towards the tough decision we are in a situation in which the Court's ruling is determined by stare decisis. Whenever the precedents do not bind, the case at hand is sufficiently idiosyncratic to escape the doctrine of stare decisis.<sup>14</sup>

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. Since the parties and the Court

<sup>14</sup> 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.

both observe the full past history of decisions, the parties 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.<sup>15</sup>

### 3.3. 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 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  $t$ ,  $\tau^t \in [0, 1]$ . 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  $t + 1$ . Hence  $\tau^{t+1} = \tau^t$ .

When a Court is *not* constrained by precedents, with probability  $(1 - \tau^t)$ , it can choose the tough or weak decision at its discretion.

The discretionary ruling  $\mathcal{R}^t \in \{\mathcal{T}, \mathcal{W}\}$  of the  $t$ -th Court determines how  $\tau^t$  is modified to yield  $\tau^{t+1}$ , on the basis of which the  $t + 1$ -th Court will operate. Therefore, the precedents technology can be viewed as a map  $\mathcal{J} : [0, 1] \times \{\mathcal{T}, \mathcal{W}\} \rightarrow [0, 1]$ , so that  $\tau^{t+1} = \mathcal{J}(\tau^t, \mathcal{R}^t)$ .

Mostly for simplicity, we assume that when a Court takes the weak decision the body of precedents is unchanged so that again  $\tau^{t+1} = \tau^t$ . In other words, we take  $\tau^{t+1} = \tau^t = \mathcal{J}(\tau^t, \mathcal{W})$ . This simplifying assumption can also be interpreted as Courts taking weak decisions with “narrow” rulings so that their effect on future Courts is as small as possible.

We denote by  $\tau$  the (projection) map  $[0, 1] \rightarrow [0, 1]$  that is defined by fixing  $\mathcal{R} = \mathcal{T}$  in  $\mathcal{J}$ . Hence, if the Court takes the tough decision:  $\tau^{t+1} = \tau(\tau^t) = \mathcal{J}(\tau^t, \mathcal{T})$ . In what follows we make the following assumption on  $\tau(\cdot)$ .

**Assumption 1** (*Well-behaved precedents technology*). The map  $\tau(\cdot)$  is continuous, strictly increasing and concave, and satisfies  $\tau(0) = 0$  and  $\tau(1) = 1$ .

Typically, the map  $\tau(\cdot)$  will embody the workings of a complex set of legal mechanisms and constitutional arrangements, which may entail complex interaction effects among its many arguments. Obviously, our model is simplified in the extreme. This, on the other hand, allows us to proceed with a simple and transparent set of assumptions.

### 3.4. 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.<sup>16</sup> 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  $\tau^t$  from the past. Given  $\tau^t$ , it observes the outcome of the draw that determines the legal characteristics of the case ( $\ell$ ). Together with  $\tau^t$ , this determines whether the  $t$ -th period Court has discretion or not. If it has discretion, the  $t$ -th Court then chooses  $\mathcal{R}^t$ . Together with  $\tau^t$  this determines  $\tau^{t+1}$ , 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  $\tau^{t+1} = \tau^t$ .

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 the  $t$ -th Court depends on the state of precedents  $\tau^t$  and the realization  $\ell^t$ . Although the  $t$ -court *does* observe  $\ell^t$  before it makes its decision, it turns out to be convenient to omit it from the formal notation of the Court's strategy of period  $t$ . This is inessential since the realization of  $\ell^t$  determines whether the  $t$ -th Court actually makes a decision or not (if  $\ell^t \leq \tau^t$  the Court is constrained to take the  $\mathcal{T}$  decision). Provided the Court knows  $\tau^t$ , we can think of it as making a decision *contingent* on  $\ell^t > \tau^t$  interchangeably with it making a decision *after* if observes the actual realization of  $\ell^t$ .

<sup>15</sup> 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 (with the notable exception of Spier, 1992). The question of why, in equilibrium (and therefore with “correct” expectations), contracting parties go to Court is a key question, well beyond the scope of this paper.

<sup>16</sup> 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 Section 4 below).

When the Court is not constrained by precedents, the choice of ruling in period  $t$ ,  $\mathcal{R}^t$ , depends on  $\tau^t$ . Notice that, in principle, the choices of the  $t$ -th Court could depend on the entire history of past rulings, legal characteristics (including the ones at time  $t$ ) and parties' behavior. We restrict attention to behavior that depends only on the body of precedents  $\tau^t$ . 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* (Maskin and Tirole, 2001). We will do so throughout the rest of the paper.

The strategy of the Court in period  $t$  is then a map  $\sigma : \tau^t \rightarrow [0, 1]$ , that determines the (possibly mixed)  $t$ -th Court's choice of ruling. By convention we set  $\sigma(\tau^t) \in [0, 1]$  to be the probability that the  $t$ -th Court's ruling is  $\mathcal{T}$ . This strategy is time independent given our restriction to Markov strategies. This implies that  $\sigma$  denotes also the entire strategy profile for all Courts.

Given  $\tau^t$ , the expected (as of the beginning of period  $t$ ) payoff accruing in period  $t$  to the  $t$ -th Court from using strategy  $\sigma$  is denoted  $\Pi^A(\tau^t, \sigma)$ . This can be written as

$$\Pi^A(\tau^t, \sigma) = \tau^t \Pi^A(\mathcal{T}) + (1 - \tau^t) \{ \sigma(\tau^t) \Pi^A(\mathcal{T}) + [1 - \sigma(\tau^t)] \Pi^A(\mathcal{W}) \} \tag{3}$$

The interpretation of (3) is straightforward. The first term refers to the cases in which the Court is constrained to a tough decision. The second term is the Court's payoff given its discretionary ruling  $\sigma(\tau^t)$ .

