Democracy, Judicial Attitudes and Heterogeneity: the Civil Versus Common Law Tradition.[*]

Carmine Guerriero

Faculty of Economics, University of Cambridge.

May 19, 2009

Abstract

A key issue in the design of a legal system is the choice of the mechanism aggregating preferences over the level of deterrence. While under Case law appellate judges' biases offset one another at the cost of volatility of precedents, under Statute law the Legislator chooses certain rules that are biased only when bribes are accepted: i.e., when political institutions are weak and/or the preference heterogeneity is sufficiently high. Thus, only in the last scenario, Case law can outperform Statute law. Also, institutions fostering limited discretion by lower courts improve the performance of Case law. Instrumental variables estimates based on historical data from 156 countries confirm this prediction.

Keywords: legal origins, culture, democracy, economic development.
JEL classification: K40; Z1; H11; 010; P16.

"A legal tradition [. . .] is not a set of rules of law [. . .] rather it is a set of deeply rooted, historically conditioned attitudes about [. . .] the proper organization and operation of a legal system. The legal tradition relates the legal system to the culture of which it is a partial expression" *The Civil Law Tradition*, Merryman (1969), page 2.

I.   Introduction

The laws and the institutions entrusted with their creation are deeply influenced, in a great part of the world, by either the Civil or the Common law tradition. While the latter originated in England and has, then, been transplanted through colonization into England's ex-colonies, the former has its roots in the Roman law and was exported by Napoleon to continental European countries and by the latter to Latin America, Africa and Asia. Structurally the two traditions operate in different ways: while Common law recognizes a crucial role to precedents and allows more procedural discretion to lower adjudicating courts, Civil law relies on legal codes and bright line adjudication rules. Exploiting these differences and assuming that transplanted institutions remained in place unaffected after independence, the "legal origins" project has provided evidence suggesting that French and socialist Civil law countries display more inefficient governments and courts, less secure investor rights and a stricter market regulation than Common law countries do (see for a review La Porta, Lopez-de-Silanes and Shleifer, [2008]).

Yet, legal scholars have recently documented a great wave of convergence in such a way that, for instance, countries where the Common law tradition was transplanted are relying more heavily on statutes (see Zweigert and Kötz, [1998]; Roe, [2004]). What is,

therefore, the set of forces justifying, despite the above mentioned evidence in favor of Common law, the existence of such dissimilar institutions and their evolution?

This paper lays out a theoretical framework for thinking about these issues, and explores its empirical implications using data on the history of the legal order of 156 countries that received their initial legal tradition externally by a fairly exogenous transplantation process—transplanted, therein. The model merges a recent body of research on endogenous lobbying (Felli and Merlo, 2006) with a lively literature on judicial activism and Case law (Gennaioli and Shleifer 2007a and b; Fernandez and Ponzetto 2008a and b), and studies a world of social groups separated by their cultural preference over the level of deterrence that laws impose. This society compares the welfare properties of two law-making institutions. Under Case law appellate judges, randomly picked from the population and bearing variable costs for overruling the precedent, choose the law. Under Statute law, instead, the prevailing rule is selected by a Legislator who bargains with coalitions of the social groups facing fixed collective action costs. While the Legislator selects certain rules which are unbiased only if she refuses bribes, judicial biases balance one another and induce the long run optimal rule at the cost of everlasting volatility. If the cultural distance among groups is small, collective action costs discourage bribes and Statute law outperforms Case law. Yet, as the cultural distance rises, the Legislator eventually accepts bribes and statutes move away from the optimal rule the more the lower is the quality of political process—i.e., democracy. In this case, there is a level of democracy below which Case law prevails over Statute law. The intuition remains robust if I consider alternatives to the basic assumptions on the functioning of the two institutions. Moreover, rules allowing the use of limited discretion by lower adjudicating courts increase appellate

3

judges' overruling costs and make statutes volatile. Thus, a pure Common (Civil) law tradition where Case (Statute) law is used along with discretion-enhancing (bright line) adjudication rules endogenously arises.

To test the model's predictions I have collected data tracking the evolution of the legal system of a cross section of transplanted countries. I record whether Statute or Case law was in place in the year of independence and in 2000 and four proxies for the level of discretion in adjudication for the same two points in time. Consistent with the model, in countries in which Civil law—Statute law and bright line adjudication rules— was transplanted, reforms toward Common law—Case law and flexible adjudication—are more likely the weaker are the constraints on the executive and the broader is cultural heterogeneity. Symmetrically moves toward Civil law in countries to which Common law was imposed are found where the constraints on the executive are the strictest. I measure cultural heterogeneity with both the genetic distance between the population who chose the law-making rule and that of the transplanted country and the ethno-linguistic fractionalization in the transplanted country. These results survive if I employ an instrument suggested by evolutionary biology (cross-cultural psychology) to deal with the endogeneity of the ethno-linguistic fractionalization (constraints on the executive) proxy. I also show that the likelihood of reforms toward Common law and the risk of coercion by strong elites are positively related. This result casts doubts on the Glaeser and Shleifer's (2002) explanation of the rise of Common (Civil) law in England (France).[1]

The paper most closely related to mine is Fernandez and Ponzetto (2008a). Building on

---

[1] Glaeser and Shleifer (2002) study the trade-off between vulnerability of judges to private subversion and the possible capture by the central power. In the 13th century, England was more peaceful than France and it was optimal for the former (latter) to adopt adjudication by independent juries (state judges).

similar adjudication microstructures, they also compare Statute law and Case law. They obtain that judge-made law has the same reversion toward the optimal rule property. Yet, the hypothesis that the set of coalitions bargaining with the Legislator is random produces the result that statutes are unpredictable and eventually biased and so suboptimal with respect to Case law in the long run. The assumption is not only empirically unreasonable (see Wright, [1996]), but it also fails to highlight the role of preference heterogeneity and its relation with adjudication institutions. Yet, this is a general drawback of the literature. Indeed, although the comparative merits of Statute law and Case law have been debated for centuries, the present paper constitutes the first attempt to link the welfare properties of the two traditions to the cultural and political environment.[2] Moreover, a major contribution of my analysis is to provide a first empirical test of different theories justifying the choice of the legal institutions typical of each tradition as the result of maximizing behaviors. Finally, my results cast several doubts on the supposed primacy of Common law stressed by the "legal origins" literature. To this extent, the new "legal origins" project should consider the evolution of both law making and adjudication institutions as driven by the dispersion of preferences and the quality of the political process, both of which affect performances as well. The rest of the paper is organized as follows.

Section II identifies the extent of democracy and cultural heterogeneity conducive to a primacy of either Statute law or Case law. Section III studies the relation between law making and adjudication institutions. Section IV discusses the empirical results, and section V concludes. The appendix gathers proofs and tables.

---

[2]Proponents of Statute law (Hobbes 1681; Bentham, 1891) stress the certainty of statutes; supporters of Case law (Hayek, 1960; Cardozo, 1921; Posner, 2007) exalt the evolutionary properties of precedents.

II.   The Optimal Law-making Institution

Next, I will first introduce the analysis of the underlining regulation problem, which takes its approach from Gennaioli and Shleifer (2007a); then, I will explain how it is performed under Case law and Statute law. Finally, I will compare the two institutions.

The regulation problem: first and second best.—Consider a society composed by a continuum of citizens whose mass is normalized to one. This society is interested in regulating a harmful action affecting two parties: an offender O and a victim V. The former can take precautions at a cost C. Precautions reduce the probability of an accident from p_N to $p_P = 0$. The harm suffered by the victim is normalized to unity, and precautions are socially optimal if and only if $p_N - p_P = p_N > C$. The conditional probability of an accident depends on the attributes $a \in [0, 1]$ and $u \in [0, 1]$ which are independently and uniformly distributed in the population of cases. Assume that $p_N - p_P = \overline{\Delta P} > C$ if $a + u \geq 1$ and $p_N - p_P = \underline{\Delta P} < C$ if $a + u < 1$. Therefore, the offender should be held liable whenever $a + u \geq 1$. The legal system has two tiers. In the lower tier adjudicating courts observe only a and act as a perfect agent of the law-maker—to be relaxed in section II.B. Damages are so high that they induce the offender to take precautions when she is held liable; so, in the upper tier, the law-maker's problem reduces to finding the optimal level of a—call it A—such that lower courts hold the offender liable if and only if $a \geq A$.

Imperfect information implies statistical errors, and the offender is mistakenly held liable if $a \geq A \wedge a < 1 - u$, which happens with probability $\int_A^1 (1 - a)\, da = (1/2) (1 - A)^2$, and mistakenly held not liable if $a < A \wedge a \geq 1 - u$ which happens with probability $\int_0^A a\, da = (1/2) A^2$. Over (under) precautions bring the fixed marginal cost $\bar{\Lambda} = C - \underline{\Delta P}$

$(\underline{\Lambda} = \overline{\Delta P} - C)$; $\lambda = \bar{\Lambda}/\underline{\Lambda}$ is the relative cost of a false positive or of over-precautions.

A world of biases.—The population is equally split in two groups $i \in I = \{L, H\}$. They differ in the relative perceived cost of a false positive $\beta_i = \beta_{O,i}/\beta_{V,i}$. If the groups have different mass, the model's features are similar but the algebra more cumbersome. The preference intensity is normalized in order to have $\beta_{O,i} + \beta_{V,i} = 1, \forall i$. The unconcerned group has $\beta_L = \lambda\pi$ and the concerned one has $\beta_L = \lambda\pi$. Accordingly, $\pi \in (1, \infty)$ represents a measure of cultural heterogeneity: i.e., the extent of disagreement about the perceived danger of the action between citizens belonging to groups with different ethnic, religious or political long-lasting tastes or ideologies.[3] As a result, the loss of welfare relative to the first best for a citizen belonging to group i is

$$\Lambda_i(A) = (1/2)\left[\beta_{O,i}(1-A)^2 + \beta_{V,i}A^2\right], \tag{1}$$

which can be fully characterized by group i's favorite threshold $\hat{A}_i = \beta_{O,i} = \Gamma(\beta_i)$ with $\Gamma(x) \equiv x(1+x)^{-1}$.[4] The unconcerned group prefers the lenient $\hat{A}_L = \lambda\pi(1+\lambda\pi)^{-1} \equiv \bar{A}$ while the concerned one the stricter $\hat{A}_H = \lambda(\pi+\lambda)^{-1} \equiv \underline{A}$. Thus, (1) rewrites as:

$$\Lambda_i(A) = (1/2)\left[\hat{A}_i(1-A)^2 + \left(1-\hat{A}_i\right)A^2\right] \propto (1/2)\left(A - \hat{A}_i\right)^2. \tag{2}$$

Under rule A, the loss of social welfare—relative to the first best—is

---

[3]Running public good experiments on comparable participant pools around the world, Herrmann, Thöni, and Gächter (2008) document the existence of large differences in individual willingness to punish free riders and their relation with cultural biases. Glaeser and Sacerdote (2003) provide evidence compatible with the idea that sentences respond to citizenry tastes for vengeance as driven by ethnic differences.

[4]The first best implies that in half of the cases precautions are not taken and the expected cost of accident is $\underline{\Delta P}$ and that, in the other half, precautions are optimally taken and the only cost is C.

$$\tilde{\Lambda}(A) = (1/2)\left[\Lambda_L(A) + \Lambda_H(A)\right] \propto (1/2)\left(A - E\left(\hat{A}_i\right)\right)^2, \tag{3}$$

and the optimal threshold $\tilde{A}$ is equal to the mean of the $\hat{A}_i$.[5] In evaluating the generality of the foregoing, two observations should be borne in mind. First, the model applies not only to tort law but to every conceivable legal branch as long as a is interpreted as the observed between two dimensions proving the culpability of the defendant. In this case, the cost of a false positive (negative) is given by the disutility from punishment—or the foregone utility from acting—net of the expected harm (the opposite) measured when conviction is suboptimal (optimal).[6] Second, I can relax—see section II.B—the assumption according to which all citizens are biased and the one according to which they do not care about the technological efficiency of the law—i.e., the distance between A and A^*.

The timing.—The timing is as follows (see also Fernandez and Ponzetto [2008a]):

1. At the Constitutional table—$t = 0$—society chooses between Statute and Case law comparing the expected long run losses of welfare under the two institutions;

2. A If Statute law has been chosen, the Legislator selects in $t = 1$ a rule A_S at the end of a bargaining game to be further discussed below. Then A_S sticks forever.

   B Under Case law an appellate judge randomly picked from the population selects in $t = 1$ a $A_{c,1}$ which guides adjudication until another judge overrules it in t, turning $A_{c,t-1} = A_{c,1}$ into $A_{c,t}$, and possibly giving rise to a new round of changes.

---

[5]$\tilde{A}$ is weakly greater (strictly lower) than the technologically efficient $A^* = \Gamma(\lambda)$ for $\lambda \leq 1$ ($\lambda > 1$)—see the appendix. Without normalization, $\tilde{A}$ differs from both $E\left(\hat{A}_i\right)$ and A^* for all $\lambda \neq 1$, and the lobbying game described below has more equilibria. Yet, the main qualitative message of the model still holds.
[6]The set up also applies to those situations in which ex-ante contracting and ex-post observability lead to conflicts on the optimal allocation: constitutional design, contracting and insurance (Gennaioli, 2008).

II.A   Statute Law versus Case Law

Let me start from the characterization of the equilibrium under Case law.

Case law.—Stare decisis only binds in so far as it is costly for judges to change the precedent. A judge inheriting precedent $A_{c,t-1}$ and setting a $A_{c,t} \neq A_{c,t-1}$ can introduce any change in the law. Yet, the need to justify his decision implies a persuasive effort whose cost rises with the relevance of the doctrine of *stare decisis* $K > 0$—see the institutional analysis at the end of section III—and the magnitude of the legal revision. This last detail formalizes "the idea that the greater the effective deviation from precedent, the greater the difficulty to reconcile it with the rhetorical demands of *stare decisis*" (Fernandez and Ponzetto, 2008b). Judge i 's strictly concave problem in t is:

$$\hat{A}_{c,t}(i) = \arg \max_{A_{c,t}(i)} - (1/2) \left(A_{c,t}(i) - \hat{A}_i \right)^2 - (K/2) \left(A_{c,t}(i) - \hat{A}_{c,t-1} \right)^2.$$

A type i judge minimizes both the loss she bears as a part of group i when $A_{c,t}$ differs from $\hat{A}_i$ and the one of justifying a $A_{c,t}$ far from $A_{c,t-1}$. The unique and global solution $\hat{A}_{c,t} = \Gamma(K) A_{c,t-1} + [1 - \Gamma(K)] \hat{A}_i$ implies that Case law follows an autoregressive process converging to the ergodic distribution with mean $E\left(\hat{A}_i\right)$ and variance $V\left(\hat{A}_c\right) = V\left(\hat{A}_i\right)(1 + 2K)^{-1}$, where $V\left(\hat{A}_i\right)$ is the variance of the groups' favorite rules and $\hat{A}_c$ the long run judge-made law. The long run loss of social welfare is:

$$E\left(\tilde{\Lambda}\left(\hat{A}_c\right)\right) \propto (1/2) V\left(\hat{A}_c\right) + (1/2) \left(E\left(\hat{A}_c\right) - \tilde{A}\right)^2 = V\left(\hat{A}_i\right)[2(1 + 2K)]^{-1}.$$

As already stressed by Cardozo (1921), Gennaioli and Shleifer (2007b), and Ponzetto and Fernandez (2008a, b), Case law assures long run unbiasedness at the cost of everlasting volatility.[7] Lemma 1 summarizes these results:

---

[7]Fernandez and Ponzetto (2008a) and Gennaioli and Shleifer (2007b) clarify that this basic pattern survives when the selection of disputed for litigation is endogenous (but see also Depoorter, Fon and Parisi, [2005]) or only a subset of adjudicated cases can possibly give rise to a precedent change.

Lemma 1 (Proposition 5, Gennaioli and Shleifer, 2007b): *Let u be unobservable. Case law evolves as a first-order autoregressive process which converges to the ergodic distribution* $N\left(E\left(\hat{A}_i\right), V\left(\hat{A}_i\right)(1+2K)^{-1}\right)$. *The asymptotic variance* $\mathrm{V}\left(\hat{A}_c\right)$ *rises (falls) with* $\pi\left(K\right)$ *and it is always strictly smaller than the variance of the judges' bliss points* $V\left(\hat{A}_i\right)$ *for* $K > 0$.

