

Collegio Carlo Alberto

Statute Law or Case Law?

Luca Anderlini
Leonardo Felli
Alessandro Riboni

Working Paper No. 83
December 2008
www.carloalberto.org

Statute Law or Case Law?*

LUCA ANDERLINI
(Georgetown University)

LEONARDO FELLI
(London School of Economics)

ALESSANDRO RIBONI
(Université de Montréal)

August 2008

Abstract. In a Case Law regime Courts have more flexibility than in a Statute Law regime. Since Statutes are inevitably incomplete, this confers an advantage to the Case Law regime over the Statute Law one. However, all Courts rule *ex-post*, after most economic decisions are already taken. Therefore, the advantage of flexibility for Case Law is unavoidably paired with the potential for time-inconsistency. Under Case Law, Courts may be tempted to behave myopically and neglect *ex-ante* welfare because, *ex-post*, this may afford extra gains from trade for the parties currently in Court.

The temptation to behave myopically is traded off against the effect of a Court's ruling, as a precedent, on the rulings of *future* Courts. When Case Law matures this temptation prevails and Case Law Courts succumb to the time-inconsistency problem. Statute Law, on the other hand pairs the lack of flexibility with the ability to commit in advance to a given (forward looking) rule. This solves the time-inconsistency problem afflicting the Case Law Courts. We conclude that when the nature of the legal environment is sufficiently heterogeneous and/or changes sufficiently often, the Case Law regime is superior: flexibility is the prevailing concern. By the same token, when the legal environment is sufficiently homogeneous and/or does not change very often, the Statute Law regime dominates: the ability to overcome the time-inconsistency problem is the dominant consideration.

JEL CLASSIFICATION: C79, D74, D89, K40, L14.

KEYWORDS: Statute Law, Case Law, Flexibility, Rigidity, Time-Inconsistency, Precedents.

ADDRESS FOR CORRESPONDENCE: Leonardo Felli, London School of Economics, Houghton Street, London WC2A 2AE, UK, lfelli@econ.lse.ac.uk

*Part of the research work for this paper was carried out while Luca Anderlini was visiting the "Ente Einaudi" 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, Gillian Hadfield, Andrea Mattozzi, Jean-Laurent Rosenthal, Alan Schwartz, David Scoones, Jean Tirole and audiences at the American Law and Economics Association (2006), Caltech, CIFAR, Ente Einaudi Rome, IIES Stockholm, LBS, LSE Law and Economics Forum, Northwestern University, SED Meetings 2008, Stanford GSB, University of Glasgow, University of Edinburgh, University of Illinois at Urbana-Champaign, UCLA, UQAM, WZB Berlin, and the 2006 BP Lecture at the LSE. Copyright 2008 by Luca Anderlini, Dino Gerardi and Roger Lagunoff. Any opinions expressed here are those of the authors and not those of the Collegio Carlo Alberto.

1. Introduction

1.1. Motivation

Law never is, but is always about to be. It is realized only when embodied in a judgment, and in being realized, expires. There are no such things as rules or principles: there are only isolated dooms. [...]

[...] *No doubt the ideal system, if it were attainable, would be a code at once so flexible and so minute, as to supply in advance for every conceivable situation the just and fitting rule. But life is too complex to bring the attainment of this ideal within the compass of human powers. — Benjamin Cardozo (1921).*

At face value, US Supreme Court Justice Benjamin Cardozo is a lot wiser than Italian legislators in the following attempt to prescribe rules well beyond the powers of their “compass.”

If the birth takes place during a railway trip, the declaration must be rendered to the railroad officer responsible for the train, who will prepare a transcript of verbal declarations, as prescribed for birth certificates. Said railroad officer will hand over the transcript to the head of the railroad station where the train next stops. The head of such station will transmit the documents to the local registrar’s office to be appropriately recorded. — Law of the Republic of Italy (2000)¹

Even abstracting from such misguided attempts to fine-tune legislation, a key question remains. Is the pragmatism of Case Law simply always superior to the rigidity of Statute Law? Are there universes in which Statute Law is instead superior to Case Law?

¹This the text of Article 40 of the regulations for registrar’s offices, issued as Decree Number 393 of November 3rd 2000 of the President of the Republic of Italy. Regulations being issued to ensure the streamlining of procedures, as prescribed by Article 2, comma 12, of Law Number 15 of May 1997 of the Republic of Italy. Translation by the authors.

The original Italian text is: “Se la nascita avviene durante un viaggio per ferrovia, la dichiarazione deve essere fatta al responsabile del convoglio che redige un processo verbale con le dichiarazioni prescritte per gli atti di nascita e lo consegna al capo della stazione nella quale si effettua la prima fermata del convoglio. Il capo della stazione lo trasmette all’ufficiale dello stato civile del luogo, per la trascrizione.” The original reference in Italian Law is: “Articolo 40, Decreto del Presidente della Repubblica 3 Novembre 2000 n. 396. Regolamento per la revisione e la semplificazione dell’ordinamento dello stato civile, a norma dell’articolo 2, comma 12, della legge 15 maggio 1997, n. 127.” See, for instance, <http://www.normeinrete.it/>

After all Statute Law was *the* prevailing system throughout a substantial part of organized human societies for many centuries after 529 AD.² Is it then that the ascent of Case Law is like a scientific discovery? It just was not known before the 11th or 12th century,³ and those who discovered it and started using it became unambiguously better off; just like, say, penicillin after 1929. Once one poses the question in these terms, surely the unambiguous dominance view seems too simplistic to be final.

Our goal here is to build a simple stylized model in which, depending on the value of some significant parameters, which can be interpreted as embodying the speed of social and/or technological change, Case Law sometimes performs better than Statute Law while the reverse can also be true. Our analysis also affords us insights into the *dynamics of precedents* in a Case Law regime.

There does not seem to be a general consensus as to whether the distinction we analyze here between Statute Law and Case Law corresponds in any general way to the distinction between Civil and Common Law, and we do not purport to resolve, or even fully describe, the debate here. It is tempting, however, to draw a parallel in this way since at least historically Common Law relied on few, if any, statutes while Civil Law starts from a large body of statutes rooted in Roman Law dating back to the sixth century. In both Common and Civil Law the body of statutes has expanded dramatically through time (Calabresi, 1982), which makes the parallel problematic, and “pure” forms of either system hard to identify (Von Mehren, 1957, Ch. 16).⁴

However, we believe that our analysis has at least some *normative* implications concerning the distinction between Civil and Common Law. This is because the gaps left open by the Statute Book are filled by the Courts according to different criteria in the two systems. In a Common Law regime the gaps are filled utilizing the body of *applicable precedents*, which is what we model below. In a Civil Law system the gaps are filled by *interpretation* of the code.

²The “Corpus Juris Civilis,” almost universally regarded as the origin of Statute Law, was issued by the Byzantine Emperor Justinian I between 529 and 534 AD.

³The English system of Common Law (originally almost pure Case Law) is universally agreed to have become fully established between the beginning of the reign of William the Conqueror as King of England in 1066 and the end of the reign of Henry II of England in 1189.

⁴Hadfield (2007, 2008) systematically scrutinize the key differences between Common and Civil Law systems. These work address the question of how these differences affect the “dynamic quality” of the Law. We prefer to stay with the safer distinction between Case and Statute Law, which suffices for the purposes of this paper.

At least in the world we model here, the use of precedents stands out as a more (economically) efficient way to fill the gaps. Common Law *adapts* via the use of precedents, while Civil Law changes little unless the Statute Book itself is changed. If one were designing Civil Law and Common Law from scratch, then it would be efficient to strive for more detailed legislation in the Civil Law than in the Common Law world. If this were the case, in this re-designed world, the distinction we make between Statute and Case Law *would* broadly correspond to the distinction between Civil and Common Law.

Before we move on, it is also important to mention a large and influential body of empirical literature known as “Law and Finance” which examines the relative performance of Common and Civil Law in Financial and related markets.⁵ We believe that our results lend support to its main finding — namely that Common Law dominates Civil Law in this fast-paced section of the economy. We return to this point in Section 6 below where we analyze the conditions under which the Case Law regime dominates the Statute Law regime.

1.2. Preview

We model both the Statute Law regime and the Case Law regime in a way that is designed to bring the differences into stark relief, more than capture the fact that the distinction between the two can often be subtle and hard to pinpoint precisely. Our model comprises an ex-ante heterogeneous “pool” of cases; a draw from this pool materializes each period. Under Case Law, in each period a Court of Law can, in principle, either take a forward looking, *tough*, or a myopic, *weak* decision.

Under Statute Law, *all* Courts are constrained to behave in the same way (by the relevant part of the “Statute Book”).

Under Case Law, each Court may be either constrained by precedents (which evolve according to a dynamic process specified below) or unconstrained.⁶ In the latter case the

⁵See for instance La Porta, Lopez-de-Silanes, and Shleifer (2005), La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1997, 1998, 1999, 2002), Lombardo and Pagano (1999, 2002).

⁶In reality, of course, it is seldom the case that a Case Law 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.

Case Law Court has complete *discretion* to either take the tough or the weak decision.⁷

Our point of departure is the observation that under Case Law, whenever a Court of Law exercises *discretion* it does so necessarily *ex-post*. It is hard to argue with the view that Courts (if at all) intervene *ex-post* in the parties' relationship. This affords the Case Law Courts the flexibility to fine tune its rulings to the realized state of nature. Statute Law, on the contrary, by its assumed incompleteness, does not have the possibility to make rulings contingent on the realized state and hence commits all Courts to the same, predetermined *ex-ante*, decision.

This has far reaching implications for the behavior of Courts under the two regimes. Under Case Law, when a Court exercises discretion on whether to take a tough or a weak decision, the *ex-ante* actions of the parties no longer matter because the parties' strategic decisions are *sunk*. This biases the Court's decision away from *ex-ante* efficiency (in our stylized model always towards weak decisions). In short, under Case Law, because Courts exercise discretion (when they in fact do) *ex-post*, they suffer from a *time-inconsistency* problem. If they just maximize the (*ex-post*) welfare of the parties in Court, they would choose the weak decision. What affords Case Law Court the possibility to make contingent rulings—the fact that a Court decides only at an *ex-post* stage—is also the source of the *time-inconsistency* (*present-bias*) problem that affects these Courts.⁸

Under Case Law, the Courts' bias towards weak decisions is mitigated, although not entirely resolved, by the *dynamics of precedents*. Each Court is tempted to take the weak decision even when it should not do so. However, 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 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

⁷We 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 Case Law 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.

⁸The 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 Case Law Courts in our set-up. We use the two terms in a completely interchangeable way.

gain from a tough one, via the dynamics of precedents.

The time inconsistency problem that characterizes the Case Law regime simply does not arise under Statute Law. In this regime, all Courts are committed in advance to a given decision, determined by the Statute Book. In our stylized world, under Statute Law the Courts are completely inflexible. They cannot tailor their decision to the case drawn each period from a heterogeneous pool. Inflexibility is the cost that the Statute Law regime bears while solving the time inconsistency problem.

Our key finding is that the time-inconsistency problem prevents the Case Law regime from reaching full efficiency. This, surprisingly, is true under very general conditions on the dynamics of precedents, and regardless of the rate at which the future is discounted, provided it is positive. Eventually, under Case Law, 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. It is then relatively straightforward to argue that if the heterogeneity of the pool of cases that come before the Courts is “sufficiently small,” then Statute Law will be superior to Case Law. As the pool of cases becomes more and more homogeneous, the loss from the inflexibility of Statute Law eventually becomes smaller than the loss from the time-inconsistency problem under Case Law.

Under some further restrictions on the mechanics that govern the dynamics of precedents, we are able to characterize more stringently the equilibrium behavior of the model. Besides being of independent interest as we will argue below, this gives us the chance to establish conditions of both a legal and an economic nature under which Case Law dominates Statute Law. In particular, we verify that when the degree of heterogeneity of the pool of cases is high the optimal regime is in fact Case Law.

The degree of heterogeneity of the pool of cases in our model can easily be reinterpreted as the “rate of legal innovation.” Our results then lead us to conclude that when this is high — for instance in finance — Case Law dominates, while when it is low — for instance ownership rights, or inheritance law — Statute Law is instead the optimal regime.

Our findings rely on a characterization of the evolution of precedents through time in a Case Law regime. At least since Cardozo (1921), the economic efficiency properties of

this process have been the subject of intense scrutiny.⁹ In these writings, we often find a hypothesized “convergence” toward efficient rules under Case Law (Posner, 2004). How do our results stack against this hypothesis then? Roughly speaking, we find that, in our model, on the one hand the evolution of precedents *improves welfare through time* (see Lemma B.1 below) but on the other *it does not yield efficient rules* in the limit.

1.3. Related Literature

We begin our review of related literature with some papers that are directly related to ours. The hypothesis that Common Law is efficient (and, possibly, superior to Statute Law) has been widely investigated by the literature on Law and Economics. According to Posner (2003), the most influential scholar to endorse this view, judge-made laws are more efficient than statutes mainly because Courts, unlike legislators, have personal incentives to maximize efficiency.¹⁰ 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.

More recently, Gennaioli and Shleifer (2007b) have taken up the claim by (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 the efficiency of Case Law on average, since judges add new dimensions to the adjudication by distinguishing the case from precedents. But on the other, Gennaioli and Shleifer (2007b) show that Case Law reaches full efficiency only under very implausible conditions.¹¹

Similarly to us, two recent papers have explicitly compared judge-made laws and statutes. In a pioneering paper, Glaeser and Shleifer (2002) analyze Common Law (Independent Juries) and Civil Law (Bright Line Rules) in a static model with particular emphasis to the ability of each system to control law enforcers. Ponzetto and Fernandez (2008) compare Case Law and Statute Law in a dynamic setting with a focus on the evolution of precedents and statutes

⁹We return to this point in Subsection 1.3. when we review some related literature.

¹⁰In Hadfield (1992), however, efficiency-oriented Courts may fail to make efficient rules because of the bias in the sample of cases observed by Courts.

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

over time. In a model where judges have idiosyncratic preferences and overruling is costly, they show that Case law converges to an asymptotic distribution with mean equal to the efficient rule. In the long run, as precedents become more consistent, Case Law eventually dominates Statute Law by making better and more predictable decisions.

Aside from a variety of modeling choices, one key difference between Ponzetto and Fernandez (2008) and our work is our focus on the potential time-inconsistency generated by ex-post Court intervention. Compared to judicial bias, the present-bias temptation has quite different implications in terms of dynamics of precedents. Moreover, a central ingredient of our model of the Case Law regime is the disciplinary role of stare decisis. In Ponzetto and Fernandez (2008) 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), the rule of precedent induces more extremism, which is welfare reducing. 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 cost of overruling the precedent.

