The Quality of Law: Judicial Incentives, Legal Human Capital and the Evolution of Law

Gillian K. Hadfield
University of Southern California

December 18, 2007

Abstract

Much of the existing literature investigating the relationship between legal regimes and economic growth focuses on the agency problem of aligning judicial incentives with social welfare. In this paper I look instead at the factors that influence the quality of law when judges have incentives to promote social welfare but they have limited knowledge about the environment in which law is to be applied. The key insight is that the capacity for a legal regime to generate value-enhancing legal adaptation to local and changing conditions depends on its capacity to generate and implement adequate expertise about the environment in which law is applied. The central mechanism of adaptation is the interaction among three factors: 1) judicial incentives for rule-following and rule-adaptation, 2) litigant incentives for investing in costly evidence and innovative legal argument and 3) the accumulation of *shared legal human capital*–defined as the sum of litigant investments in evidence and argument– which determines the systemic likelihood of judicial error.

1 Introduction

As developing countries and countries transitioning from planned economies struggle to develop the institutions that support market democracy, there has been increased attention from economists and legal scholars directed to the question of what legal environments best promote economic growth and stability. Much of this work focuses on the substance of legal rules: the provisions of a constitution, the elements of a corporations or antitrust statute, the law governing the enforcement of contracts or property rights. Relatively little attention is paid to the institutions of the legal system, such as the organization of courts, the judiciary and the legal profession. A recent exception is the legal origins literature (La Porta et al 1997, 1998, 2004, Mahoney 2001, Djankov et al 2002, 2003, Botero et al 2004), which identifies an empirical relationship between economic variables and legal families broadly identified as those rooted in

1

civil law (French, German, Scandinavian) and common law (English). While some strands in this literature are explicitly focused on differences in substantive law (financial or administrative regulations for example) that appear correlated with legal origin, others suggest that differences arise from the institutional features of different legal families, apart from the substantive law they implement. Some writing in the comparative literature on the common law and the civil law suggest, for example, that these regimes differ in the extent to which judges (or juries) are independent of distortionary political control (Glaeser and Shleifer 2002, Mahoney 2001). Others have explored differences in the information available to and the incentives facing judges as opposed to legislators or regulators. Shavell (2005) analyzes the value of judicial discretion when judges have better (ex post) case-specific information than legislators or regulators but preferences that may diverge from social welfare. Although Shavell's work is framed as a choice between more and less detailed rules (and thus related to an earlier literature comparing the costs and benefits of regulation by rules versus regulation by standards (Diver 1983, Rose 1988, Kaplow 1992)), it is easily interpreted to address the institutional question of how the judicial role should be structured. Anderlini, Felli and Riboni (2006) engage in a very similar type of analysis, supposing that judicial and legislative incentives diverge because although better informed, judges acting ex post face time-inconsistency in their preferences and cannot commit to implementing an ex ante efficient rule. Anderlini, Felli and Riboni explicitly place this work in the context of the choice between a legal regime based on codes versus one based on judge-made precedents.

The existing literature thus tends to focus heavily on the agency problem: what is the likelihood that judges will implement established legal rules and forego private benefits that might be derived from personal policy preferences or corrupting political or social influences? Missing from the existing literature is an equally important question: what is the likelihood that judges will accurately identify, interpret and implement legal rules in a complex environment? This is the problem of judicial competence and the quality of legal rules as they developed and applied in practice. Even judges with socially-aligned incentives and access to better information than that available to legislators may make good faith errors in interpreting evidence and exercising discretion in socially optimal ways. Moreover judges are vulnerable to being misled by strategic litigants who may distort the evidence they present or the arguments they make about how a judge should exercise his or her discretion to interpret or adapt law. Several papers have explored strategic revelation of private information in models of adjudication (Milgrom and Roberts 1986, Shin 1994, Shin 1998, Dewatripont and Tirole 1999, Daughety and Reinganum 2000.) By and large, however, this work has not been integrated into the institutional frameworks of different legal regimes.[1]

---

[1] Shin (1998) compares what he calls adversarial and inquisitorial regimes, but these do not map onto the institutional differences between, for example, common law and civil law. In particular, the inquisitorial regime he analyzes is one in which judges bear the cost of discovering evidence as opposed to parties. Although it is true that judges in many civil law regimes are empowered to seek out evidence, by contacting a title registry or a bank for

2

In this paper I focus specifically on a mechanism by which the structure of judicial incentives, in light of the potential for judicial error, affects the dynamic quality of law. Positive economic analysis of the common law has, since Posner (1977), been organized around the claim that the value of the common law is its ability to work out, over time, efficient legal rules. Some authors have rested this claim on the premise that common law judges inherently seek efficiency; often this literature has framed the analysis as an investigation of the different incentives influencing parties interacting with courts and legislatures and as a debate about whether judges or legislators are more susceptible to rent-seeking (Posner 1977, Rubin 1982, Tullock 1980). A recent model by Gennaioli and Shleifer (2007) focuses on the impact of judicial bias on the capacity of common law to evolve to efficiency. Hadfield (1992) considers the capacity of common law to evolve to efficiency when the information generated through litigation is necessarily a biased sample, in contrast to the potential for legislative investigation to be comprehensive. Other authors have rested predictions about the likelihood that the common law will evolve to efficiency on the incentives of litigants to continue challenging inefficient rules (Rubin 1977, Priest 1977, Goodman 1978, Cooter, Kornhauser and Lane 1979). Despite differences, these models of the evolution of the common law all share a common recognition that courts do not start out with the right rules. Rather, they move towards them based on the information learned directly or indirectly from litigants who bring cases to them.

The dynamic quality of law is especially important for the evaluation of alternative legal regimes. As Berkowitz, Pistor and Richard (2003) and Botero et al (2003) have emphasized in evaluating the empirical evidence of the growth generated by transplanted legal regimes, the value of a legal regime depends on its ability to adapt to local conditions. Intuitively also, in a changing environment, law must adapt to changing conditions in order to continue to promote economic value in the organization of activities and trade. At its core, a market economy is decentralized in its response to information about the environment. The more important this is to the organization of economic activity, the more important we can expect the adaptability of a legal system to be. This is the insight of the rules versus standards literature, recognizing the value of what judges learn about a heterogeneous environment through adjudication. The point has also been recognized in the legal origins literature: Johnson et al (2000), for example, attribute the differential success in controlling "tunnelling" (the removal of assets from a company by controlling shareholders at the expense of minority shareholders) in civil law and common law countries to the capacity of common law courts to develop more refined regulation of opportunistic behavior based on what is learned in litigation from specific instances of abuse. Given the importance of such 'grass-roots' information, it is essential to understand the capacity of different institutional environments to support the dynamic evolution of a legal regime, and to direct its development to optimal adaptation to

---

example, much of the cost of evidence production is still borne by litigants, as is the case under an adversarial regime.

local and changing circumstances.

The central mechanism of adaptation in this paper is the equilibrium interaction among three factors: 1) judicial incentives for rule-following and rule-adaptation, 2) litigant incentives for investing in costly evidence and innovative legal argument and 3) the accumulation of *shared legal human capital*–defined as the sum of litigant investments in evidence and argument– which determines the systemic likelihood of judicial error. Landes and Posner (1976) also develop a model which conceptualizes precedent as a stock of legal capital produced by the investments of lawyers, litigants and judges. The services provided by the stock of legal capital in their model consist primarily in the information precedents provide to future disputants about the likely outcome of their disputes. Here I am focused on the value of legal human capital in generating more precise (efficient) legal rules. In addition, I explicitly address the question of how legal human capital accumulates; Landes and Posner (1976), who concentrate primarily on empirical tests of the depreciation of precedents, take the investment in precedent as exogenous.

The key insight here is that the capacity for a legal regime to generate value-enhancing legal adaptation to local and changing conditions depends on its capacity to generate and implement adequate expertise about the environment in which law is applied. The process by which this happens in a legal regime is an organic and evolutionary one, dependent on institutional design and the equilibrium coordination of the work of judges, lawyers and litigants. Efforts to develop legal regimes to support economic growth and efficiency, therefore, must take into account the impact of legal design on legal human capital and the incentives of lawyers and judges. A focus on the static content of legal rules is inadequate and misleading.

Section 2 constructs a model of legal adaptation in the context of legal error and the accumulation of legal human capital. Judicial incentives are modeled in terms of the private returns to rule-following and to error-free rule adaptation; judges differ in the returns they enjoy from these two decision-making approaches. Section 2 considers the case in which all defendants are all the same and the existing rule–which holds a defendant liable for damages–is inferior to a revised rule under which all defendants are released from liability. The analysis shows that in order for a legal regime to switch to the new, superior, rule three conditions must be met: legal costs and judicial error must not be too high and judicial rewards for rule-change must not be too low. I then examine the institutional attributes that, in a comparative analysis, would lead to the prediction that one legal regime will enjoy more widespread rule change than another: lower judicial error rates, higher judicial rewards for rule-change and more effective methods of extracting information from individual cases and converting this information into systemic error-reducing legal human capital. In the case of homogeneous defendants, by construction, more widespread rule change leads to higher social welfare.

Section 3 then turns to the case of heterogeneous defendants. Defendants–both those who should be found liable (*bad types*) and those who should not (*good types*)–choose whether to invest in costly evidence and argument to edu-

cate the court and persuade a judge to adapt the rule to their circumstances. Their incentive to do so depends on the cost of evidence and argument relative to the amount at stake, and the likelihood that the judge will interpret their evidence and argument as a basis for a finding of no liability. Judicial error is modeled as unbiased uncertainty about a defendant's true type. Because the error is unbiased, good types are more likely than bad types to be seen by judges to be good types and so to benefit from their investment in evidence and argument. Good defendants are therefore more likely to invest than bad defendants. With bad defendants in the pool, it is now possible for a regime to be mired at the existing rule not only if legal costs are too high, but also if they are too low: low legal costs encourage bad defendants to invest, increasing the risks of rule-change for judges. The presence of bad defendants also makes the welfare analysis more complex. Rule change may now be costly in the sense that judges may make the wrong decision in particular cases, mistaking bad types for good types. However, the capacity to learn from litigants–even bad litigants–may outweigh the errors incurred with rule-change, as evidence and legal argument accumulates that, over time, reduces judicial error.

Section 3 works through the comparative analysis of how different institutional attributes will affect the welfare achieved by different legal regimes. In this more complex setting, we can conclude that–with some exceptions–lower absolute or relative legal costs and better information processing will generate higher welfare; higher judicial rewards and lower initial legal errors, however, may or may not lead to better results for a legal regime. This latter ambiguity derives from a basic tradeoff that arises: initial rates of judicial error and legal costs that do not warrant, of themselves, a shift to a new rule initially may nonetheless generate legal human capital that will reduce future errors enough to compensate for the investment in error. When there is some optimal level investment in costly rule-change to generate future returns, factors that lead a regime to more extensive rule-change early on may reduce welfare relative to a regime that limits initial rule change. Finally in Section 3 I consider the implications of 'disinformation'–the case in which the evidence and legal argument generated by bad defendants degrades legal human capital, leading to higher rather than lower rates of judicial error. In this case, a regime may be better off with higher relative legal costs, which screen out bad defendants.

Section 4 relates the model to the institutional attributes of legal regimes, to suggest how the model can be used to deepen our understanding of the comparative costs and benefits of legal systems. The primary lesson of the model is to de-emphasize the importance of the common law/civil code classification in comparative analysis and to highlight the far more nuanced details of legal regimes that structure judicial incentives and the sharing of legal human capital. This section offers a brief sketch of these differences and points to the need for a more detailed approach to the empirical investigation of the impact of different legal systems on economic welfare. I also consider here some directions for further theoretical work and policy implications.

2 Homogeneous Defendants

Assume a two-period world in which there is a population of judges, plaintiffs and potential defendants and an existing legal rule, R^e. This existing rule determines the liability of defendants sued under the rule on the basis of information about a defendant's conduct that is costlessly observable to judges, plaintiffs and defendants and judges make no mistakes in applying the rule. If a defendant is held liable it pays damages, D. In this section I consider the case in which all defendants are the same and treated the same under R^e. For simplicity, I assume that all cases are tried to conclusion in each of the two periods, with no settlements and no appeal.

We are interested in the capacity of a legal regime to adapt this existing legal rule to new information, that is, to welfare-relevant changes in the environment in which the rule operates or changes in our understanding about the welfare implications of a rule. The new information could be about a technological change, for example, that reduces the cost of precautionary efforts and so makes previously non-negligent behavior negligent. It could be about a new theoretical basis for concluding that existing judicial policies to expansively interpret contracts ultimately increase contracting costs. It could be the recognition that in the context of a particular economy undergoing transition from socialism to market capitalism it is not reasonable to assume that buyers are easily able to obtain substitute goods, making the standard remedies for breach of a sales contract different from those in a more developed market. For concreteness, I look at the case in which all defendants are held liable under R^e and the new information provides a valid legal defense, implying that all defendants should be released from liability. (In Section 3 I look at the case in which defendants are heterogeneous, meaning that it would be optimal to hold some but not all defendants liable.) The analysis for the case in which the existing rule holds no defendants liable and the new information warrants a new cause of action under which all defendants are held liable is equivalent. Also, for simplicity, I assume that plaintiffs do not contemplate the possibility of a switch from R^e to R^n; they therefore sue all potential defendants.