Statute law.—Dating back to the rise of the absolute monarchy, legislation has been interpreted as the sovereign act of the legitimate political authority (Hobbes, 1681). This conception became part of the common heritage of the Enlightenment and received its normative foundations from the establishment of modern democracies. Yet, legislation is the outcome of a political process that involves also lobbies representing the special interests of groups of citizens particularly affected by the law: as a consequence, statutes reflects the society's will the more accurately the stricter the constraints on the Legislator—the government, the legislature and the president—are.

In order to capture at best such a complex interaction between political institutions and ideological biases, I use the endogenous lobbying model proposed by Felli and Merlo (2006): A_s is chosen by a Legislator who bargains with coalitions of the two groups, and maximizes a weighted average of the loss of social welfare and bribes y—$U\left(A_s, y\right) = -\left(1-\mu\right)\tilde{\Lambda}\left(A_s\right)+\mu\left|i\right|y$, where μ has to be correctly considered as the relative weight that the Legislator is constrained or prefer to place on special interests *vis-a-vis* the society's will (see also Acemoglu and Robinson, [2006]). The latter is inversely related to the quality of the political process which, in turn, has been shown to be highly persistent and mainly related to long lived cultural traits (Acemoglu et al., 2008; Tabellini, 2008; Guiso, Sapienza, and Zingales, 2008): the component of democracy considered here is exactly

this persistent one. The adjustment for the number of groups in the population takes care of the mechanical bonus to democracy brought by a rise in $|i|$, and just simplifies the algebra. The Legislator chooses the winning coalition who, in turn, has to pay a fixed non sunk "collective action" (Olson, 1965) fee $\Psi > 0$ in order to transfer the bribes.[8] These two model's features constitute the sense in which lobbying is endogenous, and avoid the empirically unsatisfactory (Wright, 1996) menu-auction models' assumption according to which all lobbies participate to the policy-making (see Grossman and Helpman, [2001]). I also maintain that the Legislator appropriates the entire willingness to pay of the winning coalition: given the model's focus on the *ratio decidendi*, the hypothesis is immaterial.

The lobbying game works as follows. Lobby $i \in I = \{L, H\}$ has utility $-(1/2)\Lambda_i(A_s) - y$, and can sign binding contracts with the Legislator over A_s in exchange for transfers y. The Legislator has the option of not signing any contract and implementing the status quo $\tilde{A}$. Let $\Omega = \{\{\phi\}, \{L\}, \{H\}, \{L, H\}\}$ denotes the set of possible coalitions l. For any A_s the Legislator may implement instead of $\tilde{A}$, l is willing to pay

$$W_l\left(A_s, \tilde{A}\right) = \sum_{i \in l} w_i\left(A_s, \tilde{A}\right) = (1/4)\sum_{i \in l}\left[\left(\tilde{A} - \hat{A}_i\right)^2 - \left(A_s - \hat{A}_i\right)^2\right] - \Psi,$$

which is the sum of the maximum individual rational bribes of each lobby $i \in l$— $w_i\left(A_s, \tilde{A}\right)$—less the fixed fee. The Legislator, first, chooses an A_s for each $l \in \Omega$

$$\hat{A}_s(l) \in \arg\max_{A_s} -(1 - \mu)\tilde{\Lambda}(A_s) + \mu|i|W_l\left(A_s, \tilde{A}\right), \tag{4}$$

and then she chooses her preferred bargaining coalition

---

[8]These are the costs of "establishing links with politicians, hiring professional lobbyists, building a communications network [...], designing a scheme of punishments for defaulting members" (Mitra, 1999).

$$\hat{l} \in \arg\max_{l \in \Omega} - (1 - \mu) \tilde{\Lambda} \left(\hat{A}_s \left(l \right) \right) + \mu \left| i \right| W_l \left(\hat{A}_s \left(l \right), \tilde{A} \right). \tag{5}$$

Therefore, a subgame perfect equilibrium of the endogenous lobbying game is given by a threshold function $\hat{A}_s (l)$, and a coalition $\hat{l}$. If the Legislator accepts bribes, the outcome of the bargaining is a compromise between the socially optimal $\tilde{A}$ and the rules preferred by the lobbies included in the winning coalition. Given the quadratic loss function, A_s takes the form of a weighted average further away from $\tilde{A}$ the less democratic is society and the higher is the cultural heterogeneity, and in particular:

Lemma 2.A: *Let u be unobservable. For any coalition $l \in \Omega$, there is a unique optimal threshold $A_s (l_s)$ that solves problem (4):*

$$\hat{A}_s \left(l \right) = \left[(1 - \mu) \tilde{A} + \mu \sum_{i \in l} \hat{A}_i \right] \left(1 - \mu + \left| l \right| \mu \right)^{-1}, \tag{6}$$

where $\left| l \right|$ is the numerosity of coalition l.

Yet, the Legislator accepts bribes only if the polarization between groups is so wide that the rewards from corruption overcome the losses of social welfare. In particular:

Lemma 2.B: *Let u be unobservable. **1.** There are two levels of heterogeneity $\overline{\pi}$ and $\overline{\overline{\pi}}$, decreasing with μ and increasing with Ψ, such that: $\overline{\overline{\pi}} \geq \overline{\pi}$; for $\pi \geq \overline{\overline{\pi}}$ the Legislator is indifferent between $\{L\}$ and $\{H\}$, and picks either $\hat{A}_s (\{L\}) = (1 - \mu) \tilde{A} + \mu \bar{A}$ or $\hat{A}_s (\{H\}) = (1 - \mu) \tilde{A} + \mu \underline{A}$, and for $\pi < \overline{\overline{\pi}}$ the Legislator does not accept bribes; for $\pi \geq \overline{\pi}$, $\{L\}$ and $\{H\}$ pay the fixed cost, and for $\pi < \overline{\pi}$ they don't. Neither $\{L\}$ nor $\{H\}$ will pay any fixed fee $\Psi > \bar{\Psi} = 1/16$. **2.** As a consequence, for $\mu > 0$ and $\Psi \leq \bar{\Psi}$, from $t = 1$ on*

12

the loss of social welfare equals 0 *for* $\pi < \bar{\bar{\pi}}$ *and* $(\mu^2/8)\left(\bar{A} - \underline{A}\right)^2$ *for* $\pi \geq \bar{\bar{\pi}}$.[9]

Thus, the constraints on the Legislator's actions cover a second role in the equilibrium. Not only if side payments are paid, the equilibrium threshold is the nearer to the second best $\tilde{A}$ the lower μ is, but also, in a more democratic society, the Legislator finds it less appealing to be bribed—i.e., $\bar{\bar{\pi}}$ is increasing with the quality of the political process.

The optimal level of centralization.—The welfare comparison between the long run losses of welfare driven by the two institutions is now straightforward. For $K < \infty$ Case law can outperform Statute law only if the policy chosen by the Legislator is at least partially biased by bribing—i.e., if $\pi \geq \bar{\bar{\pi}}$—and

$$\tilde{\Lambda}\left(\hat{A}_s\right) > \mathrm{E}\left(\tilde{\Lambda}\left(\hat{A}_c\right)\right) \leftrightarrow (\mu^2/8)\left(\bar{A} - \underline{A}\right)^2 > \left(\bar{A} - \underline{A}\right)^2 \left[8\left(1 + 2K\right)\right]^{-1} \leftrightarrow \mu^2 > \left(1 + 2K\right)^{-1},$$

which leads directly to the following proposition:

Proposition 1: *Let u be unobservable. Statute law is always strictly better than Case law for* $\pi < \bar{\bar{\pi}}$ *and* $K < \infty$ *and the two institutions induce the same expected welfare for* $\pi < \bar{\bar{\pi}}$ *and* $K = \infty$. *For* $\pi \geq \bar{\bar{\pi}}$ *and every* K, *there is a* $\bar{\mu}\left(K\right)$ *with* $\partial\bar{\mu}\left(K\right)/\partial K < 0$, $\bar{\mu}\left(\infty\right) = 0$ *and* $\bar{\mu}\left(0\right) = 1$, *such that Statute law performs strictly better (weakly worse) than Case law for every* $\mu < \bar{\mu}\left(K\right)$ *(otherwise).*[10]

The intuition for this result is immediate. Whenever bribes affect the equilibrium rule, Statute law assures certainty at the cost of a biased rule. Case law, instead, achieves the

---

[9]Implicit in (4) and (5), there is the assumption that both side transfers and collective action costs do not entail social wastes: relaxing it does not change the qualitative message of the model—see footnote 12. Only the algebra (the interpretation) but not the qualitative message of lemma 2.B changes when the fixed cost is per-group (is borne ex-ante before the Legislators moves).

[10]Such features stand if judges can be bribed. Provided that a sufficiently high weight is put on money, judges accept bribes, Case-law remains unbiased and volatile, and the proposition's idea survives. Also, the Constitutional Table's decision is collusion and renegotiation proof, being the losing group's willingness to pay for a reform of the optimal institution always lower than that of the winning group. Finally, statutes become volatile but unbiased if selected in each period by a Legislator accepting bribes: yet, the same inequality identifies the optimal institution. The proofs of these and all the robustness contained in the following footnotes are available from the author upon request.

socially optimal long run rule but at the cost of volatility.[11] Both costs are a function of the extent of disagreement between the groups—i.e., the distance $(\bar{A} - \underline{A})^2$. Therefore, for $\pi \geq \bar{\bar{\pi}}$, whether Case or Statute law prevail is only due to the comparison between the overruling cost and the strength of democratic institutions. Given a less than infinite institutional relevance of *stare decisis*, a society who better fares in assuring the probity of her representatives will embrace Statute law.[12] A fortiori, when the cultural distance is small—$\pi < \bar{\bar{\pi}}$, Statute law always outperforms Case law.[13]

This pattern conflicts with the Fernandez and Ponzetto's (2008a) claim that Case law is evolutionary superior to Statute law. The deep reason is that they model Statute law as a process of exogenous lobbying where the active ones are random. In this case, statutes are volatile and eventually biased; in this scenario, the expected long run loss of welfare is bigger than the one under Case law provided that democracy is imperfect and *stare decisis* is sufficiently relevant. Yet, a similar set up not only comes very short in explaining which coalition is willing to bargain with the Legislator but also what is the role of cultural heterogeneity.

Next, I consider a series of extensions to the basic model. First, I characterize the equilibrium of the lobbying game in a more articulated world where there is also a third cultural group with unbiased preferences for A^*. Then, I explore whether the idea laid

---

[11]With fixed overruling costs—as in Gennaioli and Shleifer (2007a), there is a π^* such that for $\pi \leq \pi^*$ (for $\pi > \pi^*$) the rule fixed by the first judge sticks forever (judges with opposite tastes overrule each other) and the loss of welfare is $(1/8)(\bar{A} - \underline{A})^2$. In this scenario, Case law is always sub-optimal for $K < \infty$.

[12]Costly lobbying only reinforces the main result, given that statutes become even less efficient when the Legislator is bribed at a high level of heterogeneity and/or at relatively—to K—low level of democracy.

[13]When citizen-candidates run to become the Legislator, the section's qualitative features stand, provided that the winner is not too insulated from the opposing party—i.e., she puts a not too small weight on her loss of welfare. The two candidates gather the same number of votes and have the same probability of winning. Again statutes perform better if no bribe is paid or if μ is relatively small with respect to K.

down in proposition 1 is upset if: 1. society trades off cultural biases with efficiency concerns; 2. appellate judges are forward-looking or (3.) can introduce new relevant information distinguishing the precedent. The message coming from these very different exercises is that the main qualitative result of the model—i.e., for fixed K, relatively more heterogeneous and/or less democratic societies should choose Case law—remains true.

II.B   Robustness

Introducing an unbiased group.—The population is equally split among the unconcerned group with $\beta_L = \lambda \pi$, the concerned one with $\beta_H = \lambda/\pi$ and a third unbiased group with $\beta_E = \lambda$. Again, the Legislator chooses an extremist coalition whenever π is sufficiently high. Yet, the presence of the unbiased group breaks the Legislator's indifference and:

Lemma 3: *Let u be unobservable.* **1.** *For any $l \in \Omega$, there exist a unique optimal $\hat{A}_s(l)$ defined by (6). For $\lambda \leq 1$ ($\lambda > 1$), the Legislator chooses $\{L\}$ ($\{H\}$) and there are levels of cultural heterogeneity $\tilde{\pi}_{\tilde{l}}$ and $\tilde{\tilde{\pi}}_{\tilde{l}}$ for $\tilde{l} = \{\{L\}, \{H\}\}$ falling with μ and rising with Ψ such that: $\tilde{\tilde{\pi}}_{\tilde{l}} \geq \tilde{\pi}_{\tilde{l}}$; for $\pi \geq \tilde{\tilde{\pi}}_L$ ($\pi \geq \tilde{\tilde{\pi}}_H$) the Legislator chooses $\{L\}$ ($\{H\}$), and for $\pi < \tilde{\tilde{\pi}}_L$ ($\pi < \tilde{\tilde{\pi}}_H$), she selects $\tilde{A}$; for $\pi \geq \tilde{\pi}_L$, $\{L\}$ (for $\pi \geq \tilde{\pi}_H$, $\{H\}$) pays Ψ, and for $\pi < \tilde{\pi}_L$ (for $\pi < \tilde{\pi}_H$) she does not. No coalition pays a fixed charge $\Psi > \tilde{\Psi} = 1/27$.* **2.** *Thus, for $\mu > 0$ and $\Psi \leq \tilde{\Psi}$, from $t = 1$ on the loss of social welfare equals 0 for $\pi < \tilde{\tilde{\pi}}_L$ ($\pi < \tilde{\tilde{\pi}}_H$) and $(\mu^2/18)\left(A^* + \underline{A} - 2\bar{A}\right)^2$ ($(\mu^2/18)\left(\bar{A} + A^* - 2\underline{A}\right)^2$) for $\pi \geq \tilde{\tilde{\pi}}_L$ and $\lambda \leq 1$ ($\pi \geq \tilde{\tilde{\pi}}_H$ and $\lambda > 1$).*[14]

Thus, if A^* is stricter than the second best, which happens for $\lambda \leq 1$, the Legislator favors $\{L\}$; when, instead, $\lambda > 1$ and so $\tilde{A} = E\left(\hat{A}_i\right) < A^*$, coalition $\{H\}$ prevails. Again, the

---

[14]For $\lambda = 1$ the Legislator is supposed to incline for $\{L\}$. The hypothesis is without loss of generality.

welfare comparison between the two institutions is immediate. When $K < \infty$ Case law can outperform Statute law if and only if $\pi \geq \tilde{\tilde{\pi}}_{\tilde{l}}$ and:

$$\tilde{\Lambda}\left(\hat{A}_s\right) > \mathrm{E}\left(\tilde{\Lambda}\left(\hat{A}_c\right)\right) = \left[3\left(\bar{A}\right)^2 + 3\left(A^*\right)^2 + 3\left(\underline{A}\right)^2 - \left(\bar{A} + A^* + \underline{A}\right)^2\right]\left[18\left(1 + 2K\right)\right]^{-1}.$$

This last inequality leads directly to the following proposition:

Proposition 2: *Let u be unobservable. Statute law is always strictly better than Case law for $\pi < \tilde{\tilde{\pi}}_{\tilde{l}}$ and $K < \infty$, and the two institutions induce the same expected welfare for $\pi < \tilde{\tilde{\pi}}_{\tilde{l}}$ and $K = \infty$. For $\pi \geq \tilde{\tilde{\pi}}_{\tilde{l}}$ and for every K, instead, there is a $\tilde{\mu}_{\tilde{l}}\left(K, \pi\right)$ with $\partial\tilde{\mu}_{\tilde{l}}/\partial K < 0$, $\partial\tilde{\mu}_{\tilde{l}}/\partial\pi < 0$, $\tilde{\mu}_{\tilde{l}}\left(\infty, \pi\right) = 0$ and $\tilde{\mu}_{\tilde{l}}\left(\bar{K}_{\tilde{l}}, \pi\right) = 1$ for $\bar{K}_{\tilde{l}} > 0$, such that Statute law performs better (weakly worse) than Case law for every $\mu < \tilde{\mu}_{\tilde{l}}$ (otherwise).*

Differently from proposition 1, cultural heterogeneity has also an intensive margin: as it increases the autocracy threshold over which Case law outperforms Statute law. As a consequence, the main model's message remains true even for $\Psi = 0$. The rest of the paper considers, for simplicity, the two groups case; however, all the results hold true also when the unbiased group is introduced.