A key contribution that is directly related to our work is a recent empirical study by Niblett, Posner, and Shleifer (2008). They refer to a specific tort doctrine known as the economic loss rule (ELR hereafter). In its most general interpretation this rule states that “one cannot sue in tort for economic loss” unless that loss is accompanied by personal injury or property damage.¹² In other words, the only way in which a plaintiff can bring, say, a contractor to Court asking to be compensated for an economic loss created by the contracted work is if such a loss is covered, for instance, by the warranty specified in their contract.¹³ Taking for granted the ex-ante efficiency of the ELR, it is however 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. In other words, when enforcing ELR the Court may be affected by a time inconsistency problem of the type we posit here.

¹² “‘Economic loss’ thus just means a personal injury or property damage.” (Niblett, Posner, and Shleifer, 2008, p. 4)

¹³In 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).

Niblett, Posner, and Shleifer (2008) provide a unique empirical test of whether precedents in the past 35 years converged to what the authors regard as the clearly ex-ante efficient rule: the ELR. They assemble and analyze a remarkable data set of 465 State Court appellate decision involving the application of the ELR within the construction industry in a number of US States from 1970 to 2005.¹⁴ The key question addressed by this pathbreaking empirical study is whether in the past 35 years Case Law in this area has in fact converged to the ELR. They conclude that it has not.

In a nutshell Niblett, Posner, and Shleifer (2008) show that while convergence 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. On the other hand, in our model, Case Law matures and settles into a regime in which the Case Law 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 at least two, not mutually exclusive, possibilities. Whether there could be other sources of change in the data analyzed by Niblett, Posner, and Shleifer (2008). And whether our model with a simple addition of other sources of noise and a further mechanism for path-dependence could yield a theoretical construct in agreement with the patterns observed in their data. For example, if 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. These issues are obviously ripe for future research, but clearly remain beyond the scope of the present paper.

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

¹⁴The ELR was first applied to disputes in the construction industry in the 1970's. See Holmes's opinion for the Supreme Court in *Robins Dry Dock & Repair Co. v. Flint*, 275 U.S. 303, 308-310 (1927).

(1921) and subsequently Posner (2003).¹⁵ It is mainly to make sure that our results can be clearly attributed to the sources we choose to focus on (rigidity versus time-inconsistency). Introducing judicial bias may well have ambiguous effects on welfare when Courts have more discretion, because it changes the incentives of the current Court to constrain future Courts via precedents.¹⁶

We also ignore the distinction between “lower” and “appellate” Courts. 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), but, again, its differential impact in the Case and Statute Law regimes is far from obvious both theoretically and empirically. As with judicial bias, we prefer to maximize the transparency of our results and leave the distinction out of the model. In our model, under Case Law, *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. Nevertheless we proceed as we do in the belief that the general flavor of our results would survive in a richer model.¹⁷

The commitment value provided by rules has not, to our knowledge, been pointed out in the context of judicial decision making. The vast legal literature on rules versus standards has instead focused on other merits of rules.¹⁸ For instance, the benefit of predictability, which is likely to result into more adherence to norms, more productive behavior, fewer disputes, and more settlements. Rules reduce arbitrariness and bias: they bind a decision maker to respond in a determinate way to some specific triggering facts.¹⁹ Finally, rules reduce the cost

¹⁵There is a flourishing literature on the effects (and remedies for) judicial bias interpreted in a broad sense that ranges from “idiosyncracies” in the judges’ preferences Bond (2007), Gennaioli and Shleifer (2007b), among others, to plain “corruption” of the Courts Ayres (1997), Bond (2008), Legros and Newman (2002), among others.

¹⁶See again Gennaioli and Shleifer (2007b), where judicial polarization may actually improve the efficiency of the Common Law.

¹⁷For instance Gennaioli and Shleifer (2007b) insist, realistically, that the Court that 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.

¹⁸See, for example, Kaplow (1992), Posner (1990), Sullivan (1993), Sunstein (1995).

¹⁹According to Glaeser and Shleifer (2002), this would explain why France, where local feudal lords were powerful and often in conflict with the king, opted for a centralized legal system where royal judges were constrained by Bright Line Rules. Conversely, England, where local lords were less powerful and less able to

of enforcement: they minimize the need of time-consuming balancing of all relevant interests and facts.²⁰

In contrast, the literature on rules versus discretion in macroeconomics has long recognized that precommitment to a rule involves the loss of flexibility to respond to unforeseen contingencies. The question of how much commitment is desirable in a stochastic world has been addressed, among many others, by Rogoff (1985), who emphasizes the cost in terms of output stabilization of delegating monetary policy to a conservative banker, and Obstfeld (1997), who studies policy rules with escape clauses, which benefit from precommitment while still retaining some flexibility in exceptional circumstances.

Athey, Atkeson, and Kehoe (2005) analyze the trade-off between commitment and flexibility when a time-inconsistent central bank has access to private information about the state of the economy. By taking a dynamic mechanism design perspective, they find that the optimal degree of monetary policy discretion takes the form of an inflation cap that allows discretion as long as the inflation rate is below a certain value.²¹ Compared to this recent literature, in our model constraints on discretion are not exogenously imposed by an optimal mechanism designer, but arise *endogenously* as a result of Courts' decisions and the system of precedents.

Phelan (2006), and Hassler and Rodríguez Mora (2007) study 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 however different from ours. 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

bias local justice, opted for a decentralized adjudication of disputes (Independent Juries).

²⁰Shavell (2007) studies the optimal scope of discretion of a rule, which should balance the informational advantage of adjudicators and the cost of delegation due to the adjudicators' bias. Kaplow (1992) argues that when the frequency with which similar cases arise is high, it is better to incur the one-time, up-front investment to create a rule.

²¹Similarly, Amador, Angeletos, and Werning (2006) find that a minimum saving rule optimally resolves the tension between discretion and time-consistency in the context of a consumption-saving model with quasi-geometric discounting.

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.

Finally, our interest is in the comparison of the regimes of Case and Statute Law in the *economic* sphere of course, particularly within the realm of what economists call Contract Theory. During the last two decades, since the seminal work of Grossman and Hart (1986) and Hart and Moore (1990), much energy has been devoted to the analysis of *ex-ante* contracting under an *incompleteness* constraint.²² The focus is on a situation in which *ex-ante* contracting is critical to the parties’ incentives to undertake relationship-specific investments that enhance economic efficiency. The parties’ ability to contract on the relevant variables is assumed to be incomplete. This has proved to be an extremely fertile ground to address a variety of issues of first-order economic importance.²³ Our paper focuses on how the underlying legal regime can also help overall efficiency in the presence of incomplete laws.

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 Appendix A, we give the formal details of one of the three.

In Section 3 we set up the model; we first describe the basic structure of our model the Statute Law regime, and solve for the equilibrium in this case (Proposition 1). We then proceed to lay down the model of precedents and hence our dynamic model of the Case

²²See Kaplow and Shavell (2002a), Section 4, for a general discussion of incomplete contracts and enforcement.

²³To cite but a few contributions, this literature has shed light on vertical and lateral integration (Grossman and Hart, 1986), the allocation of ownership over physical assets (Hart and Moore, 1990), the allocation of authority (Aghion and Tirole, 1997) and power (Rajan and Zingales, 1998) in organizations.

Law regime. In Section 4 we report our first two results that highlight the effect of time-inconsistency on the evolution of precedents (Proposition 2) and the possibility of Statute Law dominance (Proposition 3).

In Section 5 we impose some further restrictions on the precedent technology that allow us to characterize further the equilibrium behavior of our model of the Case Law regime (Proposition 4). Among other things, this construction allows us in Section 6 to identify the conditions under which Case Law dominates Statute Law (Proposition 5). Section 7 concludes the paper.

For ease of exposition, all proofs are in Appendix B. In the numbering of equations, Lemmas, and so on, a prefix of “A” or “B” indicates that the relevant item is in the corresponding 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.²⁴ It is also important that this is not always the case: the considerations that impact the *ex-ante* decision making it differ from the *ex-post* one may, sometimes, be unimportant and, hence, the *ex-post* optimal Court ruling is optimal when viewed from *ex-ante* as well. Optimal decisions that vary across different cases are what gives positive value to the *flexibility* of the Courts.

There are many examples of spheres in which the potential time-inconsistency we work with occurs. Here, we briefly describe two of these that we think are both important and fit well our setup. In the Appendix, we report more extensively on a third example (Anderlini, Felli, and Postlewaite, 2006) that also fits the bill. This is also briefly described below.

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

²⁴The 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 (2002b) distinguish between “welfare” (*ex-ante*) and “fairness” (*ex-post*). Summers (1992) distinguishes between “goal reasons” (*ex-ante*) and “rightness reasons” (*ex-post*).

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.

To complete the example, we observe that in some instances allowing debt restructuring may be optimal both ex-ante and ex-post. This is the case, for instance, if all potential projects are of high enough quality (or the proportion of low quality projects is sufficiently low).

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 it 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.²⁵ 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.²⁶

The model in Anderlini, Felli, and Postlewaite (2006) involves a buyer and a seller in a multiple-widgit model with relationship-specific investment, asymmetric information and incomplete contracting.²⁷ In this world, it may be optimal for a Court to actively intervene

²⁵See 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).

²⁶As before, in some cases the optimal decision is the same. When R&D investment is unimportant the socially optimal breadth of the patent is zero. In other words, it is optimal both ex-ante and ex-post to rule in favor of the infringer.

²⁷See Appendix A for a fully specified numerical version of this model.

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.²⁸

3. The Model

3.1. The Static Environment

The environment in which the Court operates can be either “simple” (denoted \mathcal{S}) or “complex” (denoted \mathcal{C}). Which environment occurs is determined by the realization of a random variable $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$.²⁹ We denote by ρ the probability $\mathcal{E} = \mathcal{C}$.

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. The Court’s payoffs are determined by the ruling it chooses and by the environment, and, critically, they may be different viewed from ex-ante and ex-post. Let $\Pi(\mathcal{R}, \mathcal{E})$ and $\hat{\Pi}(\mathcal{R}, \mathcal{E})$, with $\mathcal{R} \in \{\mathcal{W}, \mathcal{T}\}$ and $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$, denote the ex-ante and ex-post payoffs respectively.

When the environment is simple ($\mathcal{E} = \mathcal{S}$) the optimal ruling is \mathcal{W} both ex-ante and ex-post. In other words

$$\Pi(\mathcal{W}, \mathcal{S}) > \Pi(\mathcal{T}, \mathcal{S}) \quad \text{and} \quad \hat{\Pi}(\mathcal{W}, \mathcal{S}) > \hat{\Pi}(\mathcal{T}, \mathcal{S}) \quad (1)$$

When the environment is complex ($\mathcal{E} = \mathcal{C}$) 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, when the realized state is \mathcal{C} the

²⁸As in the previous two examples, it is possible that the optimal decision both ex-ante and ex-post is that the Court should not intervene. In Appendix A, we argue that this is the case if the number of potential widgets is reduced.

²⁹With a small abuse of terminology, we will sometimes refer to \mathcal{E} as the “state.”

time-inconsistency problem arises. Formally, we have

$$\Pi(\mathcal{W}, \mathcal{C}) < \Pi(\mathcal{T}, \mathcal{C}) \quad \text{and} \quad \widehat{\Pi}(\mathcal{W}, \mathcal{C}) > \widehat{\Pi}(\mathcal{T}, \mathcal{C}) \quad (2)$$

3.2. The Statute Law Regime

We model the Statute Law regime in a deliberately stark and simple way. Since one of our punch-lines is that it can in fact dominate the more flexible Case Law regime this, besides having the virtue of being simple, strengthens our results.

We assume that in the Statute Law regime the legislators constrain all Courts by selecting ex-ante, and once and for all, the ruling that Courts will take. In this world laws are also highly *incomplete*. Under Statute Law, all Courts are constrained to choose the *same* ruling, *regardless* of the state $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$ determining the nature of the environment. The legislators only have one choice. Either they constrain all Courts to choose \mathcal{W} , or they constrain all Courts to choose \mathcal{T} .

The analysis of the Statute Law regime is sufficiently straightforward to allow us to move directly to the full blown dynamic version of the model. Time is indexed by $t = 0, 1, 2, \dots$. A sequence of Courts face a stream of (i.i.d) environments and one set of parties in each period. The planner's (the legislature's) discount factor is $\delta \in (0, 1)$. The optimal Statute Law regime is obtained by picking a *single* ruling $\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}$ so as to solve

$$\max_{\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} (1 - \delta) \sum_{t=0}^{\infty} \delta^t [(1 - \rho)\Pi(\mathcal{S}, \mathcal{R}) + \rho\Pi(\mathcal{C}, \mathcal{R})] \quad (3)$$

The maximization problem in (3) is extremely simple since it is not genuinely dynamic.³⁰ In fact, given the inequalities in (1) and (2) it can easily be solved. We state the following without proof.

Proposition 1. *Statute Law Equilibrium Welfare:* The maximized value of (3) is denoted by $\mathbf{\Pi}_{SL}(\rho)$. We refer to this value as the equilibrium welfare of the Statute Law regime. The ruling that solves the maximization problem (3) is denoted by $\mathcal{R}_{SL}(\rho)$. We refer to this as the equilibrium ruling under Statute Law.

³⁰Problem (3) is clearly equivalent to $\max_{\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} (1 - \rho)\Pi(\mathcal{S}, \mathcal{R}) + \rho\Pi(\mathcal{C}, \mathcal{R})$.

The equilibrium ruling $\mathcal{R} = \mathcal{R}_{SL}(\rho)$ is \mathcal{W} for ρ between 0 and a threshold value $\rho_{SL}^* \in (0, 1)$ and is \mathcal{T} for ρ between ρ_{SL}^* and 1. The threshold ρ_{SL}^* is such that:³¹

$$(1 - \rho_{SL}^*) \Pi(\mathcal{S}, \mathcal{W}) + \rho_{SL}^* \Pi(\mathcal{C}, \mathcal{W}) = (1 - \rho_{SL}^*) \Pi(\mathcal{S}, \mathcal{T}) + \rho_{SL}^* \Pi(\mathcal{C}, \mathcal{T}) \quad (4)$$

The intuition behind Proposition 1 is straightforward. Given the structure of payoffs in (1) and (2) the payoff to \mathcal{W} is larger in the \mathcal{S} environment, while the payoff to \mathcal{T} is larger in the \mathcal{C} environment. It then follows that choosing \mathcal{W} is optimal if the probability of the \mathcal{S} environment is sufficiently large, while choosing \mathcal{T} is optimal if the probability of the \mathcal{C} environment is sufficiently large.

3.3. The Case Law Regime: Time-Inconsistent Courts

As we mentioned before, our model of the Case Law regime 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 statutes, or by precedents) and that only considers the *present* case, without looking at any effect that its ruling might have on future Courts. Then, from (1) and (2) it is immediate that its ruling will be \mathcal{W} regardless of the realized state.