Given the (stationary) preferences we have postulated, the overall payoffs to each Court can be expressed in a familiar recursive form. Let a  $\sigma$  be given. Let  $\mathcal{Z}(\tau^t, \sigma)$  be the expected (as of the beginning of the period) overall payoff to the  $t$ -th Court, given  $\tau^t$  and  $\sigma$ .<sup>17</sup> We can then write this payoff as follows.

$$\mathcal{Z}(\tau^t, \sigma) = (1 - \delta) \Pi^A(\tau^t, \sigma) + \delta \{ \tau^t \mathcal{Z}(\tau^t, \sigma) + (1 - \tau^t) [ \sigma(\tau^t) \mathcal{Z}(\tau(\tau^t), \sigma) + (1 - \sigma(\tau^t)) \mathcal{Z}(\tau^t, \sigma) ] \} \tag{4}$$

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  $\tau^{t+1} = \tau^t$ . The second term that multiplies  $\delta$  is the Court's continuation payoff if the realization of the Court's strategy at  $t$  is  $\mathcal{T}$  while the last term that multiplies  $\delta$  is the Court's continuation payoff if the realization of the Court's strategy at  $t$  is  $\mathcal{W}$ .

Now recall that the  $t$ -th Court decides whether to take a tough or a weak decision (if it is given discretion) 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  $\tau^t$  and  $\sigma$ , the decisions of the  $t$ -th Court can be characterized as follows. Suppose that the  $t$ -th Court is *not* constrained by precedents to a tough decision,<sup>18</sup> then, the values of  $\sigma = \sigma(\tau^t)$  must solve

$$\max_{\sigma \in [0,1]} (1 - \delta) [ \sigma \Pi^P(\mathcal{T}) + (1 - \sigma) \Pi^P(\mathcal{W}) ] + \delta \{ \sigma \mathcal{Z}(\tau(\tau^t), \sigma) + (1 - \sigma) \mathcal{Z}(\tau^t, \sigma) \} \tag{5}$$

It is now straightforward to define what constitutes an equilibrium.

**Definition 1** (*Equilibrium behavior*). An equilibrium is a strategy profile  $\sigma^*$  such that, for every possible  $\tau^t$ , the value  $\sigma = \sigma^*(\tau^t)$  is a solution to problem (5).<sup>19</sup>

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  $\tau^0$ . Using the notation we already established, this is denoted by  $\mathcal{Z}(\tau^0, \sigma^*)$ .

#### 4. The evolution of Case Law

##### 4.1. Case Law without stare decisis

To appreciate fully the effect of stare decisis on the evolution of Case Law it is instructive to characterize what can happen in a world in which precedents never constrain the decision of Courts. This is a counterfactual polar case in which *all* Courts have full discretion.

<sup>17</sup> The function  $\mathcal{Z}(\cdot, \cdot)$  is independent of  $t$  because we are restricting attention to stationary Markov-Perfect Equilibria.

<sup>18</sup> 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  $\tau^{t+1} = \tau^t$ .

<sup>19</sup> It should be noted that in equilibrium the decision of the  $t$ -th Case Law Court is required to be optimal given *every possible*  $\tau^t$ , and not just those that have positive probability given  $\sigma^*$  and  $\tau^0$ . This is a standard "perfection" requirement.

The model we have set up above contains this as special case. From [Assumption 1](#) above, if the initial state of precedents is  $\tau^0 = 0$ , then  $\tau^t = 0$  for every  $t > 0$  as well. In other words, every Court is constrained by precedents with probability 0 and hence has discretion with probability 1. The evolution of precedents never gets off the ground.

In this polar case, Courts cannot affect the future through their ruling, hence their optimal strategy in each period is to maximize the ex-post payoff of the parties. In other words, in every period the Court chooses  $\mathcal{W}$ .

**Proposition 1** (No stare decisis equilibrium). *Suppose that  $\tau^0 = 0$  and let any  $\delta \in [0, 1)$  be given. Then the unique equilibrium of the model is such that the Court chooses a  $\mathcal{W}$  ruling in every period of every subgame. The payoff of every Court is then  $\Pi^P(\mathcal{W})$ .<sup>20</sup>*

#### 4.2. The evolution of Case Law under stare decisis

Consider now the case where the initial state of precedents is  $\tau^0 > 0$ . It is then evident from [Assumption 1](#) that precedents will evolve every time a Court chooses  $\mathcal{T}$ . Every tough decision increases the probability that future Courts are constrained by precedents to rule  $\mathcal{T}$ . With weakly monotonically increasing probability, stare decisis forces each Court to rule  $\mathcal{T}$ , even in the face of the immediate temptation to be excessively lenient and rule  $\mathcal{W}$ .

Stare decisis imposes a very tight structure on the set of outcome paths that can be observed in equilibrium. As in the absence of stare decisis, our model displays a *unique* equilibrium under stare decisis (that is, whenever  $\tau^0 > 0$ ). In general, the equilibrium will involve mixed strategies, hence in contrast to the case of no stare decisis a multiplicity of outcome paths can obtain with positive probability. However, for any given  $\tau^0 > 0$  and  $\delta$  the equilibrium strategy profile is in fact unique.

Below, we present three incarnations of the same result – the unique equilibrium of the model under stare decisis.