The novelty of the welfare-relevant information about defendants' conduct implies that conveying this information to courts is costly: defendants, who are sued under R^e, can present factual evidence and legal argument about the change at a cost k, seeking to persuade the court to switch to R^n, releasing them from liability for damages. (If we were looking at the case in which no defendants are held liable under R^e, k would be the cost to plaintiffs of presenting evidence and argument to the court to persuade the court to recognize a new cause of action and hold defendants liable for damages.) As I will discuss below, the information presented by defendants may or may not be reliable or properly understood, from a social welfare point of view, by judges. Defendants are only interested in presenting information, accurate or not, that promotes their private goal of avoiding liability. The fact that there has been a change in the social-welfare calculus relevant to the choice between rules simply creates the potential for defendants to make use of that change to argue for a change

that generates the private benefit of avoiding liability. We are interested in whether these private efforts can lead judges to make changes, and only those changes, that are socially optimal. This includes all the relevant considerations in the social welfare calculus, including the benefits of predictability in the law and the need to evaluate legal rules from an ex ante perspective.

Suppose that a defendant has presented the case for switching to R^n. How will a judge decide whether or not to change rules? Our model of judges needs to capture the incentives for a judge to shift from established rules to novel rules. We have few satisfactory economic models of what motivates judges.[2] Models that look exclusively to the effort costs of judging are in tension with the occupation of judging: judges are expected to expend effort to decide cases and to do so with care. Models that specify financial incentives run into trouble because of the difficulty of describing the relationship between particular decisions and income, particularly for judges with life-tenure or other protections to achieve independence. Models that assume judges act exclusively to satisfy their preferences over policy seem not to capture basic judicial norms requiring judges not to pursue a private policy agenda and to follow rules instead. I avoid these specific problems by focusing on the particular aspect of judging in which I am interested–the decision whether to follow an existing rule or to adopt a new rule that may increase social welfare–and by constructing a model of judicial incentives that is sufficiently broad to allow for a wide mixture of incentives and motivations. Indeed, part of my goal in constructing this model is to explore how judicial incentives with respect to rule-following are a function of a number of parameters, and importantly related to the risk of judicial error.

Let γ be the reward judges ultimately enjoy when they follow existing rules. The source of the reward in any system of judicial incentives may vary. In a judicial bureaucracy such as those that characterize many civil code countries, the reward might be promotion within the ranks to a superior position or transfer to a more desirable court or city. In a common law country, the reward might be re-election, the prestige or perks derived from being well-thought of by professional peers and/or the public, or (rarely) appointment to a higher court.[3] In any regime, the reward might be the personal satisfaction that is derived from adhering to a professional norm of rule-following. For simplicity I assume this reward is the same for all judges. Now consider the reward for adopting a new rule. I assume that the reward enjoyed by a judge who changes the rule varies among judges and depends on whether those who determine judicial rewards (the judge's audience–superiors in a civil service bureaucracy, the public, politicans, etc.) conclude that the rule change generated an increase in social welfare. Let a_j be the reward judge j expects to enjoy if he or she *correctly* adapts or changes a rule; a_j is distributed with positive density everywhere on the interval $[0, \bar{\alpha}]$ according to the cumulative distribution function G.

---

[2] For a discussion and citations to other sources, see Posner (1993).

[3] For an example of the reward structure in civil code systems, see Ramseyer and Rasmussen (1997) on Japan. For an analysis of reputation effects for judges in common law systems see Levy (2005). On the rarity of promotion to a higher court in the U.S. federal system see Klerman (1999).

I normalize the return judges obtain if they are found to have made an error and incorrectly changed the rule to be zero. Note that the concept of a correct rule and judicial error here is based on the social welfare effects of the rule, and thus is not the same as the jurisprudential concept of legal error meaning an appellate judgment that a lower court made an error in the application of a legal rule. To keep matters simple, I assume that those who assess the judge (including the judge him or herself) are accurate in their assessment of the welfare effects of the rule. This assessment is made at some point in the future, when all judges have rendered their decisions, there is time for reflection, and expert commentators on the law may have had an opportunity to weigh in. The assumption that this assessment is ultimately accurate is a strong one, and emphasizes that this is not a model of judicial corruption: judges do not obtain benefits or pay penalties based on promoting the interests of any sub-group at the expense of social welfare, nor do they have biased preferences for outcomes that do not promote social welfare. My reason for making this assumption is not that judicial corruption (whether of the policy preference variety or the more venal judge-for-sale variety) is not a problem–it clearly is–but in order to focus on the problem of generating judicial competence, an issue that has received far less attention in the literature on comparative legal evolution. Finally, note that I assume that judges do not pay a penalty for getting it wrong from a welfare perspective when they stick with the existing rule. This assumption makes the analysis simpler, but it also reflects a key idea that responsibility for the errors in an existing rule are the responsibility of the entire judiciary and/or the legislature, not any particular judge; the judge that takes it on him or herself to change the rule runs the risk of separating from the crowd and the background norm of following rules. Even if there are judicial penalties associated with following rules when they should be changed, the analysis is the same so long as the expected penalty for mistaken rule change are greater.

A judge will choose whether to change a rule or stick with an existing rule, then, based on a comparison between the rule-following payoff γ and the expected payoff for rule change, $(1-\hat{\sigma}_1)\alpha_j$, where $\hat{\sigma}_1$ is the judge's belief in period 1 about the likelihood that switching to R^n will be a mistake. Note that in this context the judge is concerned about the risk of a type 2 error–finding no liability when liability is the correct outcome–since by assumption all defendants are released from liability when R^n is applied. In fact, also by assumption, there is no risk of a type 2 error, but the judge does not know this when making his or her decision; this is what it means to say that the situation facing the judge is novel and that it is only over time that the welfare implications of the rule can be assessed. Let $\tilde{\alpha}_1 = \frac{\gamma}{(1-\hat{\sigma}_1)}$. Then a judge in period 1 will accept a defendant's argument to switch to R^n if $\alpha_j > \tilde{\alpha}_1$. I will call judges who satisfy this criterion rule-changers.

Now consider defendants' incentives to invest in the effort to persuade a judge to switch rules. I assume defendants are randomly assigned to judges and do not observe the judicial reward parameter, α_j, although they are aware of the systemic belief among judges about the likelihood that a rule change would

be a mistake, $\hat{\sigma}_1$ and know the distribution $G(\alpha)$. Defendants therefore have to decide whether the investment is worth it given the likelihood that their judge is not a rule-changer. This likelihood is given by $G(\tilde{\alpha}_1) = \sigma_1$, the true probability of an error from the defendant's perspective given that the defendant has full information and, by assumption, social welfare is higher under R^n. Note that this true error is a type 1 error: holding a defendant liable who should not pay damages. Defendants will invest in seeking a rule change in period 1 when $\sigma_1 D + k \leq D$ that is, when $\frac{k}{D} \leq (1 - \sigma_1) = (1 - G(\tilde{\alpha}_1))$. It is straightforward to see that there are three independent conditions for rule change to occur in period 1.

Proposition 1 *Rule change will occur in period 1 with homogeneous defendants if and only if (1) k is not too high relative to D, and in particular $\frac{k}{D} \leq (1 - G(\tilde{\alpha}_1)) \leq 1$; (2) rewards to correct rule change are sufficiently high, and in particular for any given level of expected error $\hat{\sigma}_1$, $\bar{\alpha} > \frac{\gamma}{(1 - \hat{\sigma}_1)}$; and (3) judicial expectations about the likelihood that rule change is an error are sufficiently low and in particular for a given $\bar{\alpha}$, $\hat{\sigma}_1 \leq 1 - \frac{\gamma}{\bar{\alpha}}$. Assuming these three conditions are met, R^n will be adopted by a proportion of judges equal to $(1 - G(\tilde{\alpha}_1))$.*

Proposition 1 contains an important insight. In order for rule change to occur, it is necessary for *both* defendants *and* at least some judges to face incentives to incur the costs associated with rule change. These are the costs of legal evidence and argument in the case of defendants and the costs associated with risking a mistaken decision for judges. Thus a legal regime in this simple world can easily be stuck at R^e, despite the existence of an optimal R^n known to defendants, either because the costs of persuading courts to change their rule are too high ($k > D$, for example) or because no judges perceive an adequate reward to risking rule change ($\bar{\alpha} < \gamma$, for example.) Note the important role of judicial expectation of error in the decision of both defendants and judges: if judges perceive this error to be significant, then fewer judges (if any) perceive the risk of rule change to be worth taking on. The smaller the number of judges that perceive rule change to be worth the risk, the lower the likelihood that a defendant is facing a rule-changing judge and the lower the returns to investing in the effort to persuade the judge to change the rule. If this proportion of rule-changing judges is small enough, no defendants will invest in the evidence and legal argument necessary to present the case to any judges, and no rule change will occur.

What determines judicial beliefs about the welfare consequences of changing the rule? I assume that judicial competence–expertise about the relationship between rules and social welfare in a given environment–is determined by *shared legal human capital*, the accumulated level of knowledge within the legal system as a whole. I abstract here from individual differences between judges to focus on what judges learn systemically. Moreover, to analyze the importance of information generated within the legal system and made available to other judges, I assume that judges can only learn from the evidence and legal arguments that

are presented to individual courts.[4] This learning occurs with a lag. A judge in period 1 makes decisions based on the pre-existing belief $\hat{\sigma}_1$. A judge in period 2 however has had a chance to reflect on the information learned in period 1. Moreover, and more importantly, the judge has had a chance to compare what was learned in his or her own case (which may have been nothing) with what other judges might have learned and shared through publication or some other means, and to review what other commentators (law professors, journalists, litigants, experts, regulators, lawyers) might have had to say about the period 1 decisions.

Let Δ be the total amount the pool of defendants invested in evidence and legal argument in period 1. Note that because defendants are homogeneous, either all defendants invest or none do. Then let $K_2 = K_1 + i(\Delta)$ be the shared legal human capital available to all judges in period 2 and $\hat{\sigma}_t = f(K_t)$. By definition, legal human capital reduces legal error: $\frac{\partial \hat{\sigma}_t}{\partial K_t} < 0$. i is an information processing function that captures the extent to which information in particular cases is shared among judges and other commentators and converted into legal human capital. I will say that i is *informative* if $i' > 0$ and *disinformative* if $i' < 0$. This function allows us to recognize the potential divergence between the private goals of defendants–which are to present evidence and argument that will induce a judge to release them from liability–and the social welfare goals of judges–which are to distill the information presented by defendants to determine whether releasing defendants from liability (as a rule) promotes social welfare.

Defendants' and judges' decisions in period 2 depend on the level of shared legal human capital generated by deicisons in period 1. Suppose the conditions in Proposition 1 were not met and thus there was no evidence and argument presented and no rule change in period 1. The important consequence of uniform rule-following in period 1 is that there is no legal human capital accumulation: $\Delta = 0$. Then judicial beliefs are unchanged, $\hat{\sigma}_2 = \hat{\sigma}_1$, and it is still the case in period 2 that no judges have an incentive to risk changing the rule and/or individual defendants perceive that the expected return to presenting evidence and argument is not worth the cost. There is no rule change in period 2 and the regime remains permanently stuck at R^e.

If, however, there was some rule change in period 1, then Δ is positive. Provided information processing by judges is informative, $i' > 0$, some of what was presented as evidence and argument in period 1 is now available to all judges as increased legal human capital, reducing the systemic level of error such that $\hat{\sigma}_2 < \hat{\sigma}_1$: judges are less likely to entertain the erroneous belief that there is a risk that rule change reduces welfare. Consequently, more judges are willing to risk rule change in period 2 and R^n becomes more widely adopted in period 2 than in period 1.   Note that this occurs without any institutional rule of *stare decisis* (courts are bound to follow the decisions of another court, the common law practice) or *jurisprudence constante* (courts are expected to follow

---

[4]This is reflected in the Roman legal maxim, *quod non est in actis, non est in mundo*: what is not in the documentary record does not exist.

the decisions of other courts when several have generated the same decision, the civil code practice). Such rules (which would modify the reward structure for judges, generating a payoff based on following other courts) could increase the rate at which the period 1 rule changers generated more widespread rule change in period 2. (A rule of jurisprudence constante could also decrease the spread of rule change in period 2, if there was not a critical mass of rule-changers in the first period and judicial incentives created a penalty for persisting in a minority rule.) The important point here is that the spread of rule change is a function of the accumulation of shared legal human capital within the system, which reduces the risk of error for judges. Even a small population of rule-changing judges in period 1 can thus generate more widespread rule change. Conversely, without these initial rule-changers, there is no legal human capital accumulated in the first period and rule evolution will not occur. We can summarize these period 2 results as follows:

Proposition 2 *If no rule change occurs in period 1, then no rule change occurs in period 2 and a legal regime remains at R^e. If some rule change occurs in period 1 ($\Delta > 0$) and judicial information processing is informative ($i' > 0$), then $\hat{\sigma}_2 < \hat{\sigma}_1$ and more judges are willing to allow rule change in period 2 ($\tilde{\alpha}_2 < \tilde{\alpha}_1$): rule change becomes more widespread.*

This simple model also allows us to see what factors will influence the degree of rule evolution in different regimes. Proposition 1 tells us that any regime with legal costs that are too high relative to damages ($\frac{k}{D}$), or initial beliefs about legal error, $\hat{\sigma}_1$ that are too high or judicial rewards for rule-changing α that are inadequate will remain stuck at R^e whereas one with sufficiently lower legal costs and/or perceived error and/or higher judicial rewards will experience at least some rule change. As between two regimes both of which experience some rule change, we can say the following:

Proposition 3 *Consider two regimes that satisfy the conditions of Proposition 1. Ceteris paribus, rule change in period 1 will be more widespread in the regime in which the initial perceived rate of legal error $\hat{\sigma}_1$ is lower or the density of judges with judicial rewards above the initial threshold value, $(1 - G(\tilde{\alpha}_1))$ is higher. Ceteris paribus, rule change in period 2 will be more widespread in the regime in which the rate at which evidence and legal argument in individual cases is converted into shared legal human capital ($i(\Delta), i' > 0$) is higher. Even if both regimes have the same process for generating shared legal human capital, $i(\Delta)$, provided this function converts evidence and argument presented to rule-changing judges into shared legal human capital at a higher rate than evidence and argument presented to rule-following judges, any gap between the two regimes in the extent of rule change generated in period 1 by differences in $\hat{\sigma}_1$ or $G(\tilde{\alpha}_1)$ will increase in period 2.*

The first part of Proposition 3 gives us the relatively unremarkable result that there will be more widespread rule change in period 1 in a regime in which there are more judges who expect to receive rewards for rule-change. The more

subtle result shows up in period 2 and stems from the relationships between rule change, shared legal human capital and judicial error. Even if judges in the two regimes are the same, facing the same set of rewards and beliefs about the likelihood of error, differences in the extent of rule change can emerge in period 2 if one regime is more effective than another at converting the presentations of evidence and legal argument in individual cases into informative shared legal human capital. This can happen if, for example, one regime publishes legal opinions while the other does not, or if judges in one regime write more extensive opinions than the other. It could happen if the output of judges in one regime is analyzed more carefully by outside experts, be they other judges, law professors or experts in the facts relevant to a particular type of case. Regimes that are more effective at capturing the information value of individual cases and distributing that within the judiciary will reduce the rate of judicial error and thus increase the incentive of judges to risk rule-change.