When the environment is \mathcal{C} , an unconstrained Court that intervenes *ex-post* in a contractual dispute, when it considers the present payoffs, will therefore choose the weak ruling \mathcal{W} . Viewed 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 in a Case Law regime.

Before proceeding further it is worth recalling here that from (1) and (2) we immediately know that the time-inconsistency problem does not arise when the environment is \mathcal{S} . This simplifies our analysis but is by no means essential to the basic flavor of our results.

3.4. The Case Law Regime: The Nature of Precedents

Consider the Case Law regime. In each period the environment is \mathcal{S} (simple) with probability $1 - \rho$ and is \mathcal{C} (complex) with probability ρ . Each case (simple or complex) comes equipped

³¹Clearly, when $\rho = \rho_{SL}^*$ both the \mathcal{T} and \mathcal{W} rulings solve problem (3).

with *its own* specific *legal characteristics*, which determine, as we will explain shortly, whether the current *body of precedents* apply.

We model the legal characteristics of the case as random variables ℓ_S and ℓ_C , each uniformly distributed over $[0, 1]$, describing the legal characteristics of the case in the \mathcal{S} and \mathcal{C} environments respectively.³² This allows us to specify the body of precedents in a particularly simple way.

The body of precedents \mathcal{J} is represented by four numbers in $[0, 1]$ so that $\mathcal{J} = (\tau_S, \omega_S, \tau_C, \omega_C)$ with the restriction that $\tau_S \leq \omega_S$ and $\tau_C \leq \omega_C$. Once the nature of the environment (\mathcal{S} or \mathcal{C}) is determined, the legal characteristics of the case are determined as well (ℓ_S or ℓ_C as appropriate).

The interpretation of $\mathcal{J} = (\tau_S, \omega_S, \tau_C, \omega_C)$ is straightforward. The body of precedents is seen to either apply or not apply and in which direction. Say that the environment is \mathcal{S} , then if $\ell_S \leq \tau_S$ the body of precedents constrains the Court to a *tough* decision, if $\ell_S \geq \omega_S$ the body of precedents constrains the Court to a *weak* decision, while if $\tau_S < \ell_S < \omega_S$ the Court has *discretion* over the case. A similar interpretation applies if the environment is \mathcal{C} so that the Court is constrained to take a tough decision, a weak decision, or has discretion according to whether $\ell_C \leq \tau_C$, $\ell_C \geq \omega_C$ or $\tau_C < \ell_C < \omega_C$.

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 do not bind, the case at hand is sufficiently idiosyncratic to escape the doctrine of stare decisis. The assumption that stare decisis either does or does not apply, without intermediate possibilities is obviously an extreme one. It seems a plausible first cut in the modeling of the role of precedents that we wish to pursue here.

Finally, note that in each period the contracting parties *observe* the nature of the environment, 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, under Case Law, in each of the environments, the parties anticipate correctly whether the

³²The 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.

Court will take a tough or weak decision.³³

3.5. *The Case Law Regime: The Dynamics of Precedents*

The present-bias or time-inconsistency problem that afflicts the Courts in a Case Law regime 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 environment and ruling will determine the body of precedents in the next period.

Consider a body of precedents for date t , $\mathcal{J}^t = (\tau_S^t, \omega_S^t, \tau_C^t, \omega_C^t)$. Let $d_S^t = \omega_S^t - \tau_S^t$ and $d_C^t = \omega_C^t - \tau_C^t$ be the probabilities that the t -th Case Law Court has discretion in each of the two possible environments, so that the t -th Case Law Court is constrained by precedents with probability $1 - d_S^t$ and $1 - d_C^t$ in each environment respectively. 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 Case Law Court is *not* constrained by precedents (with probability d_S^t and with probability d_C^t in the two environments respectively), in either environment it can choose the tough or weak decision at its discretion.

A key feature of our model is that a Case Law 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.

³³It 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. Nevertheless, it clearly remains well beyond the scope of this paper.

The discretionary ruling $\mathcal{R}^t \in \{\mathcal{T}, \mathcal{W}\}$ of the t -th Case Law Court and the state of the environment $\mathcal{E}^t \in \{\mathcal{S}, \mathcal{C}\}$, 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 Case Law Court will operate. Therefore, the precedents technology in the Case Law regime can be viewed as a map $\mathcal{J} : [0, 1]^4 \times \{0, 1\} \times \{\mathcal{T}, \mathcal{W}\} \times \{\mathcal{S}, \mathcal{C}\} \rightarrow [0, 1]^4$, so that $\mathcal{J}^{t+1} = \mathcal{J}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$. We will later use the notation $\tau_{\mathcal{S}}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$, $\omega_{\mathcal{S}}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$, $\tau_{\mathcal{C}}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$ and $\omega_{\mathcal{C}}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$ to denote the first, second, third and fourth element of $\mathcal{J}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$.

Typically, the map \mathcal{J} will embody the workings of a complex set of legal mechanisms and constitutional arrangements. It may also embody complex interaction effects that go, for instance across the two environments. 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 of the Case Law regime requires more hypotheses. We return to these at length below.

Next, we turn to what constitutes an equilibrium for our model of the Case Law regime

3.6. The Case Law Regime: Dynamic Equilibrium

We assume that all Case Law 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.³⁴ Critically, when the current Court takes into account the payoffs of future Courts it does so using the *ex-ante* payoffs satisfying (1) and (2) above.

The t -th Case Law Court inherits \mathcal{J}^t from the past. Given \mathcal{J}^t , it first observes the state of the environment $\mathcal{E}^t \in \{\mathcal{S}, \mathcal{C}\}$, then it observes the outcome of the draw that determines the legal characteristics of the case ($\ell_{\mathcal{S}}$ or $\ell_{\mathcal{C}}$, as appropriate). Together with \mathcal{J}^t , this determines whether the t -th period Case Law Court has discretion or not. If it has discretion, the t -th Case Law Court then chooses \mathcal{R}^t and b^t — the ruling and its breadth. Together with \mathcal{E}^t and \mathcal{J}^t this determines \mathcal{J}^{t+1} , and hence the decision problem faced by the $t+1$ -th Case Law

³⁴We interpret δ as the common discount factor shared by the Court and the parties. Notice, however, that δ 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 δ should be part of the legal characteristics of the case. However, this reinterpretation would yield no changes to the role that δ plays in the equilibrium characterization under Case Law (see Sections 4 and 5 below).

Court. If the Case Law 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 Case Law Courts when they are not constrained by precedents. The choice of ruling \mathcal{R}^t depends on both \mathcal{J}^t and \mathcal{E}^t . We let $\mathcal{R}^t = \mathbf{R}(\mathcal{J}^t, \mathcal{E}^t)$ denote this part of the Court's strategy. Similarly, we let the Court's (contingent) choice of breadth be denoted by $b^t = \mathbf{b}(\mathcal{J}^t, \mathcal{E}^t)$. Notice that, in principle, the choices of the t -th Case Law Court could depend on the entire history of past rulings, breadths, environments, 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 and the type of environment \mathcal{E}^t . These are clearly the only "payoff relevant" state variables for the t -th Case Law Court. In this sense our restriction is equivalent to saying that we are restricting attention to the set of *Markov-Perfect Equilibria*.³⁵ We will do so throughout the rest of the paper.

With this restriction, we can simply refer to *the* strategy of the Case Law Court, regardless of the time period t . This will sometime be written concisely as $\boldsymbol{\sigma} = (\mathbf{R}, \mathbf{b})$. 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 Case Law Court, can be written as follows.

$$\begin{aligned} \mathbf{\Pi}(\mathcal{J}^t, \boldsymbol{\sigma}, \rho) = & \\ & (1 - \rho) \{ \tau_S^t \mathbf{\Pi}(\mathcal{S}, \mathcal{T}) + (1 - \omega_S^t) \mathbf{\Pi}(\mathcal{S}, \mathcal{W}) + d_S^t \mathbf{\Pi}(\mathcal{S}, \mathbf{R}(\mathcal{J}^t, \mathcal{S})) \} + \\ & \rho \{ \tau_C^t \mathbf{\Pi}(\mathcal{C}, \mathcal{T}) + (1 - \omega_C^t) \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) + d_C^t \mathbf{\Pi}(\mathcal{C}, \mathbf{R}(\mathcal{J}^t, \mathcal{C})) \} \end{aligned} \quad (5)$$

The interpretation of (5) is straightforward. The first two terms that multiply $(1 - \rho)$ refer to the cases in which the Court is constrained (to a tough and weak decision respectively) in the \mathcal{S} environment. The third term that multiplies $(1 - \rho)$ is the Court's payoff in the \mathcal{S} environment given its discretionary ruling $\mathbf{R}(\mathcal{J}^t, \mathcal{S})$. Similarly, the first two terms that multiply ρ refer to the cases in which the Court is constrained (to a tough and weak decision respectively) in the \mathcal{C} environment. The third and final term that multiplies ρ is the Court's payoff in the \mathcal{C} environment given its discretionary ruling $\mathbf{R}(\mathcal{J}^t, \mathcal{C})$.

Given the (stationary) preferences we have postulated, the overall payoffs to each Case Law Court can be expressed in a familiar recursive form. Let a $\boldsymbol{\sigma}$ be given. Let $\mathbf{Z}(\mathcal{J}^t, \boldsymbol{\sigma}, \rho)$

³⁵See Maskin and Tirole (2001) or Fudenberg and Tirole (1991), Ch.13.

be the expected (as of the beginning of the period) overall payoff to the t -th Case Law Court, given \mathcal{J}^t and σ .³⁶ We can then write this payoff as follows.

$$\begin{aligned} \mathbf{Z}(\mathcal{J}^t, \sigma, \rho) &= (1 - \delta) \mathbf{\Pi}(\mathcal{J}^t, \sigma^t, \rho) + \\ &\delta [(1 - \rho)(1 - d_S^t) + \rho(1 - d_C^t)] \mathbf{Z}(\mathcal{J}^t, \sigma, \rho) + \\ &\delta (1 - \rho) d_S^t \mathbf{Z}(\mathcal{J}(\mathcal{J}^t, \mathbf{b}(\mathcal{J}^t, \mathcal{S})), \mathcal{R}(\mathcal{J}^t, \mathcal{S}), \mathcal{S}), \sigma, \rho) + \\ &\delta \rho d_C^t \mathbf{Z}(\mathcal{J}(\mathcal{J}^t, \mathbf{b}(\mathcal{J}^t, \mathcal{C})), \mathcal{R}(\mathcal{J}^t, \mathcal{C}), \mathcal{C}), \sigma, \rho) \end{aligned} \quad (6)$$

The interpretation of (6) is also straightforward. The first term on the right-hand side is the Court's period- t payoff. The first term that multiplies δ 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 δ is the Court's continuation payoff if the environment at t turns out to be \mathcal{S} and the Court's choices at t are $[\mathcal{R}(\mathcal{J}^t, \mathcal{S}), \mathbf{b}(\mathcal{J}^t, \mathcal{S})]$, while the third term that multiplies δ is the Court's continuation payoff if the environment at t turns out to be \mathcal{C} and the Court's choices at t are $[\mathcal{R}(\mathcal{J}^t, \mathcal{C}), \mathbf{b}(\mathcal{J}^t, \mathcal{C})]$.

Now recall that the t -th Case Law Court decides whether to take a tough or a weak decision (if it is given discretion) and chooses the breadth of its ruling ex-post, after the nature of the environment (\mathcal{S} or \mathcal{C}) is known and the parties' actions are *sunk*. Hence the t -th Case Law 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 (1) and (2) reflecting the Court's present-bias in the \mathcal{C} environment. 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 (6) since all the relevant decisions lie ahead of when the t -th Case Law Court makes its choices.

It follows that, given \mathcal{J}^t and σ , the decisions of the t -th Case Law Court can be characterized as follows. Suppose that the t -th Case Law Court is *not* constrained by precedents to either a tough or a weak decision.³⁷ Then, the values of $\mathcal{R}^t = \mathcal{R}(\mathcal{J}^t, \mathcal{E}^t) \in \{\mathcal{T}, \mathcal{W}\}$ and $b^t =$

³⁶The function $\mathbf{Z}(\cdot)$ is independent of t because we are restricting attention to stationary Markov-Perfect Equilibria.

³⁷ Recall that if the ruling turns out to be constrained by precedents, the t -th Case Law Court does not make any choice and the body of precedents remains the same so that $\mathcal{J}^{t+1} = \mathcal{J}^t$.

$\mathbf{b}(\mathcal{J}^t, \mathcal{E}^t) \in \{0, 1\}$ must solve

$$\max_{\mathcal{R}^t \in \{\mathcal{T}, \mathcal{W}\}, b^t \in \{0, 1\}} (1 - \delta) \widehat{\Pi}(\mathcal{E}^t, \mathcal{R}^t) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t), \boldsymbol{\sigma}, \rho) \} \quad (7)$$

It is now straightforward to define what constitutes an equilibrium in the Case Law regime.

Definition 1. *Case Law Equilibrium Behavior:* An equilibrium under the Case Law regime is a $\boldsymbol{\sigma}^* = [\mathbf{R}^*, \mathbf{b}^*]$ such that, for every $t = 0, 1, 2, \dots$, for every $\mathcal{E}^t \in \{\mathcal{S}, \mathcal{C}\}$ and for every possible \mathcal{J}^t , the pair $[\mathbf{R}^*(\mathcal{J}^t, \mathcal{E}^t), \mathbf{b}^*(\mathcal{J}^t, \mathcal{E}^t)]$ is a solution to the following problem.³⁸

$$\max_{\mathcal{R}^t \in \{\mathcal{T}, \mathcal{W}\}, b^t \in \{0, 1\}} (1 - \delta) \widehat{\Pi}(\mathcal{E}^t, \mathcal{R}^t) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t), \boldsymbol{\sigma}^*, \rho) \} \quad (8)$$

For any given equilibrium behavior as in Definition 1 we can compute the value of the expected payoff to the Case Law Court of period $t = 0$, as a function of the initial value \mathcal{J}^0 . Using the notation we already established, this is denoted by $\mathbf{Z}(\mathcal{J}^0, \boldsymbol{\sigma}^*, \rho)$.

We denote by $\mathbf{\Pi}_{CL}(\mathcal{J}^0, \rho)$ the supremum of $\mathbf{Z}(\mathcal{J}^0, \boldsymbol{\sigma}^*, \rho)$ taken over all possible equilibria of the Case Law regime. With optimistic terminology, we refer to $\mathbf{\Pi}_{CL}(\mathcal{J}^0, \rho)$ as the equilibrium welfare of the Case Law regime given \mathcal{J}^0 .

4. The Possibility of Statute Law Dominance

4.1. Residual Discretion and Zero Breadth

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

Assumption 1. *Residual Discretion:* Assume that \mathcal{J}^t is such that $d_c^t > 0$ and $d_s^t > 0$. Then $\mathcal{J}^{t+1} = \mathcal{J}(\mathcal{J}^t, b^t, \mathcal{R}^t, \mathcal{E}^t)$ is such that $d_c^{t+1} > 0$ and $d_s^{t+1} > 0$, whatever the values of b^t , \mathcal{R}^t and \mathcal{E}^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 a Case Law regime.