The first is a property that predicts the onset of what we call “Mature Precedents.” In equilibrium, the Courts eventually stop ruling  $\mathcal{T}$  and precedents stops evolving as a result. The state of precedents no longer changes, and all Courts from that point on are constrained to rule  $\mathcal{T}$  with the same probability, while all Courts that have discretion rule  $\mathcal{W}$ . One of the direct implications of this property is that, while precedents do evolve towards efficiency, they always stop short of constraining the Courts to *always* take the ex-ante efficient decision  $\mathcal{T}$ . We examine this property separately (even though it is implied by the characterization that follows it) because we think it is an interesting property in its own right and because it helps streamline the intuition for the results as we go along.

The second take on the unique equilibrium of the model is a general characterization of the equilibrium strategy profile for each  $\tau^0$  and  $\delta$ . As it turns out, if  $\delta$  is sufficiently large this is a richer structure than one would expect. The reason is that, in general, the equilibrium will involve mixed strategies which are not completely straightforward to characterize.

The third and last angle we use on the unique equilibrium is that of computing the actual equilibrium strategies for a *specific* form of the precedents technology  $\tau(\cdot)$  satisfying [Assumption 1](#) above.

#### 4.3. Mature Case Law

Given an initial body of precedents  $\tau^0 > 0$  and a profile  $\sigma$ , as the randomness in each period is realized (whether the precedents bind or not, and possibly the outcome of mixed strategies) a sequence of Court rulings will also be realized.

The realized number of times that the Courts will take the tough decision ( $\mathcal{T}$ ) has an *upper bound*. Precedents eventually *mature*, and, after they do, all discretionary Courts succumb to the (time-inconsistent) temptation to rule  $\mathcal{W}$  instead of  $\mathcal{T}$ .

**Proposition 2** (Mature precedents). *Fix a  $\delta$  and a  $\tau^0 > 0$ , and let any equilibrium strategy profile  $\sigma^*$  be given.*

*Then, there exists an integer  $q$ , which depends on  $\delta$  and  $\tau^0$  but not on the particular equilibrium  $\sigma^*$ , 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.

The 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  $\tau^t$  is monotonically weakly increasing and, from [Assumption 1](#), satisfies “decreasing returns” in the future gains stemming from a tough decision today, eventually, precedents mature. That is, 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  $\mathcal{W}$ .

<sup>20</sup> We state this result without proof, since it is an immediate consequence of  $\tau(0) = 0$  and (2) above.

#### 4.4. Mixed strategies

The characterization of the equilibrium behavior is somewhat intricate. To appreciate some of the difficulties involved, recall that from [Proposition 2](#) we know that there is an upper bound on the number of times the tough decision  $\mathcal{T}$  can possibly be taken by a discretionary Court in any equilibrium.

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  $\mathcal{T}$  to constrain future Courts to do the same with higher probability. Now consider “the last” Court to rule  $\mathcal{T}$  exercising its discretion to do so. In other words, suppose that the equilibrium prescribes that some Court that has discretion rules  $\mathcal{T}$ , knowing that from that point on all future Courts will rule  $\mathcal{W}$  whenever they are not bound by precedents. That is to say, 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.

The situation we have just described, faced by the last tough Court cannot in fact take place in equilibrium. To see this, consider the possibility that the last tough Court, say at time  $t$ , deviates and takes the weak decision  $\mathcal{W}$  instead of the tough one. If it does so, the next Court that has discretion will face the same body of precedents, and, because of the upper bound of [Proposition 2](#) it will be the last tough Court.

This must then be an appealing deviation for the time  $t$  Court, for two separate but complementary reasons. First of all, the deviating Court 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:  $\Pi^{\mathcal{A}}(\mathcal{T}) > \Pi^{\mathcal{A}}(\mathcal{W})$ .

The solution to the puzzle we have just outlined is that a typical equilibrium of our model may require *mixed strategies*. Before precedents mature, Courts randomize between the  $\mathcal{T}$  decision and the  $\mathcal{W}$  decision. This in turn allows Case Law to begin with tough discretionary decisions without violating [Proposition 2](#) (precedents 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 each Court that acts before precedents mature can be kept indifferent between the two decisions by an appropriate choice of the mixing probabilities employed by future Courts.

Before any further discussion, we now proceed with a detailed characterization of the equilibrium behavior of our model.

Denote, for convenience  $\kappa(\tau) = \tau(\tau) - \tau$  the *increment* in  $\tau$ , stemming from a tough decision today. Suppose that there exist some value(s) of  $\tau \in [0, 1]$  such that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] < \delta\kappa(\tau)[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \quad (6)$$

Notice that from [Assumption 1](#), we know that  $\kappa(\tau)$  is non-negative, and equal to zero if  $\tau$  is either 1 or 0, and  $\kappa(\cdot)$  is concave. It is then immediate that if condition (6) holds, there exist precisely two values of  $\tau \in (0, 1)$  such that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] = \delta\kappa(\tau)[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \quad (7)$$

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. Clearly, all else given,  $\tau_*$  and  $\tau^*$  are defined whenever  $\delta$  is sufficiently large.

**Proposition 3** (Equilibrium characterization). *Let a precedents technology  $\tau(\cdot)$  satisfying [Assumption 1](#) be given. Suppose also that  $\delta$  is large enough so that  $\tau_*$  and  $\tau^*$  are well defined.*

*For any given initial state of precedents  $\tau^0 > 0$  the model has a unique equilibrium as in [Definition 1](#) as follows. There exists a threshold value  $\underline{\tau} \in (0, \tau_*)$  such that:*

- (i) *If  $\tau \in [0, \underline{\tau}] \cup [\tau^*, 1]$ , then, if it has discretion, the ruling of the Court is  $\mathcal{W}$  with probability 1 in every period.*
- (ii) *If instead  $\tau$  is such that  $\underline{\tau} < \tau < \tau^*$ , the Court, when it has discretion, randomizes between a ruling of  $\mathcal{T}$  and a ruling of  $\mathcal{W}$ . We denote the probability of a  $\mathcal{T}$  by  $p(\tau) \in (0, 1)$ , so that the probability of a  $\mathcal{W}$  is  $1 - p(\tau) \in [0, 1)$ .*

According to [Proposition 3](#), if  $\tau^0 \leq \underline{\tau}$  then Case Law does not evolve,  $\tau^t = \tau^0$  in every period, and every discretionary Court chooses the time-inconsistent ruling  $\mathcal{W}$ .