Accounting for tastes versus efficiency.—Let me now consider a world in which technological efficiency is weighted against preferences aggregation, in such a way that group i's losses are $\left[(1 - T)/2\right]\hat{A}_i\left(1 - A^E\right)^2 + \left[(1 - T)/2\right]\left(1 - \hat{A}_i\right)\left(A^E\right)^2 + (T/2)\left(A^E - A^*\right)^2$ where T is the common relative technological concern in the efficiency—notice the apex E—regime. Group i's bliss point is now $(1 - T)\hat{A}_i + TA^*$. Consequently, $\Lambda_i^E\left(A^E\right)$ is proportional to $(1/2)\left(A^E - \hat{A}_i^E\right)^2$ and the optimal second best rule is a compromise between each group's bliss point and A^*. Under Case law, judge i's objective function at time t is $-(1/2)\left(A^E - \hat{A}_i^E\right)^2 - (K/2)\left(A_{c,t}^E - A_{c,t-1}^E\right)^2$. The unique and global solution is $\Gamma\left(K\right)A_{c,t-1}^E + \left[1 - \Gamma\left(K\right)\right]\hat{A}_i^E$, and judge-made law follows an auto-regressive process

16

converging to the ergodic distribution $N\left(E\left(\hat{A}_i^E\right), V\left(\hat{A}_i\right)(1-T)^2(1+2K)^{-1}\right)$. So the

loss of social welfare is proportional to $(1-T)^2 V\left(\hat{A}_i\right)[2(1+2K)]^{-1}$. The Legislator,

instead, selects $\hat{A}_s^E = \left[(1-\mu)\tilde{A}^E + \mu\sum_{i\in l}\hat{A}_i^E\right](1-\mu+|l|\mu)^{-1}$ and the same patterns

described by lemma 2.B when $\tilde{A}$ and $\hat{A}_i$ are replaced by $\tilde{A}^E$ and $\hat{A}_i^E$ respectively arise—see

appendix. For $K < \infty$, Case law outperforms Statute law if only if $\pi \geq \bar{\bar{\pi}}_E$ and

$$\mu^2(1-T)^2\left(\bar{A}-\underline{A}\right)^2 > (1+2K)^{-1}(1-T)^2\left(\bar{A}-\underline{A}\right)^2 \leftrightarrow \mu^2 > (1+2K)^{-1}.$$

Thus, proposition 1 applies unchanged, and the evolution of the legal system will again

be driven by the preference heterogeneity and the quality of the political process only,

provided that the population has the same relative concerns for technology. Thus, when

we look at the impact of legal institutions on economic outcomes, if the measure of devel-

opment we employ account for both tastes and technological efficiency, we should find no

significant difference among countries that have reached their preferred legal tradition—

i.e. that did not remain trapped due to high switching costs with the tradition that they

received with transplantation: see section IV.A below.[15]

Forward-looking judges.—As stressed by Gennaioli and Shleifer (2007b), another source

of judicial moderation is given by the forward looking concerns of appellate judges. To

reduce the possibility of future overruling, appellate judges may choose to set a less biased

precedent today. Such concerns produce patterns similar to those illustrated in proposi-

tion 1 in the same environment studied by Gennaioli and Shleifer (2007b). In particular,

let me assume that there are only two period, appellate judges discount the future by a

---

[15]Yet, the existing literature employ measures of development capturing the impact of the law on tech-
nological efficiency only—i.e., a function of $\left(A^E - A^*\right)^2$. In this case, at low level of π Statute law
outperforms Case law because it produces rules that are certain and unbiased. If instead π is sufficiently
high and only the long run bias of the law (both the long run bias of the law and her volatility) affects the
dimension studied—e.g. regulation of entry (creditor rights), Case (Statute) law is welfare maximizing
because unbiased (provided that the quality of institutions is sufficiently high because certain).

factor $\delta \leq 1$ and bear a fixed cost of overruling equal to $\tilde{K}$. In the second period judge i overrules the precedent replacing it with his desired rule when $\left(\hat{A}_{c,1} - \hat{A}_i\right)^2 > 2\tilde{K}$. As a result, if $-i$ is the group with preference opposite to i, the first period judge maximizes:

$$-\left(1/2\right)\left(A_{c,1}\left(i\right) - \hat{A}_i\right)^2 - \delta\left[\left(1/2\right)\left(A_{c,2}\left(i\right) - \hat{A}_i\right)^2 + \left(1/2\right)\left(A_{c,2}\left(-i\right) - \hat{A}_i\right)^2\right],$$

where the dependence of second period rule from the identity of the second period appellate judge is made explicit. Forward-looking concerns can induce a first period compromise that also a judge with similar preferences will overrule. The appendix shows that, under mild regularities on $\tilde{K}$, this is the only equilibrium when π is sufficiently high. Indeed:

Lemma 4: *Let u be unobservable and $\tilde{K} < 1/8$.* **1.** *There exists $\underline{\hat{\underline{\pi}}}$, $\hat{\underline{\pi}}$, $\hat{\bar{\pi}}$ such that: $\underline{\hat{\underline{\pi}}} < \hat{\underline{\pi}} < \hat{\bar{\pi}}$; for $\pi < \underline{\hat{\underline{\pi}}}$ $\left(\hat{\underline{\pi}} < \pi < \hat{\bar{\pi}}\right)$ the first period judge selects $\hat{A}_{c,1} = \hat{A}_i$ which is not (is) overruled (by a judge of type $-i$ who chooses $\hat{A}_{-i}$); for $\underline{\hat{\underline{\pi}}} \leq \pi \leq \hat{\underline{\pi}}$ $\left(\pi \geq \hat{\bar{\pi}}\right)$ a type L judge selects $\underline{A} + \sqrt{2\tilde{K}}$ and a type H selects $\bar{A} - \sqrt{2\tilde{K}}$ which is not (is) overruled (by a judge of type i who chooses $\hat{A}_i$).* **2.** *The second period loss of social welfare is $\left(1/8\right)\left(\bar{A} - \underline{A}\right)^2$ for $\pi < \underline{\hat{\underline{\pi}}}$ and $\hat{\underline{\pi}} < \pi < \hat{\bar{\pi}}$, $\left(1/2\right)\left[\left(1/2\right)\left(\bar{A} - \underline{A}\right) - \sqrt{2\tilde{K}}\right]^2$ for $\underline{\hat{\underline{\pi}}} \leq \pi \leq \hat{\underline{\pi}}$, and $\left(1/16\right)\left(\bar{A} - \underline{A}\right)^2 + \left(1/4\right)\left[\left(1/2\right)\left(\bar{A} - \underline{A}\right) - \sqrt{2\tilde{K}}\right]^2$ for $\pi \geq \hat{\bar{\pi}}$.* For $\pi \geq \bar{\bar{\pi}}$, the second period loss of welfare under Statute law is $\left(\mu^2/8\right)\left(\bar{A} - \underline{A}\right)^2$ and:

Proposition 3: *Let u be unobservable and $\tilde{K} < 1/8$. Statute law is weakly better than Case law for $\pi < \min\left\{\bar{\bar{\pi}}, \underline{\hat{\underline{\pi}}}\right\}$. For $\pi \geq \min\left\{\bar{\bar{\pi}}, \underline{\hat{\underline{\pi}}}\right\}$, there is a $\hat{\mu}$ with $\partial\hat{\mu}\big/\partial\tilde{K} \geq 0$ such that Statute law outperforms (performs weakly worse than) Case law for $\mu < \hat{\mu}$ (otherwise).*

Thus, even when appellate judges are forward-looking the model's main result holds.

Distinguishing.—*Stare decisis* requires judges to abide by the holding of the first court, but still allows them to distinguish the precedent by introducing a new dimension b into consideration. "A judge can claim that the previous rule applies to an arbitrarily small

portion of the infinitely larger space that he is now mapping: he thereby "reduces the range of the binding *ratio decidendi* to a vanishing point" (Stone 1964)."

This ability to create "complex balancing tests based on marginal trade-offs between different factors" (Fernandez and Ponzetto, 2008b) can be easily captured in the model of section II.A supposing, as Fernandez and Ponzetto (2008b) do, that eventually the second dimension u becomes observable and appellate judges can implement the optimal two-dimensional rule $B = 1 - a$ at a fixed cost $\tilde{K}^I > 0$—where I stands for information.[16] In order to highlight in the clearest way the effect of this flexibility advantage of Case law, I shut down overruling assuming that $K = 0$. Lemma 5 summarizes my findings:

Lemma 5: *Let $K = 0$ and suppose that u becomes observable, then for $\tilde{K}^I < 1/2$ there exists a π^I increasing with $\tilde{K}^I$ such that for $\pi \geq \pi^I$ the optimal two-dimensional rule $B = 1 - a$ is introduced, and the first best prevails thereafter.*[17]

Conflicts and overruling persist so long as ignorance on u does; if the latter becomes observable all judges agree on the reduction at zero of the errors if the cultural heterogeneity is sufficiently high to justify the payment of the fixed charge. As a consequence, again the main model's main feature stands. In point of fact:

Proposition 4: *Let $K = 0$ and $\tilde{K}^I < 1/2$, and suppose that u becomes observable, for $\pi < \pi^I$ ($\pi \geq \pi^I$) Statute law outperforms Case law (the opposite is true).*

---

[16]I obtain similar results when, as in Gennaioli and Shleifer (2007a), distinguishing judges choose two levels of b respecting the precedent in the sense that O is held liable if $a < \hat{A}_c$ and $b \geq \underline{B}$ or $a \geq \hat{A}_c$ and $b \geq \bar{B}$.

[17]It is worth to notice that all the analysis holds true when appellate judges' objective function is instead $-\Lambda_i^E \left(A_{c,t}^E \right) - (K/2) \left(A_{c,t}^E - A_{c,t-1}^E \right)^2$. With these preferences, in order to have lemma 5 holding I also need that $\tilde{K}^I$ is not too small in such a way that the judges' willingness to pay $\tilde{K}^I$ is rising with π.

III.   Rules versus Discretion in Adjudication

Scholars in comparative law (Merryman, 1969; Damaška, 1986; Zweigert and Kötz, 1998; Glendon, Gordon, and Carozza, 1999) see the reliance on broad principles rather than specific "bright line" rules of adjudication, as a defining feature of the Common as opposed to the Civil law tradition. In section II, I solved the model under the maintained assumption that bright line rules of adjudication were always in place—i.e., assuming that the rule selected by the law-maker is strictly respected by lower adjudicating courts. In general, instead, under more flexible principles of adjudication—*discretion* hereafter, lower courts have the right to adjudicate a case according to a *ratio-decidendi* different from the one chosen by the law-maker by, for instance, manipulating the evidence (Gennaioli and Shleifer, 2008). In this section, I show that this procedurally accepted form of discretion emerges as an efficiency attempt to blunt the losses driven by the volatility of Case law.

I consider a mild form of discretion. Bearing a cost $\theta > (\alpha/2)\left(\bar{A} - \underline{A} - \alpha\right) + \varepsilon$ with $\varepsilon \geq 0$ very small, lower courts can use any *de facto* $A_{j,t}^{F}$ with $j = \{c, s\}$ whose distance α from the *de jure* $A_{j,t}^{J}$ is lower than $\bar{\alpha} \equiv \left(\bar{A} - \underline{A}\right) \min\left\{1 - (1 + 2K)^{-1/2}, 2K\left(1 + 3K + 4K^2\right)^{-1}\right\}$. The lower courts' objective function is $(1/2)\left(A_{j,t}^{F} - \hat{A}_{i}\right)^{2}$ and they are randomly picked from the population. This assures that a H (L) lower court uses $\hat{A}_{j,t}^{J} - \alpha$ $(\hat{A}_{j,t}^{J} + \alpha)$ if $\hat{A}_{j,t}^{J} > \tilde{A} = E\left(\hat{A}_{i}\right)$ $(\hat{A}_{j,t}^{J} < \tilde{A})$ and sticks to the *de jure* rule $\hat{A}_{j,t}^{J}$ if $\hat{A}_{j,t}^{J} \leq \tilde{A}$ $(\hat{A}_{j,t}^{J} \geq \tilde{A})$—see the appendix—with $\hat{A}_{j,t}^{J}$ equal to the $\hat{A}_{j,t}$ found in section II. This time, the loss of social welfare is proportional to $(1/2)\,\mathrm{E}\left(A_{j,t}^{F} - \hat{A}_{i}\right)^{2}$ where A_{j}^{F} is the long run *de facto* rule. Introducing discretion entails a cost $\Theta > 0$.

First, let me characterize the equilibrium under Case law. I posit that Case law is

in its steady state and, in each period, first an appellate judge fixes a precedent and then a lower court can use discretion. A type i appellate judge not only bears the usual justification cost, but also cares about the expected distance between the *de facto* rule in period t and the bliss point of her own cultural group $\hat{A}_i$; thus, she maximizes:

$$-\left(1/2\right)\mathrm{E}\left(A_c^F\left(i\right)-\hat{A}_i\right)^2-\left(K/2\right)\left(A_{c,t}^J\left(i\right)-\hat{A}_{c,t-1}^J\right)^2,$$

where $\mathrm{E}\left(A_c^F\left(i\right)-\hat{A}_i\right)^2=\mathrm{V}\left(A_c^F\left(i\right)\right)-\left[\mathrm{E}\left(A_c^F\left(i\right)\right)-\hat{A}_i\right]^2$. As shown in the appendix, under the hypotheses on α, the prevailing *de jure* rule is: $\hat{A}_{c,t}^J\left(i\right)=\Gamma\left(K\right)A_{c,t-1}^J+\left[1-\Gamma\left(K\right)\right]\left(\hat{A}_i\left(i\right)+\tilde{\alpha}/2\right)$, with $\tilde{\alpha}$ equal to α ($-\alpha$) when $i=L$ ($i=H$). As a result, it follows that $\hat{A}_{c,t}^J\left(L\right)>E\left(\hat{A}_i\right)$, $\hat{A}_{c,t}^J\left(H\right)<E\left(\hat{A}_i\right)$ and that the *de jure* rule converges to the ergodic distribution with mean $E\left(\hat{A}_i\right)$ and variance:

$$\left[V\left(\hat{A}_i\right)+\alpha^2/4+\left(\bar{A}-\underline{A}\right)\left(\alpha/2\right)\right]\left[\left(1+2K\right)\right]^{-1}.$$

Each judge moves $\hat{A}_{c,t}^J\left(i\right)$ nearer to her own bias in order to counteract the action of a lower court with opposite bias. Yet, discretion more than compensates this strategic bias, and $A_c^F\left(i\right)$ equals $\hat{A}_{c,t}^J\left(i\right)$ if the lower court's type is i and $\hat{A}_{c,t}^J\left(i\right)-\tilde{\alpha}$ if the lower court's type is $-i$. This not only implies that the *de facto* rule is unbiased, being $\mathrm{E}\left(\hat{A}_{c,t}^F\right)=E\left(\hat{A}_{c,t}^J\right)+\left(1/2\right)\Pr\left[\hat{A}_{c,t}^J>E\left(\hat{A}_i\right)\right]\left(-\alpha\right)+\left(1/2\right)\Pr\left[\hat{A}_{c,t}^J<E\left(\hat{A}_i\right)\right]\alpha=E\left(\hat{A}_i\right)$, but also that its variance is strictly lower than $V\left(\hat{A}_i\right)\left[\left(1+2K\right)\right]^{-1}$. This means that, provided that the extent of discretion is limited—$\alpha\leq\bar{\alpha}$, the volatility of the prevailing rule under the pure Common law tradition—i.e., Case law plus discretion in adjudication—is lower than the one under the mixed system where bright line rules are used. Discretion will be introduced if $\Theta<\bar{\Theta}$—with $\bar{\Theta}$ found in the appendix, while for $\Theta>\bar{\Theta}$, the hybrid tradition prevails. Germany, Switzerland and the Scandinavian countries are examples of this tradition (see footnote 20 and Zweigert and Kötz, [1998]).