The knock-on effect of rule-change is thus increased judicial competence systemwide and hence even further rule-change. This effect will magnify initial differences between regimes in terms of the distribution of judicial rewards or beliefs if more is learned systemwide from the cases in which rule-change occurs than from the cases in which rule change is rejected. (In this simple model, if any defendants present evidence, all defendants invest in evidence, whether they appear before rule-changing or rule-following judges.) There are several ways in which this difference could arise. Judges who are rule-followers–meaning that they perceive the expected rewards for rule-change to be too low–may simply disallow the presentation of evidence and legal argument that is intended to cause an abandonment of the existing rule. The stronger the role of the judge in controlling the presentation of evidence and legal argument–judges in civil code regimes, for example, typically play a greater role in shaping the issues and evidence than do common law judges (see Hadfield 2008 for a discussion)–the more likely it is that evidence and argument proffered by defendants will simply never make its way into the case file or decision in the first place. (Note that if the regime operates in this way, defendants are likely to hold off investing in the evidence and argument until they have a green light from the judge that he or she is interested in hearing it.) Even if some evidence and argument is allowed by rule-following judges, the quantity may be less than when the defendant's presentation is allowed by a rule-changing judge. The judge may not need to hear much before deciding to deny the request for rule-change; the judge who ultimately changes a rule based on the defendant's presentation, on the other hand, may well want to hear as much evidence and argument as the defendant can muster, in order to overcome his or her concerns about the risk of error and/or in order to have as much material as possible with which to craft an opinion that justifies the rule-change. And even if the same amount of information and argument is received by the court in both cases, the judge who denies the request for rule-change may feel little need to provide an exhaustive treatment of the evidence and argument in the written decision; the judge who changes the rule may be more likely to face an incentive to provide detailed facts and reasons for the change. Finally, even if the same amount

of information is conveyed in the written decision of the rule-changing and the rule-following judge, the rule-changing decision may attract more attention and analysis from outside commentators than the mere confirmation of the status quo rule. This external analysis can increase the information content of what was presented in a particular case in which the rule was changed relative to one in which it was not. Regimes that treat the information proffered by defendants differently depending on whether the rule was changed or not then will show an important positive feedback relationship between rule-change and legal human capital accumulation: more rule-change will lead to greater shared legal human capital and increased judicial competence, leading to more rule-change.

2.1 Optimality: Adaptive and maladaptive legal change

Is a regime with more widespread rule change and higher levels of shared legal human capital better off than one with more limited rule change and learning? By construction, social welfare in a given case is higher under R^n than R^e. Assume in particular that the gain in social welfare under R^n in a particular case is worth the expenditure k needed to obtain the new rule. Then it follows that the more widespread is rule change, the higher is social welfare: the right outcome is obtained in more cases. This rather trivial observation identifies an important feature of the case of homogeneous defendants: if the problem with the existing rule is that it gets it wrong for all defendants, then it can only be welfare-improving for judges to adopt the new rule. Although judges are concerned about the risk of error in deviating from the established rule, in fact there is no risk of error.

Although I have framed the analysis as an inquiry into the capacity for a legal regime to adapt to local and changing condition–an important attribute for a vital legal system–the important role of judicial error identifies the potential for maladaptive legal change. What if judges are misled by defendants into changing an optimal existing rule? Limited judicial competence would then mean that judges, at least initially, underestimate the risk of a type 2 error, the true probability of which equals 1. Defendants now will invest k when the likelihood that they will face a rule-changing judge who will make an error is sufficiently great. As before, it may be the case that the rewards to rule-change are sufficiently low that few or no judges will entertain rule-change and so there may be no movement from the existing rule. If judicial competence is sufficiently low–now meaning that the perceived risk of error $\hat{\sigma}_1$ is sufficiently low (the actual risk is high)–however, some investment and rule-change in period 1 will occur. This will reduce social welfare. But the error here may be self-correcting. This depends on whether the presentations of evidence and argument by individual defendants–which are intended to lead courts into making mistakes–are *ultimately* informative or disinformative, that is, whether the shared legal human capital that is generated over time increases or decreases judicial competence. If the process of reflection on earlier decisions–by judges and other legal commentators–is able to identify the mistakes of rule-changing judges, then the misadventures of these judges in period 1 will partially correct

the systemic underestimation about the likelihood that rule-change is a mistake. Consequently there will be less rule-change in period 2 and perhaps none. This is the case in which all evidence and legal argument, even when intended to generate errors, is ultimately informative.

It may be the case, however, that the accumulation of shared legal human capital based on misleading evidence and legal argument is disinformative, such that $i' < 0$. In this case, the process of sharing what is learned in one case does not correct the error, it compounds it. We might expect this to be more likely if the information is shared only among judges whose knowledge is limited to what is learned from litigants and does not include external commentators, or if pro-defendant external commentators have a greater impact on the shared understanding of judges about the quality of first period rule-changes than pro-plaintiff commentators. The assumption that all judges have the same level of judicial competence and hear the same evidence and argument reinforces the likelihood that sharing the information among judges may not lead to the discovery of error, simply the replication of error. If this is the case, then period 1 errors are reinforced and made worse in period 2, as maladaptive rule-change spreads.

The potential for a new rule to be maladaptive–to generate results worse than under the existing rule–brings us to the generic problem of heterogeneity among defendants, where it is optimal to hold some liable but not others. I turn to this case now.

3   Heterogeneous Defendants

Suppose there are two types of defendants, *high* types and *low* types; defendants know their type. The probability that a defendant is a high type is p; this probability is known to judges and plaintiffs but the defendant's type in a given case is unobservable to judges and plaintiffs without proof. Under the existing rule R^e all defendants are held liable, without regard to type. Under the optimal rule, R^n, if types could be costlessly and perfectly identified, low types would be held liable and high types would be entitled to a defense, releasing them from liability. I will therefore refer to high types as good defendants and low types as bad defendants. Judges can make either of two types of errors if they recognize the availability of a defense, adopt R^n and attempt to apply it in a given case: they could hold a high type liable (type 1 error, σ^1) or release a low type from liability (type 2 error, σ^2). There are two interpretations about the source of these errors. Judges may know that under the optimal rule only high types should be entitled to the defense, but make errors in intepreting the defendant's evidence about type, misidentifying high as low or low as high. Alternatively, judges may be able to observe type but not know whether the defense is appropriately made available to only high types, both types or neither. That is, the judge does not know for sure whether high types are good defendants or low types are bad defendants. (The judge's recognition of the potential availability of a defense to some third type is thus theoretical; although this

seems strained in the two type model, it has a natural interpretation if there is are multiple good and bad types or a continuum. Multiple types or a continuum, however, merely makes the exposition more complicated without changing the basic results.)

To streamline the analysis, I will change the basic assumptions from the homogeneous case in two ways. First, I assume that defendants do not invest in producing evidence and legal argument until they have received approval from a judge; legally, we would say that the defendant proffers the defense and the judge indicates whether he or she is willing to consider the defense. The defense can be considered and then rejected, after the judge has seen the evidence and heard the argument. I will say that the judge has adopted R^n if the proffer is accepted; adopting the new rule, however, does not mean that the defendant will be released from liability.[5] Second, I now assume that judges know the true probability of type 1 and type 2 errors in each period, σ_t^1 and σ_t^2. As in the homogeneous case, I assume defendants also know these true probabilities. Let $\sigma_t^1 = f^1(K_t)$ and $\sigma_t^2 = f^2(K_t)$, with $\frac{df^i}{dK_t} < 0$, $\frac{d^2 f^i}{dK_t^2} > 0$.

Consider first defendants' incentives to invest in the cost of producing evidence and argument to support a claimed defense. Defendants know before they make this decision whether they are facing a rule-changing or a rule-following judge, but analogously to the homogeneous case, they do not know whether their presentation of a defense will be successful. For high types, the investment in the defense is worth it if $\sigma_t^1 D + k \leq D$, that is, if relative costs $\frac{k}{D} \leq (1-\sigma_t^1)$. For low types, the investment is worth it if $(1 - \sigma_t^2)D + k \leq D$, that is, if $\frac{k}{D} \leq \sigma_t^2$. Notice that good types are discouraged from investing by the risk of type 1 errors; bad types are encouraged by the risk of type 2 errors. I assume that defendants' true type has some impact on the information content of what they present as evidence and argument and in particular that good types are more likely than bad types to be identified as good types, and bad types are more likely than good types to be identified as bad types. Formally, $(1 - \sigma_t^1) > \sigma_t^2$. Given this, we have the following result:

Lemma 4 *If relative legal costs are high relative to the likelihood that a judge will accurately identify a high type as a good defendant, that is, if $\frac{k}{D} > (1-\sigma_t^1)$, no defendants are willing to invest in presenting a defense. If relative legal costs are low relative to the likelihood that a judge will accurately identify a high type as a good defendant, but high relative to the likelihood that a judge will mistakenly identify a low type as a good defendant, that is, if $\sigma_t^2 < \frac{k}{D} \leq (1-\sigma_t^1)$, only good defendants are willing to invest in presenting a defense. If relative legal costs are low relative to the likelihood that a judge will mistakenly identify*

---

[5] In some legal regimes, such as the American, strictly speaking a decision does not establish a rule of law unless the rule is necessary to account for the result in the case. Thus the announcement of R^n would not in fact work a change away from R^e unless the defendant was released from liability. The availability of the defense, assuming it is discussed in the opinion, would be 'mere dicta.' In other legal regimes, and in practice in the American system, the expression of the availability of the defense in theory, even if not in the instant case, will work a change in the rule.

a low type as a good defendant, that is if $\frac{k}{D} \leq \sigma_t^2 < (1 - \sigma_t^1)$, both good and bad types are willing to invest in presenting a defense.

This result tells us about the impact of legal costs on the potential quantity of evidence and argument that might be presented to judges that are willing to consider a defense. If legal costs are relatively high (damages are relatively low), there may be no defendants willing to invest. Under such circumstances, as with homogeneous defendants, a legal regime may be mired at R^e, even if the likelihood that courts could correctly apply the rule is high. As relative legal costs fall, good defendants are first encouraged to invest; judges may know that they are looking only at high types but still make errors in assessing the optimal rule. At still lower cost levels, bad defendants join good defendants in the pool of those proffering a defense; courts now face the difficulty of distinguishing between good and bad defendants, risking both type 1 and type 2 errors. Moreover, the mix of evidence and argument received by courts includes that produced both by good defendants seeking to help courts get it right and bad defendants seeking to induce courts to get it wrong.

Lemma 4 also shows the complicated impact of judicial error on the total amount of evidence and argument that may be presented to a court. A regime with higher rates of judicial error may have *more* defendants willing to invest in presenting a defense than an otherwise similar regime. Provided good defendants are willing to invest in both regimes, the regime with higher error, in particular, higher σ_t^2, may have both good and bad defendants proffering a defense, while the regime with lower error has only good defendants. A regime with a higher rate of type 1 error, σ_t^1, however, can never have a higher number of defendants willing to invest than an otherwise similar regime.