If instead  $\tau^0 \in (\underline{\tau}, \tau^*)$ , Case Law undergoes two phases: a transition state, 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 taken by discretionary Courts with strictly positive probability. 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 take instead the weak decision.

Finally, if  $\tau^0 \geq \tau^*$  once again Case Law does not evolve and all discretionary Courts take the time-inconsistent ruling  $\mathcal{W}$ .

The rich behavior captured by [Proposition 3](#) is due to the temptation to entertain the time-inconsistent behavior. This is obvious when  $\tau_* < \tau^0 < \tau^*$ . 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 – inequality (6) holds.

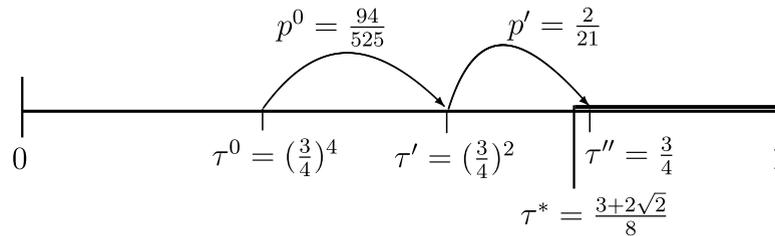


Fig. 1. Dynamics of precedents.

However, when inequality (6) holds, for the reasons we described above, a *pure strategy* equilibrium in which a finite sequence of  $\mathcal{T}$  decisions are taken may not be viable. The equilibrium then involves the Courts who have discretion mixing between a  $\mathcal{T}$  ruling and a  $\mathcal{W}$  ruling. 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 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  $\mathcal{W}$ .

We conclude noting that the region ( $\tau \leq \underline{\tau}$ ) near zero in which the equilibrium prescribes that precedents do not evolve is essentially an artifact of the regularity conditions (namely [Assumption 1](#)) we have imposed on the precedents technology. More complex equilibria that do not display this region can be constructed when these conditions are dropped, for example by choosing a high enough value of  $\tau(0)$ . 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 precedents can be made arbitrarily small as discounting decreases –  $\tau_*$ , and  $\underline{\tau}$ , approach zero as  $\delta$  approaches one.

#### 4.5. 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 [Assumption 1](#). Consider then the following specification for  $\tau(\cdot)$ .

$$\tau^{t+1} = \tau(\tau^t) = \sqrt{\tau^t} \quad (8)$$

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.

$$\delta = 4/5, \quad \Pi^A(\mathcal{T}) = 15, \quad \Pi^A(\mathcal{W}) = 5, \quad \Pi^P(\mathcal{T}) = 15, \quad \Pi^P(\mathcal{W}) = 20 \quad (9)$$

The initial state of precedents is

$$\tau^0 = \left(\frac{3}{4}\right)^4 = \frac{81}{256} \quad (10)$$

Eq. (6) is then satisfied for some values of  $\tau \in (0, 1)$  and the two thresholds  $\tau_*$  and  $\tau^*$  are well defined and given by

$$\tau^* = \frac{3 + 2\sqrt{2}}{8} \quad \text{and} \quad \tau_* = \frac{3 - 2\sqrt{2}}{8} \quad (11)$$

Using the values in (9), (10) and (11) we can now proceed to compute the specifics of the equilibrium in [Proposition 3](#) in this case. [Fig. 1](#) 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 ruling  $\mathcal{T}$  and  $\mathcal{W}$ . The numerical values we have picked imply that precedents 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  $\mathcal{T}$  twice, the equilibrium prescribes that no further randomization on the part of the Court will occur. The court will always take the weak decision  $\mathcal{W}$  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 (10), and the fact that two tough decisions have been taken. Similarly, the intermediate value of  $\tau$  after one tough decision has been taken is  $\tau' = (3/4)^2 = 9/16$ .

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 indifferent between the tough and the weak decision. The value of the first probability of tough decision is  $p^0 \simeq 0.18$  while the value of the second probability is  $p' \simeq 0.10$ .<sup>21</sup>

<sup>21</sup> The exact values are  $p^0 = p(\tau^0) = \frac{94}{525}$  and  $p' = p(\tau') = \frac{2}{21}$ .

To see how these probabilities are computed, notice that we can proceed backwards from the steady state so as to compute  $p'$ . The probability  $p^0$  can then be computed proceeding backwards one more step.<sup>22</sup>

### 5. 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. We argue that this is one of the roots of the principle of stare decisis that disciplines Common 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 precedents “mature” 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.

We characterize the unique equilibrium under stare decisis. In the more interesting cases this equilibrium requires the Courts to 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.