Consider now Statute law, the objective function in (4) rewrites as:

$$-\left[(1-\mu)/2\right]\left(A_s^J + \tilde{\alpha}/2 - \tilde{A}\right)^2 + (\mu/2)\sum_{i\in l}\left[\left(\tilde{A}-\hat{A}_i\right)^2 - \left(A_s^J + \tilde{\alpha}/2 - \hat{A}_i\right)^2\right] - 2\mu\Psi.$$

Given the restrictions on α, the Legislator is always able to foresee the lower courts' actions and neutralize—in expectation—the use of discretion choosing the threshold

$$\hat{A}_s^J(l) = \left[(1-\mu)\,\tilde{A} + \mu\sum_{i\in l}\hat{A}_i\right](1-\mu+|l|\,\mu)^{-1} - \tilde{\alpha}/2.$$

The expected *de facto* rule $A_s^J + \tilde{\alpha}/2$ under the pure Civil law tradition equals the *de jure* rule prevailing when discretion is forbidden whether or not bribes are paid. Also, if biased, *de facto* statutes are also volatile with variance $\alpha^2/2$. Thus, the pure Civil law system is retained for every $\Theta > 0$. The following proposition summarizes my findings:

Proposition 5: *Let u be unobservable, and $\alpha < \bar{\alpha}$. The pure Civil law tradition is always optimal. If $\Theta \leq \bar{\Theta}$ ($\Theta > \bar{\Theta}$) the pure (hybrid) Common law tradition prevails.*[18]

From theory to evidence.—Scholars in comparative law have classically considered the institutional relevance of precedents as a crystallized by-product of the origins' legal writings and High Courts' jurisprudence. The British model of precedent is a case in point. At the beginning of the nineteenth century, the runaway of legal positivism over legal naturalism built the consensus position that, in overruling or distinguishing a precedent, appellate courts should always consider the hierarchical rank of the court which set it, whether or not the precedent was chosen *per incuriam*, and the precedent's age (see Bankowski, MacCormick, and Marshall, [1997]: pp. 315-353). These precise and stable factors have guided the judicial law-making activities of those countries to which the English Case law was transplanted or that adopted it after independence (Abdesselem et al., 1995: U-107,

---

[18]Notice that proposition 1 continues to hold. If α is sufficiently high, the volatility of Case law becomes higher under the pure system; besides, the heterogeneity threshold under which statutes prevail rises with α leading to a supremacy of the mixed Civil law system when the extent of discretion is wide enough. Yet, this uncontrolled use of discretion does not match the procedurally accepted level studied here.

U-108).[19] The German "free law" movement had a similar impact on other non British models of precedent (footnote 20; Grechenig and Gelter, 2007). Thus, this wide legacy of legal and historical evidence suggests that K can be considered as country-invariant.

Common law originated from the laws of England and was transplanted through conquest or colonization to England's ex-colonies—e.g., U.S.A., Australia, Canada, and several countries in Africa and Asia. Instead, Statute and in general Civil law, as embedded in the French "Code Civil des Française", brought into force in 1804, and the related Austrian (1811), Portuguese (1867), Spanish (1889) and Soviet Union (1922) Civil codes has been exported first to Latin America, Africa and Asia and then to the former Sovietic republics (see Zweigert and Kötz, [1998]). Differently from these two groups, Denmark, Germany, Sweden and Switzerland have approved substantive pieces of codification entrusting, at the same time, a key role to judge-made law as a key instrument for the maintenance of those parts of the statutes becoming with time obscure and/or inadequate.[20] This last experience was especially relevant for some countries—e.g., Ethiopia and Thailand—which, after independence—mostly in the aftermath of the Second World War, modelled their legal system on the shape of these European examples.

Starting from the historical features of the transplantation wave, Berkowitz, Pistor, and Richard (2003) have introduced the distinction between countries that are *origins*—

---

[19]This homogenization was institutionalized in 1833 when the Judicial Committee of the Privy Council was entrusted with the power of unifying the action of the Commonwealth appellate courts. Even if abolished by eight countries in my sample, the Committee's decisions maintain a highly recognized value (Abdesselem et al., 1995). Considering only the ex-English colonies as using Case law or excluding the U.S.A. which use slightly different distinguishing procedures does not affect the main results.

[20]I coded these countries as using Case law, yet embracing a different strategy does not affect the main results. Germany and Switzerland are two controversial cases. Yet, "the maintenance of the general structure of the BGB [German Civil code, 1900] is really the work of *courts*" (Zweigert and Kötz, 1998: pp. 153). Courts have relied on the BGB general clauses—number 157, 242 and 826—to adequate the codes to social and technological innovations. Similar rules were enhanced in the Swiss Civil Code (1881).

because they have developed their legal order internally—and countries that are *trans-planted*—because they have received their legal order either through colonization or by adoption of the legal tradition considered most advanced at the time. At the same time, they have stressed the essential randomness of the initial matching between legal traditions and national legal systems at independence in the transplanted countries. Yet, from independence on countries penalized by this match—e.g., culturally heterogeneous ex-Spanish colonies plagued by inefficient institutions—had all the incentives to reform their institutions toward their ideal legal order—see, for instance, footnote 23: the empirical exercise takes exactly advantage of this unique historical experiment. Indeed, provided that the restrictions on the exogenous parameters hold and with K invariant, the model produces the following predictions on the likelihood of legal reforms in transplanted countries:

Testable Predictions: *The likelihood of a reform toward Case and, in general, Common (Statute and, in general, Civil) law in countries to which Civil (Common) law was transplanted will increase (decrease) with the degree of cultural heterogeneity and fall (rise) with the quality of democratic institutions.*

IV.   Civil Law versus Common Law: Evidence

Next, I will test these predictions, comparing first Statute law and Case law.

IV.A   Data and Strategy

I require both a sample of reforms from one to the other law making institution and proxies for the extent of cultural heterogeneity and the quality of political institutions.

Law-making reforms.—In order to accomplish the first task, I made use of two data

sources: 1. the first volume of the *International Encyclopedia of Comparative Law* written by some of the most illustrious scholars of comparative law of the last century—Conard, David, Von Mehren, and Zweigert—between 1969 and 1995, and listed among the references as Abdesselem et al. (1995); 2. the appendix accompanying the Djankov et al.'s (2003) dataset on legal procedures, which is listed among the references as Acartürk et al. (2005). I considered as *origins* Austria, Denmark, France, Germany, Russia, Sweden, Switzerland and, of course, the United Kingdom and, as transplanted, the 156 countries studied in the above mentioned publications for which enough details on the history of law-making institutions are reported.[21] I observe the law-making process in place in each transplanted country at independence and in 2000, distinguishing countries using Case law or Statute law according to the following criterion:[22]

Definition: *A country uses Case law, in a given year, if the decisions of a subset of appellate courts but not only of the Constitutional court are considered as a source of law and treated as binding by lower courts. Otherwise, a country employs Statute law.*

Next, I can define *Statute_Ind* (*Statute_2000*) as the dummy variable which equals one if a country used Statute law in the first year of independence (in 2000). Each country switches at most once. Thus, the long run evolution of the law-making institutions in each country can be summarized by the indicator *Law_Making* which equals four if *Statute_Ind* equals one and *Statute_2000* zero, one if the opposite is true, three if both *Statute_Ind* and *Statute_2000* equal one, and two if the two dummies both equal zero. There are

---

[21]When I restrict the sample to the ex-colonies, the main results remain unaffected. The timing of each reform is unobserved. Identifying the latter is a relevant agenda for future research.

[22]A more (less) restrictive criterion is to consider as using Statute (Case) law also the countries in which only the decisions of a subset of appellate courts (Constitutional-Cassation court) are treated as binding. If I switch to one of these two alternatives, the main qualitative pattern of the empirical exercise stand.

ten countries that switched from Statute law to Case law, sixteen that reformed in the opposite direction, and 60 (70) countries that stick to Case (Statute) law. Switching the two for the three in the definition of *Law_Making* does not affect the main results.[23]

Proxying cultural heterogeneity and democracy.—Testing the model's predictions also requires finding fairly exogenous measures of both the extent of cultural heterogeneity and the strength of democracy available for the whole sample.

With regard to the former, I make use of two variables. The first captures the cultural distance between the principal population in the country which chose the law-making institution and the one in the transplanted country. The second proxy measures the cultural fractionalization within the transplanted country.[24] The first of the two dimensions helps in clarifying the impact on the evolution of the legal system of the initial inefficiency of the transplantation match and, in general, of the preference polarization when one group can completely exploit the other—as in footnote 11 and lemma 4 where Case law does not converge to optimality.

Yet, creating a meaningful proxy for the cultural distance between populations is not an easy task. Cavalli-Sforza, Menozzi, and Piazza (1994)—CMP, hereafter—suggest that an index of genetic distance, the coancestry coefficient—*Genetic_Distance*, is a natural proxy. CMP justify this claim noticing the strong link between population's linguistic roots and *Genetic_Distance*. "Both genetic and cultural contacts [...] take place by the same routes [e.g., marriage]; they respond to the same geographic and ecological barriers;

---

[23]A key example is Mexico that inherited Statute law from the Spanish colonizers but then introduced a set of "compulsory case law" rules (Abdesselem et al., 1995: pp. M-68) because of the chronic inefficiency of democratic institutions (Gonzales, 2002) emphasized by the Mexican revolution (1910-1917).

[24]The main results are essentially unaffected if I use the distance between the principal population in the transplanting country and that in the *transplanted* one. More intuitive proxies like the polarization along the political spectrum or differences in value surveys' scores are available for a significantly smaller period.

and they can also reinforce each other" (CMP, pp. 23).[25] Besides, looking at European

nations formation, Desmet et al. (2008) have documented a strong relation between a

measure of opinion poll distances based on answers from the section on perceptions of

"life, family and religion and moral" of the World Value Survey and *Genetic_Distance*.

The coancestry coefficient is calculated as the sum of the differences in frequencies of ge-

netic or DNA polymorphisms—situations in which a gene or a DNA sequence exists in at

least two different forms or alleles.[26] Gene frequencies change over time mainly because of

migration, natural selection and genetic drift—i.e., random sampling at each generation.

Yet, describing the evolution of selection and the different migration waves is an essentially

unfruitful undertaking. To the latter extent, CMP sampled only genetic markers unaf-

fected by evolution from those aboriginal populations in place before 1500—i.e., before

colonization had fuelled the greatest migration episode in human history. Thus, not only

do higher values of *Genetic_Distance* imply wider genetic differences, but the coancestry

coefficient is both immune to measurement errors and exogenous to all those unobserved

historical patterns which affected cultural fractionalization and political institutions dur-

ing and after the transplantation experience (see also Spolaore and Wacziarg, [2009]).

Data on genetic distance are available for 42 populations groups. Following Spolaore and

Wacziarg (2009),[27] I matched populations to countries using the labels listed in the ap-

---

[25]Building on this intuition, Guiso, Sapienza, and Zingales (2009) use the coancestry coefficient as an instrument for bilateral trust in trade gravity regressions. Also, Spolaore and Wacziarg (2009) propose *Genetic_Distance* as a proxy of family-transmitted characteristics, including cultural and moral values.

[26]For example, even if the ABO blood group alleles are present in all populations, the frequency of each allele varies substantially across populations so that the estimated frequency of the O allele is 61% within Africans and 98% in American Natives. These differences arise for many other genes or DNA sequences.

[27]The resulting matching is equal to that used by Spolaore and Wacziarg (2009) except for a few entries—Egypt, Finland, Hungary, Kenya, Lybia, Malawi, Poland, Tanzania, Tunisia and Uganda—for which the histories of demography in the Encyclopedia Britannica suggest different classifications. CS provide a similar measure of genetic distance—Nei's distance—which delivers very similar results.

pendices two and three of CMP and the historical information reported in Encyclopedia Britannica (2008). The strong connection among genetic pool, socioeconomic distance, linguistic families and culture described above makes the ethnic and linguistic fractionalization in a country a natural proxy for the "within" heterogeneity. The most widely used measure of ethno-linguistic fractionalization is the probability that two individuals randomly chosen from the population differ in the characteristic under consideration. The results presented below are based on *Av_Elf*, which is built from 1960 data coming from a Soviet ethnographic source—the Atlas Narodov Mira (Apenchenko and Solomon, 1964).[28]

Turning to democracy, I will make use of a proxy for the strength of the constraints on the executive—*Constraint_Executive*—available from the Polity IV dataset for most of the countries of the sample starting from 1800.[29] *Constraint_Executive* is defined as the average constraints on the executive over the period elapsing between independence and the year 2000, ignoring missing data. This index is mainly related to the existence of checks and balances among different layers of the decision-making process and not on other Civil liberties, like the rule of law or the freedom of the press. The latter could be correlated with unobserved country-specific characteristics driving, for instance, also the population tolerance for different cultural preferences (see Acemoglu et al., [2008]).

Conditional independence.—Yet, as made clear by a recent literature on endogenous linguistic and ethnic group formation (see Ahlerup and Olsson, [2008]; Michalopoulos, [2008]), neither *Av_Elf* nor *Constraint_Executive* are immune to several other reverse causation and self selection biases. This line of research formalizes the intuitions proposed

---

[28]The qualitative results are in all similar when I use the linguistic, ethnic and religious fractionalization indexes proposed by Alesina et al. (2003) or the absolute number of ethnic or linguistic groups.

[29]Different measures of democracy available for the same period (see, for instance, Vanhanen, [2003]; Marshall and Jaggers, [2004]) deliver essentially similar results.

by two major theories of social evolution: the primordial and the constructive views. While the former contends that ethnicity is mostly a primordial by-product of extended kinship,[30] the latter has proposed two more recent factors: 1. the rise of the stratified society—i.e., the state and the nation with its formal legal order—deprived the extended kinships of their *raison d'être*, pushing for a more homogeneous citizenship; 2. the European colonizers, instead, pursued a "divide-and-rule" approach in order to increase the chance of successful exploitation of the colonies (Michalopoulos, 2008). Besides, this latter exploitation was pursued to different degrees depending on the adaptability of the colonizer to the climate and to pathogen loads (see Acemoglu, Johnson, and Robinson [2001]), which, in turn, affect the species richness as stressed by a large legacy of biology and ecology research (Ahlerup and Olsson, 2008).[31]

Therefore, it is not difficult to think of unobserved social and physical factors fostering, at the same time, the extent of unobserved heterogeneity—and, thus, driving also the reforms of the legal system—and either the preferences for ethnic homogenization or the quality of political process. In the following, I will first maintain the conditional independence assumption; then, I will relax it considering also other biological and geographical factors which most likely belong to the set of unobserved forces just discussed.

IV.B   Statute Law versus Case Law: What do the Data Say?

Empirical strategy.—I need a model for more than two non-ordered alternatives where the controls do not vary over alternatives. The natural choice is the multinomial logit

---

[30]Languages differentiate the members of the extended family from "the others" facilitating the provision of public good and the preservation of the group physical and human capital (Ahlerup and Olsson, 2008).

[31]Acemoglu et al. (2001) document that Europeans refrained from developing extractive societies, and set up efficient institutions protecting property rights only where they were more likely to successfully settle.

$$\Pr\left(y_i = \tilde{k} \,|z_i\right) = \exp\left(\beta_{\tilde{k}}' z_i\right) \Big/ \sum_{k=1}^{4} \exp\left(\beta_k' z_i\right), \tag{7}$$

where y_i is *Law_Making* and z_i gathers *Genetic_Distance*, *Av_Elf*, *Constraint_Executive*, the log of the real GDP per capita—*GDP*—and the population size—*Population*—averaged over the 1950-2000 period, and the number of years between independence and the year 2000—*Time_Independence*.[32] The latter should crudely proxy for lower switching costs or stronger preferences for homogenization. Table I (II) reports sources (summary statistics) of all the controls. *Genetic_Distance* and *Constraint_Executive* are normalized in order to range between zero and one.

Results under conditional independence.—Panel A of Table III lists the marginal effects obtained when z_i gathers only the cultural heterogeneity proxies. These figures give the percentage variation in the likelihood of the outcome considered when the control rises by one percentage point, and they are strongly consistent with the model's predictions. A one-standard-deviation rise in the normalized genetic distance—0.29—implies a little more than four—six in column (3)—percentage point increase in the likelihood of a reform from Statute to Case law and a little more than a five percentage point decrease in the likelihood of a reform from Case to Statute law. Also a one-standard-deviation rise in the ethno-linguistic fractionalization—0.30—is associated to a three percentage point increase—1.5 in column (3)—in the likelihood of a reform from Statute to Case law, and a six percentage point decrease in the likelihood of a reform from Case to Statute law. The effect of the "between" cultural heterogeneity is definitely stronger than that

---

[32]Using the mean of GDP and population over the period 1900-2000 does not affect the analysis.

of the "within" heterogeneity. Besides, only the former is always statistically significant at a level nowhere lower than 5%. Finally, a longer spell of political autonomy raises the likelihood of both reforms while *GDP* (*Population*) significantly decreases (increases) the probability of reform toward Statute law only. This last set of results is common to the remaining tables and therefore not reported.

Panel B adds to the picture *Constraint_Executive*. Here, the evidence is more mixed. Stricter constraints on the executive decrease the likelihood of both reforms. In column (4), a one-standard-deviation rise in the constraints on the executive—0.32—implies a little less than a seven percentage point fall in the likelihood of a reform from Statute to Case law and a seven percentage decrease in the likelihood of a reform from Case to Statute law. Yet, while the first effect tends to become more significant when cultural heterogeneity is taken into consideration, the second vanishes.

Relaxing conditional independence.—In order to relax conditional independence, I first require a discrete choice model with eventually endogenous covariates and, then, instruments related to the endogenous controls but not to the unobserved determinants of the evolution of the legal system. For what concerns the former, I will use the IV probit with dependent variable *Case_Law*, which equals one when *Law_Making* is equal to four and zero otherwise. Given the differences in the underlining model, the relative marginal effects cannot be compared with those produced by the multinomial logit.