Now consider judicial incentives when facing heterogeneous defendants. Judges need to decide whether to allow the defendant who proffers a defense to present it; allowing the defense amounts to adopting the new rule, R^n. Again this is a tradeoff for the judge, between the safe return of rule-following, γ, and the risky return of rule-change. These risks depend on the mix of defendants who proffer a defense. From Lemma 4 we know that if $\sigma_t^2 < \frac{k}{D} \leq (1 - \sigma_t^1)$, only good defendants proffer a defense. This means the judge faces a risk of making a type 1 error only and will take this risk if $\gamma \leq \alpha(1 - \sigma_t^1)$, that is, if $\alpha \geq \frac{\gamma}{(1 - \sigma_t^1)}$. For lower relative legal costs, or higher rates of type 2 error (subject to the constraint that $\sigma_t^2 < (1 - \sigma_t^1)$,) the judge knows that the pool of defendants proffering a defense includes both good and bad types and will take the risk of rule-change if $\gamma \leq \alpha(1 - p\sigma_t^1 - (1-p)\sigma_t^2)$, that is, if $\alpha \geq \frac{\gamma}{(1 - p\sigma_t^1 - (1-p)\sigma_t^2)}$. This gives us first the analog to Proposition 1 with heterogeneous defendants:

Proposition 5 *Rule change will occur in period 1 with heterogeneous defendants if and only if (1) k is neither too high relative to D (in particular $\frac{k}{D} \leq (1 - \sigma_1^1) \leq 1$) nor too low (in particular, $\frac{k}{D} \geq \sigma_1^2$ if $\frac{\gamma}{(1 - p\sigma_1^1 - (1-p)\sigma_1^2)} > \bar{\alpha}$); (2) rewards to accurate rule change are sufficiently high, and in particular for any given level of error σ_1^1 and σ_1^2, $\bar{\alpha} > \tilde{\alpha}_1$ where $\tilde{\alpha}_1 = \frac{\gamma}{(1 - \sigma_t^1)}$ if $\sigma_1^2 < \frac{k}{D} \leq (1 - \sigma_1^1)$ and $\tilde{\alpha}_1 = \frac{\gamma}{(1 - p\sigma_1^1 - (1-p)\sigma_1^2)}$ if $\frac{k}{D} \leq \sigma_1^2$; and (3) judicial errors are sufficiently low*

and in particular for a given $\bar{\alpha}$, $\sigma_1^1 \leq 1 - \frac{\gamma}{\alpha}$ if $\frac{k}{D} > \sigma_1^2$ and $p\sigma_1^1 + (1-p)\sigma_1^2 \leq 1 - \frac{\gamma}{\alpha}$ if $\frac{k}{D} \leq \sigma_1^2$. Assuming these three conditions are met, R^n will be adopted by a proportion of judges equal to $p(1 - G(\tilde{\alpha}_1))$ if $\frac{k}{D} > \sigma_1^2$ and $(1 - G(\tilde{\alpha}_1))$ if $\frac{k}{D} \leq \sigma_1^2$. If these three conditions are not met, R^e will persist in periods 1 and 2.

Again, we can see that a regime will remain stuck at R^e if defendants and judges do not both face incentives to bear the costs of changing the rule, that is if legal costs are too high relative to damages, if errors are too high or if judicial rewards for welfare-promoting rule-change are inadequate for all judges. If there is no rule change in period 1, there cannot be any in period 2, as shared legal human capital does not accumulate and judicial errors remain unchanged at a level that discourages defendants and/or judges. Note the surprising new result, however, that legal costs that are too *low* can also prevent the system from moving away from R^e. Low relative legal costs encourage bad defendants to present evidence, introducing the risk of type 2 errors; if these errors are sufficiently likely, then judges may refuse to take the risk of rule change with bad defendants in the pool.

3.1 Optimality: The costs and benefits of error

The risk of error also raises the possibility that, while better in theory, R^n is not better in practice. To see this, let W^* be the level of social welfare enjoyed in a given case when either type of defendant is treated correctly: a good defendant is released from liability and a bad defendant is held liable. I assume that $W^* > k$, meaning that the cost of information about a defendant's true type is justified by the social benefits of correct legal treatment. Normalize the level of social welfare when either a type 1 or a type 2 error is made to zero. Assume that there are N defendants. Social welfare in a given period under R^e is then $W^e = N(1 - p)W^*$. Social welfare under R^n depends on whether only good types are willing to invest in evidence or both good and bad types. Moreover, under R^n social welfare in period 2 depends on what happens in period 1, as first period investments in legal human capital affect the error rates in period 2. If only good types invest, same period social welfare under R^n is

$$W_t^n = N \left[G(\tilde{\alpha}_t)(1-p)W^* + (1 - G(\tilde{\alpha}_t)) \left(p(1 - \sigma_t^1) W^* + (1-p)W^* - pk \right) \right].$$

If both good and bad types invest,

$$W_t^n = N \left[G(\tilde{\alpha}_t)(1-p)W^* + (1 - G(\tilde{\alpha}_t)) \left(p(1 - \sigma_t^1) W^* + (1-p)(1 - \sigma_t^2)W^* - k \right) \right].$$

Rearranging allows us to state the following:

Proposition 6 *If only good defendants are willing to invest in presenting a defense, W_t^n is higher than W^e and more widespread rule change is welfare-improving iff $(1 - \sigma_t^1)W^* - k > 0$. If both good and bad defendants are willing to invest in presenting a defense, W_t^n is higher than W^e and more widespread rule change is welfare-improving iff $(p(1 - \sigma_t^1) - (1-p)\sigma_t^2)W^* - k > 0$. If*

$W_1^n < W^e$, *rule change in period 1 is optimal only if* $W^e - W_1^n \geq W^e - W_2^n > 0$. *In this case, more widespread rule change in period 1 is not necessarily better than more restricted rule change but is weakly better in period 2.*

Proof. See Appendix. ■

Proposition 6 is straightforward and intuitive. If only good types seek rule change, the change is valuable if the cost of presenting the defense, k, is not wasted too often, meaning that judges do not often make a type 1 error and fails to release the defendant from liability. If both good and bad types seek rule change, the change is valuable if the expected error–the weighted average of type 1 and type 2 errors–does not exceed the cost of presenting the defense. Thus if errors are high, even a theoretically valuable rule change may be undesirable. This is clearly true if we are looking at high errors in the terminal period, 2. But there is another consideration in assessing the desirability of rule change in period 1. This is the potential for learning in period 1 which reduces the level of error in period 2. Provided period 1 error rates relative to the learning benefits realized in period 2 are not too high, it can be optimal for a regime to engage in rule change in period 1 even if $W_1^n < W^e$.

Proposition 6 emphasizes an essential relationship between legal change and judicial competence or expertise. In a given case, or at a given point in time, the risk of error may not justify legal modification. But from a dynamic perspective, errors may be necessary in order to generate long-term benefits from what is learned from those errors. During the learning phase, modest rule adaptation only may be warranted; so long as what is learned from modest rule change is distributed throughout the legal system, and is on net informative, more widespread rule change will then spread the benefits of what is learned. This focuses the analysis on the determinants of legal error and the systemic accumulation of shared legal human capital. The more effectively a legal regime translates what is learned in individual cases into shared legal human capital, available to all judges, the more likely it is that the system can bear the costs of moving to the higher welfare levels produced by lower errors. If, on the other hand, information gathered in particular cases is not shared (perhaps because courts do not publish within the judiciary extensive discussion of their findings and analysis) or is not subject to objective critique and commentary (perhaps because judicial findings and reasoning are not shared outside the judiciary or there are distortions in the market of ideas in which legal arguments are distributed), then a regime will be less likely to benefit from short-term error costs and may, optimally, remain at a lower level of social welfare.

3.1.1   Comparing Legal Regimes

I turn now to comparisons across legal regimes to determine the welfare impact of changing attributes: the quality of judicial information processing, the distribution of judicial rewards, the initial levels of legal human capital and the

relative cost of legal evidence and argument. The case in which rule change is never justified in either period 1 or period 2 is one in which the policy implications are clear, namely that a legal regime should structure judicial incentives that do not reward rule change. The more interesting cases from a policy perspective involve circumstances in which some rule change is warranted. I look first at the case in which rule change is justified even at initial levels of judicial error ($W_1^n > W^e$) and then at the case in which rule change in period 1 is justified only by gains in period 2 ($W_1^n < W^e$), both under the assumption that all evidence and legal argument, whether collected from bad or good defendants, is informative. I then look at the comparative analysis of a special case in which bad defendants are eliminated from the pool of those presenting evidence in one regime but not another and widen the analysis to include consideration of the potential for what is learned from bad defendants to be disinformative.

Welfare gains from rule change in period 1 If initial levels of judicial error are sufficiently low to justify period 1 rule change without regard to gains in period 2, the level of social welfare achieved in a legal regime depends on how widespread rule change becomes and what further reductions in judicial error might be enjoyed in period 2. We are interested in determining what features of a legal regime may lead to higher social welfare through rule change and error reduction.

Consider two legal regimes, A and B, both of which would benefit from rule change in period 1: $W_1^n > W^e$. In order to simplify the analysis initially, assume that whatever differences might arise between these two regimes, it is never the case that the mix of defendants seeking rule change diverges such that only good types proffer a defense in one regime but both good and bad types proffer a defense in the other. (This assumption is relaxed, in section 3.1.1, below.)

Information processing Suppose first that A and B are identical in all respects except that they have different information processing functions, with $i^A(\Delta) > i^B(\Delta)$. Regime A is thus more effective at extracting valuable error-reducing information out of the accumulated evidence and legal argument presented to courts; it may be characterized by greater publication of judicial opinions, more effective feedback from experts, or higher levels of judicial specialization or training. Regimes A and B will experience the same degree of rule change in period 1. Assuming this involves some positive accumulation of evidence and argument ($\Delta > 0$), regime A will experience lower rates of both type 1 and type 2 errors in period 2, provided judicial information processing in both regimes is informative ($i' > 0$). This implies that judges in A are more willing to change rules than judges in B. Social welfare will then be higher in A than B if period 2 rule change is optimal in A ($W_2^{nA} > W^e$), which is implied by the optimality of rule change in period 1 ($W_1^n > W^e$). Higher period 2 social welfare in A results both from more widespread rule change and lower judicial errors.

Judicial incentives Now consider two regimes that differ only in the distribution of judicial incentives. Suppose that in regime A there are more judges with higher judicial rewards for accurate rule adaptation than in regime B and in particular $G^A(\tilde{\alpha}_1) < G^B(\tilde{\alpha}_1)$. This implies that in period 1, for a given level of judicial error, more judges in A are willing to accept a proffered defense than is the case in B and rule change is more widespread in A. Period 1 social welfare is thus higher in A. Assuming that judicial information processing is informative ($i' > 0$), regime A also begins period 2 with a higher level of shared legal human capital and lower type 1 and type 2 errors. Higher judicial rewards then magnify the results we obtained with respect to improved judicial processing in period 2: rule change is more widespread in A than B, both because of lower error and because more judges at a given level of error are willing to risk rule change. Taken together, social welfare in A is higher in both periods and the gains come both directly from increased willingness to risk welfare-improving rule change and indirectly from the reductions in legal error and the further expansion of rule-changing that initial rule-change produces.

Initial judicial error What if regime A begins period 1 with a higher level of legal human capital and lower errors than regime B? The effect here is analogous to what happens with higher judicial rewards. Lower errors imply that judges in A are more willing to entertain rule change in period 1. This implies more widespread rule change in A, resulting in even higher gains in A in period 2, through a widening gap between the levels of judicial error in the two regimes.

Legal costs Consider finally a comparison between regimes with different legal costs. Even if there are no differences in relative legal costs, and thus no differences in behavior between a low-cost regime and a high-cost regime, lower absolute legal costs reduce the cost of achieving rule change and reductions in judicial error. Social welfare is then higher in the low-cost regime. Now consider regimes that differ in relative legal costs, either because of a difference in absolute legal costs, damages or both. Relative legal costs affect defendants' decisions about whether or not to invest in proposing rule-change. As relative legal costs increase, there are fewer circumstances in which bad defendants are willing to invest; if relative legal costs get high enough, even good defendants are unwilling to invest. If relative legal costs in two regimes lead the same mix of defendants to invest (all good in both or all good and bad in both), then the legal regimes will see no difference in social welfare coming from differences in the extent of rule change or judicial error; a social welfare differential will arise only through differences in absolute legal costs. If relative legal costs in A are low enough to encourage at least good defendants to invest, while those in B are too high, the relatively low-cost regime A will clearly be better off unless lower relative legal costs are sustained with higher absolute legal costs that outweigh the differential gains from rule change and error reduction. In general, any gains from behavioral changes caused by lower relative legal costs

will be augmented by expenditure savings if the lower relative costs are achieved at least in part through lower absolute costs, and will be reduced by increased expenditures if the lower relative costs are achieved with higher absolute costs offset by higher damages.

Proposition 7 summarizes the above results.

Proposition 7 *Consider two regimes, A and B, both of which would benefit from rule change in period 1: $W_1^n > W^e$. Ceteris paribus, and assuming it is not the case that these conditions lead to a divergence in defendant behavior such that all defendants in one regime but only good defendants in the other seek rule change in some period, regime A will enjoy a weakly higher level of social welfare under any of the following conditions:*

$$a. \quad i^A(\Delta) > i^B(\Delta)$$
$$b. \ G^A(\tilde{\alpha}_1) < G^B(\tilde{\alpha}_1)$$
$$c. \quad K_1^A > K_1^B$$
$$d. \ k^A < k^B$$
$$e. \ (\frac{k}{D})^A < (1 - \sigma_1^1) < (\frac{k}{D})^B$$

Proof. See Appendix. ∎

Welfare gains from rule change only in period 2   The more interesting

case in which to examine the comparative benefits of different regime attributes is the case in which initial levels of judicial error are sufficiently high that rule change is not warranted on the basis of outcomes in period 1 alone, but only on the basis of gains achieved in period 2 through accumulation of shared legal human capital and reductions in judicial error. Again I initially restrict attention to cases in which differences between the regimes do not lead to a divergence in the mix of defendants seeking rule change such that good types in one regime but both good and bad types in the other seek rule change.

Consider two regimes A and B in both of which $W_1^n < W^e$. In this case, more widespread rule change in period 2 will increase total social welfare, but more widespread rule change in period 1 will have ambiguous effects: increasing total legal human capital and period 2 error reductions but also increasing the costs of achieving these effects. Total social welfare will then depend on balancing the costs incurred in period 1 with benefits generated in period 2. Regime A will benefit from rule change in a wider range of circumstances than B and higher social welfare when rule change is justified if it enjoys more informative judicial information processing: If $i^A(\Delta) > i^B(\Delta)$, regime A can justify greater losses in period 1 to achieve benefits in period 2. Regime A may benefit from rule change in a wider range of circumstances than B and enjoy higher total social welfare when rule change is justified if it enjoys lower absolute legal costs or higher initial legal human capital. If $k^A < k^B$, gains in period 2 do not

have to be as high in A as in B to justify period 1 investments in legal human capital; if $K^A > K^B$, a smaller accumulation of legal human capital in period 1 is needed to generate sufficiently low errors in period 2. Similarly, if A enjoys higher initial legal human capital, the lower incidence of judicial error will lead to more widespread rule change. More widespread rule change in period 1, however, may again lead to higher total costs not justified by period 2 gains.