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

³⁸It 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.

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

Assumption 2 states that, regardless of the ruling it issues and of the environment, any Case Law 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 of the Case Law regime 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 in the Case Law regime.

Finally, the possibility that a Case Law regime Court might decide to narrow down on purpose the precedential effect that its ruling has 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 precedential value.³⁹

4.2. Mature Case Law

Given σ^* and an initial body of precedents \mathcal{J}^0 , as the randomness in each period is realized (the nature of the environment, 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 Case Law Courts have discretion and will take the tough decision (\mathcal{T}) has an *upper bound*. Case Law eventually “matures,”

³⁹We 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)

and, after it does, all discretionary Case Law Courts succumb to the (time-inconsistent) temptation to rule \mathcal{W} instead of \mathcal{T} .

Proposition 2. *Evolving and Mature Case Law: Let any equilibrium σ^* for the Case Law regime be given.*

Then, there exists an integer m , which depends on δ but not on \mathcal{J}^0 , on ρ , or on the particular equilibrium σ^ , with the following property.*

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

Intuitively, each time a Case Law 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 Case Law Courts cannot decrease through time (see Lemma B.1 in Appendix B). 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} when the environment is \mathcal{C} . In other words, future gains stem from the upwards effect on $\tau_{\mathcal{C}}^t$ that a tough decision today may have. It is then apparent that, since $\tau_{\mathcal{C}}^t$ cannot reach 1, 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} should the environment be \mathcal{C} , 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 improve 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 in state \mathcal{C} ; 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.

Finally, note that stare decisis plays a key disciplinary role in our model of Case Law. In the absence of the rule of precedents, Case Law Courts would always take the present-bias decision whenever they have discretion. Notice the particular mechanisms through which stare decisis provides (limited) commitment in our model. Case Law Courts are not deterred from taking the present-bias decision by the perspective of worsening future precedents as a result of a weak ruling. In fact, Courts can always claim that the case at hand is an idiosyncratic exception and choose the weak decision together with zero breadth. Instead, stare decisis disciplines Case Law Courts (at least for a limited number of periods) by rewarding tough decisions and, more specifically, with the perspective of increasing the probability that future judges are constrained to choose the efficient ruling.

4.3. Dominance of Statute Law

The present-bias temptation to take the weak decision when the environment is \mathcal{C} lowers the equilibrium welfare under the Case Law Regime. Eventually Case Law becomes mature as in Proposition 2. This effect is obviously larger when ρ is larger so that the environment \mathcal{C} is more likely to obtain.

At the same time, when ρ is extreme (near 0 or near 1), the lack of flexibility of the Statute Law regime becomes less and less important. The decision that is optimal for the environment that obtains almost all the time is also almost optimal in expected terms.

Putting these two considerations together leads us to our next result.

Proposition 3. *Statute Law Welfare Dominance: The Statute Law regime yields strictly higher equilibrium welfare than the Case Law regime for high values of ρ .*

More specifically, let any \mathcal{J}^0 be given, and assume that this leaves positive discretion to the first Case Law Court. In other words assume that \mathcal{J}^0 is such that both d_S^0 and d_C^0 are strictly positive.

Then there exists a $\rho_{CL}^ \in (0, 1)$ such that for every $\rho \in (\rho_{CL}^*, 1]$ we have that:⁴⁰*

$$\Pi_{SL}(\rho) > \Pi_{CL}(\mathcal{J}^0, \rho).$$

Proposition 3 establishes that when the pool of cases faced by the Courts is sufficiently homogeneous (ρ is sufficiently close to one) Statute Law dominates Case Law. In other

⁴⁰In general, ρ_{CL}^* depends on \mathcal{J}^0 .

words, the lack of commitment due to the time-inconsistency problem that afflicts the Case Law Courts outweighs the lack of flexibility of Statute Law.⁴¹

5. Equilibrium Characterization

5.1. Mixed Strategies

While the characterization of the equilibrium behavior of our model of the Statute Law regime is extremely straightforward (see Subsection 3.2 above), the same cannot be said of the Case Law regime. To appreciate some of the difficulties involved, recall that from Proposition 2 we know that along any path of resolved uncertainty the Case Law Courts can only take the tough decision \mathcal{T} when they have discretion in environment \mathcal{C} a finite number of times m .

Suppose now that we are in a configuration of parameters (a δ not “too low” is, for instance, necessary) such that in equilibrium the Case Law Courts initially rule \mathcal{T} in environment \mathcal{C} with $b = 1$ to constrain future Courts to do the same with higher probability. Now consider “the last” Court to rule \mathcal{T} with $b = 1$ in environment \mathcal{C} 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 in environment \mathcal{C} , knowing that from that point on all future Courts will rule \mathcal{W} (with $b = 0$) when the environment is \mathcal{C} and they are not bound by precedents. In other words, suppose that the equilibrium involves a state of precedents that generates the “last tough Case Law Court,” with all subsequent Courts succumbing to the time-inconsistency problem. Assume also, that the equilibrium involves all Case Law Courts ruling \mathcal{W} whenever the environment is \mathcal{S} with breadth equal to 1.

This “natural conjecture” as to how a typical Markov perfect equilibrium of the Case Law regime might play out is in fact contradictory in some cases. To see this, consider the possibility that the last tough Case Law 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 operates in environment \mathcal{C} and has discretion will face the same body of precedents, and (by stationarity) *it* will be the last tough Court. To

⁴¹It is worth noting that when ρ is sufficiently small, provided that $\tau_{\mathcal{S}}^0 > 0$, we also know that $\mathbf{\Pi}_{SL}(\rho) > \mathbf{\Pi}_{CL}(\mathcal{J}^0, \rho)$. This result does not seem so interesting since it stems from the following rather obvious observations. When ρ is near zero the Statute Law Courts will be taking the optimal decision — namely \mathcal{W} in environment \mathcal{S} — almost all the time. On the other hand, since $\tau_{\mathcal{S}}^0 > 0$, the Case Law Courts at least initially will be constrained by precedents to take the wrong decision — namely \mathcal{T} in environment \mathcal{S} — a non-vanishing fraction of the time.

make the argument more straightforward, suppose that $\omega_{\mathcal{C}}^t = 1$ so that the current precedent does not constrain Courts to choose \mathcal{W} in the \mathcal{C} environment. Note that the t -th period Case Law 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 $\hat{\Pi}(\mathcal{W}, \mathcal{C}) > \hat{\Pi}(\mathcal{T}, \mathcal{C})$. Second, it puts one of the future Courts (the first one to face environment \mathcal{C} and to have discretion) in the position of being the *last* tough Case Law 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 $\Pi(\mathcal{T}, \mathcal{C}) > \Pi(\mathcal{W}, \mathcal{C})$.⁴²

The solution to the puzzle we have just outlined is that a typical Markov perfect equilibrium of our model of the Case Law regime 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 $b = 1$ in environment \mathcal{C} , without violating Proposition 2 (Case Law eventually must mature), and without running into the difficulty we have outlined. No Case Law 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 Case Law 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.

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

⁴²If instead our expository assumption that $\omega_{\mathcal{C}}^t = 1$ does not hold (so that $\omega_{\mathcal{C}}^t < 1$) 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} in the \mathcal{C} environment, the deviation we are describing may not be profitable. In this case, besides the 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.

5.2. Well-Behaved Precedents

The regularity conditions on the precedents technology that we work with are summarized next. We comment on each condition immediately after their statement.

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

- (i) Continuity and Monotonicity: For any ruling \mathcal{R}^t and environment \mathcal{E}^t , $\mathcal{J}(\mathcal{J}^t, 1, \mathcal{R}^t, \mathcal{E}^t)$ is continuous in \mathcal{J}^t . Moreover, if $\mathcal{R}^t = \mathcal{T}$, $b^t = 1$ and $\tau_{\mathcal{E}}^t < 1$, then $\tau_{\mathcal{E}}^{t+1} > \tau_{\mathcal{E}}^t$ and $\omega_{\mathcal{E}}^{t+1} \geq \omega_{\mathcal{E}}^t$.⁴³ If instead $\mathcal{R}^t = \mathcal{W}$, $b^t = 1$ and $\omega_{\mathcal{E}}^t > 0$, then $\tau_{\mathcal{E}}^{t+1} \leq \tau_{\mathcal{E}}^t$ and $\omega_{\mathcal{E}}^{t+1} < \omega_{\mathcal{E}}^t$.⁴⁴
- (ii) Decreasing Returns from \mathcal{T} Decisions in State \mathcal{C} : Suppose that $\mathcal{E}^{t-1} = \mathcal{E}^t = \mathcal{C}$, $\mathcal{R}^{t-1} = \mathcal{R}^t = \mathcal{T}$, and $b^{t-1} = b^t = 1$. Then $\tau_{\mathcal{C}}^{t+1} - \tau_{\mathcal{C}}^t < \tau_{\mathcal{C}}^t - \tau_{\mathcal{C}}^{t-1}$.
- (iii) Independent Impact of \mathcal{T} Decisions in State \mathcal{C} : Consider $\mathcal{J}^t = (\tau_{\mathcal{S}}^t, \omega_{\mathcal{S}}^t, \tau_{\mathcal{C}}^t, \omega_{\mathcal{C}}^t)$ and $\tilde{\mathcal{J}}^t = (\tilde{\tau}_{\mathcal{S}}^t, \tilde{\omega}_{\mathcal{S}}^t, \tau_{\mathcal{C}}^t, \tilde{\omega}_{\mathcal{C}}^t)$ with $(\tau_{\mathcal{S}}^t, \omega_{\mathcal{S}}^t, \omega_{\mathcal{C}}^t) \neq (\tilde{\tau}_{\mathcal{S}}^t, \tilde{\omega}_{\mathcal{S}}^t, \tilde{\omega}_{\mathcal{C}}^t)$. Then $\tau_{\mathcal{C}}(\mathcal{J}^t, 1, \mathcal{T}, \mathcal{C}) = \tau_{\mathcal{C}}(\tilde{\mathcal{J}}^t, 1, \mathcal{T}, \mathcal{C})$.
- (iv) No Cross-State Effects: If $\mathcal{E}^t = \mathcal{S}$ then $\tau_{\mathcal{C}}^t = \tau_{\mathcal{C}}^{t+1}$ and $\omega_{\mathcal{C}}^t = \omega_{\mathcal{C}}^{t+1}$. If $\mathcal{E}^t = \mathcal{C}$ then $\tau_{\mathcal{S}}^t = \tau_{\mathcal{S}}^{t+1}$ and $\omega_{\mathcal{S}}^t = \omega_{\mathcal{S}}^{t+1}$.⁴⁵
- (v) Reversibility: If $\mathcal{R}^t = \mathcal{T}$, $b^t = 1$ and $\mathcal{R}^{t+1} = \mathcal{W}$, $b^{t+1} = 1$ or alternatively $\mathcal{R}^t = \mathcal{W}$, $b^t = 1$ and $\mathcal{R}^{t+1} = \mathcal{T}$, $b^{t+1} = 1$ then, for any given $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$, $\tau_{\mathcal{E}}^t = \tau_{\mathcal{E}}^{t+2}$ and $\omega_{\mathcal{E}}^t = \omega_{\mathcal{E}}^{t+2}$.
- (vi) No Interior Asymptotes: Consider the sequence $\{\tau_{\mathcal{S}}^t\}_{t=1}^{\infty}$ generated assuming that $\mathcal{E}^t = \mathcal{S}$, $\mathcal{R}^t = \mathcal{W}$, and $b^t = 1$ for every t . Then $\lim_{t \rightarrow \infty} \tau_{\mathcal{S}}^t = 0$. Symmetrically, consider the sequence $\{\omega_{\mathcal{C}}^t\}_{t=1}^{\infty}$ generated assuming that $\mathcal{E}^t = \mathcal{C}$, $\mathcal{R}^t = \mathcal{T}$, and $b^t = 1$ for every t . Then $\lim_{t \rightarrow \infty} \omega_{\mathcal{C}}^t = 1$.

Besides continuity (useful for technical reasons), (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 1 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 $\tau_{\mathcal{C}} = 1$ the mapping $\tau_{\mathcal{C}}$ necessarily satisfies a form of decreasing returns to scale in $\tau_{\mathcal{C}}^t$. Condition (ii) extends this property to the whole interval $[0, 1]$. In other words, the impact of the first instance of a Court’s tough decision with positive breadth is greater than the impact of later instances of the same decision.

Conditions (iii) and (iv) are both assumptions that guarantee “de-coupling” of the effects of Courts’ decisions. Condition (iii) guarantees that the effect of a tough decision on the

⁴³If $\mathcal{R}^t = \mathcal{T}$, $b^t = 1$ and $\tau_{\mathcal{E}}^t = 1$, then just set $\tau_{\mathcal{E}}^{t+1} = \tau_{\mathcal{E}}^t$ and $\omega_{\mathcal{E}}^{t+1} = \omega_{\mathcal{E}}^t$.

⁴⁴If $\mathcal{R}^t = \mathcal{W}$, $b^t = 1$ and $\omega_{\mathcal{E}}^t = 0$, then just set $\tau_{\mathcal{E}}^{t+1} = \tau_{\mathcal{E}}^t$ and $\omega_{\mathcal{E}}^{t+1} = \omega_{\mathcal{E}}^t$.

⁴⁵In both cases, regardless of \mathcal{R}^t and b^t .

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. Condition (iv) ensures that decisions taken in a given environment have no effect on the constraints faced by future Courts in the other environment.

Condition (v) is technically extremely convenient. It guarantees that opposite consecutive decisions by Courts (with positive breadth) cancel each other out. In effect, this allows us to narrow down dramatically the cardinality of the set of possible quadruples $(\tau_S^t, \omega_S^t, \tau_C^t, \omega_C^t)$ that need to be considered overall along the equilibrium path.

Finally, condition (vi) guarantees that a sufficiently long sequence of ex-ante efficient decisions (with positive breadth) by the Case Law Courts will eliminate any initial precedent that forces judges to take the inefficient decision with positive probability.⁴⁶

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 of the Case Law regime.