More complex is to select reasonable instruments. The primordial theory of social evolution provides a natural instrument for *Av_Elf*. In particular, Michalopoulos (2007) argues that, starting with the Neolithic agricultural revolution in the Near East—10,500 BP, heterogeneous land endowments generated region specific human capital, limiting

population mobility and leading to the formation of localized ethnicities and languages. This intuition is empirically supported by cross-regional data. Thus, I use an index of richness of soil types—*Soil_Diversity*—proposed by Ahlerup and Olsson (2008), as an instrument for the within heterogeneity. The index aggregates qualitative measures of soil types unrelated to both human intervention and the average quality of the land: this assures the exogeneity of *Soil_Diversity* to unobserved driving also legal reforms.[33]

Turning to the quality of the political institutions, a recent literature informed by cross-cultural psychology studies (Licht, Goldschmidt, and Schwartz, 2007; Tabellini, 2008) has linked the preservation of the democratic accountability to the emphasis of the prevailing culture on autonomy. "Societies whose cultures emphasize individual uniqueness and view individual persons as moral equals are likely to develop norms that promote societal transparency as a means for social coordination [...]. In contrast, societies [that] view the individual as an embedded part of hierarchically organized groups [...] accommodate exercise of power from above" (Licht, Goldschmidt, and Schwartz, 2007). Languages are a key factor constraining the transmission of culture, and Kashima and Kashima (1998) suggest that a stable role in this diffusion mechanism is covered by the set of rules on pronouns. Languages that have several second person pronouns modulated according to the social distance among the speakers instill in the subject the respect for hierarchies—*Second_Person*. Thus, I use *Second_Person* as an instrument for the long run extent of democracy; given the very slow evolution of language traits (see CMP), this is a particularly appropriate choice.

---

[33]The characteristics from which the index is calculated refers to the texture and mineral components of the soil: i.e., to the quality in terms of erosion and water retention of a piece of land (Zobler, 1986).

Table IV reports in the lower panel the first stage results and in the upper panel the second stage IV probit marginal effects. The figures confirm the evidence about *Genetic_Distance* and *Constraint_Executive*, and also suggest that the weak significance of *Av_Elf* was driven by a failure of conditional independence. Indeed, the coefficient attached to *Av_Elf* is now significant in column (3) at 1 percent.

This pattern is robust to the introduction of other relevant covariates. Recent papers have proposed that a less unpredictable climate has a negative impact on species richness (Ahlerup and Olsson, 2008) and a positive effect on the quality of the political institutions (Acemoglu, Johnson, and Robinson, 2001). Column (4) lists the results when the maximum difference in land altitude—*Altitude_Diff*—is considered; the latter proxies for more unpredictable climate or higher soil diversity. In column (6), I also test the Glaeser and Shleifer's (2002) hypothesis adding a proxy for lower risk of coercion by powerful elites. The index *Economic_Power* assumes higher values the more diffuse is the economic power within the population (see Vanhanen, [2003]).[34] While the coefficient attached to *Altitude_Diff* shows the correct sign and is highly significant, the one attached to *Economic_Power* is significant but unexpected. The pattern could be driven by the fact that *Economic_Power* is linked to unobserved determinants of stronger political institutions; yet, this result seems to confirm the doubts on the Glaeser and Shleifer's (2002) analysis cast by a recent economic history literature which has been looking at the medieval experiences of both England and France (Klerman and Mahoney, 2007; Roe, 2007). Finally, the exogeneity test—which is nowhere lower than 0.89 whenever both *Av_Elf* and

---

[34]The evidence is in all similar when I use a two stage multinomial logit obtained bootstrapping a routine that substitutes the fitted values from the first stages to the endogenous in (7). Yet, this model is less stable than the IV probit due to the lack of sufficient variation in the data. Besides, interaction effects are usually not significant at a probability of the reform level of 0 and 0.5 (Ai and Norton, 2003).

Constraint_Executive are considered as endogenous—and the first stage R^2 confirm the quality of the chosen instruments. These covariates, moreover, enter into the first stages in a nice separable form whereby soil diversity affects only the within heterogeneity, and the lack of individualism only affects the strength of the constraints on the executive. Given the correlation between *Av_Elf* and *Constraint_Executive*—16%, this assures that the attenuation bias induced by the variable with the greatest measurement error does not load on to the other one (see Acemoglu and Johnson, [2005]). Since the first stages in this and subsequent tables are very similar, I do not report them to save space.[35]

IV.C   Civil Law versus Common Law: What do the Data Say?

Next, I will attack the second set of testable predictions concerning the convergence of the whole bundle of institutions characterizing the two traditions. I make use of two data sources: 1. Acartürk et al. (2005); 2. Several treaties, whose list is available from the Author, detailing the history of the adjudication procedures of the countries that transplanted the legal order into the nations studied in Acartürk et al. (2005). Djankov et al.'s (2003) analyze procedural rules governing the adjudication of simple legal disputes—the eviction of a non-paying tenant and the collection of a bounced check—for 109 countries in 2000. Among other data for procedural formalism, Djankov et al. (2003) collect the following four dummies which I use as proxies for the lack of discretion in the adjudication of civil disputes: 1. *Appeal_l&f_2000* equals one if issues of both law and fact can be reviewed in appeal and zero if only new evidence or issues of law can be reviewed,

---

[35]I obtain mirror results if the dependent equals 1 if *Law_Making* equals 1 and 0 otherwise. Using the difference in temperature (an index of intellectual power) in spite of *Altitude_Diff* (*Economic_Power*), or considering the rescaled latitude or a measure of malaria risk, the evidence remains quite similar.

or if there is not appeal; 2 *Judge_Law_2000* equals one if judgment must be on law only, and zero when they may be based on equity grounds; 3. *Inquisitorial_2000* equals one if the evidence gathering procedure can be considered inquisitorial and zero otherwise;[36] 4. *Written_2000* has value one if the evidence is mostly submitted to the court in written form—attachments, affidavits, or other—and zero otherwise.

Such rules are consistently (Merryman, 1969; Zweigert and Kötz, 1998; Glendon, Gordon, and Carozza, 1999) considered as discriminating among procedures typical of the hierarchical, inquisitorial and technical Civil law tradition and those of the coordinate, adversarial and substantive Common law tradition.[37] Similar indexes—*Appeal_l&f_Ind*, *Judge_Law_Ind*, *Inquisitorial_Ind*, *Written_Ind*—can be built for the first year of independence for 98 countries, starting from the history of the civil procedure governing similar commercial cases in the transplanting country. As in Balas et al. (2008), for each of the four indicator measured at independence and in 2000, I stack one over the other the observations for the eviction of a non-paying tenant and the collection of a bounced check in order to obtain eight indexes with two observations for each of the 98 countries for which data at independence are available. Next, I build four multiple outcomes indicators similar to *Law_Making*: *Appeal*, *Equity*, *Procedure*, *Evidence*. For instance *Appeal* equals four if *Appeal_l&f_Ind* is one and *Appeal_l&f_2000* is zero; one if the opposite is true; three if both *Appeal_l&f_Ind* and *Appeal_l&f_2000* equal one; and two if both dummies equal

---

[36]I consider a procedure to be inquisitorial if judges can freely request or take evidence that has not been introduced by the parties, and refuse to collect or admit requested evidence (see Damaška, [1986]).

[37]The link with the degree of discretion by lower adjudicating courts can be summarized as follows. While extensive supreme review and written records assure a credible check on discretionary lower courts; inquisitorial procedural requirements curbs the parties' incentive to report noise and contrasting evidence which calls, in turn, for interpretation of the law (see Damaška, [1986]: pp. 3-6). Finally, an explicit limitation to equitable judgments is of similar use (see Merryman, [1969]: pp. 123-127).

zero. Two crucial background hypotheses make this exercise more fragile than the one discussed above: 1. rules in place in a transplanted country at independence are fairly well approximated by those in place at the same time in the transplanting country; 2. the procedure for small commercial cases at independence and the rules governing the eviction of a non-paying tenant and the collection of a bounced check in 2000 are good proxies of the underlying "discretion in adjudication" concept.

Recovering legal traditions.—The analysis of these sub-indices generally point in the same direction. Yet, as Rosenthal and Voeten (2007) point out, this is not a convincing argument in favor of the approach usually embraced by the "legal origins" literature. Similar works by Djankov et al. (2003) and Balas et al. (2008) have chosen as study objective the sum of all the indexes. Such a methodology relies on some untested assumptions which could be defended only if we had particularly strong a priori about the relevance of each indicator. Indeed, the methodology assumes that each indicator is equally informative about the underlying concept, that the employed indicators measure a single dimensional concept and, finally, that measurement error in the observations is irrelevant (Rosenthal and Voeten, 2007). Yet, while several scholars emphasize *Appeal*, *Equity* and *Evidence* (see Merryman, [1969]: pp. 52, pp. 123-127; Glendon, Gordon, and Carozza, [1999]: pp. 127) as differentiating the two traditions, others (Damaška, 1986; Zweigert and Kötz, [1998]: pp. 271-272) stress that the defining feature of the adjudication procedures of each tradition is their adversarial or inquisitorial nature.

To the extent of solving the diatribe, I follow Rosenthal and Voeten (2007), and I use factor analysis tools in order to model the relationship among *Law_Making*, *Appeal*, *Equity*, *Procedure* and *Evidence* and the latent "convergence" construct. Let $j = 1, ..., J$ denote

countries and $i = 1, ..., I$ summarize the evolution of the considered characteristics of the legal system. Each observation $x_{i,j}$ is the value for country j on aspect i of its history of the legal system. If, as scholars of comparative law argue, the degree of convergence toward a particular tradition was the defining characteristic underlying these individual qualities, then we have I imperfect indicators of a single unobserved variable. More generally, the indicators arise probabilistically from a set of $\tilde{I}$ unobserved fundamental historical aspects with $\tilde{I}$ smaller than I. The analysis is implemented mapping the observed matrix of indicators $\mathbf{X}$ into an $I \times J$ matrix of latent variables $\mathbf{X}^*$ with $x_{i,j}^*$ representing the log odds of country j evolving in the characteristic and in the sense described by i. Next, I can define the two-parameter multinomial logit item response model:

$$x_{i,j}^* = \beta_i + \eta_j \lambda_i,$$

where $-\beta_i$ represents the difficulty parameters, the dummy or "factor" η_j represents the country specific latent construct and λ_i the extent to which aspect i captures the underlining construct. The factor loading for *Law_Making* is constrained to equal 1 in order to identify the scale of the factor.

Table V reports both coefficients and standard errors of the factor loadings and the likelihood ratio test that their values do not differ significantly from 1 for a specification including or not *Evidence*. The main message is that the latter does not discriminate between Common law and Civil law countries while all the other indicators are all relevant and equally informative about the underlying "convergence" construct. This result confirms the key relevance of a preliminary study of the relative importance of each of the structures that sustain an institutional architecture—in this case, the legal tradition. Next I define the index *Convergence* as the sum of *Law_Making, Appeal, Equity, Procedure*

and, then, the indicator *Tradition* equal to three if *Convergence* is weakly greater than thirteen—at least three indicators equal to four, equal to one if *Convergence* is weakly lower than seven—at most one indicator equal to four, and two otherwise. There are sixty observations equal to three, sixty equal to two, and 76 equal to one.

Results under conditional independence.—An analysis in all similar to the precedent can be devised: this time, the dependent of (7) is *Tradition*. Panel A (B) of Table VI lists the marginal effects when the controls comprehend only the extent of cultural heterogeneity (cultural heterogeneity and democracy). A one-standard-deviation rise—0.32—in democracy implies a little more than a 24 percentage fall in the likelihood of a reform from institutions typical of Civil law to those typical of Common law and more than a 33 percentage point increase in the likelihood of a reform in the opposite direction in column (4) of panel B. Cultural heterogeneity, instead, enters often with an unexpected sign; yet, the relative coefficients are generally insignificant at the usual levels.

Relaxing conditional independence.—This last pattern could be driven by a failure in the conditional independence assumption. Let me define, this time, *Common_Law* as the dummy equal to one when *Tradition* is equal to three and zero otherwise. As columns (1) - (3) of Table VII show, while the evidence on *Constraint_Executive* remains robust, the one regarding the proxies for cultural heterogeneity becomes more in tune with the above model. Both *Genetic_Distance* and *Av_Elf* enter with coefficients that are sometimes negative and significant; but the one attached to *Genetic_Distance* (*Av_Elf*) becomes insignificant (positive and significant) when *Economic_Power* is considered. To this extent, the model's failures in describing the impact of cultural heterogeneity over the convergence of the whole bundle of institutions typical of the two traditions could be simply driven by

the lack of relevant covariates as, for instance, adequate proxies for the society's aversion to discretion, which—as shown above—is responsible of the origin of hybrid systems.

All in all, it is reasonable to summarize the evidence saying that the key institutions of actual legal systems cannot be considered as randomly assigned but they are the result of a unique historical shock and of the welfare maximizing choice of the whole society.

V.   Concluding Comments

This paper has moved some steps toward a theory of "endogenous legal systems" (see also Niblett, Posner, and Shleifer, [2008]; and Guerriero, [2008]). I have focused on a crucial aspect of the design of a legal system, namely the choice of those institutions that all together can be classified as "Civil" and "Common" law traditions. Rather than reviewing my results, I highlight several avenues for further research in the field.

My results leave to the applied researcher the key task of rewriting the "legal origins" literature. In the lights of the actual analysis, the empirical literature on comparative legal and economic systems should not only consider all the different institutions—i.e., law making and adjudication—characterizing the two traditions, but also take care of the endogeneity of such rules to the level of cultural heterogeneity and democracy in a society.[38] For what concerns the former point, the present analysis also weakens those defenses of the "legal origins" literature claiming that the La Porta et al.'s (1999) classification just captures some persistent feature of a legal system possibly shaping their

---

[38]These last two forces drive also the performances that we want to explain. Persson and Tabellini (2009) show the effect of the accumulated democratic capital on the long run level of output. Alesina et al. (2003) provide evidence documenting the relevance of ethnolinguistic fractionalization as a determinant of economic success—i.e., output, literacy rate and infant mortality.

present performances. Not only the law-making institutions switches studied in section IV.B are not considered by La Porta et al. (1999) but, among the countries studied in section IV.C, 92% (68%) of the observations coded by La Porta et al. (1999) as having an English (French) law origin have a value of *Tradition* of one (three).[39] This amounts to say that these countries have reformed at least three out of the four rules clearly differentiating the two traditions in a direction showing a *de facto* interest in solving the bias versus volatility trade off with rules different from those transplanted.

Furthermore, more work should be pointed to understand to what extent legal institutions constitute an efficiency response to the economic and political environment, and to what extent they are, instead, long-lasting transplantation constraints left in place by huge switching costs. As seen in section II.B, if the latter are irrelevant, no difference among countries in the two traditions should be appreciated, provided that the measures studied take into full consideration both cultural tastes and technological efficiency.[40]

All in all, exploring further the selection, evolution and impact on performances of legal institutions is a crucial but challenging task, which requires both a solid statistical work and a careful collection of data. Progress on this task will not only advance the research frontier in economics, but also deliver huge gains in terms of quality and quantity of life to many developed and underdeveloped countries in the world.[41]

---

[39] Also three non Scandinavian countries coded as having some form of Civil law origin (three countries assigned to the English law group) were using Case (Statute) law over the whole independence period.

[40] Another key link to be uncovered is the one relating law making institutions and the ex-post degree of insulation of political leaders (see Damaška, [1986]; Aghion, Alesina and Trebbi, [2004]).

[41] A completely similar routine should be applied to a related literature focusing on the link between legal traditions and long-run economic growth, and claiming that, contrary to the incomplete markets intuition (Coase, 1960), contracting institutions—as driven by "legal origins"—have no first-order effect on long-run economic growth, investment, and financial development (Acemoglu and Johnson, 2005).