The complexity of these comparisons derives from a basic feature of the world in which rule change is not justified by period 1 returns alone. In such a world, there is an optimal level of first period rule change and in general this optimum will be less than 100%. More extensive rule change and higher levels of legal human capital accumulation in period 1 are therefore not necessarily welfare-promoting. So whereas reduced legal costs or higher levels of initial legal human capital clearly increase the range of circumstances in which first period investments will generate compensating second period benefits, these benefits will be offset in those circumstances in which too much first period rule change is encouraged. For this reason, it is not possible to say in general whether a regime with better judicial incentives for rule change will justify rule change in a wider set of circumstances or achieve higher levels of social welfare or not. More extensive rule change in period 1 will generate higher levels of legal human capital–but at a higher total cost. If judicial information processing is weak and this results in only small reductions in judicial error, the regime may still not warrant rule change or may achieve lower gains from rule change where it is warranted.

Proposition 8 summarizes these results.

Proposition 8 *Consider two legal regimes, A and B in both of which $W_1^n < W^e$. Ceteris paribus, and assuming it is not the case that these conditions lead to a divergence in defendant behavior such that all defendants in one regime but only good defendants in the other seek rule change in some period, rule change will be justified by sufficient period 2 welfare gains in a wider range of circumstances and lead to higher total social welfare in A under any of the following conditions:*

$$(a)\ i^A(\Delta) > i^B(\Delta)$$
$$(b)\ k^A < k^B$$
$$(c)\ (\frac{k}{D})^A < (1 - \sigma_1^1) < (\frac{k}{D})^B.$$

Rule change may or may not be justified in a wider range of circumstances and lead to higher total social welfare in A under either of the following conditions

$$(d)\ G^A(\tilde{\alpha}_1) \quad < \quad G^B(\tilde{\alpha}_1)$$
$$(e)\ K_1^A \quad > \quad K_1^B.$$

Proof. See Appendix. ∎

Bad Defendants and Disinformation Now consider the case I excluded from the analysis in the above two sections, namely the case in which the mix of defendants seeking rule change diverges such that only good types seek rule change in one regime (A) but both good and bad types seek change in the other (B). This divergence appears when either an initial or induced difference in either type 2 errors or the relative cost of presenting evidence causes bad defendants to drop out of the pool of defendants proffering a defense in one regime while they remain in the other.

Eliminating bad defendants from the pool of those seeking to induce judges to change rules would seem to be unambiguously good from a social welfare point of view: society avoids the expense of hearing evidence and argument in those cases and the losses that arise from mistakenly identifying a bad defendant as a good one and releasing the defendant from liability. But there are two reasons why keeping bad defendants in the pool of defendants seeking rule change might generate social benefits.

First, if the risk of type 2 errors is lower than the risk of type 1 errors, having bad defendants in the pool can encourage judges to entertain a proffered defense. This occurs because judges considering a request to change the rule may be more willing to do so if they believe that at least some of those who are seeking the change are bad defendants whom the judge will have little trouble identifying as such. If only good defendants request rule change in regime A, and type 1 errors are harder to avoid, judges in A may be less willing to open the door to the proffered evidence and legal argument than judges in B. Formally, the critical value $\tilde{\alpha}_t^A = \frac{\gamma}{1-\sigma_t^{1A}} > \tilde{\alpha}_t^B = \frac{\gamma}{1-p\sigma_t^{1B}-(1-p)\sigma_t^{2B}}$, which can occur even if $\sigma_t^{1A} < \sigma_t^{1B}$. So, what appears as an welfare-promoting advantage for regime A–lower judicial error rates due either to higher levels of initial legal human capital in the start of period 1 or to greater accumulation and/or better judicial processing of information during period 1–can reduce the extent of rule change. This means that the welfare gains available from changing rules when requested to do so by good defendants are reduced. If the impact on the willingness of judges to hear evidence is large enough, we could reach the result, contrary to what is shown in Propositions 7 and 8, that regime B–with higher initial legal errors or lower less effective information processing or smaller judicial rewards for rule change–could achieve higher levels of social welfare than regime A arising from more extensive rule change in period 1 and/or period 2.

The potential for this case to arise as a practical matter seems, *a priori*, small. Against a backdrop of a safe existing rule that holds all defendants liable, it seems strained to believe that judges will see less risk in deviating from the established rule when they know there are defendants who are trying to mislead them than when they believe that, at worse, they will reach the same result for a good defendant that they would under the existing rule. (If we were looking at the case in which the existing rule holds no defendants liable, the analogous claim would be that it is hard to believe judges will see less risk in considering a new cause of action when they know that there are bad plaintiffs who are trying to mislead them than when at worse they will refuse

worthy plaintiffs who are out of luck any way under the established rule.) From a social welfare point of view, given that any gain in the likelihood of judges allowing a proffered defense must be offset against the real losses associated with increased expenditures on legal costs and type 2 errors, it seems unlikely that this exceptional case, where a regime B with lower initial legal human capital, less effective judicial information processing or smaller judicial rewards for rule change does better than regime A, obtains in practice. It seems more likely that judges are discouraged knowing that there are bad defendants (plaintiffs) in the mix of those urging them to recognize an exception to the existing rule. In this case, higher relative legal costs or initial legal human capital, or any factors that lead to lower second period type 2 errors–differences that screen out bad defendants–improve the level of welfare achieved by a legal regime relative to one that keeps bad defendants in the mix.

The second, probably more significant, reason why keeping bad defendants in the pool of defendants seeking rule change might generate social benefits is that a regime that entertains rule-change in cases with both bad and good defendants may generate higher levels of legal human capital accumulation and thus enjoy greater reductions in error than a regime that restricts its reception of evidence and legal argument to cases in which good defendants appear. (Note that this effect only concerns a potential benefit to having bad defendants in the pool in period 1; having bad defendants in the pool in period 2 is uniformly bad unless the special case described in the above paragraph obtains.) We have maintained thus far the assumption that all evidence and legal argument presented to courts, whether from good defendants or bad defendants, is ultimately informative. When this is true, differences in initial conditions that lead only good defendants to proffer a defense in one regime (because of higher initial legal human capital or higher relative legal costs) but both types to proffer a defense in the other, can lead to more learning and error reduction in the regime with lower initial legal human capital or lower relative legal costs. This will result so long as the increase in the number of defendants proffering evidence and legal argument is not outweighed by reductions in the willingness of judges to consider rule change in the regime that includes bad defendants in the pool of those seeking rule change. (This is the opposite effect to the one described in the previous paragraph and arises, for example, if type 2 errors are more likely than type 1 errors.)

This result brings to the fore the assumption about the impact of evidence and legal argument presented by bad defendants on systemic legal human capital. Whether information collected from bad defendants is ultimately informative is an interesting and difficult question of epistemology, cognitive psychology and social learning. It may be the case that even the successful efforts of bad defendants to lead a particular court into error contribute, in the long run, to the ability of judges to avoid error. Reviewing these misleading presentations alongside those of in other cases, and in light of the analysis, feedback and commentary generated from other judges, lawyers and legal commentators, may enhance the ability of judges to identify their errors. This could be the benefit of both hindsight and context, as well as external expertise. Indeed, it may

be that understanding what makes a defendant 'good' is only possible when one has seen how such a defendant differs from a 'bad' one. (This is the sense in which variance improves our ability to estimate coefficients using econometrics.) It seems entirely likely, for example, that in the absence of any evidence and argument from bad defendants, the evidence and argument from good defendants may be overgeneralized—"all defendants are good." This would be a common mistake, of failing to take into account sample size and selection biases (Kahneman et al 1982) in interpreting data. If every design defect case, for example, reached the conclusion that the cost of designing around defects was not warranted by a reduced risk of harm, many judges might easily conclude that redesign costs are never justified.

It is just as easy, of course, to suppose that the accumulation of evidence and legal argument from bad defendants, which leads particular courts into error in any given instance, has no value as legal human capital or, indeed, that it degrades the capacity of all judges in the long run to distinguish good from bad. Here we might need not epistemology—a theory of knowledge—but, as Proctor (2005) suggests, agnotology—a theory of ignorance and doubt. Proctor's work on cancer and litigation suggests ways in which the efforts of, for example, tobacco and asbestos manufacturers to defend against tort liability claims result in the production of evidence (expert studies, testimony and reports) that degrade what is reliably known. As one defendant's internal memo claimed with respect to the goals of its research funding: "Doubt is our product." (Proctor 1995) In such cases, the information generated by bad defendants may be on net disinformative. The proportion of bad defendants in the pool of defendants proffering evidence and legal argument would then have an important impact on how legal error changes in period 2; if there are more bad defendants than good, for example, it is conceivable that judicial error increases.

If the information generated by bad defendants is either uninformative or, worse, disinformative, then the regime that eliminates bad defendants from the pool in period 1–through higher initial legal human capital or higher relative legal costs–is generally better off than the one that includes them (with the probably rare exception of the case in which keeping bad defendants in the pool has a substantial positive impact on the willingness of judges to entertain rule change.) This regime learns as much or more and incurs lower costs in doing so, provided the higher relative legal costs are not produced by absolute legal costs that are so high as to swamp the gains from improved learning and the elimination of type 2 errors.

Proposition 9 summarizes these results.

Proposition 9 *A legal regime B in which both good and bad defendants seek rule change may enjoy higher social welfare than legal regime A in which only good defendants do if evidence and legal argument collected from bad defendants is informative and either $\sigma_1^{1A} > p\sigma_1^{1B} + (1-p)\sigma_1^{2B}$ or, if $\sigma_1^{1A} < p\sigma_1^{1B} + (1-p)\sigma_1^{2B}$, $p < \frac{1-G(\tilde{\alpha}_1^B)}{1-G(\tilde{\alpha}_1^A)}$. Even if evidence from bad defendants is uninformative or disinformative, B may enjoy higher social welfare than A if $W_t^{nB} > W^e$ and $\sigma_1^{1A} > p\sigma_1^{1B} + (1-p)\sigma_1^{2B}$.*

Proof. See Appendix. ∎

Note that thinking about the impact of bad defendants on social welfare generates two surprising insights. First, if the information generated by bad defendants is ultimately informative and helpful to judges in period 2, it can be a good thing for a legal regime to encourage rule change despite the presence of bad defendants and the risk of type 2 errors. Second, if the information generated by bad defendants is uninformative or disinformative or if judges are discouraged from rule change by the risk of type 2 errors, a legal regime can be better off with *higher* relative legal costs that screen out bad defendants. In period 2, except in the rare case in which the presence of bad defendants leads judges to be more willing to entertain a defense, any regime is better off if it manages to screen out bad defendants through reduced type 2 errors and higher relative legal costs.

4 Conclusion

We began with the question of which legal regimes better support economic growth and the development of markets. The analysis in this paper suggests that making progress on that question will require moving beyond the simple dichotomy between common law and civil code regimes that has thus far dominated the literature. This model suggests that the important distinctions between legal regimes are found not in the reliance on code versus caselaw but rather in the institutional determinants of judicial incentives and the capacity for a legal regime to generate investments in legal human capital that reduce legal error. Furthermore, there is more to the development of an effective legal regime than the legislation of optimal (static) legal rules and the alignment of judicial incentives with social welfare. Even good faith judges face the problem of solving the difficult challenge of correctly analyzing and implementing legal rules, particularly when the environment is changing over time and space.

I have identified five key parameters that shape the capacity of a legal regime to adapt law to local and changing conditions so as to promote social welfare: (1) the distribution of judicial rewards for rule-adaptation; (2) the cost of producing evidence and legal argument for presentation to a court; (3) the level of damages; (4) the initial or exogenous level of judicial error and (5) the extent to which the legal regime transforms individual case information into informative–error-reducing–shared legal human capital. Each of these parameters is linked in important ways to the institutional environment that defines a legal regime. Judicial rewards for rule-adaptation, for example, are a function of the organization of the judiciary and the nature of the information available to those who make up a judge's audience. A civil service judiciary for example (common in civil code countries) arguably has a more insular, professional audience that is focused on legal reasoning–senior judges–than does a common law judiciary that is evaluated by a wide public audience that is more interested in the consequences of legal outcomes than the competence of legal argument. Judges in a civil service judiciary also enter the judiciary as junior magistrates, fresh out

of law school, and face more systematic and frequent evaluation and promotion than do common law judges, arguably generating a system which rewards higher levels of rule-following and less individual creativity (Ramseyer and Rasmussen 1997, Posner 2005). Certainly this is the way in which the philosophy of judging is thought to differ as between common law and civil code regimes (Merryman 1985)[6]. My analysis here suggests that there are institutional attributes that generate the judicial behavior otherwise explained by ideology or culture.