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

The second piece of additional notation identifies a critical threshold for the value of τ_C . Suppose that

$$(1 - \delta)[\widehat{\Pi}(\mathcal{W}, \mathcal{C}) - \widehat{\Pi}(\mathcal{T}, \mathcal{C})] < \delta \rho \boldsymbol{\kappa}(0) [\Pi(\mathcal{T}, \mathcal{C}) - \Pi(\mathcal{W}, \mathcal{C})] \quad (9)$$

And notice that from (i) of Assumption 3 (monotonicity), we know that $\boldsymbol{\kappa}(\tau_C)$ is non-negative, and equal to zero (only) if $\tau_C = 1$.⁴⁷ By (i) of Assumption 3 again (continuity) it is then immediate that if (9) holds, there exists a value $\tau_C^* \in (0, 1)$ such that

$$(1 - \delta)[\widehat{\Pi}(\mathcal{W}, \mathcal{C}) - \widehat{\Pi}(\mathcal{T}, \mathcal{C})] = \delta \rho \boldsymbol{\kappa}(\tau_C^*) [\Pi(\mathcal{T}, \mathcal{C}) - \Pi(\mathcal{W}, \mathcal{C})] \quad (10)$$

⁴⁶This condition is not necessary to characterize the equilibrium and will be used only in Proposition 5

⁴⁷Note that by (ii) of Assumption 3, we also know that $\boldsymbol{\kappa}(\tau_C)$ is monotone decreasing.

If instead (9) does not hold we set $\tau_C^* = 0$.

Proposition 4. *A Markov Perfect Equilibrium: Suppose that Assumptions 1, 2 and 3 hold. Then our model of the Case Law regime has an equilibrium (see Definition 1) $\sigma^* = (\mathcal{R}^*, \mathbf{b}^*)$ as follows.*

- (i) For every \mathcal{J}^t , $\mathcal{R}(\mathcal{J}^t, \mathcal{S}) = \mathcal{W}$ and $\mathbf{b}(\mathcal{J}^t, \mathcal{S}) = 1$, so that the Court's ruling in state \mathcal{S} is always \mathcal{W} with breadth 1.
- (ii) If \mathcal{J}^t is such that $\tau_C^t \geq \tau_C^*$, then $\mathcal{R}(\mathcal{J}^t, \mathcal{C}) = \mathcal{W}$ and $\mathbf{b}(\mathcal{J}^t, \mathcal{C}) = 0$. In other words, if $\tau_C^t \geq \tau_C^*$, then the Court's ruling in state \mathcal{C} is \mathcal{W} with breadth 0.
- (iii) Finally, if $\tau_C^* > 0$ and \mathcal{J}^t is such that $\tau_C^t < \tau_C^*$, in state \mathcal{C} the Court randomizes between a ruling of \mathcal{T} with breadth 1 and a ruling of \mathcal{W} with breadth 0. The mixing probabilities only depend on τ_C^t and ω_C^t . We denote the probability of a \mathcal{T} ruling with breadth 1 by $p(\tau_C^t, \omega_C^t)$, so that the probability of a \mathcal{W} ruling with breadth 0 is $1 - p(\tau_C^t, \omega_C^t)$.

In the environment \mathcal{S} , in which time-inconsistency is not a problem, the Court takes the ex-ante efficient decision \mathcal{W} with breadth 1 when it has discretion. The evolution of precedents simply works to eliminate the potential inefficiency that may arise from initial conditions ($\tau_S^0 > 0$) that force the Case Law Courts to issue the inefficient ruling \mathcal{T} . A (potentially infinite) sequence of \mathcal{W} rulings with positive breath eventually reduce to zero the probability that Case Law Courts may be forced by precedent to take the inefficient decision \mathcal{T} . In other words, $\lim_{t \rightarrow \infty} \tau_S^t = 0$. At least in the limit Case Law reaches full efficiency, conditional on the environment being \mathcal{S} (where the time-inconsistency problem does not arise).

The equilibrium behavior captured by Proposition 4 is considerably richer in the \mathcal{C} environment where the temptation of time-inconsistent behavior is present. This can be seen focusing on the case in which $\tau_C^* > 0$ and the initial \mathcal{J}^0 has $\tau_C^0 < \tau_C^*$. 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 δ) from the increase in τ_C stemming from a \mathcal{T} decision with $b = 1$ — inequality (9) holds.

However, when inequality (9) 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 in environment \mathcal{C} may not be viable. The equilibrium then involves the Case Law Courts who

have discretion in environment \mathcal{C} mixing between a \mathcal{T} ruling with $b = 1$ and a \mathcal{W} ruling with $b = 0$. Each Case Law 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 $\tau_{\mathcal{C}}$ increases stochastically through time when the tough ruling with breadth 1 is chosen. Eventually, this process puts the value of $\tau_{\mathcal{C}}$ over the threshold $\tau_{\mathcal{C}}^*$. At this point Case Law is mature. All Case Law Courts from this point on, if they have discretion in environment \mathcal{C} , issue ruling \mathcal{W} with breadth 0.

We conclude this Section by noticing that the randomization involved in the equilibrium described in Proposition 4 has an appealing interpretation: the body of precedents is sufficiently ambiguous so as to leave the parties with genuine uncertainty at the ex-ante stage as to which decision the Court will actually take.

6. The Possibility of Case Law Dominance

Proposition 3 above establishes that in some cases the Statute Law regime is superior to the Case Law one. While it is clear that in many cases there will also be a region of parameters in which instead Case Law dominates, the regularity conditions on the precedents technology embodied in Assumption 3 allow us to be more precise and identify an actual parameter region in which this is the case.

Recall that (see (4) above) ρ_{SL}^* is the value of ρ — the probability of the complex state \mathcal{C} — for which committing to either the tough or the weak decision yields the same expected payoff under the Statute Law regime. Note that there is an obvious sense in which when $\rho = \rho_{SL}^*$, the Statute Law regime incurs the maximum cost for its lack of flexibility in the face of a heterogeneous pool of cases. Therefore, the Case Law regime may dominate (yielding higher welfare) when ρ is in a neighborhood of ρ_{SL}^* .

To see why this is the case in some more detail, assume that ρ is near ρ_{SL}^* , and note that to evaluate the welfare under Statute Law at $\rho = \rho_{SL}^*$ we can use either the weak decision or the tough one interchangeably. It will be convenient to use the weak one — $\mathcal{R}_{SL}(\rho_{SL}^*) = \mathcal{W}$.

A first case in which the conclusion is straightforward is when the initial body of precedents \mathcal{J}^0 is such that $\tau_{\mathcal{S}}^0 = 0$ and $\tau_{\mathcal{C}}^0 > 0$. In this case, the two regimes are equivalent in state \mathcal{S} since in either case all Courts will take the \mathcal{W} decision. In state \mathcal{C} on the other hand, because $\tau_{\mathcal{C}}^0 > 0$, the Case Law Courts will take the ex-ante optimal decision \mathcal{T} with positive

probability. As mentioned above, this is not the case under Statute Law. Hence, overall welfare will be higher in the Case Law regime.

A second case in which the conclusion is not hard to reach is when \mathcal{J}^0 and the other parameters of the model (per-period payoffs and δ) are such that $\tau_{\mathcal{S}}^0 = 0$ and condition (9) above is satisfied. Just as in the first case, in state \mathcal{S} the two regimes are now equivalent because $\tau_{\mathcal{S}}^0 = 0$ — all Courts take the \mathcal{W} decision. Moreover, in state \mathcal{C} the Case Law Courts will start by taking the ex-ante optimal tough decision with positive probability because inequality (9) is satisfied. Therefore, again, overall welfare is higher in the Case Law regime.

In general, for ρ in a neighborhood of ρ_{SL}^* , overall welfare will be higher in the Case Law regime when δ is sufficiently large. Note that in this case, because of (i) of Assumption 3, for δ large inequality (9) must necessarily be true. Therefore the only real difference between this general case and the second case we have just considered above is that it could be that $\tau_{\mathcal{S}}^0 > 0$ so that the initial body of precedents \mathcal{J}^0 forces the Case Law Courts to take the inefficient decision \mathcal{T} in the simple state with positive probability. However, we know from Proposition 4 that this inefficiency will be progressively eliminated by the evolution of Case Law so that $\lim_{t \rightarrow \infty} \tau_{\mathcal{S}}^t = 0$. Hence, for δ close to 1, the initial inefficient behavior in state \mathcal{S} has a negligible effect on overall welfare. Hence, the case Law regime yields overall higher welfare again.

Proposition 5. *Case Law Welfare Dominance:* Suppose that Assumptions 1, 2 and 3 hold, and let ρ_{SL}^* be as in Proposition 1.

Let any initial body of precedents \mathcal{J}^0 be given. Then there exists a $\delta_{CL}^* \in (0, 1)$ such that for every $\delta \in (\delta_{CL}^*, 1]$ we have that

$$\Pi_{CL}(\mathcal{J}^0, \rho_{SL}^*) > \Pi_{SL}(\rho_{SL}^*).$$

In a neighborhood of ρ_{SL}^* , the flexibility of the Case Law Courts dominates the costs associated with the time-inconsistency problem. In this case, the Case Law regime dominates the Statute Law regime in terms of ex-ante welfare.

Figure 1 below compares the welfare of the two legal regimes. Welfare under Statute Law can be easily obtained from Proposition 1. Equilibrium welfare under Case Law is instead computed for given δ and a particular (well behaved) precedents technology that satisfies Assumptions 1, 2 and 3. The details of the computation are presented in Appendix C below.

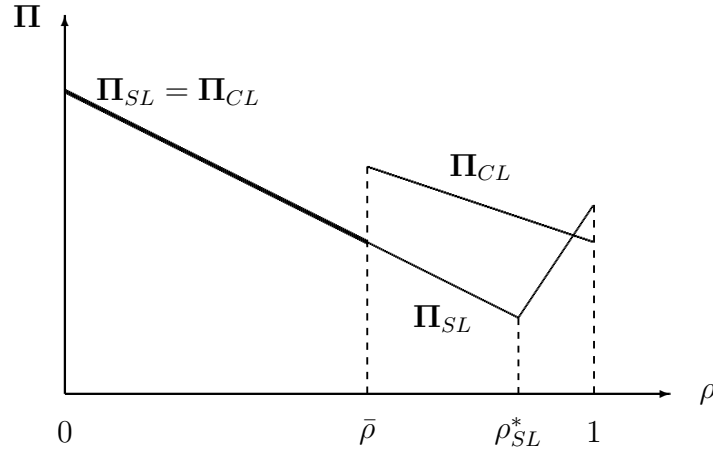


Figure 1: Welfare under Statute Law and Case Law.

When $\rho < \bar{\rho}$ the welfare in the two regimes is the same. Courts in both regimes take the weak decision regardless of the environment. Welfare in Case Law has a discontinuous increase at $\bar{\rho}$. Precisely at this value of ρ Case Law Courts that have discretion begin to take a \mathcal{T} decision when the state is \mathcal{C} . As long as $\bar{\rho} < \rho_{SL}^*$ Case Law dominates Statute Law in a neighborhood of ρ_{SL}^* . On the other hand, as we are told by Proposition 3, Statute Law dominates when ρ is close to 1.

7. Conclusions

Courts intervene in economic relationships at the *ex-post* stage (if at all). Because of sunk strategic decisions this might generate a time-inconsistent present-bias in the decisions of Courts that exercise discretion.

This observation has wide-ranging implications for the hypothesis that Case Law evolves towards efficient decisions. In a Case Law regime, each Court will trade off its current temptation to take an inefficient decision dictated by its present bias with the effect that its decision has on *future* Case Law Courts, via precedents. We find that under general circumstances this effectively prevents Case Law from reaching full efficiency. Eventually, the effect via precedents must become small since it is a marginal one. The temptation to take the inefficient decision on the other hand remains constant through time. Hence, at some point Case Law “matures” in the sense that precedents are already very likely to constrain

future Courts to take the efficient 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, inefficient decisions whenever precedents do not bind.

Once the propensity of Case Law to succumb to time-inconsistency is established, it is natural to ask the question of whether an inflexible regime of Statute Law in which Courts never have any discretion is superior in some cases. Even though we model the Statute Law regime in an extreme way — no flexibility at all — we find that Statute Law does indeed dominate in some cases. In particular if the environment changes sufficiently slowly (is sufficiently homogeneous) so that the inflexibility of Statute Law carries a sufficiently low cost, then it will dominate the Case Law regime in welfare terms. Conversely, Case Law does dominate Statute Law in an environment that changes sufficiently quickly (is sufficiently heterogeneous). In this case, the lack of flexibility of Statute Law is large, the rewards to flexibility reaped by the Case Law regime are also large thus mitigating the effects of the time-inconsistency problem.

Our findings are consistent with the idea that Case Law would dominate in highly dynamic sectors of the economy (e.g. Finance, Information Technology),⁴⁸ while Statute Law would yield higher welfare in the slow-paced sectors of the economy (e.g. Agriculture, Inheritance, Ownership Rights).

Appendix A: Disclosure and Time-Inconsistency

A.1. Disclosure and Time-Inconsistency

In Anderlini, Felli, and Postlewaite (2006) (henceforth AFP) we study a multiple-widget contracting model with asymmetric information in which the Court optimally voids some of the parties’ contract in order to obtain separation in equilibrium. The key features of the AFP model are that Court intervention is beneficial in some cases and that *ex-post* the Court’s incentives are *not* to intervene when it in fact should from the

⁴⁸Rajan and Zingales (2003) present evidence that Common Law countries only develop better financial systems than Civil Law ones *after* 1913. Our model would be consistent with this observation if one could argue that the rate at which the financial sector environment changes accelerated sufficiently around that time. This is not the explanation put forth by Rajan and Zingales (2003) (who focus on the political economy of the problem), but in our view one worthy of future research. Along the same lines, Lamoreaux and Rosenthal (2004) argue that US Law during the nineteenth century was neither more flexible nor more responsive than French law to the businesses needs.

point of view of ex-ante welfare. We present here a numerical version of the parametric model in AFP, with the added possibility that the environment may in fact be such that the Court should uphold all contracts.

In the terminology used in the paper, we refer to the latter as the simple environment (with fewer widgets) \mathcal{S} and to the former as the “complex” environment (with more widgets) \mathcal{C} . The environment is \mathcal{S} with probability $1 - \rho$ and is \mathcal{C} with probability ρ .

In both environments there is a buyer and a seller, both risk-neutral. The buyer has private information on the costs and values of the relevant widgets. He can be of a “high” type (denoted \mathcal{H}) or of a “low” type (denoted \mathcal{L}), with equal probability. The buyer knows his type at the time of contracting, while the seller does not. As standard, there is an ex-ante contracting stage, followed by an investment stage, followed by the ex-post trading stage. For simplicity, at the ex-ante contracting stage the buyer has all the bargaining power, while the seller has all the bargaining power ex-post.