Appendix: Proofs of Lemmas and Propositions

Socially Optimal versus Technologically Efficient Rules

For $\pi > 1$, $E\left(\hat{A}_i\right) \geq A^*$ implies that

$(1/2)\left(\bar{A} + \underline{A}\right) \geq A^* \leftrightarrow \Gamma\left(\lambda\pi\right) + \Gamma\left(\lambda/\pi\right) \geq 2\Gamma\left(\lambda\right) \leftrightarrow$

$\left(\pi + \lambda\right)\left(\pi + \pi\lambda - 2 - 2\pi\lambda\right) + \lambda\left(1 + \lambda\pi\right)\left(1 + \lambda\right) \geq 0 \leftrightarrow \lambda\left(\pi - 1\right)^2\left(1 - \lambda\right) \geq 0.$

Therefore $E\left(\hat{A}_i\right) \geq A^*$ ($E\left(\hat{A}_i\right) < A^*$) whenever $\lambda \leq 1$ ($\lambda > 1$). $\qquad\square$

Proof of Lemma 1

The distribution of judges' bliss points has expectation $(1/2)\left(\bar{A} + \underline{A}\right)$ and variance $V\left(\hat{A}_i\right) = (1/2)\left(\bar{A} - E\left(\hat{A}_i\right)\right)^2 + (1/2)\left(\underline{A} - E\left(\hat{A}_i\right)\right)^2 = (1/4)\left(\bar{A} - \underline{A}\right)^2$. By the properties of the AR(1), Case law converges to the distribution $\hat{A}_c = N\left(E\left(\hat{A}_i\right), V\left(\hat{A}_i\right)\left(1 + 2K\right)^{-1}\right)$. The comparative statics with respect K holds by inspection while the one with respect π is proved by the fact that $\partial V\left(\hat{A}_i\right)\big/\partial\pi = (\lambda/2)\left(\bar{A} - \underline{A}\right)\left[\Gamma'\left(\lambda\pi\right) - \Gamma'\left(\lambda/\pi\right)\left(-1/\pi^2\right)\right] > 0.$ $\qquad\square$

Proof of Lemma 2.A and 2.B

From (2) and (3) it is immediate to see that the objective function in (4) is strictly concave and the relative necessary and sufficient first order condition is

$$(1 - \mu)\left(\hat{A}_s\left(l\right) - \tilde{A}\right) + \mu\left[|l|\,\hat{A}_s\left(l\right) - \sum_{i \in l}\hat{A}_i\right] = 0 \tag{A1}$$

Then, the unique solution to (A1) is (6) which rewrites for every possible l as

$\hat{A}_s\left(\{\phi\}\right) = \tilde{A}; \qquad \hat{A}_s\left(\{L\}\right) = (1 - \mu)\tilde{A} + \mu\bar{A};$

$\hat{A}_s\left(\{H\}\right) = (1 - \mu)\tilde{A} + \mu\underline{A}; \quad \hat{A}_s\left(\{L, H\}\right) = \left[(1 - \mu)\tilde{A} + \mu\left(\underline{A} + \bar{A}\right)\right]\left(1 + \mu\right)^{-1} = \tilde{A}.$

The equilibrium willingness to pay for each of the last three coalitions are

$$W_l\left(\hat{A}_s\left(\{L\}\right),\tilde{A}\right) = \frac{1}{4}\left\{\left(\tilde{A}-\bar{A}\right)^2 - \left[(1-\mu)\left(\tilde{A}-\bar{A}\right)\right]^2\right\} - \Psi =$$

$$= \frac{1}{4}\mu\left(2-\mu\right)\left(\tilde{A}-\bar{A}\right)^2 - \Psi = \frac{1}{16}\mu\left(2-\mu\right)\left(\bar{A}-\underline{A}\right)^2 - \Psi = \qquad\text{(A2)}$$

$$= \frac{1}{4}\mu\left(2-\mu\right)\left(\tilde{A}-\underline{A}\right)^2 - \Psi = W_l\left(\hat{A}_s\left(\{H\}\right),\tilde{A}\right);$$

$$W_l\left(\hat{A}_s\left(\{L,H\}\right),\tilde{A}\right) = \frac{1}{4}\left\{\left(\tilde{A}-\bar{A}\right)^2 + \left(\tilde{A}-\underline{A}\right)^2 - \left(\tilde{A}-\bar{A}\right)^2 - \left(\tilde{A}-\underline{A}\right)^2\right\} - \Psi = -\Psi < 0.$$

Being $2\tilde{A} = \bar{A} - \underline{A}$, coalition $\{L,H\}$ will never pay the fixed cost $\Psi > 0$. Therefore, the Legislator chooses a coalition among $\{\phi\}$, $\{L\}$ and $\{H\}$:

$$U\left(\hat{A}_s\left(\{\phi\}\right),\tilde{A}\right) = -\frac{1-\mu}{2}\left(\tilde{A}-\tilde{A}\right)^2 = 0$$

$$U\left(\hat{A}_s\left(\{L\}\right),\tilde{A}\right) = -(1-\mu)\frac{\mu^2}{2}\left(\tilde{A}-\bar{A}\right)^2 + \frac{\mu^2}{2}\left(2-\mu\right)\left(\tilde{A}-\bar{A}\right)^2 - 2\mu\Psi = \qquad\text{(A3)}$$

$$= \frac{\mu^2}{8}\left(\bar{A}-\underline{A}\right)^2 - 2\mu\Psi = U\left(\hat{A}_s\left(\{H\}\right),\tilde{A}\right)$$

Clearly the Legislator chooses either $\{L\}$ or $\{H\}$—it is strictly indifferent between the two, and both her utility in favoring one of these two coalitions and the coalitions' bribing schedules rise with both μ and π, being μ^2 and $\mu\left(2-\mu\right)$ increasing with μ and $\partial\left(\underline{A}-\bar{A}\right)^2\big/\partial\pi > 0$.

Thus, for $\Psi > 0$ sufficiently small, it does always exist a level of π implicitly defined by $(\mu/16)\left(2-\mu\right)\left(\bar{A}\left(\bar{\pi}\right)-\underline{A}\left(\bar{\pi}\right)\right)^2 = \Psi$, such that—as (A2) suggests—for $\pi \geq \bar{\pi}$ both groups become organized and for $\pi < \bar{\pi}$ they remain dormant. Similarly, there is a π implicitly defined by $(\mu/16)\left(\bar{A}\left(\bar{\bar{\pi}}\right)-\underline{A}\left(\bar{\bar{\pi}}\right)\right)^2 = \Psi$, such that—as (A3) suggests—for $\pi \geq \bar{\bar{\pi}}$ the Legislator chooses either $\{L\}$ or $\{H\}$, and for $\pi < \bar{\bar{\pi}}$ $\tilde{A}$ prevails. For every π and μ, $(\mu/16)\left(2-\mu\right)\left(\bar{A}\left(\pi\right)-\underline{A}\left(\pi\right)\right)^2 \geq (\mu/16)\left(\bar{A}\left(\pi\right)-\underline{A}\left(\pi\right)\right)^2$: as a result, $\bar{\bar{\pi}} \geq \bar{\pi}$. From $t = 1$ on and for $\pi \geq \bar{\bar{\pi}}$, the loss of social welfare is $\left(\mu^2/8\right)\left(\bar{A}-\underline{A}\right)^2$. No group pays a $\Psi > 1/16$.

The following comparative statics complete the proof:

$(1/16)\,\mu\,(2-\mu)\left[\partial\left(\bar{A}\left(\bar{\pi}\right)-\underline{A}\left(\bar{\pi}\right)\right)^{2}\Big/\partial\pi\right]d\bar{\pi}-d\Psi=0\to d\bar{\pi}/d\Psi>0;$

$(1/16)\,\mu\,(2-\mu)\left[\partial\left(\bar{A}\left(\bar{\pi}\right)-\underline{A}\left(\bar{\pi}\right)\right)^{2}\Big/\partial\pi\right]d\bar{\pi}+(1/8)\,(1-\mu)\left(\bar{A}\left(\bar{\pi}\right)-\underline{A}\left(\bar{\pi}\right)\right)^{2}d\mu=0\to\frac{d\bar{\pi}}{d\mu}<0;$

$(1/8)\,\mu^{2}\left[\partial\left(\bar{A}\left(\bar{\bar{\pi}}\right)-\underline{A}\left(\bar{\bar{\pi}}\right)\right)^{2}\Big/\partial\pi\right]d\bar{\bar{\pi}}-2\mu d\Psi=0\to d\bar{\bar{\pi}}/d\Psi>0;$

$(1/8)\,\mu^{2}\left[\partial\left(\bar{A}\left(\bar{\bar{\pi}}\right)-\underline{A}\left(\bar{\bar{\pi}}\right)\right)^{2}\Big/\partial\pi\right]d\bar{\bar{\pi}}+(1/4)\,\mu\left(\bar{A}\left(\bar{\bar{\pi}}\right)-\underline{A}\left(\bar{\bar{\pi}}\right)\right)^{2}d\mu=0\to d\bar{\bar{\pi}}/d\mu<0.$ □

Proof of Lemma 3

The set of possible coalitions is $\Omega=\{\{\phi\},\{L\},\{H\},\{E\},\{L,H\},\{L,E\},\{E,H\},\{L,H,E\}\}$ and $\tilde{A}=(1/3)\left(\bar{A}+\underline{A}+A^{*}\right)$. Therefore, for every possible l, (6) rewrites as:

$$\hat{A}_{s}\left(\{\phi\}\right)=\tilde{A}; \qquad\qquad \hat{A}_{s}\left(\{L\}\right)=(1-\mu)\,\tilde{A}+\mu\bar{A};$$

$$\hat{A}_{s}\left(\{H\}\right)=(1-\mu)\,\tilde{A}+\mu\underline{A}; \qquad\qquad \hat{A}_{s}\left(\{E\}\right)=(1-\mu)\,\tilde{A}+\mu A^{*};$$

$$\hat{A}_{s}\left(\{L,H\}\right)=\frac{(1-\mu)\,\tilde{A}+\mu\left(\underline{A}+\bar{A}\right)}{1+\mu}; \qquad \hat{A}_{s}\left(\{L,E\}\right)=\frac{(1-\mu)\,\tilde{A}+\mu\left(A^{*}+\bar{A}\right)}{1+\mu};$$

$$\hat{A}_{s}\left(\{E,H\}\right)=\frac{(1-\mu)\,\tilde{A}+\mu\left(\underline{A}+A^{*}\right)}{1+\mu}; \quad \hat{A}_{s}\left(\{L,E,H\}\right)=\frac{(1-\mu)\,\tilde{A}+\mu\left(\underline{A}+\bar{A}+A^{*}\right)}{1+2\mu}=\tilde{A}.$$

The equilibrium willingness to pay for each of the last seven coalitions are:

$$W_{l}\left(\hat{A}_{s}\left(\{L\}\right),\tilde{A}\right)=\left(\frac{\mu}{6}\right)(2-\mu)\left(\tilde{A}-\bar{A}\right)^{2}-\Psi=\left(\frac{\mu}{6}\right)(2-\mu)\left(\frac{A^{*}+\underline{A}-2\bar{A}}{3}\right)^{2}-\Psi; \quad (A4)$$

$$W_{l}\left(\hat{A}_{s}\left(\{H\}\right),\tilde{A}\right)=\left(\frac{\mu}{6}\right)(2-\mu)\left(\tilde{A}-\underline{A}\right)^{2}-\Psi=\left(\frac{\mu}{6}\right)(2-\mu)\left(\frac{\bar{A}+A^{*}-2\underline{A}}{3}\right)^{2}-\Psi; \quad (A5)$$

$$W_{l}\left(\hat{A}_{s}\left(\{E\}\right),\tilde{A}\right)=(\mu/6)\,(2-\mu)\left(\tilde{A}-A^{*}\right)^{2}-\Psi=(\mu/6)\,(2-\mu)\left[\left(\bar{A}+\underline{A}-2A^{*}\right)/3\right]^{2}-\Psi;$$

$$W_l\left(\hat{A}_s\left(\{L,H\}\right),\tilde{A}\right) = (1/6)\left[\left(2\mu+\mu^2\right)\big/(1+\mu)^2\right]\left[(2A^* - \bar{A} - \underline{A})/3\right]^2 - \Psi;$$

$$W_l\left(\hat{A}_s\left(\{L,E\}\right),\tilde{A}\right) = (1/6)\left[\left(2\mu+\mu^2\right)\big/(1+\mu)^2\right]\left[(2\underline{A} - A^* - \bar{A})/3\right]^2 - \Psi;$$

$$W_l\left(\hat{A}_s\left(\{E,H\}\right),\tilde{A}\right) = (1/6)\left[\left(2\mu+\mu^2\right)\big/(1+\mu)^2\right]\left[(2\bar{A} - A^* - \underline{A})/3\right]^2 - \Psi;$$

$$W_l\left(\hat{A}_s\left(\{L,E,H\}\right),\tilde{A}\right) = -\Psi < 0.$$

Coalition $\{L,E,H\}$ will never pay a $\Psi > 0$. Thus, in order to identify the winner, it is enough to compare the Legislator's utility for the remaining coalitions:

$$U\left(\hat{A}_s\left(\{\phi\}\right),\tilde{A}\right) = 0; \qquad U\left(\hat{A}_s\left(\{E\}\right),\tilde{A}\right) = (\mu^2/2)\left[(\bar{A} + \underline{A} - 2A^*)/3\right]^2 - 3\mu\Psi;$$

$$U\left(\hat{A}_s\left(\{L\}\right),\tilde{A}\right) = (\mu^2/2)\left[(A^* + \underline{A} - 2\bar{A})/3\right]^2 - 3\mu\Psi; \tag{A6}$$

$$U\left(\hat{A}_s\left(\{H\}\right),\tilde{A}\right) = (\mu^2/2)\left[(\bar{A} + A^* - 2\underline{A})/3\right]^2 - 3\mu\Psi; \tag{A7}$$

$$U\left(\hat{A}_s\left(\{L,H\}\right),\tilde{A}\right) = (1/2)\left[\left(2\mu^3+\mu^2\right)\big/(1+\mu)^2\right]\left[(2A^* - \bar{A} - \underline{A})/3\right]^2 - 3\mu\Psi;$$

$$U\left(\hat{A}_s\left(\{L,E\}\right),\tilde{A}\right) = (1/2)\left[\left(2\mu^3+\mu^2\right)\big/(1+\mu)^2\right]\left[(2\underline{A} - A^* - \bar{A})/3\right]^2 - 3\mu\Psi;$$

$$U\left(\hat{A}_s\left(\{E,H\}\right),\tilde{A}\right) = (1/2)\left[\left(2\mu^3+\mu^2\right)\big/(1+\mu)^2\right]\left[(2\bar{A} - A^* - \underline{A})/3\right]^2 - 3\mu\Psi.$$

For every π, $\underline{A} \leq A^* \leq \bar{A}$ then $\min\left\{\left|\tilde{A} - \bar{A}\right|, \left|\tilde{A} - \underline{A}\right|\right\} > \left|\tilde{A} - A^*\right|$ which implies that

coalition $\{L\}$ and $\{H\}$ will always be preferred to $\{E\}$. Thus, it remains to compare coalition

$\{L\}$, $\{H\}$, $\{L, H\}$, $\{L, E\}$ and $\{E, H\}$. The Legislator prefers $\{L\}$ to $\{H\}$ if $\left|A^* + \underline{A} - 2\bar{A}\right| =$

$2\bar{A} - A^* - \underline{A} > \bar{A} + A^* - 2\underline{A}$ which is true for $\lambda \leq 1$, while the opposite is true for $\lambda > 1$. Notice also

that $\left|2\bar{A} - A^* - \underline{A}\right| = 2\bar{A} - A^* - \underline{A}$, $\left|2\underline{A} - A^* - \bar{A}\right| = A^* + \bar{A} - 2\underline{A}$, $\left|2A^* - \bar{A} - \underline{A}\right| = 2A^* - \bar{A} - \underline{A}$

for $\lambda > 1$ and $\left|2A^* - \bar{A} - \underline{A}\right| = \bar{A} + \underline{A} - 2A^*$ for $\lambda \leq 1$. For $\lambda \leq 1$, $\{E, H\}$ is preferred to $\{L, H\}$—

being $2\bar{A} - A^* - \underline{A} > \bar{A} + \underline{A} - 2A^*$—and to $\{L, E\}$—because $2\bar{A} - A^* - \underline{A} \geq \bar{A} + A^* - 2\underline{A}$. Yet, the

Legislator weakly (strictly) prefers $\{L\}$ to $\{E, H\}$ for every $\mu \geq 0$ ($\mu > 0$) being $2\mu + 1 \leq (1 + \mu)^2$.