Moreover, the analysis emphasizes that the phenomenon of judicial orientation to rule-following and rule-adaptation is not merely explained by the institutional determinants of judicial rewards. Even judges who face excellent incentives to adapt the law to new information may, quite rationally, elect not to do so in light of what they recognize is their limited knowledge about what constitutes a welfare-improving change in the law. And even if judges are willing to risk rule-adaptation, they are dependent on the willingness of litigants to make the costly investments (at risk of being wasted if they do not secure the desired legal rule change) in information and legal argument on which the accuracy of judicial decisionmaking depends. It is this dynamic interaction between the incentives of judges and litigants that shapes the evolution of legal human capital and judicial error and thus the capacity of a legal regime to move towards welfare-improving rule changes. Thus what we observe behaviorally– the extent to which judges in fact stick to existing rules or adapt them to new information–is also importantly dependent on the institutions that determine legal costs–such as the organization and regulation of the legal profession (Hadfield 2000); exogenous judicial error–such as the use of formal judicial training or requirements of extensive practice experience as a lawyer; and judicial information-processing–such as opinion-writing, publication and expert commentary practices. (Common law regimes, for example, tend to produce extensive, publicly available judicial opinions laying out factual findings and legal reasoning; civil code regimes tend to produce detailed academic commentary, published alongside short, relatively opaque, legal opinions with extensive judidical analysis sometimes confined to documents distributed only within the judiciary (Lasser 2004).) In a companion paper (Hadfield 2008), I explore in more detail what we know about the institutional attributes of existing regimes in light of these five parameters. Unfortunately our knowledge is thin.

My review of the institutional landscape highlights the need for two impor-

---

[6] As Merryman explains: "The basic difference is epitomized in [a] quotation from the German legal scientist Rudolph Sohm: "A rule of law may be worked out either by developing the consequences that it involves, or by developing the wider principles that it presupposes...The more important of these two methods of procedure is the second, i.e., the method by which, from given rules of law, we ascertain the major premises they presuppose." . . . An American legal realist would resist the implication that rules of law should be the principal objects of his study or the suggestion that there are only these two ways of studying them. But if pushed to Sohm's choice, most law professors, judges, and lawyers in the United States would easily and quickly choose the first of his two methods. Most civil lawyers would still choose the second." (p. 67) Elsewhere Merryman states simply: "The civil law has [ ] sacrificed flexibility for certainty. In contrast, the common law tends to strike the balance between them more equally."

tant empirical projects. First, we clearly need to deepen our attention to the specifics of the institutional environments in different countries that affect judicial incentives and the accumulation of legal human capital. Classifying regimes as either civil code or common law is not likely to prove helpful. Rather, we need to know far more, country-by-country, about the structure of judicial rewards and the information available to those who judge the performance of judges and hence influence the structure of judicial rewards and penalties. This suggests a far more refined comparative project than the one that currently engages comparative scholars. The model in this paper suggests that the key variables include the identity of those who evaluate judges and thus determine their reward structure (senior judges? politicians? lawyers? journalists?) and the information available to those evaluators (are decisions published? with what level of detail on factual findings and reasoning? is the information filtered by a judge or available in its original form as verbatim testimony and exhibits?). The structure of courts is important (are judges identified? do they sit alone or in panels? how collegial are courts? are opinions attributable to individual judges? who determines evidentiary questions?) The exposure of judges to the welfare effects of their decisions may also be important (have judges been exposed to the practical problems of clients? do they enter the judiciary directly from their legal education or only after a period of practice? what training do judges have in evaluating evidence about the impact of legal rules and assessing policy questions?) And, critically, how is information learned by judges in a particular case diffused through the system (again, are decisions published and how detailed is the presentation of facts and reasoning?)

With a more refined descriptive catalogue of differences between legal regimes, we will be in a position to conduct a second important empirical project: more careful study of the relationship between these institutional variables and economic growth. As many have noted, the classification of regimes on the basis of legal origins is somewhat crude and makes it difficult to sort out the effect of a particular legal history from other cultural or human capital imports. The analysis in this paper suggests more specific legal variables–which undoubtedly vary across countries that are otherwise classified as belonging to a particular legal family and over time within countries–on which empirical work can focus in the effort to assess the role of legal factors in economic growth and development. Not only might this help disentangle confounding effects from the inheritance of legal rules bundled with human capital and other cultural attributes, but it may also help to increase the precision of our estimation techniques, as we can make use of the substantial variability in legal regimes, variability that is masked by the macro division into legal families.

Further theoretical work is also clearly needed. I have only been able to sketch how the more complex dynamics at work may play out, both in positive and normative terms, in particular environments. Moreover, in order to simplify the analysis, I have suppressed several features of litigation and the response to litigation that clearly will have an impact on the incentives of judges, the accumulation of legal human capital and the path of legal evolution. Settlement behavior is obviously a critical component of litigation and as many have

noted, settlement is not random. It has a systematic effect on the nature of the cases that reach final decision in a court and thus on the information available to courts. Hylton (2006) considers some of these effects of settlement on evolution. Legal rules also affect activity levels, the behavioral choices plaintiffs and defendants make about the conduct implicated by a legal rule. As I have argued elsewhere (Hadfield 1992), this will also affect the information set reaching a court. A more general model would also relax the assumption that only defendants present evidence and argument, and analyze the strategic behavior that surrounds information revelation to a court. Several economists have explored in particular the impact of signaling, strategic revelation and the competition between plaintiff and defendants on the nature and amount of the evidence presented to a court. (See, for example, Milgrom and Roberts 1986, Shin 1994, Shin 1998, Posner 1999, Daughety and Reinganum 2000) This work has focused on the impact of strategic behavior on the accuracy of a court's determination of the facts in a given case. The model presented in this paper identifies another important effect that may flow from strategic evidentiary behavior, namely the effect on the informative quality of the legal human capital stock and thus on the likelihood of error in the legal system as it evolves. And although I have de-emphasized the importance of the relative reliance on statutes or regulation as opposed to judge-made law because of the exaggeration of this difference in the existing literature, it will ultimately be important in a fuller model to situate the analysis of learning through litigation in the context of legislative determinations about the extent to which regulation will be accomplished through courts as opposed to legislatures and agencies. If, for example, there are significant obstacles to the accumulation of legal human capital that do not confront the development of bureaucratic expertise in legislatures and agencies, optimal legal regulation may involve heavier reliance on statutes and regulations. In the end however, even the most refined statutes require interpretation and application and hence depend on the quality of legal human capital available to judges. Finally, it will be important for further work to assess more carefully the trade-offs between controlling corruption within courts and facilitating the capacity of judges to engage in welfare-promoting rule adaptation. The effort to control corruption is a key reason for many of the institutional features that this analysis identifies as critical to the capacity of a legal regime to learn and adapt over time. But the trade-offs may not be as stark as they first appear. The capacity of a system to detect corruption is also dependent on the level of legal human capital: indeed, this is one of the principal justifications for public and reasoned decisionmaking. Like rule adaptation, the elimination of corruption may be best analyzed as a dynamic problem of structuring the mechanisms that contribute to the organic accumulation of legal human capital.

The policy prescriptions that flow from the analysis I have presented suggest that the choice facing transition and developing economies is not between writing codes or borrowing volumes of caselaw. Rather it is a series of choices about institutional attributes such as the publication and expansiveness of legal opinions, the institutional structuring of judicial incentives for rule adaptation and the mechanisms by which information about the welfare effects of particular

rules (or, more to the point, particular interpretations of statutory provisions) makes its way to judges and those who evaluate judges. The model also links the effectiveness of courts to the organization and regulation of the legal profession. Lawyers play a key role in the generation and transmission of specialized legal human capital, specifically expertise about the relationship between legal rules and welfare. As the model makes clear, the adaptation of law to local and changing circumstances over time requires that litigants face incentives to invest in lawyers' efforts to produce evidence and innovate legal arguments. The organization and regulation of the legal profession–the extent to which the market for lawyers is competitive, for example–will influence the path of the law, both through the cost of legal services and the cost of generating a certain level of expertise. Rules governing the organization of legal practice–limitations on firm size or prohibitions on employment[7] for example–influence the extent to which legal human capital is shared among those in the profession.[8] Professional control over legal ethics will also have an impact on the potential for disinformation in courts. The model also suggests that countries attempting to transition quickly to a legal regime that supports economic growth and market development may need to take specific steps to overcome both inadequate judicial incentives and an initially high level of legal error. Particularly in systems transitioning from socialist or communist governance to market democracy, it is likely that the shared level of legal human capital about the relationship between legal rules and outcomes will be low by virtue of the lack of experience with markets. In these settings, policy efforts to effectively import legal human capital into the profession and judiciary may be necessary. This has implications, for example, for the rules governing the access of foreign lawyers and law firms to practice in the new regime as well as for the access the profession and judiciary has to the work of lawyers and courts in other jurisdictions.

The principal lesson is that law that supports economic growth and market development has to be seen in dynamic terms, as an organic entity that evolves over time in response to local and changing conditions. In order for that process to take place, it is necessary for judges to face incentives that support welfare-improving rule adaptation and for litigants to invest in presenting to courts the evidence and arguments they need to evaluate proposed rules or statutory interpretations. Whether a system is denominated a code system or a common law system, it is the institutions that structure incentives for judges and litigants to learn over time and the mechanisms by which this learning is translated into shared legal human capital that determine the quality of a legal regime.

---

[7]Slovakia, for example, has a rule that lawyers may not be employed by other lawyers. In the interests of independence, lawyers must personally represent a client.

[8]Hadfield (2007) examines how the regulation of the legal profession at the national level inhibits the growth of cross-national law firms and hence the development of legal human capital tailored to resolving the unique problems of global cross-border trade.

References

[1] Anderlini, Luca, Leonardo Felli and Alessandro Riboni. 2006. "Statute Law or Case Law?" Manuscript.

[2] Berkowitz, Daniel, Katharina Pistor and Jean-Francois Richard. 2003. "The Transplant Effect." American Journal of Comparative Law. Vol. 51, pp. 163-203.

[3] Botero, Juan, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer. 2003. "Judicial Reform." The World Bank Research Observer.

[4] Botero, Juan, Simeon Djankov, Rafael La Porta and Florencio Lopez-de-Silanes. 2004. "The Regulation of Labor." Quarterly Journal of Economics. Vol. 119, pp. 1339-1382.

[5] Cooter, Robert, Lewis Kornhauser and David Lane. 1979. "Liability Rules, Limited Information and the Role of Precedent." The Bell Journal of Economics. Vol. 10, pp. 366-373.

[6] Daughety, Andrew and Jennifer Reinganum. 2000. "On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias." Journal of Law, Economics and Organization. Vol. 16, pp. 365-394.

[7] Dewatripont, Mathias and Jean Tirole. 1999. "Advocates." Journal of Political Economy. Vol. 107, pp. 1-39.

[8] Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer. 2002. "The Regulation of Entry." Quarterly Journal of Economics. Vol. 117, pp. 1-37.

[9] _____. 2003. "Courts: The Lex Mundi Project." The Quarterly Journal of Economics. Vol. 118, pp. 453.-517.

[10] Diver, Colin. 1983. "The Optimal Precision of Administrative Rules." Yale Law Journal. Vol. 93, pp. 65-109.

[11] Gennaioli, Nicola and Andrei Shleifer. 2007. "The Evolution of Common Law." Journal of Political Economy. Vol. 115, pp. 43-68.

[12] Glaeser, Edward L. and Andrei Shleifer. 2002. "Legal Origins." Quarterly Journal of Economics. Vol. 117, pp. 1193-1229.

[13] Goodman, John C. 1978. "An Economic Theory of the Evolution of Common Law." Journal of Legal Studies. Vol. 7, 393-406.

[14] Hadfield, Gillian K. 1992. "Bias in the Evolution of Legal Rules." Georgetown Law Journal Vol. 80, pp. 583-616.

[15] _____. 2000. "The Price of Law: How the Market for Lawyers Distorts the Justice System." Michigan Law Review Vol. 98, pp. 953-1006.

[16] _____. 2008. "The Levers of Legal Design: Institutional Determinants of the Quality of Law" Forthcoming, Journal of Comparative Economics.

[17] Hylton, Keith N. 2006. "Information, Litigation and Common Law Evolution." American Law and Economics Review. Vol. 8, pp. 33-61.

[18] Johnson, Simon, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer. 2000. "Tunneling." American Economic Review. Vol. 90, pp. 22-27.

[19] Kaplow, Louis. 1992. "Rules versus Standards: An Economic Analysis." Duke Law Journal. Vol. 42, pp. 557-623.

[20] Klerman, Daniel. 1999. "Nonpromotion and Judicial Independence." Southern California Law Review. Vol. 72, pp. 455-463.

[21] Kahneman, Daniel, Paul Slovic and Amos Tversky. 1982. Judgment Under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press.

[22] Landes, William M. and Richard A. Posner. 1976. "Legal Precedent: A Theoretical and Empirical Analysis." Journal of Law and Economics. Vol. 19 pp. 249-307.

[23] La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert Vishny. 1997. "Legal Determinants of External Finance." Journal of Finance. Vol. 52, pp. 1131-1150.

[24] _____. 1998. "Law and Finance." Journal of Political Economy Vol. 106, pp. 1113-1155.

[25] La Porta, Rafael, Florencio Lopez-de-Silanes, Cristian Pop-Eleches, and Andrei Shleifer. 2004. "Judicial Checks and Balances." Journal of Political Economy. Vol. 112, pp. 445-470.

[26] Lasser, Mitchel de S. -O. -L'E. 2004. Judicial Deliberations: A Comparative Analysis of Judicial Transparency and Legitimacy. Oxford: Oxford University Press.

[27] Levy, Gilat. 2005. "Careerist Judges and the Appeals Process." The RAND Journal of Economics. Vol. 36, pp. 275-297.

[28] Mahoney, Paul G. 2001. "The Common Law and Economic Growth: Hayek Might Be Right." Journal of Legal Studies. Vol. 30, pp. 503-525.

[29] Merryman, John Henry. 1985. The Civil Law Tradition. Stanford, CA: Stanford University Press.

[30] Milgrom, Paul and John Roberts. 1986. "Relying on the Information of Interested Parties." The RAND Journal of Economics. Vol. 17, pp. 18-32.

[31] Posner, Richard A. 1977. Economic Analysis of Law.

[32] _____. 1993. "What Do Judges Maximize? (The Same Thing Everybody Else Does.)" Supreme Court Economic Review. Vol. 3, pp. 1-41.