In the simple environment there are two widgets, w_1 and w_2 . These two widgets are mutually exclusive because they require a widget- and relationship-specific investment of $I = 1$ on the part of the buyer. The buyer can only undertake *one* investment, and the cost and value of either widget without investment are zero. The cost and value of w_i ($i = 1, 2$) (net of the ex-ante investment) if the buyer’s type is $\theta \in \{\mathcal{L}, \mathcal{H}\}$ are denoted by c_i^θ and v_i^θ respectively. When investment takes place we take each of them to be as follows

	w_1	w_2	
Type \mathcal{H}	$v_1^{\mathcal{H}} = 21, \quad c_1^{\mathcal{H}} = 1$	$v_2^{\mathcal{H}} = 25, \quad c_2^{\mathcal{H}} = 1$	(A.1)
Type \mathcal{L}	$v_1^{\mathcal{L}} = -1, \quad c_1^{\mathcal{L}} = 0$	$v_2^{\mathcal{L}} = 3, \quad c_2^{\mathcal{L}} = 1$	

The complex environment is the same as the simple environment, save for the fact that a third widget w_3 is available. This widget is *not contractible* at the ex-ante stage, and does not require any investment.⁴⁹ Widget w_3 can be traded ex-post via a “spot” contract. Trading w_3 yields a positive surplus only if the buyer’s type is \mathcal{L} . We take the cost and values of the three widgets in the complex environment to be

	w_1	w_2	w_3	
Type \mathcal{H}	$v_1^{\mathcal{H}} = 21, \quad c_1^{\mathcal{H}} = 1$	$v_2^{\mathcal{H}} = 25, \quad c_2^{\mathcal{H}} = 1$	$v_3^{\mathcal{H}} = 76, \quad c_3^{\mathcal{H}} = 100$	(A.2)
Type \mathcal{L}	$v_1^{\mathcal{L}} = -1, \quad c_1^{\mathcal{L}} = 0$	$v_2^{\mathcal{L}} = 3, \quad c_2^{\mathcal{L}} = 1$	$v_3^{\mathcal{L}} = 65, \quad c_3^{\mathcal{L}} = 3$	

The Court may intervene in the parties’ contractual relationship by *voiding* contracts for either w_1 or w_2 .⁵⁰ The Court’s decision to void corresponds to the *tough decision* in the paper. Because of the hold-up problem generated by the the widget- and relationship-specific investment, if the Court voids contracts for either w_1 or w_2 or both, then the corresponding widget will not be traded.

In the simple environment, the Court has no welfare-enhancing role to play. When all contracts are

⁴⁹In AFP we argue that the ex-ante non-contractibility of w_3 is without loss of generality for the class of contracts we consider here.

⁵⁰In AFP we argue that not allowing the Court to void contracts for w_3 is without loss of generality.

enforced, in equilibrium both types of buyer invest in and trade w_2 . This yields full social efficiency. The total expected surplus from trading (net of investment) is 13.

Equilibria in the complex environment are fully characterized in AFP. When the Court enforces all contracts, there is a *unique* equilibrium, which involves inefficient pooling. Both types of buyer invest in and trade w_2 , and, since the buyer's type is not revealed, they also trade w_3 ex-post. The total expected surplus from trading (net of investment) in this case is 32. This outcome is clearly short of social efficiency since the type \mathcal{H} buyer trades w_3 , which generates *negative* surplus (-24).

If instead the Court intervenes and voids contracts for w_2 , the two types of buyer separate: behaving differently, they reveal their private information at the ex-ante contracting stage. The unique equilibrium outcome is that type \mathcal{H} buyer invests in and trades w_1 , but does not trade w_3 , while the type \mathcal{L} buyer does not invest in and does not trade either w_1 or w_2 ; he only trades w_3 ex-post. In this case the total expected surplus from trading (net of investment) is 41. While this outcome does not achieve full social efficiency it dominates the pooling outcome since it avoids the inefficient trade of w_3 from the part of the type \mathcal{H} buyer.⁵¹

In AFP it is also shown that voiding contracts for w_2 is the best that the Court can do in the complex environment.⁵²

To sum up, if the environment is simple a welfare-maximizing Court can do no better than not intervening at all and enforcing the contract. This is the Court's *weak decision* \mathcal{W} in the terminology of the paper. Intuitively, Court intervention has no value since disclosure of the buyer's private information itself has no social value.

If instead the environment is complex then an active Court that intervenes and voids contracts for w_2 will enhance social welfare. This is the *tough decision* \mathcal{T} . By intervening, the Court induces the two types of buyer to *disclose information* at the ex-ante contracting stage. This disclosure has positive social value in the complex environment.

Translating the numbers above into the payoffs used in the main body of the paper (Subsection 3.1) yields:

$$\Pi(\mathcal{W}, \mathcal{S}) = 13 > \Pi(\mathcal{T}, \mathcal{S}) = 10 \quad \text{and} \quad \hat{\Pi}(\mathcal{W}, \mathcal{S}) = 14 > \hat{\Pi}(\mathcal{T}, \mathcal{S}) = 0 \quad (\text{A.3})$$

and

$$\Pi(\mathcal{W}, \mathcal{C}) = 32 < \Pi(\mathcal{T}, \mathcal{C}) = 41 \quad \text{and} \quad \hat{\Pi}(\mathcal{W}, \mathcal{C}) = 33 > \hat{\Pi}(\mathcal{T}, \mathcal{C}) = 0 \quad (\text{A.4})$$

⁵¹Full social efficiency in the complex environment would entail that both types of buyer invest in and trade w_2 , while only the type \mathcal{L} buyer trades w_3 ex-post. The total expected surplus from trading (net of investment) in this case would be 44.

⁵²Recall that the Court can choose between voiding no contracts, voiding contracts for w_1 , voiding contracts for w_2 and voiding contracts for both w_1 and w_2 . In AFP the case of mixed strategies for the Court is also considered. We do not allow probabilistic Court choices in the present set-up.

Appendix B: Proofs

Lemma B.1: *Let σ^* be an equilibrium for the Case Law regime. Then expected welfare is weakly monotonically increasing in the sense that for any $\mathcal{J} \in [0, 1]^4$ and any $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$ we have that*

$$\mathbf{Z}(\mathcal{J}(\mathcal{J}, \mathbf{b}^*(\mathcal{J}, \mathcal{E}), \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \geq \mathbf{Z}(\mathcal{J}, \sigma^*, \rho) \quad (\text{B.1})$$

Proof: By Definition 1 for every $\mathcal{J} \in [0, 1]^4$ and any $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$ the values $b = \mathbf{b}^*(\mathcal{J}, \mathcal{E})$ and $\mathcal{R} = \mathcal{R}^*(\mathcal{J}, \mathcal{E})$ must solve

$$\max_{\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}, b \in \{0, 1\}} (1 - \delta) \widehat{\Pi}(\mathcal{E}, \mathcal{R}) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, b, \mathcal{R}, \mathcal{E}), \sigma^*, \rho) \} \quad (\text{B.2})$$

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

$$\mathbf{Z}(\mathcal{J}, \sigma^*, \rho) = \mathbf{Z}(\mathcal{J}(\mathcal{J}, 0, \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) > \mathbf{Z}^*(\mathcal{J}(\mathcal{J}, \mathbf{b}^*(\mathcal{J}, \mathcal{S}), \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \quad (\text{B.3})$$

and hence

$$\begin{aligned} \widehat{\Pi}(\mathcal{E}, \mathcal{R}^*(\mathcal{J}, \mathcal{E})) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, 0, \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \} > \\ \widehat{\Pi}(\mathcal{E}, \mathcal{R}^*(\mathcal{J}, \mathcal{E})) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, \mathbf{b}^*(\mathcal{J}, \mathcal{E}), \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \} \end{aligned} \quad (\text{B.4})$$

which contradicts the fact that $\mathbf{b}^*(\mathcal{J}, \mathcal{E})$ and $\mathcal{R}^*(\mathcal{J}, \mathcal{E})$ must solve (B.2). ■

Lemma B.2: *Let σ^* be an equilibrium for the Case Law regime. Suppose that for some $\mathcal{J} \in [0, 1]^4$ and $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$ we have that*

$$\mathcal{R}^*(\mathcal{J}, \mathcal{E}) = \mathcal{T} \quad (\text{B.5})$$

then it must be that

$$\mathbf{Z}(\mathcal{J}(\mathcal{J}, \mathbf{b}^*(\mathcal{J}, \mathcal{E}), \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) - \mathbf{Z}(\mathcal{J}, \sigma^*, \rho) \geq \frac{1 - \delta}{\delta} \left[\widehat{\Pi}(\mathcal{E}, \mathcal{W}) - \widehat{\Pi}(\mathcal{E}, \mathcal{T}) \right] \quad (\text{B.6})$$

Proof: From (8) of Definition 1, we know that for every $\mathcal{J} \in [0, 1]^4$ and any $\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}$ the values $b = \mathbf{b}^*(\mathcal{J}, \mathcal{E})$ and $\mathcal{R} = \mathcal{R}^*(\mathcal{J}, \mathcal{E})$ must solve

$$\max_{\mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}, b \in \{0, 1\}} (1 - \delta) \widehat{\Pi}(\mathcal{E}, \mathcal{R}) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, b, \mathcal{R}, \mathcal{E}), \sigma^*, \rho) \} \quad (\text{B.7})$$

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

$$\begin{aligned} (1 - \delta) \widehat{\Pi}(\mathcal{E}, \mathcal{T}) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, \mathbf{b}^*(\mathcal{J}, \mathcal{E}), \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \} \geq \\ (1 - \delta) \widehat{\Pi}(\mathcal{E}, \mathcal{R}^*(\mathcal{J}, \mathcal{W})) + \delta \{ \mathbf{Z}(\mathcal{J}(\mathcal{J}, 0, \mathcal{R}^*(\mathcal{J}, \mathcal{E}), \mathcal{E}), \sigma^*, \rho) \} \end{aligned} \quad (\text{B.8})$$

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

Proof of Proposition 2: Let m be the smallest integer that satisfies

$$m \geq \frac{\max_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}, \mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} \Pi(\mathcal{E}, \mathcal{R}) - \min_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}, \mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} \Pi(\mathcal{E}, \mathcal{R})}{\min_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}} \frac{1 - \delta}{\delta} [\widehat{\Pi}(\mathcal{E}, \mathcal{W}) - \widehat{\Pi}(\mathcal{E}, \mathcal{T})]} + 1 \quad (\text{B.9})$$

Notice next that $\mathbf{Z}(\mathcal{J}, \boldsymbol{\sigma}^*, \rho)$ is obviously bounded above by $\max_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}, \mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} \Pi(\mathcal{E}, \mathcal{R})$ and below by $\min_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}, \mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} \Pi(\mathcal{E}, \mathcal{R})$.

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

$$\mathbf{Z}(\mathcal{J}^{t-1}, \boldsymbol{\sigma}^*, \rho) \geq m \min_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}} \frac{1 - \delta}{\delta} [\widehat{\Pi}(\mathcal{E}, \mathcal{W}) - \widehat{\Pi}(\mathcal{E}, \mathcal{T})] + \min_{\mathcal{E} \in \{\mathcal{S}, \mathcal{C}\}, \mathcal{R} \in \{\mathcal{T}, \mathcal{W}\}} \Pi(\mathcal{E}, \mathcal{R}) \quad (\text{B.10})$$

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

Proof of Proposition 3: Fix an initial body of precedents \mathcal{J}^0 and a $\delta \in (0, 1)$. Fix a $\underline{\rho} \in (0, 1)$, and for every $\rho \in [\underline{\rho}, 1]$ fix an equilibrium (given \mathcal{J}^0) for the Case Law regime $\boldsymbol{\sigma}^*(\rho)$.

Let m be as in Proposition 2. Consider a possible realization of uncertainty $\{\mathcal{E}^t\}_{t=0}^{m-1}$, $\{\ell_{\mathcal{S}}^t\}_{t=0}^{m-1}$ and $\{\ell_{\mathcal{C}}^t\}_{t=0}^{m-1}$ with the following properties. First $\mathcal{E}^t = \mathcal{C}$ for every $t = 0, \dots, m-1$. Second, if we let $h^m(\rho) = (\mathcal{J}^0, \mathcal{J}^1(\rho), \dots, \mathcal{J}^{m-1}(\rho))$ be the associated realized history in the $\boldsymbol{\sigma}^*(\rho)$ equilibrium, then $\ell_{\mathcal{C}}^t \in (\tau_{\mathcal{C}}^t(\rho), \omega_{\mathcal{C}}^t(\rho))$ for every $t = 0, \dots, m-1$ and for every $\rho \in [\underline{\rho}, 1]$. In other words, along the realized path, the environment is \mathcal{C} and the Case Law Court has discretion in every period up to and including $t = m-1$, for every equilibrium $\boldsymbol{\sigma}^*(\rho)$ with $\rho \in [\underline{\rho}, 1]$.

Next, we argue that this path has positive probability, bounded away from zero, provided that $\rho \in [\underline{\rho}, 1]$. To see this, observe first that the probability that the environment is \mathcal{C} in periods $t = 0, \dots, m-1$ is given by ρ^m . The probability that the Case Law Court has discretion in every period in $\boldsymbol{\sigma}^*(\rho)$ is $\mathbf{d}(m, \rho) = \prod_{t=0}^{m-1} d_{\mathcal{C}}^t(\rho)$, where $d_{\mathcal{C}}^t(\rho)$ is given by the realized history $h^m(\rho)$. Therefore, if we let $\mathbf{d}(m) = \inf_{\rho \in [\underline{\rho}, 1]} \mathbf{d}(m, \rho)$ the probability of the entire path with the requisite properties is bounded below by $\rho^m \mathbf{d}(m)$. Trivially, the first term of this product is bounded away from zero, provided that $\rho \in [\underline{\rho}, 1]$. To see that $\mathbf{d}(m) > 0$, notice that since \mathcal{J}^0 by assumption has $d_{\mathcal{C}}^0 > 0$, and m is finite, then using Assumption 1 we have that for some $\underline{d} > 0$ it must be that $d_{\mathcal{C}}^t(\rho) > \underline{d}$ for every $t = 0, \dots, m-1$ and every $\rho \in [\underline{\rho}, 1]$. It follows that the entire path with the requisite properties must have probability, call it ξ , that is no smaller than $\rho^m \underline{d}^m > 0$.

Next, we consider two cases. Fix a $\rho \in [\underline{\rho}, 1]$. Along the positive probability path we have identified, in the equilibrium $\boldsymbol{\sigma}^*(\rho)$, either all the Case Law Courts' rulings are \mathcal{T} or they are not. Suppose first that all the rulings are \mathcal{T} . Then by Proposition 2 it must be that in the $\boldsymbol{\sigma}^*(\rho)$ equilibrium we have $\mathcal{R}^*(\mathcal{J}^{m-1}(\rho), \mathcal{C}) =$

\mathcal{W} . If one or more rulings along the path are different from \mathcal{T} then clearly in the $\sigma^*(\rho)$ we have $\mathcal{R}^*(\mathcal{J}^t(\rho), \mathcal{C}) = \mathcal{W}$ for some $t = 0, \dots, m-1$.

We can now conclude that in any $\sigma^*(\rho)$ equilibrium with $\rho \in [\underline{\rho}, 1]$, with probability $\xi > 0$, some Case Law Court at time $t \leq m-1$ issues a ruling of \mathcal{W} in environment \mathcal{C} .