For $\lambda > 1$, $\{L, E\}$ is preferred to $\{E, H\}$—because $\bar{A} + A^* - 2\underline{A} > 2\bar{A} - A^* - \underline{A}$—which, in turn, is

preferred to $\{L, H\}$—being $2\bar{A} - A^* - \underline{A} > 2A^* - \bar{A} - \underline{A}$. Yet again the Legislator picks an extrem-

ist coalition because she weakly (strictly) prefers $\{H\}$ for every $\mu \geq 0$ ($\mu > 0$). The Legislator's

utility in favoring $\{L\}$ or $\{H\}$ and the two coalitions' bribing schedules increase with both

μ and π being $\partial\left(A^* + \underline{A} - 2\bar{A}\right)^2 \Big/ \partial\pi = 2\left(A^* + \underline{A} - 2\bar{A}\right)\left[\Gamma'\left(\lambda/\pi\right)\left(-1/\pi^2\right) - 2\Gamma'\left(\lambda\pi\right)\pi\right] =$

$\partial\left(\bar{A} + A^* - 2\underline{A}\right)^2 \Big/ \partial\pi > 0$. As (A4) ((A5)) suggests, for $\lambda \leq 1$ ($\lambda > 1$), provided that $\Psi > 0$ is

sufficiently small, there exists a $\tilde{\pi}_L$ ($\tilde{\pi}_H$) defined by $(1/54)\,\mu\,(2 - \mu)\left(A^* + \underline{A}\left(\tilde{\pi}_L\right) - 2\bar{A}\left(\tilde{\pi}_L\right)\right)^2 =$

Ψ $((1/54)\,\mu\,(2 - \mu)\left(\bar{A}\left(\tilde{\pi}_H\right) + A^* - 2\underline{A}\left(\tilde{\pi}_H\right)\right)^2 = \Psi)$ such that for $\pi \geq \tilde{\pi}_L$, $\{L\}$ ($\pi \geq \tilde{\pi}_H$, $\{H\}$)

pays the fixed fee and for $\pi < \tilde{\pi}_L$ (for $\pi < \tilde{\pi}_H$) it remains dormant. Also, as (A6) ((A7))

suggests, there is a $\tilde{\tilde{\pi}}_L$ ($\tilde{\tilde{\pi}}_H$) implicitly defined by $(1/54)\,\mu^2\left(A^* + \underline{A}\left(\tilde{\tilde{\pi}}_L\right) - 2\bar{A}\left(\tilde{\tilde{\pi}}_L\right)\right)^2 = \Psi$

$((1/54)\,\mu^2\left(\bar{A}\left(\tilde{\tilde{\pi}}_H\right) + A^* - 2\underline{A}\left(\tilde{\tilde{\pi}}_H\right)\right)^2 = \Psi)$, such that for $\pi \geq \tilde{\tilde{\pi}}_L$ ($\pi \geq \tilde{\tilde{\pi}}_H$) the Legislator

chooses $\{L\}$ ($\{H\}$) and, for $\pi < \tilde{\tilde{\pi}}_L$ ($\pi < \tilde{\tilde{\pi}}_H$), $\tilde{A}$ prevails. By a reasoning similar to that devel-

oped in the proof of lemma 2, it is easy to see that $\tilde{\tilde{\pi}}_{\tilde{l}} \geq \tilde{\pi}_{\tilde{l}}$, $\forall \tilde{l} = L, H$. For $\pi \geq \tilde{\tilde{\pi}}_L$ ($\pi \geq \tilde{\tilde{\pi}}_L$) and

$\lambda \leq 1$ ($\lambda > 1$) the loss of social welfare is $\left(\mu^2/18\right)\left(A^* + \underline{A} - 2\bar{A}\right)^2$ $\left(\left(\mu^2/18\right)\left(\bar{A} + A^* - 2\underline{A}\right)^2\right)$.

No group pays a $\Psi > 1/27$. The comparative statics on $\tilde{\pi}_{\tilde{l}}$ and $\tilde{\tilde{\pi}}_{\tilde{l}}$ are similar to those discussed

45

above and, accordingly, omitted. □

Proof of Proposition 2

For $\lambda > 1$ the expected losses of welfare under the two rules are equal for $0 \leq \tilde{\mu}_H^2 =$

$$[2(1+2K)]^{-1}\left\{\left[6\left(\bar{A}\right)^2 + 6\left(A^*\right)^2 + 6\left(\underline{A}\right)^2 - 2\left(\bar{A} + A^* + \underline{A}\right)^2\right]\left(\bar{A} + A^* - 2\underline{A}\right)^{-2}\right\} \quad \text{(A8)}$$

The difference between the numerator and the denominator of the ratio in the brackets is

$$6\left(\bar{A}\right)^2 + 6\left(A^*\right)^2 + 6\left(\underline{A}\right)^2 - 2\left(\bar{A} + A^* + \underline{A}\right)^2 - \left[\left(\bar{A} + A^* + \underline{A}\right) - 3\underline{A}\right]^2 =$$

$$6\left(\bar{A}\right)^2 + 6\left(A^*\right)^2 - 3\left[\left(\bar{A} + A^* + \underline{A}\right)^2 - 2\underline{A}\left(\bar{A} + A^* + \underline{A}\right) + \left(\underline{A}\right)^2\right] = 3\left(\bar{A} - A^*\right)^2 > 0,$$

so that (A8) can be written as $\tilde{\mu}_H^2 = (1/2)(1+2K)^{-1}\left\{1 + 3\left(\bar{A} - A^*\right)^2\left(\bar{A} + A^* - 2\underline{A}\right)^{-2}\right\}.$

As a result, the right hand side of (A8) is decreasing with π because

$$\frac{\partial}{\partial\pi}\left(\frac{\bar{A} - A^*}{\bar{A} + A^* - 2\underline{A}}\right)^2 < 0 \leftrightarrow \frac{\partial\bar{A}}{\partial\pi}\left(\bar{A} + A^* - 2\underline{A}\right) - \left(\frac{\partial\bar{A}}{\partial\pi} - 2\frac{\partial\underline{A}}{\partial\pi}\right)\left(\bar{A} - A^*\right) < 0 \leftrightarrow$$

$$\leftrightarrow \frac{2\lambda}{(1+\pi\lambda)^2}\left(\frac{\lambda}{1+\lambda} - \frac{\lambda}{\pi+\lambda}\right) < \frac{2\lambda}{(\pi+\lambda)^2}\left(\frac{\lambda\pi}{1+\lambda\pi} - \frac{\lambda}{1+\lambda}\right) \leftrightarrow \pi + \lambda < 1 + \lambda\pi \leftrightarrow \lambda > 1.$$

To see why $\tilde{\mu}_H = 1$ for a $K_H > 0$, notice that $\partial\tilde{\mu}_H/\partial K < 0$ and that for $K = 0$ the difference between the loss of social welfare under Case law and the one under Statute law is $3\left(\bar{A} - A^*\right)^2 > 0$.

A similar analysis applies to the $\lambda \leq 1$ case and it is, therefore, omitted. In this case, the expected loss of welfare is equal under the two rules for

$$\tilde{\mu}_L^2 = (1/2)(1+2K)^{-1}\left[1 + 3\left(A^* - \underline{A}\right)^2\left(A^* + \underline{A} - 2\bar{A}\right)^{-2}\right] \geq 0. \qquad\qquad □$$

Accounting for Tastes versus Efficiency

The willingness to pay for $\{L\}$, $\{H\}$ and $\{L, H\}$ are:

$$W_l\left(\hat{A}_s^E\left(\{L\}\right), \tilde{A}^E\right) = (1/16)\mu(2-\mu)(1-T)^2\left(\bar{A} - \underline{A}\right)^2 - \Psi = W_l\left(\hat{A}_s^E\left(\{H\}\right), \tilde{A}^E\right),$$

$$W_l\left(\hat{A}_s^E\left(\{L, H\}\right), \tilde{A}^E\right) = -\Psi.$$

Again coalition $\{L, H\}$ is not willing to pay the collective action cost and the Legislator is indifferent between $\{L\}$ and $\{H\}$. Both her utility in favoring one of these two coalitions— $\left(\mu^2/8\right)(1-T)^2\left(\bar{A} - \underline{A}\right)^2 - 2\mu\Psi$—and the coalitions' bribing schedules rise with μ and π. For $\Psi > 0$ sufficiently small, it exists a $\bar{\pi}^E$ defined by $(1/16)\,\mu\,(2-\mu)\,(1-T)^2\left(\bar{A}\left(\bar{\pi}^E\right) - \underline{A}\left(\bar{\pi}^E\right)\right)^2 = \Psi$, such that for $\pi \geq \bar{\pi}^E$ both groups become organized and for $\pi < \bar{\pi}^E$ they remain dormant, and there is a $\bar{\bar{\pi}}^E$ implicitly defined by $(1/16)\,\mu\,(1-T)^2\left(\bar{A}\left(\bar{\bar{\pi}}^E\right) - \underline{A}\left(\bar{\bar{\pi}}^E\right)\right)^2 = \Psi$, such that for $\pi \geq \bar{\bar{\pi}}^E$ the Legislator chooses either $\{L\}$ or $\{H\}$ and, for $\pi < \bar{\bar{\pi}}^E$, $\tilde{A}^E$ is chosen. Again $\bar{\bar{\pi}}^E > \bar{\pi}^E$. For $\pi \geq \bar{\bar{\pi}}^E$, the loss of social welfare is $\left(\mu^2/8\right)(1-T)^2\left(\bar{A} - \underline{A}\right)^2$. No group pays a $\Psi > (1/16)(1-T)^2$; besides, the usual comparative statics apply. $\qquad\square$

Proof of Lemma 4 and Proposition 3

Being $\tilde{K} < 1/8 < 1/2$ it does always exist a $\hat{\underline{\pi}}$ implicitly defined by $\bar{A}\left(\hat{\underline{\pi}}\right) - \underline{A}\left(\hat{\underline{\pi}}\right) = \sqrt{2\tilde{K}}$ such that for $\pi < \hat{\underline{\pi}}$ a first period judge selects $\hat{A}_{c,1} = \hat{A}_i$ and she is never reversed. In this case, the loss of social welfare under Case law is $(1/8)\left(\bar{A} - \underline{A}\right)^2$. When π grows over $\hat{\underline{\pi}}$, the first period judge can either fix her own preferred threshold—which will be overruled in the second period—or a compromise that is not overruled by a judge of type $-i$ but could eventually be overruled by a judge of type i. Such a compromise is given by $\tilde{\underline{A}} = \bar{A} - \sqrt{2\tilde{K}}$ $(\tilde{\bar{A}} = \underline{A} + \sqrt{2\tilde{K}})$—which rises with π and it is implicitly defined by $\left(\tilde{\underline{A}} - \bar{A}\right)^2 = 2\tilde{K}$ $(\left(\tilde{\bar{A}} - \underline{A}\right)^2 = 2\tilde{K})$—if the first period judge is of type H (L). A second period judge of type H does not overrule $\tilde{\underline{A}}$ if $\left(\bar{A} - \sqrt{2\tilde{K}} - \underline{A}\right)^2 < 2\tilde{K} \leftrightarrow (1/2)\left(\bar{A} - \underline{A}\right) > \sqrt{2\tilde{K}}$, which is the case for $\hat{\underline{\pi}} \leq \pi \leq \hat{\bar{\pi}}$ with $\bar{A}\left(\hat{\bar{\pi}}\right) - \underline{A}\left(\hat{\bar{\pi}}\right) = 2\sqrt{2\tilde{K}}$. At the same time, the first period H judge will prefer the compromise when $(\delta/2)\left(\bar{A} - \underline{A}\right)^2 > (1+\delta)\left(\tilde{\underline{A}} - \underline{A}\right)^2$, which is the case if: $(2+\delta)\left(\bar{A} - \underline{A}\right)^2 + 2(1+\delta)\,2\tilde{K} - 4(1+\delta)\left(\bar{A} - \underline{A}\right)\sqrt{2\tilde{K}} < 0$. The left hand side of this last inequality is increasing with $\left(\bar{A} - \underline{A}\right)$ for $\left(\bar{A} - \underline{A}\right) > [2(1+\delta)/(2+\delta)]\sqrt{2\tilde{K}}$, negative for the

47

lowest possible $\left(\bar{A} - \underline{A}\right)$ in the relevant range—$\sqrt{2\tilde{K}}$ at $\hat{\underline{\pi}}$—and positive for the highest possible value of $\left(\bar{A} - \underline{A}\right)$—$2\sqrt{2\tilde{K}}$ at $\hat{\bar{\pi}}$. As a result, it does exist a $\hat{\pi}$ with $\hat{\underline{\pi}} < \hat{\pi} < \hat{\bar{\pi}}$ such that for $\hat{\underline{\pi}} \leq \pi \leq \hat{\pi}$ the first period judge prefers $\tilde{\underline{A}}$ and for $\hat{\pi} < \pi < \hat{\bar{\pi}}$ she chooses $\underline{A}$. If $\pi \geq \hat{\bar{\pi}}$—which can be the case only if $\tilde{K} < 1/8$—she always prefers the compromise and being overruled by a judge of the same type being $(\delta/2)\left(\bar{A} - \underline{A}\right)^2 > (\delta/2)\left(\tilde{A} - \underline{A}\right)^2$. By symmetry, the same analysis holds true when the first period judge is of type L. The loss of social welfare is $(1/8)\left(\bar{A} - \underline{A}\right)^2$ for $\pi < \hat{\pi}$ and for $\hat{\pi} < \pi < \hat{\bar{\pi}}$, $(1/2)\left[(1/2)\left(\bar{A} - \underline{A}\right) - \sqrt{2\tilde{K}}\right]^2$ for $\hat{\underline{\pi}} \leq \pi \leq \hat{\pi}$ and $(1/4)\left[(1/2)\left(\bar{A} - \underline{A}\right) - \sqrt{2\tilde{K}}\right]^2 + (1/16)\left(\bar{A} - \underline{A}\right)^2$ for $\pi \geq \hat{\bar{\pi}}$. $\qquad\square$

Proof of Lemma 5 and Proposition 4

A type $\{H\}$ judge faced with a type $\{L\}$'s precedent distinguishes to the first best if $-\tilde{K}^I \geq -(1/2)\left(\bar{A} - \underline{A}\right)^2 \leftrightarrow 2\tilde{K}^I \leq \left(\bar{A} - \underline{A}\right)^2$. By the same token, a type $\{L\}$ judge faced with a type $\{H\}$'s precedent distinguishes to the first best if $2\tilde{K}^I \leq \left(\bar{A} - \underline{A}\right)^2$. Therefore, being $\partial\left(\underline{A} - \bar{A}\right)^2 / \partial\pi > 0$ and provided that $\tilde{K}^I \leq 1/2$, there exists a π^I with $\partial\pi^I / \partial\tilde{K}^I > 0$ such that every judge finds it worthy to introduce the second dimension for $\pi \geq \pi^I$. The loss of welfare under Case law for $\pi < \pi^I$ is:

$$\tilde{\Lambda}\left(\hat{A}_c\right) = (1/2)\left(\bar{A} - E\left(\hat{A}_i\right)\right)^2 = (1/2)\left(\underline{A} - E\left(\hat{A}_i\right)\right)^2 = (1/8)\left(\bar{A} - \underline{A}\right)^2 \geq \mu^2(1/8)\left(\bar{A} - \underline{A}\right)^2,$$

$\forall\mu \in [0,1]$. Therefore, for $\pi < \pi^I$ Statute law is weakly better; for $\pi > \pi^I$, instead, the first best is achieved under Case law which is therefore strictly better. $\qquad\square$

Proof of Proposition 5

Lower courts of type H and L will make use of discretion respectively when:

$$\left(\hat{A}_{j,t}^J - \underline{A}\right)^2 \geq \left(\hat{A}_{j,t}^J - \alpha - \underline{A}\right)^2 + 2\theta \leftrightarrow \hat{A}_{j,t}^J \geq (2\alpha)^{-1}\left(2\theta + \alpha^2\right) + \underline{A}; \qquad (A9)$$

$$\left(\hat{A}_{j,t}^{J} - \bar{A}\right)^{2} \geq \left(\hat{A}_{j,t}^{J} + \alpha - \bar{A}\right)^{2} + 2\theta \leftrightarrow \hat{A}_{j,t}^{J} \leq -\left(2\alpha\right)^{-1}\left(2\theta + \alpha^{2}\right) + \bar{A}. \qquad \text{(A10)}$$