### Appendix A

**Lemma A.1.** *Let any  $\tau^0 > 0$  and any equilibrium strategy profile  $\sigma^*$  be given. Denote  $h^t$  a realized path of the uncertainty associated with strategy  $\sigma^*$  up to period  $(t - 1)$ , that is a realization of the  $t$  random variables  $(\sigma^{*0}, \dots, \sigma^{*(t-1)})$ . Let  $\mathcal{T}(h^t)$  be the number of times that along the realized path  $h^t$  a Court that has discretion rules  $\mathcal{T}$ . Then  $\tau^t$  is weakly increasing in  $t$  and strictly increasing in  $\mathcal{T}(h^t)$ . Moreover, as  $\mathcal{T}(h^t)$  increases without bound,  $\tau^t$  converges to one.*

**Proof.** The details are omitted since, using standard arguments, the claim is a direct consequence of Assumption 1. □

**Lemma A.2.** *Fix a  $\delta$  and a  $\tau^0$ , and let any equilibrium strategy profile  $\sigma^*$  be given. Fix also any realized path  $h^t$ . Denote  $\tau(h^t)$  the threshold  $\tau^t$  associated with the realized path  $h^t$ . Suppose that  $\sigma^{*t}$  assigns strict positive probability to a ruling of  $\mathcal{T}$ . Then it must be the case that*

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] \leq \delta\{[1 - \tau[\tau(h^t)]]\}[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \tag{A.1}$$

**Proof.** Using (1), at the *ex-post* stage when the time  $t$  Court is called upon to rule, the worse payoff it can obtain by ruling  $\mathcal{W}$  consists of  $\Pi^{\mathcal{P}}(\mathcal{W})$  at time  $t$ , plus a continuation of play of Courts that all rule  $\mathcal{W}$  when they are unconstrained by precedents, and of course  $\mathcal{T}$  when stare decisis dictates such ruling. Since by Lemma A.1 it must be that  $\tau^{t+m} \geq \tau[\tau(h^t)]$  for every  $m \geq 1$  we can conclude that a lower bound on the time  $t$  Court payoff if it rules  $\mathcal{W}$  with probability one is given by

$$(1 - \delta)\Pi^{\mathcal{P}}(\mathcal{W}) + \delta\{\tau[\tau(h^t)]\}\Pi^{\mathcal{A}}(\mathcal{T}) + [1 - \tau[\tau(h^t)]]\Pi^{\mathcal{A}}(\mathcal{W}) \tag{A.2}$$

If the time  $t$  Court rules  $\mathcal{T}$  with probability one, then, using (1), its payoff is bounded above by

$$\Pi(\mathcal{T}) = (1 - \delta)\Pi^{\mathcal{P}}(\mathcal{T}) + \delta\Pi^{\mathcal{A}}(\mathcal{T}) \tag{A.3}$$

Simple algebra is now sufficient to check that if (A.1) is violated then the quantity in (A.2) is strictly greater than the quantity in (A.3). Hence if (A.1) is violated, the time  $t$  Court after the realized path  $h^t$  strictly prefers a ruling of  $\mathcal{W}$  with

<sup>22</sup> The exact computation is presented in Appendix A.

probability one to a ruling of  $\mathcal{T}$  with probability one, and hence cannot be made indifferent. This clearly suffices to prove the claim.  $\square$

**Proof of Proposition 2.** Fix a  $\delta$  and a  $\tau^0$ , and let any equilibrium strategy profile  $\sigma^*$  be given. Begin by noting that if  $\delta$  is such that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] > \delta[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \quad (\text{A.4})$$

then (A.1) cannot possibly be satisfied for any  $\tau(h^t) \in [0, 1]$ . Hence if (A.4) is satisfied, there is nothing more to prove. Using Lemma A.2, no Court that is given discretion can ever give positive probability to a ruling of  $\mathcal{T}$ . The  $q$  of the statement of the proposition is in fact equal to zero.

Assume next that (A.4) does not hold. Note that using (1) and (2) we know that there is a unique value  $\hat{\tau}$  such that when substituted for  $\tau(h^t)$  guarantees that (A.1) holds with equality. Notice further that this value does not depend on the equilibrium we fixed, but only on  $\delta$  (and the other payoff parameters of the model of course).

From Lemma A.1, for a given  $\tau^0$  we know that there exists a number  $q$  such that for any realized path  $h^t$ , if  $\mathcal{T}(h^t) > q$  then  $\tau^t = \tau(h^t) > \hat{\tau}$ .

Now suppose that along any history we have that  $\mathcal{T}(h^t) > q$ . Then the time  $t$  Court, takes its decision facing a realized path  $h^t$  such that (A.1) is violated. Hence, if it has discretion, its ruling must be  $\mathcal{W}$  with probability one. This clearly suffices to prove the claim.  $\square$

**Lemma A.3.** Fix a large enough  $\delta$  so that  $\tau_*$  and  $\tau^*$  satisfying Eq. (7) are well defined. Consider the sequence of numbers in  $[0, \tau^*]$  obtained repeatedly applying the inverse mapping  $\tau^{-1}(\cdot)$  that is well defined by Assumption 1. Formally, let  $\tau^* = \hat{\tau}_0$  and then define recursively  $\hat{\tau}_n = \tau^{-1}(\hat{\tau}_{n-1})$  for  $n = 1, \dots, \infty$ , and note that because of Assumption 1 this defines a unique sequence  $\{\hat{\tau}_n\}_{n=0}^{\infty}$ .

Denote by  $\mathcal{I}_n$  each of the half open intervals  $[\hat{\tau}_n, \hat{\tau}_{n-1})$ , and by convention set  $\mathcal{I}_0 = [\tau^*, \tau(\tau^*))$ . Then,

- (i) The sequence  $\{\hat{\tau}_n\}_{n=0}^{\infty}$  is strictly decreasing and  $\lim_{n \rightarrow \infty} \hat{\tau}_n = 0$ .
- (ii) For every  $n = 1, \dots, \infty$ , if  $\hat{\tau}_n \in \mathcal{I}_n$  then  $\tau^{-1}(\hat{\tau}_n) \in \mathcal{I}_{n+1}$  and  $\tau(\hat{\tau}_n) \in \mathcal{I}_{n-1}$

**Proof.** Claim (i) and (ii) follow directly from Assumption 1, in particular, continuity, monotonicity and concavity of  $\tau(\cdot)$  and  $\tau(0) = 0$ .  $\square$

**Proof of Proposition 3.** We proceed in 8 steps to verify that, under the hypotheses of the proposition, we can find a  $\sigma^*$  that satisfies (i) and (ii).