Using (1) and (2) it is immediate that the welfare of any Case Law Court equilibrium cannot go above that generated by a sequence of rulings that are \mathcal{W} whenever the environment is \mathcal{S} and \mathcal{T} whenever the environment is \mathcal{C} . Therefore, we can conclude that in any $\sigma^*(\rho)$ equilibrium with $\rho \in [\underline{\rho}, 1]$ the welfare of the Case Law regime is bounded above as follows⁵³

$$\begin{aligned} \mathbf{\Pi}_{CL}(\rho) \leq (1-\delta) & \left\{ \sum_{t=0}^{m-2} \delta^t [(1-\rho)\Pi(\mathcal{S}, \mathcal{W}) + \rho\Pi(\mathcal{C}, \mathcal{T})] + \right. \\ & \delta^{m-1} [(1-\rho)\Pi(\mathcal{S}, \mathcal{W}) + (\rho-\xi)\Pi(\mathcal{C}, \mathcal{T}) + \xi\Pi(\mathcal{C}, \mathcal{W})] + \\ & \left. \sum_{t=m}^{\infty} \delta^t [(1-\rho)\Pi(\mathcal{S}, \mathcal{W}) + \rho\Pi(\mathcal{C}, \mathcal{T})] \right\} \end{aligned} \quad (\text{B.11})$$

Now consider any $\rho > \max\{\underline{\rho}, \rho_{SL}^*\}$ where ρ_{SL}^* is as in equation (4) of Proposition 1. Using Proposition 1 the equilibrium welfare $\mathbf{\Pi}_{SL}(\rho)$ of the Statute Law Regime is

$$\mathbf{\Pi}_{SL}(\rho) = (1-\delta) \sum_{t=0}^{\infty} \delta^t [(1-\rho)\Pi(\mathcal{S}, \mathcal{T}) + \rho\Pi(\mathcal{C}, \mathcal{T})] \quad (\text{B.12})$$

Using (B.11) and (B.12) it is a matter of straightforward algebra to then show that if we set⁵⁴

$$\rho_{CL}^* = \max \left\{ 1 - \frac{(1-\delta)\delta^{m-1}\xi [\Pi(\mathcal{C}, \mathcal{T}) - \Pi(\mathcal{C}, \mathcal{W})]}{\Pi(\mathcal{S}, \mathcal{W}) - \Pi(\mathcal{S}, \mathcal{T})}, \underline{\rho}, \rho_{SL}^* \right\} \in (0, 1) \quad (\text{B.13})$$

then for every $\rho > \rho_{CL}^*$ it is the case that $\mathbf{\Pi}_{SL}(\rho) > \mathbf{\Pi}_{CL}(\rho)$, as required. ■

Proof of Proposition 4: Let σ^* be as described in Proposition 4. We proceed in four steps to verify that it does in fact constitute an equilibrium of the model.

Step 1: In state \mathcal{S} there is no profitable deviation from $\sigma^* = (\mathcal{R}, \mathbf{b})$ such that: $\mathcal{R}(\mathcal{J}^t, \mathcal{S}) = \mathcal{W}$ and $\mathbf{b}(\mathcal{J}^t, \mathcal{S}) = 1$.

Proof: The period- t deviation $\mathcal{R}^t = \mathcal{W}$, $b^t = 0$ —where the continuation equilibrium for $k > t$ coincides with σ^* as specified in Proposition 4—is clearly not profitable given inequalities (1), Assumption 2, Conditions (i) and (iv) of Assumption 3 and the fact that $\delta \in (0, 1)$. Consider now the period- t deviation $\mathcal{R}^t = \mathcal{T}$ —where the continuation equilibrium for $k > t$ coincides with σ^* as specified in Proposition 4. This deviation is clearly not profitable whatever b^t by inequalities (1), Assumption 2 and Conditions (i) and (iv) of Assumption 3.

⁵³The first sum of terms in (B.11) is understood to be zero if $m = 1$.

⁵⁴Notice that the first term in the curly brackets in (B.13) in general depends on \mathcal{J}^0 . This is because ξ depends on \mathcal{J}^0 .

Step 2: In state \mathcal{C} let $\tau_{\mathcal{C}}^t \geq \tau_{\mathcal{C}}^*$, then there exists no profitable deviation from $\sigma^* = (\mathcal{R}, \mathbf{b})$ such that: $\mathcal{R}(\mathcal{J}^t, \mathcal{S}) = \mathcal{W}$ and $\mathbf{b}(\mathcal{J}^t, \mathcal{S}) = 0$.

Proof: Consider first the period- t deviation $\mathcal{R}^t = \mathcal{W}$, $b^t = 1$ —where the continuation equilibrium for $k > t$ coincides with σ^* as specified in Proposition 4. We need to distinguish two cases: $\tau_{\mathcal{C}}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}) \geq \tau_{\mathcal{C}}^*$ and $\tau_{\mathcal{C}}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}) < \tau_{\mathcal{C}}^*$. Consider first $\tau_{\mathcal{C}}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}) \geq \tau_{\mathcal{C}}^*$. The period- t deviation $\mathcal{R}^t = \mathcal{W}$, $b^t = 1$ is not profitable given inequalities (2), Assumption 2 and Conditions (i) and (iv) of Assumption 3. The case $\tau_{\mathcal{C}}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}) < \tau_{\mathcal{C}}^*$ will be dealt with in the proof of Step 4. Consider next the period- t deviation $\mathcal{R}^t = \mathcal{T}$, $b^t = 1$ —where the continuation equilibrium for $k > t$ coincides with σ^* as specified in Proposition 4. This deviation is not profitable given that, for $\tau_{\mathcal{C}}^t > \tau_{\mathcal{C}}^*$ we must have that

$$(1 - \delta) \left[\widehat{\Pi}(\mathcal{C}, \mathcal{W}) - \widehat{\Pi}(\mathcal{C}, \mathcal{T}) \right] > \delta \rho \kappa(\tau_{\mathcal{C}}^t) [\Pi(\mathcal{C}, \mathcal{T}) - \Pi(\mathcal{C}, \mathcal{W})] \quad (\text{B.14})$$

Finally, consider the period- t deviation $\mathcal{R}^t = \mathcal{T}$, $b^t = 0$ —where the continuation equilibrium for $k > t$ coincides with σ^* as specified in Proposition 4. This deviation is clearly not profitable given the second inequality in (2) and Assumption 2.

Step 3: Assume condition (9) is satisfied and \mathcal{J}^0 is such that $\tau_{\mathcal{C}}^0 < \tau_{\mathcal{C}}^*$. Given σ^* as in Proposition 4 let $n > 0$ be such that $\tau_{\mathcal{C}}^{n-1} < \tau_{\mathcal{C}}^*$ and $\tau_{\mathcal{C}}^n \geq \tau_{\mathcal{C}}^*$. Then, there exist n probability distributions $p(\tau_{\mathcal{C}}^i, \omega_{\mathcal{C}}^i) \in (0, 1]$, $i \in \{0, \dots, n-1\}$ such that, at every state $(\tau_{\mathcal{C}}^i, \omega_{\mathcal{C}}^i)$, the Court rules $\mathcal{R}^i = \mathcal{T}$, $b^i = 1$ with probability $p(\tau_{\mathcal{C}}^i, \omega_{\mathcal{C}}^i)$ and $\mathcal{R}^i = \mathcal{W}$, $b^i = 0$ with probability $1 - p(\tau_{\mathcal{C}}^i, \omega_{\mathcal{C}}^i)$.

Proof: We proceed backward and construct the probabilities $p(\tau_{\mathcal{C}}^j, \omega_{\mathcal{C}}^j)$, $j \in \{0, \dots, n-1\}$. Fix $(\tau_{\mathcal{S}}, \omega_{\mathcal{S}})$ —notice that Assumption 3, Condition (iv) implies that there is no loss in generality in doing so—consider first $p(\tau_{\mathcal{C}}^{n-1}, \omega_{\mathcal{C}}^{n-1})$. According to σ^* , the ex-ante value function $\mathbf{Z}(\mathcal{J}^n, \sigma^*, \rho)$ at $\mathcal{J}^n = (\tau_{\mathcal{S}}, \omega_{\mathcal{S}}, \tau_{\mathcal{C}}^n, \omega_{\mathcal{C}}^n)$ is such that

$$\mathbf{Z}(\mathcal{J}^n, \sigma^*, \rho) = (1 - \rho) [\tau_{\mathcal{S}} \Pi(\mathcal{S}, \mathcal{T}) + (1 - \tau_{\mathcal{S}}) \Pi(\mathcal{S}, \mathcal{W})] + \rho [\tau_{\mathcal{C}}^n \Pi(\mathcal{C}, \mathcal{T}) + (1 - \tau_{\mathcal{C}}^n) \Pi(\mathcal{C}, \mathcal{W})] \quad (\text{B.15})$$

The ex-ante value function $\mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho)$ at $\mathcal{J}^{n-1} = (\tau_{\mathcal{S}}, \omega_{\mathcal{S}}, \tau_{\mathcal{C}}^{n-1}, \omega_{\mathcal{C}}^{n-1})$ is then such that:⁵⁵

$$\begin{aligned} \mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho) = & (1 - \rho) \{ (1 - \delta) [\tau_{\mathcal{S}} \Pi(\mathcal{S}, \mathcal{T}) + (1 - \tau_{\mathcal{S}}) \Pi(\mathcal{S}, \mathcal{W})] + \delta \mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho) \} + \\ & \rho \{ \tau_{\mathcal{C}}^{n-1} [(1 - \delta) \Pi(\mathcal{C}, \mathcal{T}) + \delta \mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho)] + \\ & [(1 - \omega_{\mathcal{C}}^{n-1}) + d_{\mathcal{C}}^{n-1} (1 - p(\tau_{\mathcal{C}}^{n-1}, \omega_{\mathcal{C}}^{n-1}))] [(1 - \delta) \Pi(\mathcal{C}, \mathcal{W}) + \delta \mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho)] + \\ & d_{\mathcal{C}}^{n-1} p(\tau_{\mathcal{C}}^{n-1}, \omega_{\mathcal{C}}^{n-1}) [(1 - \delta) \Pi(\mathcal{C}, \mathcal{T}) + \delta \mathbf{Z}(\mathcal{J}^n, \sigma^*, \rho)] \} \end{aligned} \quad (\text{B.16})$$

Substituting (B.15) into (B.16) we can solve for $\mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho)$ in terms of the parameters of the model and the probability $p(\tau_{\mathcal{C}}^{n-1}, \omega_{\mathcal{C}}^{n-1})$. We can then use this formula for $\mathbf{Z}(\mathcal{J}^{n-1}, \sigma^*, \rho)$ and (B.15) to explicitly

⁵⁵Recall that $d_{\mathcal{C}}^{n-1} = (\omega_{\mathcal{C}}^{n-1} - \tau_{\mathcal{C}}^{n-1})$.

derive the probability $p(\tau_C^{n-1}, \omega_C^{n-1})$ by solving the Court's indifference condition:

$$(1 - \delta) \widehat{\Pi}(\mathcal{C}, \mathcal{W}) + \delta \mathbf{Z}(\mathcal{J}^{n-1}, \boldsymbol{\sigma}^*, \rho) = (1 - \delta) \widehat{\Pi}(\mathcal{C}, \mathcal{T}) + \delta \mathbf{Z}(\mathcal{J}^n, \boldsymbol{\sigma}^*, \rho) \quad (\text{B.17})$$

First, notice that we cannot have $p(\tau_C^{n-1}, \omega_C^{n-1}) = 0$ since if we set $p(\tau_C^{n-1}, \omega_C^{n-1}) = 0$ we get that the right hand side of (B.17) is strictly greater than the left hand side since, by construction $\tau_C^{n-1} < \tau_C^*$. Next, set $p(\tau_C^{n-1}, \omega_C^{n-1}) = 1$. Two cases are possible. Either the left hand side of (B.17) is lower or equal than the right hand side or is not. In the former case just set $p(\tau_C^{n-1}, \omega_C^{n-1}) = 1$. In other words, the Court's ruling is $\mathcal{R}(\mathcal{J}^{n-1}, \mathcal{C}) = \mathcal{T}$ and $\mathbf{b}(\mathcal{J}^{n-1}, \mathcal{C}) = 1$ with probability one. In the latter case, by continuity there exists a $p(\tau_C^{n-1}, \omega_C^{n-1}) \in (0, 1)$ that solves (B.17). In other words, the Court is mixing between a \mathcal{T} ruling with $b = 1$ and a \mathcal{W} ruling with $b = 0$.

In an analogous way, we can derive, recursively, $\mathbf{Z}(\mathcal{J}^i, \boldsymbol{\sigma}^*, \rho)$ and hence construct the remaining probability distributions $p(\tau_C^i, \omega_C^i) \in (0, 1]$, $i \in \{0, \dots, n-2\}$.

Step 4: *In state \mathcal{C} assume condition (9) is satisfied and \mathcal{J}^0 is such that $\tau_C^0 < \tau_C^*$, then for every $\tau_C^t < \tau_C^*$ the t -period deviation $\mathcal{R}^t = \mathcal{W}$ and $b^t = 1$ —where the continuation equilibrium for $k > t$ coincides with $\boldsymbol{\sigma}^*$ as specified in Proposition 4—is not profitable.*

Proof: The continuation strategy $\mathcal{R}^t = \mathcal{W}$ and $b^t = 1$ with continuation equilibrium for $k > t$ as in $\boldsymbol{\sigma}^*$ is dominated by the strategy $\mathcal{R}^t = \mathcal{W}$ and $b^t = 0$ with continuation equilibrium for $k > t$ as in $\boldsymbol{\sigma}^*$. In other words, it must be the case that:

$$\mathbf{Z}(\mathcal{J}^t, \boldsymbol{\sigma}^*, \rho) \geq \mathbf{Z}(\mathcal{J}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}), \boldsymbol{\sigma}^*, \rho) \quad (\text{B.18})$$

We need to distinguish two cases: $\tau_C^t > \tau_C^0$ and $\tau_C^t = \tau_C^0$. Consider first $\tau_C^t > \tau_C^0$. Assumption 3, Condition (v) implies that there exists a period $h < t$ such that

$$\mathcal{J}^h = \mathcal{J}(\mathcal{J}^t, 1, \mathcal{W}, \mathcal{C}) \quad (\text{B.19})$$

Lemma B.1 then proves inequality (B.18).