[33] _____. 1999. "An Economic Approach to the Law of Evidence." Stanford Law Review. Vol. 51, 1477-1546.

[34] _____. 2005. "Judicial Behavior and Performance: An Economic Approach." Florida State University Law Review. Vol. 32, pp. 1259-79.

[35] Priest, George L. 1977. "The Common Law Process and the Selection of Efficient Rules." Journal of Legal Studies. Vol. 6, pp. 65-

[36] Proctor, Robert N. 1995. Cancer Wars. New York: Basic Books.

[37] Ramseyer, Mark J. and Eric B. Rasmussen. 1997. "Judicial Independence in a Civil Law Regime: Evidence from Japan." Journal of Law, Economics and Organization. Vol. 13, pp. 259-286.

[38] Rose, Carol M. 1988. "Crystals and Mud in Property Law." Stanford Law Review. Vol. 40, pp. 577-610.

[39] Rubin, Paul H. 1977. "Why is the Common Law Efficient." Journal of Legal Studies. Vol. 6, pp. 51-63.

[40] _____. 1982. "Common Law versus Statute Law." Journal of Legal Studies. Vol. 11, pp. 205-223.

[41] Shavell, Steven. 2005. "Optimal Discretion in the Application of Rules." American Law & Economics Association 15th Annual Meeting. Working Paper 25. http://law.bepress.com/alea/15th/bazaar/art25

[42] Shin, Hyun Song. 1994. "The Burden of Proof in a Game of Persuasion." Journal of Economic Theory. Vol. 64, pp. 253-264.

[43] _____. 1998. "Adversarial versus Inquisitorial Procedures in Arbitration."

[44] Tullock, Gordon. 1980. Trials on Trial.

5 Appendix

Proposition 6

Proof. If good only invest,

$$
\begin{aligned}
W_t^n(good) &= N[(1-p)W^* + p(1-G(\tilde{\alpha}_t))[(1-\sigma_t^1 W^* - k] \\
&= W^e + pN(1-G(\tilde{\alpha}_t))[((1-\sigma_t^1)W^* - k] \\
&> W^e \ \ iff \ \ (1-\sigma_t^1)W^* - k > 0.
\end{aligned}
$$

$\frac{\partial W_t^n(good)}{\partial \tilde{\alpha}} = -G'(\tilde{\alpha}_t)Np[(1-\sigma_t^1 W^* - k] \leq 0$ iff $[(1-\sigma_t^1)W^* - k] > 0$. If good and bad invest,

$$
\begin{aligned}
W_t^n(good+bad) &= N\left[G(\tilde{\alpha}_t)(1-p)W^* + (1-G(\tilde{\alpha}_t))[(p(1-\sigma_t^1)W^* + (1-p)(1-\sigma_t^2)W^* - k)]\right] \\
&= N\left[(1-p)W^* + (1-G(\tilde{\alpha}_t))\left(p(1-\sigma_t^1)W^* - (1-p)\sigma_t^2 W^* - k\right)\right] \\
&= W^e + N[(1-G(\tilde{\alpha}_t))\left(p(1-\sigma_t^1)W^* - (1-p)\sigma_t^2 W^* - k\right] \\
&\geq W^e \ \ iff \ \ p(1-\sigma_t^1)W^* - (1-p)\sigma_t^2 W^* - k > 0.
\end{aligned}
$$

$\frac{\partial W_t^n(good+bad)}{\partial \tilde{\alpha}} = -G'(\tilde{\alpha}_t)N[(p(1-\sigma_t^1)W^* - (1-p)\sigma_t^2 W^* - k] \leq 0$ iff $[p(1-\sigma_t^1)W^* - (1-p)\sigma_t^2 W^* - k] > 0$. If $W_1^n < W^e$, rule change in period 1 is optimal if Δ generates σ_2^1 and σ_2^2 such that $W_2^n > W^e + (W^e - W_1^n)$. Rule change in period 1 under these conditions will be optimal, for example, if $\sigma_2^1(K_1 + \varepsilon) = 0$ and $\frac{k}{D} > \sigma_2^2$: only good types invest in period 2 and there is no type 1 error; this is worth an arbitrarily small amount of rule change in period 1. ∎

Proposition 7 Proof. (In the following proof, superscripts A and B are dropped from parameters that are the same in both regimes.) Let $W^i = W_1^i + W_2^i$, $i = A, B$. Assume the conditions for rule change in period 1 (Proposition 5) are met in A, such that $\Delta^A > 0$. If A does not meet the conditions for rule change, each of the conditions (a) - (e) implies that B also does not meet the conditions; neither regime experiences any rule change in either period and $W^A = W^B = 2W^e$. If B does not meet the conditions for rule change in period 1 but A does, $W^A > W^B$. For all other cases, I will show that $W_1^A \geq W_1^B$ and $\sigma_2^{1A} < \sigma_2^{1B}$, $\sigma_2^{2A} < \sigma_2^{2B}$. We are restricting to cases in which if rule change is sought in both regimes the mix of defendants is the same. If there are defendants in B who seek change in period 2, lower errors in A implies $\tilde{\alpha}_2^A < \tilde{\alpha}_2^B$. Then

$$
\begin{aligned}
W_2^{nA}(good + bad) &= W^e + N[(1 - G(\tilde{\alpha}_2^A))\left(p(1 - \sigma_2^{1A})W^* - (1-p)\sigma_2^{2A}W^* - k\right] \\
&\geq W^e + N\left[(1 - G(\tilde{\alpha}_2^B)[p(1 - \sigma_2^{1B})W^* - (1-p)\sigma_2^{2B}W^* - k)\right] \\
&= W_2^{nB}(good + bad)
\end{aligned}
$$

and

$$
\begin{aligned}
W_2^{nA}(good) &= W^e + pN[(1 - G(\tilde{\alpha}_2^A))[(1 - \sigma_2^{1A})W^* - k] \\
&\geq W^e + pN[(1 - G(\tilde{\alpha}_2^B))[(1 - \sigma_2^{1B})W^* - k] \\
&= W_2^{nB}(good).
\end{aligned}
$$

If there are no defendants who seek rule change in B in period 2, then $W_2^B = W^e < W_2^A$.

(a) If $i^A(\Delta) > i^B(\Delta)$, $W_1^A = W_1^B$. Also, if $i^A(\Delta) > i^B(\Delta)$ $\sigma_2^{1A} < \sigma_2^{1B}$ and $\sigma_2^{2A} < \sigma_2^{2B}$.

(b) If $G^A(\tilde{\alpha}_1) < G^B(\tilde{\alpha}_1)$, given $W_1^n > W^e$,

$$
\begin{aligned}
W_1^A(good + bad) &= W^e + N\left[(1 - G^A(\tilde{\alpha}_1))\left(p(1 - \sigma_1^1)W^* - (1-p)\sigma_1^2 W^* - k\right)\right] \\
&> W^e + N\left[(1 - G^B(\tilde{\alpha}_1))\left(p(1 - \sigma_1^1)W^* - (1-p)\sigma_1^2 W^* - k\right)\right] \\
&= W_1^B(good + bad)
\end{aligned}
$$

and

$$
\begin{aligned}
W_1^A(good) &= W^e + pN(1 - G^A(\tilde{\alpha}_1))[((1 - \sigma_1^1)W^* - k] \\
&> W^e + pN(1 - G^B(\tilde{\alpha}_1))[((1 - \sigma_1^1)W^* - k] \\
&= W_1^B(good).
\end{aligned}
$$

$$\Delta^A(good + bad) = Nk(1 - G^A(\tilde{\alpha}_1))$$
$$> Nk(1 - G^B(\tilde{\alpha}_1))$$
$$= \Delta^B(good + bad)$$

and

$$\Delta^A(good) = pNk(1 - G^A(\tilde{\alpha}_1))$$
$$> pNk(1 - G^B(\tilde{\alpha}_1))$$
$$= \Delta^B(good).$$

Therefore $\sigma_2^{1A} < \sigma_2^{1A}$ and $\sigma_2^{2A} < \sigma_2^{2A}$.

(c) If $K_1^A > K_1^B$, $\sigma_1^{1A} < \sigma_1^{1B}$ and $\sigma_1^{2A} < \sigma_1^{2B}$, implying $\tilde{\alpha}_1^A < \tilde{\alpha}_1^B$ and as in (b) $G(\tilde{\alpha}_1^A) < G(\tilde{\alpha}_1^B)$ and $W_1^A > W_1^B$, $\Delta^A > \Delta^B$ and $\sigma_2^{1A} < \sigma_2^{1A}$ and $\sigma_2^{2A} < \sigma_2^{2A}$.

(d) If $k^A < k^B$, and $\frac{k^B}{D} > (1 - \sigma_t^1)$ then no rule change is sought in B in period t and, as shown above, $W^A > W^B$. If $\frac{k^B}{D} < (1 - \sigma_t^1)$,

$$W_t^A(good + bad) = W^e + N\left[(1 - G(\tilde{\alpha}_t))\left(p(1 - \sigma_t^1)W^* - (1 - p)\sigma_t^2 W^* - k^A\right)\right]$$
$$> W^e + N\left[(1 - G(\tilde{\alpha}_t))\left(p(1 - \sigma_t^1)W^* - (1 - p)\sigma_t^2 W^* - k^B\right)\right]$$
$$= W_t^B(good + bad).$$

and

$$W_t^A(good) = W^e + pN(1 - G(\tilde{\alpha}_t))[((1 - \sigma_t^1)W^* - k^A]$$
$$> W^e + pN(1 - G(\tilde{\alpha}_t))[((1 - \sigma_t^1)W^* - k^B]$$
$$= W_t^B(good).$$

(e) If $(\frac{k}{D})^A < (1 - \sigma_1^1) < (\frac{k}{D})^B$ no rule change is sought in B in period 1 and $W^A > W^B$. ∎

Proposition 8   **Proof.** (In the following proof, superscripts A and B are dropped from parameters that are the same in both regimes.) Rule change is justified if $W_1^n + W_2^n \geq 2W^e$. Let $W^i = W_1^i + W_2^i$, $i = A, B$. No rule change will occur in period 1, and therefore in period 2, in any regime for which $(\frac{k}{D}) > (1 - \sigma_1^1)$. Note that under the maintained assumption that $i' > 0$, $\sigma_2^1 \leq \sigma_1^1$ and $\sigma_2^2 \leq \sigma_1^2$. This implies that, for fixed $(\frac{k}{D})$, it cannot be the case that good defendants proffer a defense in period 1 but not in period 2, or that bad defendants choose not proffer a defense in period 1 but do so in period 2. If only good defendants proffer a defense in period 1 in regime A ($(1 - \sigma_1^{1A}) > (\frac{k}{D})^A > \sigma_1^{2A}$) but no defendants do so in regime B ($(\frac{k}{D})^B > (1 - \sigma_1^{1B}) > \sigma_1^{2B}$), A clearly has higher social welfare than B whenever $W_1^n + W_2^n \geq 2W^e$. This proves

(c) and any cases (a) through (e) in which only good defendants proffer a defense in A and no defendants do in B. In the remaining cases, the mix of defendants is the same in both regimes and there are three cases to consider: (i) both good and bad defendants proffer a defense in both periods $((\frac{k}{D}) < \sigma_2^2 \leq \sigma_1^2)$, (ii) both good and bad defendants proffer a defense in period 1 but only good do so in period 2 $(\sigma_2^2 < (\frac{k}{D}) < \sigma_1^2)$ and (iii) only good defendants proffer a defense in both periods $(\sigma_2^2 \leq \sigma_1^2 < (\frac{k}{D}))$. I will show that for all three cases, $W^A > W^B$ under conditions (b) and (c) and that there exist parameters such that $W^A > 2W^e > W^B$. I will show that under conditions (d) and (e) it may be the case that $W^A > W^B$ or that $W^A < W^B$.

(b) If $i^A(\Delta) > i^B(\Delta)$, $W_1^{nA} = W_1^{nB}$ and $\Delta^A = \Delta^B$. Furthermore, for $\Delta > 0$, $\sigma_2^{1A} < \sigma_2^{1B}$ and $\sigma_2^{2A} < \sigma_2^{2B}$ and this implies, given that the mix of defendants is the same in both regimes, that $\tilde{\alpha}_2^A < \tilde{\alpha}_2^B$. For any parameters, (i)

$$
\begin{aligned}
W_2^{nA}(good+bad) &= W^e + N\left[(1-G(\tilde{\alpha}_2^A))[p(1-\sigma_2^{1A})W^* - (1-p)\sigma_2^{2A}W^* - k]\right] \\
&> W^e + N\left[(1-G(\tilde{\alpha}_2^B))[p(1-\sigma_2^{1B})W^* - (1-p)\sigma_2^{2B}W^* - k]\right] \\
&= W_2^{nB}(good+bad)
\end{aligned}
$$

(ii)

$$
\begin{aligned}
W_2^{nA}(good) &= W^E + pN(1-G(\tilde{\alpha}_2^A))[(1-\sigma_2^{1A})W^* - k] \\
&> W^E + N\left[(1-G(\tilde{\alpha}_2^A))[p(1-\sigma_2^{1A})W^* - (1-p)\sigma_2^{2A}W^* - k]\right] \\
&= W_2^{nB}(good+bad)
\end{aligned}
$$

(iii)

$$
\begin{aligned}
W_2^{nA}(good) &= W^e + pN[(1-G(\tilde{\alpha}_2^A))[(1-\sigma_2^{1A})W^* - k] \\
&> W^e + pN[(1-G(\tilde{\alpha}_2^B))[(1-\sigma_2^{1B})W^* - k] \\
&= W_2^{nB}(good)
\end{aligned}
$$

Therefore, $W^A > W^B$. Furthermore, for any parameters such that $W^A > 2W^e$, $W^B < 2W^e$ in case (iii) if $\frac{\partial i^B(\Delta)}{\partial \Delta}$ is sufficiently small that $\sigma_2^{1B} \to \sigma_1^{1B}$. Then $W_2^{nB} \to W_1^{nB} < W^e$.