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

**Proof.** Recall that  $\sigma^*$  prescribes that whenever  $\tau \geq \tau^*$  then all Courts that have discretion choose  $\mathcal{W}$ . Fix any given value of  $\tau \geq \tau^*$ .

Consider, first, the deviation to setting  $\mathcal{T}$  and keeping the continuation equilibrium fixed as given by  $\sigma^*$ . By definition of  $\tau^*$ , whenever  $\tau \geq \tau^*$  we must have that

$$(1 - \delta)[\Pi^{\mathcal{P}}(\mathcal{W}) - \Pi^{\mathcal{P}}(\mathcal{T})] \geq \delta\kappa(\tau)[\Pi^{\mathcal{A}}(\mathcal{T}) - \Pi^{\mathcal{A}}(\mathcal{W})] \quad (\text{A.5})$$

with the equality holding if and only if  $\tau = \tau^*$ . Condition (A.5) directly implies that the proposed deviation cannot be profitable.  $\square$

Let a Markov Perfect Equilibrium  $\sigma^*$  as in Proposition 3 be given, and for any given  $\tau < \tau^*$ , let  $p(\tau) \in [0, 1]$  be the probability that, according to  $\sigma^*$ , the Court rules  $\mathcal{T}$ . Clearly  $(1 - p(\tau)) \in [0, 1]$  is then the probability that the Court rules  $\mathcal{W}$ . Notice that we are allowing  $p(\tau) = 0$  since (for the moment) we are not restricting  $\tau$  to be strictly greater than the  $\underline{\tau}$  of the statement of Proposition 3.

**Step 2.** The value function  $\mathcal{Z}(\tau, \sigma^*)$  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 Lemma A.3. Using (i) and (ii) of Lemma A.3, we can construct a decreasing sequence  $\{\tau_n\}_{n=0}^m$  with  $\tau_m = \tau$ ,  $\tau_0 \geq \tau^*$ , and  $\tau_{n+1} = \tau^{-1}(\tau_n)$  for every  $n = 0, \dots, m - 1$  so that  $\tau_n \in \mathcal{I}^n$  for every  $n = 0, \dots, m$ .

Since  $\tau_0 \geq \tau^*$  it is immediate that

$$\mathcal{Z}(\tau_0, \sigma^*) = \tau_0 \Pi^{\mathcal{A}}(\mathcal{T}) + (1 - \tau_0) \Pi^{\mathcal{A}}(\mathcal{W}) \quad (\text{A.6})$$

Proceeding recursively backwards from  $\tau_0$  (that is, increasing the index  $n$ ), directly from the properties of  $\sigma^*$  in Proposition 3, we get that for every  $n = 0, \dots, m - 1$

$$\begin{aligned} \mathcal{Z}(\tau_{n+1}, \sigma^*) &= \tau_{n+1}[(1 - \delta)\Pi^A(\mathcal{T}) + \delta\mathcal{Z}(\tau_{n+1}, \sigma^*)] + (1 - \tau_{n+1})\{p(\tau_{n+1})[(1 - \delta)\Pi^A(\mathcal{T}) + \delta\mathcal{Z}(\tau_n, \sigma^*)] \\ &\quad + (1 - p(\tau_{n+1}))[(1 - \delta)\Pi^A(\mathcal{W}) + \delta\mathcal{Z}(\tau_{n+1}, \sigma^*)]\} \end{aligned} \tag{A.7}$$

For future reference, we also note that, using Lemma A.1 and Lemma A.2, it is immediate that the right-hand side of (A.7) is increasing in  $p(\tau_{n+1})$ .

Let  $\tau < \tau^*$  be given. Let  $m \geq 1$  be such that  $\tau \in \mathcal{I}_m$ . Construct the associated sequences  $\{\tau_n\}_{n=0}^m$  as in Step 2. (Recall that, by construction  $\tau_m = \tau$ .)

**Step 3.** We can now construct the probability  $p(\tau) \in [0, 1]$  with which, according to  $\sigma^*$ , the Court rules  $\mathcal{T}$ . Notice that once again we are allowing  $p(\tau) = 0$  since (for the moment) we are not restricting  $\tau$  to be strictly greater than the  $\underline{\tau}$  of the statement of Proposition 3.

We construct the value function backwards as in Step 2, beginning with (A.6). Beginning with  $n = 0$  consider the equality

$$(1 - \delta)\Pi^P(\mathcal{W}) + \delta\mathcal{Z}(\tau_{n+1}, \sigma^*) = (1 - \delta)\Pi^P(\mathcal{T}) + \delta\mathcal{Z}(\tau_n, \sigma^*) \tag{A.8}$$

where of course the left-hand side depends on  $p(\tau_{n+1})$  as determined by (A.7), while the right-hand side is given, either because  $n = 0$  (and hence  $\mathcal{Z}(\tau_0, \sigma^*)$  is given by (A.6)) or because the values of  $p(\tau_h)$  with  $h = n, \dots, 1$  have been set in previous rounds of the recursive procedure we are describing here.

The values of  $p(\tau_n)$  for  $n = 1, \dots, m$  are then set proceeding recursively backwards (increasing  $n$ ). Since (A.7) is increasing as we noted in Step 2 above, by Assumption 1, the following three cases are exhaustive of all possibilities.