Consider now $\tau_C^t = \tau_C^0$. Given that $\tau_C^0 < \tau_C^*$, following the deviation $\mathcal{R}^t = \mathcal{W}$ $b^t = 1$ by $\boldsymbol{\sigma}^*$, as specified in Proposition 4, there must exist a $k > t$ such that the realization of the draw from $p(\omega_C^k, \tau_C^k)$ is such that $\mathcal{R}^k = \mathcal{T}$ and $b^k = 1$. Then, once again, by Assumption 3, Condition (v) we have

$$\mathcal{J}^t = \mathcal{J}(\mathcal{J}^k, 1, \mathcal{T}, \mathcal{C}) \quad (\text{B.20})$$

Lemma B.1 then proves inequality (B.18). ■

Proof of Proposition 5: Let $\rho = \rho_{SL}^*$. Then by (4) the Statute Law regime yields ex-ante welfare such that:

$$\mathbf{\Pi}_{SL}(\rho_{SL}^*) = (1 - \rho_{SL}^*)\Pi(\mathcal{S}, \mathcal{W}) + \rho_{SL}^*\Pi(\mathcal{C}, \mathcal{W}) \quad (\text{B.21})$$

Consider now the ex-ante welfare under Case Law. Clearly, by definition of $\mathbf{\Pi}(\mathcal{J}^0, \rho)$, we have that

$$\mathbf{\Pi}(\mathcal{J}^0, \rho) \geq \mathbf{Z}(\mathcal{J}^0, \rho, \boldsymbol{\sigma}^*) \quad (\text{B.22})$$

where as we defined in the text $\mathbf{Z}(\mathcal{J}^0, \rho, \boldsymbol{\sigma}^*)$ is the ex-ante welfare under Case Law associated with the equilibrium $\boldsymbol{\sigma}^*$ of Proposition 4. By Condition (iv) of Assumption 3, we then have that

$$\mathbf{Z}(\mathcal{J}^0, \rho_{SL}^*, \boldsymbol{\sigma}^*) = (1 - \rho_{SL}^*) \mathbf{Z}(\mathcal{J}^0, 0, \boldsymbol{\sigma}^*) + \rho_{SL}^* \mathbf{Z}(\mathcal{J}^0, 1, \boldsymbol{\sigma}^*) \quad (\text{B.23})$$

Notice first that Proposition 4 implies that along the equilibrium $\boldsymbol{\sigma}^*$ in state \mathcal{S} the state of precedents $(\tau_{\mathcal{S}}^t, \omega_{\mathcal{S}}^t)$ converges to $[0, \omega_{\mathcal{S}}^*]$ where $\omega_{\mathcal{S}}^*$ exists and is such that $\omega_{\mathcal{S}}^* \in (0, 1]$. This implies that the steady state ex-ante welfare under Case Law in state \mathcal{S} converges to $\mathbf{\Pi}(\mathcal{S}, \mathcal{W})$. In other words, asymptotically, the Statute Law ex-ante welfare and the Case Law ex-ante welfare in state \mathcal{S} coincide:

$$\lim_{\delta \rightarrow 1} \mathbf{\Pi}_{CL}(\mathcal{J}^0, 0, \boldsymbol{\sigma}^*) = \mathbf{\Pi}(\mathcal{S}, \mathcal{W}) \quad (\text{B.24})$$

Consider now state \mathcal{C} . Assume first that \mathcal{J}^0 is such that $\tau_{\mathcal{C}}^0 > 0$ then according to $\boldsymbol{\sigma}^*$ we can bound the ex-ante welfare in state \mathcal{C} in the following way.

$$\mathbf{Z}(\mathcal{J}^0, 1, \boldsymbol{\sigma}^*) \geq \tau_{\mathcal{C}}^0 \mathbf{\Pi}(\mathcal{C}, \mathcal{T}) + (1 - \tau_{\mathcal{C}}^0) \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) \quad (\text{B.25})$$

Inequality (B.25) then yields

$$\mathbf{Z}(\mathcal{J}^0, 1, \boldsymbol{\sigma}^*) > \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) \quad (\text{B.26})$$

Assume now that \mathcal{J}^0 is such that $\tau_{\mathcal{C}}^0 = 0$ and δ is sufficiently high so that condition (9) must be satisfied. Then the equilibrium $\boldsymbol{\sigma}^*$ in Proposition 4 is such that there exists a steady state value $\hat{\tau}_{\mathcal{C}} > \tau_{\mathcal{C}}^*$ such that the steady state ex-ante welfare in state \mathcal{C} is: $\hat{\tau}_{\mathcal{C}} \mathbf{\Pi}(\mathcal{C}, \mathcal{T}) + (1 - \hat{\tau}_{\mathcal{C}}) \mathbf{\Pi}(\mathcal{C}, \mathcal{W})$. This implies that

$$\lim_{\delta \rightarrow 1} \mathbf{Z}(\mathcal{J}^0, 1, \boldsymbol{\sigma}^*) = \hat{\tau}_{\mathcal{C}} \mathbf{\Pi}(\mathcal{C}, \mathcal{T}) + (1 - \hat{\tau}_{\mathcal{C}}) \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) > \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) \quad (\text{B.27})$$

Putting together definition (B.23) and inequalities (B.24), (B.26), (B.27) we can conclude that:

$$\lim_{\delta \rightarrow 1} \mathbf{Z}(\mathcal{J}^0, \rho_{SL}^*, \boldsymbol{\sigma}^*) > \rho_{SL}^* \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) + (1 - \rho_{SL}^*) \mathbf{\Pi}(\mathcal{C}, \mathcal{W}) \quad (\text{B.28})$$

Inequality (B.28), together with (B.21) and (B.22), concludes the proof. ■

Appendix C: The Example in Figure 1

Fix \mathcal{J}^0 and δ . As initial precedent we choose $\mathcal{J}^0 = (0, 1, 0, 1)$. The precedents technology is as follows. When discretionary Courts choose $\mathcal{R}^t = \mathcal{T}$ in environment \mathcal{E} , we have that

$$\tau_{\mathcal{E}}^{t+1} = \tau_{\mathcal{E}}^t + \frac{b^t}{2} (1 - \tau_{\mathcal{E}}^t) \quad (\text{C.1})$$

$$\omega_{\mathcal{E}}^{t+1} = \omega_{\mathcal{E}}^t + \frac{b^t}{2} (1 - \omega_{\mathcal{E}}^t) \quad (\text{C.2})$$

When instead they set $\mathcal{R}^t = \mathcal{W}$ in environment \mathcal{E} , we have

$$\tau_{\mathcal{E}}^{t+1} = \frac{2\tau_{\mathcal{E}}^t}{2-b^t} - \frac{b^t}{2-b^t} \quad (\text{C.3})$$

$$\omega_{\mathcal{E}}^{t+1} = \frac{2\omega_{\mathcal{E}}^t}{2-b^t} - \frac{b^t}{2-b^t} \quad (\text{C.4})$$

We also set $\Pi(\mathcal{C}, \mathcal{T}) = 15$, $\Pi(\mathcal{C}, \mathcal{W}) = 5$, $\Pi(\mathcal{S}, \mathcal{T}) = 5$, $\Pi(\mathcal{S}, \mathcal{W}) = 45$, $\widehat{\Pi}(\mathcal{C}, \mathcal{W}) = 20$ and $\widehat{\Pi}(\mathcal{C}, \mathcal{T}) = 15$. This implies that $\rho_{SL}^* = 4/5$. With these numbers it is immediate to see that for $\rho \leq 4/5$

$$\mathbf{\Pi}_{SL}(\rho) = 45 - \rho(45 - 5) \quad (\text{C.5})$$

and for $\rho \geq 4/5$

$$\mathbf{\Pi}_{SL}(\rho) = 5 + \rho(15 - 5) \quad (\text{C.6})$$

Assume that $\delta = 5/8$. This implies that discretionary Courts take at most one decision \mathcal{T} in the \mathcal{C} state. To see that this is true, consider the case $\rho = 1$ and observe that

$$\left(1 - \frac{5}{8}\right)5 > \frac{5}{8} \frac{15 - 5}{4} \quad (\text{C.7})$$

Let $\bar{\rho}$ be the solution to

$$\left(1 - \frac{5}{8}\right)5 = \bar{\rho} \frac{5}{8} \frac{15 - 5}{2} \quad (\text{C.8})$$

Hence $\bar{\rho} = 3/5$.

Note that for $\rho < 3/5$ Courts have no incentives to make a single tough decision. Hence, when $\rho < \bar{\rho}$ all Courts take the \mathcal{W} decision, regardless of the environment and therefore

$$\mathbf{\Pi}_{CL}(\rho, \sigma) = 45 - \rho(45 - 5) \quad (\text{C.9})$$

Consider now $\rho \geq 3/5$. In this range, discretionary Courts have incentives to make a single tough decision that moves $(\tau_{\mathcal{C}}, \omega_{\mathcal{C}})$ to $(1/2, 1)$. Hence

$$\mathbf{\Pi}_{CL}(\rho, \sigma) = -3 + (1 - \rho)45 + \frac{\rho}{2} [15 + 5] \quad (\text{C.10})$$

References

AGHION, P., AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1–29.

- AMADOR, M., G. ANGELETOS, AND I. WERNING (2006): "Commitment versus Flexibility," *Econometrica*, 74(2), 365–396.
- ANDERLINI, L., L. FELLI, AND A. POSTLEWAITE (2006): "Should Courts Always Enforce What Contracting Parties Write?" University of Pennsylvania, Institute for Law & Economic Research, Paper No. 06-24.
- ATHEY, S., A. ATKESON, AND P. KEHOE (2005): "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 73(5), 1431–1475.
- AYRES, I. (1997): "The Twin Faces of Judicial Corruption: Extortion and Bribery," *Denver University Law Review*, 74, 1231–1253.
- BOND, P. (2007): "Contracting in the Presence of Judicial Agency," University of Pennsylvania, *mimeo*.
- (2008): "Persistent Court Corruption," *Economic Journal*, forthcoming.
- CALABRESI, G. (1982): *A Common Law for the Age of Statutes*. Cambridge: Harvard University Press.
- CARDOZO, B. N. (1921): *The Nature of the Judicial Process*. New Haven: Yale University Press.
- 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.

- GLAESER, E. L., AND A. SHLEIFER (2002): "Legal Origins," *Quarterly Journal of Economics*, 117, 1193–1229.
- GOODMAN, J. C. (1978): "An Economic Theory of the Evolution of the Common Law," *Journal of Legal Studies*, 7, 393–406.
- GROSSMAN, S. J., AND O. D. HART (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–719.
- HADFIELD, G. K. (1992): "Bias in the Evolution of Legal Rules," *Georgetown Law Journal*, 80, 583–616.
- (2007): "The Quality of Law: Judicial Incentives, Legal Human Capital and the Evolution of Law," *The Selected Works of Gillian K. Hadfield*. Available at <http://works.bepress.com/ghadfield/6>.
- (2008): "The levers of legal design: Institutional determinants of the quality of law," *University of Southern California Law and Economics Working Paper Series*. Working Paper 78.
- HART, O. D., AND J. MOORE (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–58.
- HASSLER, J., AND J. V. RODRÍGUEZ MORA (2007): "Political Commitment and Loss Aversion," *University of Edinburgh*, *mimeo*.
- KAPLOW, L. (1992): "Rules versus Standards: An Economic Analysis," *Duke Law Journal*, 42, 557–629.
- KAPLOW, L., AND S. SHAVELL (2002a): "Economic Analysis and the Law," in *Handbook of Public Economics, Volume 3*, ed. by A.J. Auerbach and M. Feldstein. Amsterdam: Elsevier.
- KAPLOW, L., AND S. SHAVELL (2002b): *Fairness Versus Welfare*. Cambridge Massachusetts: Harvard University Press.
- KYDLAND, F. E., AND E. C. PRESCOTT (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 87, 473–492.

- LA PORTA, R., F. LOPEZ-DE-SILANES, AND A. SHLEIFER (2005): "What Works in Securities Law?," *Journal of Finance*, Forthcoming.
- LA PORTA, R., F. LOPEZ-DE-SILANES, A. SHLEIFER, AND R. W. VISHNY (1997): "Determinants of External Finance," *Journal of Finance*, 52, 1131–1150.
- (1998): "Law and Finance," *Journal of Political Economy*, 106, 1113–1155.
- (1999): "The Quality of Government," *Journal of Law Economics & Organization*, 15, 222–279.
- (2002): "Investor Protection and Corporate Valuation," *Journal of Finance*, 57, 1147–1170.
- LAMOREAUX, N. R., AND J.-L. ROSENTHAL (2004): "Legal Regime and Business's Organizational Choice: A Comparison of France and the United States," NBER Working Paper No. W10288.
- 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.
- LOMBARDO, D., AND M. PAGANO (1999): "Legal Determinants of the Return on Equity," Stanford Law and Economics Olin Working Paper, No. 193.
- (2002): "Law and Equity Markets: A Simple Model," in *Corporate Governance Regimes: Convergence and Diversity*, ed. by McCahery, J. and Moerland, P. and Raaijmakers, T. and Renneboog, L. Oxford University Press, Oxford.
- 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 (2008): "The Evolution of a Legal Rule," Harvard University, NBER Working Paper No. W13856.

- NORDHAUS, W. (1969): *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, MA: The MIT Press.
- OBSTFELD, M. (1997): “Destabilizing Effects of Exchange-Rate Escape Clauses,” *Journal of International Economics*, 43(1), 61–77.
- PHELAN, C. (2006): “Public Trust and Government Betrayal,” *Journal of Economic 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 Evolutionary Comparison,” *Journal of Legal Studies*, forthcoming.
- POSNER, R. A. (1990): *The Problems of Jurisprudence*. Harvard University Press.
- (2003): *Economic Analysis of Law*. 6th edn. New York: Aspen.
- (2004): “The Law and Economics of Contract Interpretation,” University of Chicago, mimeo.
- PRIEST, G. (1977): “The Common Law Process and the Selection of Efficient Rules,” *Journal of Legal Studies*, 6, 65–82.
- RAJAN, R., AND L. ZINGALES (1998): “Power in a Theory of the Firm,” *Quarterly Journal of Economics*, 113, 387–432.
- (2003): “The Great Reversals: The Politics of Financial Development in the 20th Century,” *Journal of Financial Economics*, 69, 5–50.
- ROGOFF, K. (1985): “The Optimal Degree of Discretion in Monetary Policy,” *Quarterly Journal of Economics*, 100(4), 1169–1189.
- 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.
- (2007): “Optimal Discretion in the Application of Rules,” *American Law and Economics Review*, 9(1), 175–194.
- 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.
- SULLIVAN, K. (1993): “Foreward: The Justices of Rules and Standards,” *Harvard Law Review*, 106, 22–123.
- SUMMERS, R. (1992): *Essays on the Nature of Law and Legal Reasoning*. Berlin: Duncker and Humblot.
- SUNSTEIN, C. R. (1995): “Problems with Rules,” *California Law Review*, 83, 853–1023.
- TIROLE, J. (2005): *The Theory of Corporate Finance*. Princeton, NJ: Princeton University Press.
- VON MEHREN, A. T. (1957): *The Civil Law System*. Englewood Cliffs: Prentice Hall.