(c) If $k^A < k^B$, given that the mix of defendants is the same in both regimes and $\sigma_1^{1A} = \sigma_1^{1B}$ and $\sigma_1^{2A} = \sigma_1^{2B}$, then $\tilde{\alpha}_1^A = \tilde{\alpha}_1^B$ and $\Delta^A = \Delta^B$. This implies $\sigma_2^{1A} = \sigma_2^{1B}$ and $\sigma_2^{2A} = \sigma_2^{2B}$, and therefore $\tilde{\alpha}_2^A = \tilde{\alpha}_2^B$. Then $W_t^{nA}(good+bad) = W_t^{nB}(good+bad) + N(1-G(\tilde{\alpha}_t))[k^B - k^A]$, $W_t^{nA}(good) = W_t^{nB}(good) + pN(1-G(\tilde{\alpha}_t))[k^B - k^A]$. This implies $W^A \geq W^B$ in cases (i), (ii) and (iii). Furthermore, for any parameters such that $W^A > 2W^e$, $W^B < 2W^e$ in case (iii) if $k^A < (1-\sigma_2^1)W^* < k^B$, which is satisfied when $\sigma_2^1 = 1 - \frac{k^A}{W^*} + \frac{k^B - k^A}{nW^*} < 1$ for n sufficiently large.

(d) If $G^A(\tilde{\alpha}_1) < G^B(\tilde{\alpha}_1)$, given that the same mix of defendants proffering a defense are present in both regimes, $\Delta^A > \Delta^B$. This implies that $\sigma_2^{1A} < \sigma_2^{1B}$ and $\sigma_2^{2A} < \sigma_2^{2B}$, leading to $\tilde{\alpha}_2^A < \tilde{\alpha}_2^B$. To see that both $W^A > W^B$ and $W^A <$

W^B are possible, consider case (iii) with $W_2^{nA} > W^e$. Recall that $W_1^n < W^e$ implies $(1 - \sigma_1^1)W^* - k < 0$. Then

$$W^A - W^B = pN[G^B(\tilde{\alpha}_1) - G^A(\tilde{\alpha}_1)][(1 - \sigma_1^1)W^* - k]$$
$$+ pN[1 - G^B(\tilde{\alpha}_2^B)](\sigma_2^{1B} - \sigma_2^{1A})W^* + pN[G^B(\tilde{\alpha}_2^B) - G^A(\tilde{\alpha}_2^A)][(1 - \sigma_2^{1A})W^* - k] \quad (1)$$

$$\underset{k \to (1 - \sigma_1^1)W^*}{Lim}(W^A - W^B) = pN[1 - G^B(\tilde{\alpha}_2^B)](\sigma_2^{1B} - \sigma_2^{1A})W^* + pN[G^B(\tilde{\alpha}_2^B) - G^A(\tilde{\alpha}_2^A)][(1 - \sigma_2^{1A})W^* - k]$$
$$> 0$$

$$\underset{\substack{\sigma_2^{1B} \to \sigma_2^{1A} \\ G^B(\tilde{\alpha}_2^B) \to G^A(\tilde{\alpha}_2^A)}}{Lim}(W^A - W^B) = pN[G^B(\tilde{\alpha}_1) - G^A(\tilde{\alpha}_1)][(1 - \sigma_1^1)W^* - k]$$
$$< 0$$

(e) If $K_1^A > K_1^B$, then $\sigma_1^{1A} < \sigma_1^{1B}$ and $\sigma_1^{2A} < \sigma_1^{2B}$. Given that the mix of defendants is the same in both regimes, this implies $\tilde{\alpha}_1^A < \tilde{\alpha}_1^B$ and therefore $\Delta^A > \Delta^B$. $\Delta^A > \Delta^B$ then implies that $\sigma_2^{1A} < \sigma_1^{1B}$, $\sigma_2^{2A} < \sigma_1^{2B}$ and $\tilde{\alpha}_2^A < \tilde{\alpha}_2^B$. To see that both $W^A > W^B$ and $W^A < W^B$ are possible, consider case (iii) with $W_2^{nA} > W^e$. Recall that $W_1^n < W^e$ implies $(1 - \sigma_1^1)W^* - k < 0$. Then

$$W^A - W^B = pN[1 - G(\tilde{\alpha}_1^B)](\sigma_1^{1B} - \sigma_1^{1A})W^* + pN[G(\tilde{\alpha}_1^B) - G(\tilde{\alpha}_1^A)][(1 - \sigma_1^{1A})W^* - k]$$
$$+ pN[1 - G(\tilde{\alpha}_2^B)](\sigma_2^{1B} - \sigma_2^{1A})W^* + pN[G(\tilde{\alpha}_2^B) - G(\tilde{\alpha}_2^A)][(1 - \sigma_2^{1A})W^* - k] \quad (2)$$

$$\underset{k \to (1 - \sigma_1^{1A})W^*}{Lim}(W^A - W^B) = pN[1 - G(\tilde{\alpha}_1^B)](\sigma_1^{1B} - \sigma_1^{1A})W^* + pN[1 - G(\tilde{\alpha}_2^B)](\sigma_2^{1B} - \sigma_2^{1A})W^*$$
$$+ pN[G(\tilde{\alpha}_2^B) - G(\tilde{\alpha}_2^A)][(1 - \sigma_2^{1A})W^* - k]$$
$$> 0$$

$$\underset{\substack{\sigma_2^{1B} \to \sigma_2^{1A} \\ G^B(\tilde{\alpha}_1^B) \to 1}}{Lim}(W^A - W^B) = pN[G(\tilde{\alpha}_1^B) - G(\tilde{\alpha}_1^A)][(1 - \sigma_1^{1A})W^* - k]$$
$$< 0$$

(Note that in this latter example, there is little rule change in B but this is sufficient to achieve almost all the gains in error reduction achieved with much more widespread (and costly) rule change in A.) ∎

Proposition 9 Proof. Taking the cases in the proposition in reverse order, I will show that under each condition, there exist parameters such that $W^B > W^A$.

(a) *Evidence from bad defendants not informative.* If only good defendants seek rule change in A while both good and bad do in B, $\sigma_t^{2A} < (\frac{k}{D})^A$ and $\sigma_t^{2B} >$

$(\frac{k}{D})^B$. $\sigma_t^{1A} > p\sigma_t^{1B} + (1-p)\sigma_t^{2B}$ implies that $\tilde{\alpha}_1^B < \tilde{\alpha}_1^A$ and $G(\tilde{\alpha}_1^B) < G(\tilde{\alpha}_1^A)$. Consider $t = 1$.

$$
\begin{aligned}
W_1^B - W_1^A &= N[(1 - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k^B] - pN[(1 - G(\tilde{\alpha}_1^A))(1 - \sigma_A^{1A})W^* - k^A] \\
&= N[(1 - G(\tilde{\alpha}_1^A))(p(\sigma_A^{1A} - \sigma_1^{1B})W^* + p(k^A - k^B) - (1-p)(\sigma_1^{2B}W^* + k^B)] \\
&\quad + N(G(\tilde{\alpha}_1^A) - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k^B)
\end{aligned}
$$

$$
\begin{aligned}
\underset{\tilde{\alpha}_1^A \to \bar{\alpha}}{Lim}(W_1^B - W_1^A) &= N(1 - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k^B) \\
&> 0 \text{ if } W_1^B > W^e
\end{aligned}
$$

Then we can conclude $\exists\ \sigma_1^{1A} \ni \tilde{\alpha}_1^A < \bar{\alpha}$ and $W_1^B > W_1^A$. If evidence from bad defendants is not informative, $\Delta^A = pN(1 - G(\tilde{\alpha}_1^A))$ and $\Delta^B = \delta pN(1 - G(\tilde{\alpha}_1^B))$, where $0 \leq \delta \leq 1$ captures the impact of evidence from bad defendants on the quantity of legal human capital obtained from good defendants. Note $\delta = 1$ if evidence from bad defendants is uninformative and $\delta < 1$ if evidence from bad defendants degrades what is learned from good defendants. Then even though $\Delta^A \geq \Delta^B$, given $W_2^B \geq W_2^B > W^e$, $W_2^B > W_2^A$ obtains provided $\sigma_2^{1A} > p\sigma_2^{1B} + (1-p)\sigma_2^{2B}$, $\tilde{\sigma}_2^{2B} > (\frac{k}{D})^B$ and, as shown for period 1, $\tilde{\alpha}_2^A$ is sufficiently close to $\bar{\alpha}$.

(b) *Evidence from bad defendants informative.* If only good defendants seek rule change in A while both good and bad do in B, $\sigma_1^{2A} < (\frac{k}{D})^A$ and $\sigma_1^{2B} > (\frac{k}{D})^B$. $\Delta^A = pN(1 - G(\tilde{\alpha}_1^A))$ and $\Delta^B = N(1 - G(\tilde{\alpha}_1^B))$. Then $\Delta^B > \Delta^A$ if $p < \frac{1 - G(\tilde{\alpha}_1^B)}{1 - G(\tilde{\alpha}_1^A)}$. This condition is always met if $\sigma_1^{1A} > p\sigma_1^{1B} + (1-p)\sigma_1^{2B}$, implying that $\tilde{\alpha}_1^B < \tilde{\alpha}_1^A$ and $G(\tilde{\alpha}_1^B) < G(\tilde{\alpha}_1^A)$. Provided $i'(\Delta) > 0$, $\Delta^B > \Delta^A$ implies $\sigma_2^{1B} < \sigma_2^{1A}$ and $\sigma_2^{2B} < \sigma_2^{2A}$ provided $\Delta^B - \Delta^A > K_1^A - K_1^B$. Under this condition only good defendants proffer a defense in both regimes in period 2 if $(\frac{k}{D})^A - (\frac{k}{D})^B < \sigma_2^{2A} - \sigma_2^{2B}$ (which implies $(\frac{k}{D})^B > (\frac{k}{D})^A - \sigma_2^{2A} + \sigma_2^{2B} > \sigma_2^{2B}$.) Then, if $k^B \leq k^A$

$$
\begin{aligned}
W_2^B &= W^e + pN[(1 - G(\tilde{\alpha}_2^B)(1 - \sigma_2^{1B})W^* - k^B] \\
&> W^E + pN[(1 - G(\tilde{\alpha}_2^A)(1 - \sigma_2^{1A})W^* - k^A] \\
&= W_2^A.
\end{aligned}
$$

If, as shown above, $W_1^B > W_1^A$, $W^B > W^A$. If $W_1^B < W_1^A$, suppose first that $\sigma_1^{1A} < \sigma_1^{1B}$ and $\sigma_1^{2A} < \sigma_1^{2B}$ ($K_1^A > K_1^B$) and $k^B = k^A$. Then

$$
\begin{aligned}
W^B - W^A &= N[(1 - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k] + pN[(1 - G(\tilde{\alpha}_2^B)(1 - \sigma_2^{1B})W^* - k] \\
&\quad - pN[(1 - G(\tilde{\alpha}_1^A)(1 - \sigma_A^{1A})W^* - k] - pN[(1 - G(\tilde{\alpha}_2^A)(1 - \sigma_2^{1A})W^* - k] \\
&= N[(1 - G(\tilde{\alpha}_1^A)p(\sigma_1^{1A} - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k] + N[(G(\tilde{\alpha}_1^A) \\
&\quad - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k)] + pN[(1 - G(\tilde{\alpha}_2^A)(\sigma_2^{1A} - \sigma_2^{1B})W^*] \\
&\quad + pN[(G(\tilde{\alpha}_2^A) - G(\tilde{\alpha}_2^B)((1 - \sigma_2^{1B})W^* - k)] \\
&> 0 \text{ if } W_t^B > W^e \text{ and } \sigma_2^{1A} - \sigma_2^{1B} > 1 - \sigma_1^{1A}.
\end{aligned}
$$

Now suppose $\sigma_1^{1A} = \sigma_1^{1A}$ and $\sigma_1^{2A} = \sigma_1^{2B}$ $(K_1^A = K_1^B)$ and $k^B < k^A$. Then

$$
\begin{aligned}
W^B - W^A &= N[(1 - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k^B] + pN[(1 - G(\tilde{\alpha}_2^B)(1 - \sigma_2^{1B})W^* - k^B] \\
&\quad -pN[(1 - G(\tilde{\alpha}_1^A)(1 - \sigma_A^{1A})W^* - k^A] - pN[(1 - G(\tilde{\alpha}_2^A)(1 - \sigma_2^{1A})W^* - k^A] \\
&= N(1 - G(\tilde{\alpha}_1^A))[p(k^A - k^B) - (1-p)(\sigma_1^{2B}W^* + k^B)] \\
&\quad +N[(G(\tilde{\alpha}_1^A) - G(\tilde{\alpha}_1^B))(p(1 - \sigma_1^{1B})W^* - (1-p)\sigma_1^{2B}W^* - k^B)] \\
&\quad +pN(1 - G(\tilde{\alpha}_2^A))[(\sigma_2^{1A} - \sigma_2^{1B})W^* + (k^A - k^B)] + pN[(G(\tilde{\alpha}_2^A) - G(\tilde{\alpha}_2^B)((1 - \sigma_2^{1B})W^* - k^B)] \\
&> 0 \text{ if } W_t^B > W^e \text{ and } k^A > (1 - \sigma_1^{1B})W^*.
\end{aligned}
$$

∎