- (i) If (A.8) cannot be satisfied because for every value of  $p(\tau_n)$  its left-hand side is strictly greater than the right-hand side we set  $p(\tau_n) = 0$ .
- (ii) If (A.8) cannot be satisfied because for every value of  $p(\tau_n)$  its left-hand side is strictly lower than the right-hand side we set  $p(\tau_n) = 1$ .
- (iii) If (A.8) can be satisfied, then we set the value of  $p(\tau_n)$  so that it in fact holds.

**Step 4.** Let  $\tau < \tau^*$  be given and consider the probabilities constructed in Step 3 and the overall  $\sigma^*$  of Proposition 3. Taking the continuation equilibrium as given, the Court has no profitable deviation available from choosing  $\mathcal{T}$  with probability  $p(\tau)$  and  $\mathcal{W}$  with complementary probability  $1 - p(\tau)$  as prescribed by  $\sigma^*$ .

**Proof.** Given  $\tau < \tau^*$ , let  $m$  and the sequences  $\{\tau_n\}_{n=0}^m$ , and  $\{p(\tau_n)\}_{n=0}^{m-1}$  be as in Step 3. Recall that by construction  $\tau = \tau_m$ .

Clearly, the left-hand side of (A.8) is the Court's (ex-post) continuation payoff after a choosing  $\mathcal{W}$  with probability one. Similarly, the right-hand side of (A.8) is the Court's (ex-post) continuation payoff after a choosing  $\mathcal{T}$  with probability one.

Hence if we are in case (i) of Step 3 it is optimal to set  $p(\tau_m) = 0$  and if we are in (ii) it is optimal to set  $p(\tau_m) = 1$ . If we are in case (iii) of Step 3, then the Court is indifferent between choosing  $\mathcal{W}$  and choosing  $\mathcal{T}$ , hence setting any value of  $p(\tau_m) \in [0, 1]$  is optimal in this case.  $\square$

**Step 5.** Consider the probabilities associated with  $\sigma^*$  computed in Step 3 above and assume  $\tau \in (\tau_*, \tau^*)$ . Then  $p(\tau) > 0$ .

**Proof.** We deal first with the claim pertaining to  $\tau \in (\tau_*, \tau^*)$ . By definition of  $\tau^*$  and  $\tau_*$  it must be that (6) holds. Rearranging terms, this implies directly that

$$(1 - \delta)\Pi^P(\mathcal{T}) + \delta[\tau(\tau)\Pi^A(\mathcal{T}) + (1 - \tau(\tau))\Pi^A(\mathcal{W})] > (1 - \delta)\Pi^P(\mathcal{W}) + \delta[\tau\Pi^A(\mathcal{T}) + (1 - \tau)\Pi^A(\mathcal{W})] \tag{A.9}$$

This, using (A.7), implies that setting  $p(\tau_{n+1}) = 0$  makes the right-hand side of (A.8) strictly greater than the left-hand side. Hence, we cannot be in case (i) of Step 3 and hence it cannot be that  $p(\tau) = 0$ .  $\square$

**Step 6.** Consider a  $\tau < \tau_*$ . Let  $\{\tau_n\}_{n=0}^m$  and  $\{p(\tau_n)\}_{n=0}^{m-1}$  be as in Step 3 associated with  $\sigma^*$ . Then, if  $p(\tau_m) > 0$ , it must be the case that  $p(\tau_n) > 0$  for every  $n = m - 1, \dots, 1$ .

**Proof.** To see that this must be the case observe that if the claim were false then, using Step 5, for some  $m - 1 \leq n \leq 1$  we should have that  $p(\tau_{n+1}) > 0$  and  $p(\tau_n) = 0$ , with  $\tau_{n+1} < \tau_*$  and  $\tau_n \leq \tau_*$ . Following the prescription of  $\sigma^*$  which entails choosing  $\mathcal{T}$  with positive probability, the Court's (ex-post) continuation payoff is

$$(1 - \delta)\Pi^P(\mathcal{T}) + \delta[\tau_n\Pi^A(\mathcal{T}) + (1 - \tau_n)\Pi^A(\mathcal{W})] \tag{A.10}$$

since after the equilibrium transition from  $\tau_{n+1}$  to  $\tau_n$ , all Courts will choose  $\mathcal{W}$  whenever they are not constrained by precedents to choose  $\mathcal{T}$ .

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

$$(1 - \delta)\Pi^P(\mathcal{W}) + \delta\{\tau_{n+1}(1 - \delta)\Pi^A(\mathcal{T}) + (1 - \tau_{n+1})(1 - \delta)\Pi^A(\mathcal{T}) + \delta[\tau_n\Pi^A(\mathcal{T}) + (1 - \tau_n)\Pi^A(\mathcal{W})]\} \quad (\text{A.11})$$

where the term that multiplies  $(1 - \tau_{n+1})$  embodies the fact that the Court following the current one will follow the equilibrium prescription and pick  $\mathcal{T}$  with positive probability (possibly equal to one), but after the precedents transition from  $\tau_{n+1}$  to  $\tau_n$ , all Courts will choose  $\mathcal{W}$  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 (6) holds as a weak inequality, it is immediate that the quantity in (A.11) is larger than the quantity in (A.10). Hence the deviation to  $\mathcal{W}$  with probability one is profitable, and this establishes our claim.  $\square$

**Step 7.** Consider the probabilities associated with  $\sigma^*$  computed as in Step 3 above. Then there exists a  $\underline{\tau} \in (0, \tau_*)$  such that  $\tau \leq \underline{\tau}$  implies that  $p(\tau) = 0$ .

Consider a  $\tau < \tau_*$ . Let  $\{\tau_n\}_{n=0}^m$  and  $\{p(\tau_n)\}_{n=0}^{m-1}$  be as in Step 3 associated with  $\sigma^*$ . Let  $q$  be the bound on the number of realized  $\mathcal{T}$  decisions along any realized equilibrium path identified in Proposition 2. Let the  $\mathcal{I}_m$  be the half-open intervals defined in Lemma A.3 above. Suppose that for some  $m > q$  we have that  $\tau \in \mathcal{I}_m$  and  $p(\tau) > 0$ . Then using Step 6, we know that  $p(\tau_n) > 0$  for every  $n = m, \dots, 1$ . Hence the equilibrium path generated by  $\sigma^*$  would have to exceed the bound  $q$  with positive probability. Since this contradicts Proposition 2 we conclude that  $\tau \in \mathcal{I}_m$  and  $m > q$  imply that  $p(\tau) = 0$ , which clearly suffices to prove the claim.  $\square$

**Step 8.** To conclude the proof of Proposition 3 we simply notice that the construction in Step 3 yields an equilibrium which, as a consequence of Steps 5 and 7, satisfies properties (i) and (ii) as required by the statement of the proposition.  $\square$

## References

- Anderlini, L., Felli, L., Riboni, A., 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.
- Backer, S., Mezzetti, C., 2012. A theory of rational jurisprudence. *Journal of Political Economy* 120 (3), 513–551.
- Black, H.C., 1886. The principle of stare decisis. *The American Law Register* (1852–1891) 34, 745–757.
- Bond, P., 2008. Persistent Court corruption. *The Economic Journal* 85, 58–80.
- Bond, P., 2009. Contracting in the presence of judicial agency. *The B.E. Journals in Theoretical Economics* 9 (1) (Advances), Article 36.
- Bustos, A., 2008. A dynamic theory of common law courts. *Northwestern Law & Econ Research Paper* No. 08-12.
- Calabresi, G., 1982. *A Common Law for the Age of Statutes*. Harvard University Press, Cambridge.
- Cardozo, B.N., 1921. *The Nature of the Judicial Process*. Yale University Press. Available via Google Books, New Haven.
- Daughety, A.F., Reinganum, J.F., 1999. Stampede to judgement: Persuasive influence and herding behavior by Courts. *American Law and Economics Review* 1, 158–189.
- Daughety, A.F., Reinganum, J.F., 2000. Appealing judgements. *The RAND Journal of Economics* 31, 502–525.
- Gennaioli, N., Shleifer, A., 2007a. Overruling and the instability of law. *Journal of Comparative Economics* 35, 309–328.
- Gennaioli, N., Shleifer, A., 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. *The Journal of Legal Studies* 7, 393–406.
- Hadfield, G.K., 1992. Bias in the evolution of legal rules. *The Georgetown Law Journal* 80, 583–616.
- Hassler, J., Rodríguez Mora, J.V., 2007. *Political commitment and loss aversion*. University of Edinburgh. Mimeo.
- Kaplow, L., Shavell, S., 2002. *Fairness Versus Welfare*. Harvard University Press, Cambridge, MA.
- Kent, J., 1896. *Commentaries on American Law*, vol. I, XIV edn. Little, Brown and Company, Boston, MA. Available via Google Books. First edn. 1826.
- Kydland, F.E., Prescott, E.C., 1977. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 87, 473–492.
- Legros, P., Newman, A.F., 2002. Courts, contracts, and interference. *European Economic Review* 46, 734–744.
- Levy, G., 2005. Carrierist judges and the appeals process. *The RAND Journal of Economics* 36, 275–297.
- Maskin, E., Tirole, J., 2001. Markov perfect equilibrium I. Observable actions. *Journal of Economic Theory* 100 (2), 191–219.
- Niblett, A., Posner, R.A., Shleifer, A., 2010. The evolution of a legal rule. *The Journal of Legal Studies* 39 (2), 325–358.
- Nordhaus, W., 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. The MIT Press, Cambridge, MA.
- Oxford English Dictionary, 2013. Online.
- Peters, C.J., 1996. Foolish consistency: On equality, integrity, and justice in stare decisis. *The Yale Law Journal* 105, 2031–2115.
- Phelan, C., 2006. Public trust and government betrayal. *Journal of Economic Theory* 127 (1), 27–43.
- Phelps, E.S., Pollak, R.A., 1968. On second-best national saving and game-equilibrium growth. *Review of Economic Studies* 35, 185–199.
- Ponzetto, G.A.M., Fernandez, P.A., 2008. Case law vs. statute law: an evolutionary comparison. *The Journal of Legal Studies* 37, 379–430.
- Ponzetto, G.A.M., Fernandez, P.A., 2012. Stare decisis: rhetoric ad substance. *Journal of Law Economics & Organization* 28 (2), 313–336.
- Posner, R.A., 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. *The Journal of Legal Studies* 6, 65–82.
- Rubin, P., 1977. Why is the common law efficient?. *The 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*. The MIT Press, Cambridge, MA.
- Shavell, S., 1995. The appeals process as a means of error correction. *The Journal of Legal Studies* 24, 379–426.
- Spier, K.E., 1992. The dynamics of pretrial negotiation. *Review of Economic Studies* 59 (1), 93–108.
- Spitzer, M., Talley, E., 2000. Judicial auditing. *The Journal of Legal Studies* 29, 649–683.
- Summers, R., 1992. *Essays on the Nature of Law and Legal Reasoning*. Duncker and Humblot, Berlin.
- Tirole, J., 2005. *The Theory of Corporate Finance*. Princeton University Press, Princeton, NJ.