From (A9) and (A10) it follows that a lower court of type H (L) uses discretion if and only if $\hat{A}_{j,t}^{J} > E\left(\hat{A}_{i}\right)$ $(\hat{A}_{j,t}^{J} < E\left(\hat{A}_{i}\right))$ whenever $\theta > \bar{\theta} \equiv (\alpha/2)\left[\left(\bar{A}\left(\pi\right) - \underline{A}\left(\pi\right)\right) - \alpha\right]$; $\bar{\theta}$ is positive because $\alpha < (1/2)\left(\bar{A} - \underline{A}\right)$. The objective function of a generic appellate judge i is now:

$$-\Pr\left[A_{c,t}^{J}\left(i\right) < \tilde{A}\right]\left\{\left[A_{c,t}^{J}\left(i\right) - A_{c,t}^{J}\left(i\right) - \alpha/2\right]^{2} + \left[A_{c,t}^{J}\left(i\right) + \alpha - A_{c,t}^{J}\left(i\right) - \alpha/2\right]^{2}\right\} +$$

$$-\Pr\left[A_{c,t}^{J}\left(i\right) > \tilde{A}\right]\left\{\left[A_{c,t}^{J}\left(i\right) - A_{c,t}^{J}\left(i\right) + \alpha/2\right]^{2} + \left[A_{c,t}^{J}\left(i\right) - \alpha - A_{c,t}^{J}\left(i\right) + \alpha/2\right]^{2}\right\} +$$

$$-\Pr\left[A_{c,t}^{J}\left(i\right) < \tilde{A}\right]\left[A_{c,t}^{J}\left(i\right) + \alpha/2 - \hat{A}_{i}\right]^{2} - \Pr\left[A_{c,t}^{J}\left(i\right) > \tilde{A}\right]\left[A_{c,t}^{J}\left(i\right) - \alpha/2 - \hat{A}_{i}\right]^{2} +$$

$$-2K\left(A_{c,t}^{J}\left(i\right) - \hat{A}_{c,t-1}^{J}\right)^{2}.$$

Consider first a type L appellate judge, she will choose an $A_{c,t}^{J}\left(L\right) > \tilde{A}$ whenever the first order condition is met and the corner $\tilde{A}$ is not a solution or

$$\hat{A}_{c,t}^{J}\left(L\right) - \alpha/2 - \bar{A} + K\left(\hat{A}_{c,t}^{J}\left(L\right) - \hat{A}_{c,t-1}^{J}\right) = 0; \qquad \text{(A11)}$$

$$\tilde{A} + \alpha/2 - \bar{A} + K\left(\tilde{A} - \hat{A}_{c,t-1}^{J}\right) < 0 \leftrightarrow \alpha < \left(\bar{A} - \underline{A}\right) - K\left(\bar{A} + \underline{A} - 2\hat{A}_{c,t-1}^{J}\right). \qquad \text{(A12)}$$

Symmetrically a type H judge will fix an $A_{c,t}^{F}\left(H\right) < \tilde{A}$ whenever

$$\hat{A}_{c,t}^{J}\left(H\right) + \alpha/2 - \underline{A} + K\left(\hat{A}_{c,t}^{J}\left(H\right) - \hat{A}_{c,t-1}^{J}\right) = 0; \qquad \text{(A13)}$$

$$\tilde{A} - \alpha/2 - \underline{A} + K\left(\tilde{A} - \hat{A}_{c,t-1}^{J}\right) > 0 \leftrightarrow \alpha < \left(\bar{A} - \underline{A}\right) - K\left(\bar{A} + \underline{A} - 2\hat{A}_{c,t-1}^{J}\right). \qquad \text{(A14)}$$

Conditions (A11) and (A13) imply that Case law follows an AR(1) process converging to the ergodic distribution with mean $E\left(\hat{A}_{c,t}^{J}\left(L\right)\right) = E\left(\hat{A}_{i}\right)$ and variance $V\left(\hat{A}_{c,t}^{J}\left(L\right)\right) =$ $(1 + 2K)^{-1}\left[(1/4)\left(\bar{A} - \underline{A}\right)^{2} + \alpha^{2}/4 + (\alpha/2)\left(\bar{A} - \underline{A}\right)\right].$

If Case law is mature conditions (A12) and (A14) are always met being $\alpha < \bar{\alpha} < \left(\bar{A} - \underline{A}\right)$, and

$\tilde{A}$ is not a solution being $\tilde{A} - \bar{A} < 0$ ($\tilde{A} - \underline{A} > 0$) when $i = L$ ($i = H$). As a consequence, $A^J_{c,t}(L) > \tilde{A}$ and $A^J_{c,t}(H) < \tilde{A}$. The variance of the *de facto* rule is $\mathrm{V}\left(\hat{A}^J_{c,t}(L) + \tilde{\alpha}\right)$—where $\tilde{\alpha}$ is the random variable equal to 0 when appellate and lower courts are of the same type and to $-\alpha$ (α) if $i = L$ ($i = H$) and the adjudicating court is of type H (L):

$$\mathrm{V}\left(\hat{A}^J_{c,t}(L) + \tilde{\alpha}\right) = \mathrm{V}\left(\hat{A}^J_{c,t}(L)\right) + \mathrm{V}(\tilde{\alpha}) + 2Cov\left(\hat{A}^J_{c,t}(L) + \tilde{\alpha}\right) =$$

$$= \mathrm{V}\left(\hat{A}^J_{c,t}(L)\right) + \mathrm{E}(\tilde{\alpha}^2) + 2\mathrm{E}\left(\hat{A}^J_{c,t}(L)\,\tilde{\alpha}\right) =$$

$$= \frac{1}{1+2K}\left[\frac{\left(\bar{A}-\underline{A}\right)^2}{4} + \frac{\alpha^2}{4} + \frac{\alpha\left(\bar{A}-\underline{A}\right)}{2}\right] + \frac{\alpha^2}{2} - \frac{\alpha\left(\bar{A}-\underline{A}\right)}{2(1+K)} - \frac{\alpha^2}{2(1+K)} = \tag{A15}$$

$$= \frac{V\left(\hat{A}_i\right)}{1+2K} + \frac{\alpha^2\left(1+3K+4K^2\right)}{4(1+K)(1+2K)} - \frac{\alpha\left(\bar{A}-\underline{A}\right)K}{2(1+K)(1+2K)} < \frac{V\left(\hat{A}_i\right)}{1+2K}.$$

The last inequality in (A15) is always true being $\alpha < \bar{\alpha} \le 2K\left(1+3K+4K^2\right)^{-1}\left(\bar{A}-\underline{A}\right)$. The pure Common law tradition is introduced whenever $\Theta < \bar{\Theta} \equiv (1/2)\left[V\left(\hat{A}_i\right)[1+2K] - \mathrm{V}\left(\hat{A}^F_{c,t}\right)\right]$. Of course, $\bar{\Theta} > 0$. If Statute law outperforms Case law

$$\mu^2 V\left(\hat{A}_i\right) < \min\left\{\left[V\left(\hat{A}_i\right) + \alpha^2/4 + \left(\bar{A}-\underline{A}\right)(\alpha/2)\right][(1+2K)]^{-1}, V\left(\hat{A}_i\right)(1+2K)^{-1}\right\}$$

and the condition $\mu < (1+2K)^{-1/2}$ has to hold true.

As a result, $\alpha < \left[1 - (1+2K)^{-1/2}\right]\left(\bar{A}-\underline{A}\right) < (1-\mu)\left(\bar{A}-\underline{A}\right)$ and the maximum (minimum) value of the *de facto* rule when coalition L (H) wins is always lower (bigger) than $\bar{A}$ ($\underline{A}$) being

$$\hat{A}^F_s(L) = (1-\mu)\tilde{A} + \mu\bar{A} + \alpha < \bar{A} \le 1 \ (\hat{A}^F_s(H) = (1-\mu)\tilde{A} + \mu\underline{A} - \alpha > \underline{A} \ge 0). \qquad \square$$

References

Abdesselem, Mustapha, etc., "National Reports," in *International Encyclopedia of Comparative Law, Vol. I*, René, David, etc., ed. (New York, NY: Oceana Publications Inc., 1995).

Acartürk, Burcu, etc., "Courts: The Lex Mundi Project. Civil Procedures Around the World," *Vol. II*, (2005) Harvard university manuscript.

Acemoglu, Daron, and Simon Johnson, "Unbundling Institutions," *Journal of Political Economy*, CXIII (2005), 949-995.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, XCI (2001), 1369-1401.

Acemoglu, Daron, and James A. Robinson, *Economic Origins of Dictatorship and Democracy* (Cambridge, UK: Cambridge University Press, 2006).

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared, "Income and Democracy," *American Economic Review*, XCVIII (2008), 808-842.

Aghion, Philippe, Alberto Alesina, and Francesco Trebbi, "Endogenous Political Institutions," *Quarterly Journal of Economics*, CXIX (2004), 565-611.

Ahlerup, Pelle, and Ola Olsson, "The Roots of Ethnic Diversity," (2008) University of Gothenburg manuscript.

Ai, Chunrong, and Edward C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters*, LXXX (2003), 123-129.

Aidt, Toke S., and Francesco Giovannoni, "Critical Decisions and Constitutional Rules," (2005) CWPE Working Paper 0523.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, "Fractionalization," *Journal of Economic Growth*, VIII (2003), 155-194.

Apenchenko, Vladimir S., and Bruk I. Solomon, *Atlas Narodov Mira*, (Moscow: 1964).

Balas, Aron, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, "The Divergence of Legal Procedures," *American Economic J.: Economic Policy*, forthcoming, 2009.

Bankowski, Zenon, Neil D. MacCormick, and Geoffrey Marshall, "Precedent in the United Kingdom," in *Interpreting Precedents: A Comparative Study*, Neil D. MacCormick, and R. S. Summers, ed. (Ashgate/Dartmouth: Aldershot, 1997).

Bentham, Jeremy, *A Fragment on Government* (Cambridge, UK: Cambridge University Press, 1891).

Berkowitz, Daniel, Katharina Pistor, and Jean-Francois Richard, "Economic Development, Legality, and the Transplant Effect," *European Economic Review*, XLVII (2003), 165-195.

Cardozo, Benjamin, N., *The Nature of the Judicial Process* (New Haven, CT: Yale University Press, 1921).

Cavalli-Sforza, Luca L., Paolo Menozzi, and Alberto Piazza, *The History and Geography of Human Genes* (Princeton, NY: Princeton University Press, 1994).

CIA, *The CIA World Factbook* (Langley, VA: CIA Press, 2008).

Coase, Ronald, "The Problem of Social Cost," *J. of Law and Economics*, III (1960), 1-44.

Damaška, Mirjan R., *The Faces of Justice and State Authority* (New Haven, CT: Yale University Press, 1986).

Depoorter, Ben, Vincy Fon, and Francesco Parisi, "Litigation, Judicial Path-dependence, and Legal Change," *European Journal of Law and Economics*, XX (2005), 43-56.

Desmet, Klaus, Michel Le Breton, Ignacio Ortuno Ortin, and Shlomo Weber, "Nation Formation and Genetic Diversity," (2008) CEPR Discussion Paper 5918.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, "Courts," *Quarterly Journal of Economics*, CXVIII (2003), 453-517.

Encyclopedia Britannica, *Encyclopedia Britannica* (http://www.britannica.com, 2008).

Felli, Leonardo, and Antonio Merlo, "Endogenous Lobbying," *Journal of the European Economic Association*, IV (2006), 180-215.

Fernandez, Patricio, A., and Giacomo A. M. Ponzetto, "Case Law vs. Statute Law: an Evolutionary Comparison," *Journal of Legal Studies*, XXXVII (2008a), 379-430.

Fernandez, Patricio, A., and Giacomo A. M. Ponzetto, "Stare Decisis: Rhetoric and Substance," (2008b) Harvard University working paper.

Gennaioli, Nicola, and Andrei Shleifer, "The Evolution of Common Law," *Journal of Political Economy*, CXV (2007a), 180-215.

Gennaioli, Nicola, and Andrei Shleifer, "Overruling and the Instability of Law," *Journal of Comparative Economics*, XXXV (2007b), 309-328.

Gennaioli, Nicola, and Andrei Shleifer, "Judicial Fact Discretion," *Journal of Legal Studies*, XXXVII (2008), 1-36.

Glaeser, Edward L., and Bruce I. Sacerdote, "Sentencing in Homicide Cases and the Role of Vengeance," *Journal of Legal Studies*, XXXII (2003), 363-383.

Glaeser, Edward, and Andrei Shleifer, "Legal Origins," *Quarterly Journal of Economics*, CXVII (2002), 1193-1229.

Glendon, Mary Ann, Michael W. Gordon, and Paolo G. Carozza, *Comparative Legal Traditions in a Nutshell* (St. Paul, MN: West Group, 1999).

Gonzales, Michael J., *The Mexican Revolution: 1910-1940* (Albuquerque: University of New Mexico Press, 2002).

Grechenig, Kristoffel, and Martin Gelter, "The Transatlantic Divergence in Legal Thought: American Law and Economics vs. German Doctrinalism," *Hastings International and Comparative Law Review*, XXXI (2008), 295-360.

Grossman, Gene, and E. Helpman, *Special Interest Politics* (Cambridge MA: MIT Press, 2001).

Guerriero, Carmine, "The Evolution of Legal Systems: on Heterogeneity, Vengeance and the Moral Sentiments," (2008) Cambridge University manuscript.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Long Term Persistence," (2008) University of Chicago manuscript.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural Biases in Economic Exchange?" *Quarterly Journal of Economics*, forthcoming, 2009.

Hayek, Friedrich, A., *The Constitution of Liberty* (South Bend, IN: Gateway, 1960).

Herrmann, Benedikt, Christian Thöni, and Simon Gächter, "Antisocial Punishment Across Societies," *Science*, CCCIXX (2008), 1362-1367.

Heston, Alan, Robert Summers, and Bettina Aten, *Penn World Table Version 6.2* (CICPIP at the University of Pennsylvania, 2006).

Hobbes, Thomas, *A Dialogue Between a Philosopher and a Student of the Common Laws of England* [1681] (New York: Oxford University Press, 2005).

Kashima, Emiko S., and Yoshihisa Kashima, "Culture and Language: The Case of Cultural Dimensions and Personal Pronoun Use," *J. of Cross-Cultural Psychology*, XXIX (1998), 461-487.

Klerman, Daniel, and Paul G. Mahoney, "Legal origins?" *Journal of Comparative Economics*, XXXV (2007), 278293.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny, "The Quality of Government," *J. of Law, Economics and Organization*, XV (1999), 222-279.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, "The Economic Consequences of Legal Origins," *J. of Economic Literature*, XLVI (2008), 285-332.

Licht, Amir N., Chanan Goldschmidt, and Shalom H. Schwartz, "Culture Rules: The Foundations of the Rule of Law and other Norms of Governance," *Journal of Comparative Economics*, XXXV (2007), 659-688.

Marshall, Monty G., and Keith Jaggers, "Political Regime Characteristics and Transitions, 1800-2002," College Park, MD: Polity IV Project, University of Maryland, 2004.

Merryman, John H., *The Civil Law Tradition* (Stanford, CA: Stanford University Press, 1969).

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity: Theory and Evidence," (2008) Brown University manuscript.

Mitra, Devashish, "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination," *American Economic Review*, LXXXIX (1999), 1116-1134.

Niblett, Anthony, Richard Posner, and Andrei Shleifer, "The Evolution of a Legal Rule," (2008) Harvard University manuscript.

Nordhaus, William et al., "The G-Econ Database on Gridded Output: Methods and Data," (2006) Yale University manuscript.

Olson, Mancur, *The Logic of Collective Action* (Cambridge, MA: Harvard University Press, 1965).

Persson, Torsten, and Guido Tabellini, *The Economic Effects of Constitutions*, (Cambridge MA: MIT Press, Munich Lectures in Economics, 2003).

Persson, Torsten, and Guido Tabellini, "Democratic capital: The Nexus of Political and Economic Change," *American Economic J.: Macroeconomics*, forthcoming (2009).

Posner, Richard A., *The Economic Analysis of Law*, [1st ed. 1973] 7th ed. (New York, NY: Aspen Publishers, 2007).

Roe, Mark, J., *Convergence and Persistence in Corporate Governance* (Cambridge UK: Cambridge University Press, 2004).

Roe, Mark, J., "Juries and the Political Economy of Legal Origin," *Journal of Comparative Economics*, XXXV (2007), 294-308.

Rosenthal, Howard, and Erik Voeten, "Measuring Legal Systems," *Journal of Comparative Economics*, XXXV (2007), 711-728.

Spolaore, Enrico, and Romain Wacziarg, "The Diffusion of Development," *Quarterly J. of Economics*, forthcoming (2009).

Stone, Julius, *Legal Systems and Lawyers Reasonings* (Stanford, CA: Stanford University Press, 1964).

Tabellini, Guido, "Institutions and Culture," *J. of the European Economic Association*, VI (2008), 255-294.

Vanhanen, Tatu, *Democratization: A Comparative Analysis of 170 Countries* (New York, NY: Routledge, 2003).

Von Mehren, Arthur T., *The Civil Law System* (Englewood Cliffs, NJ: Prentice Hall, 1957).

World Bank, *World Development Indicators* (Washington, DC: World Bank, 2002).

Wright, John R., *Interest Groups and Congress* (New York, NY: Allyn and Bacon, 1996).

Zobler, Louis, *A World Soil File for Global Climate Modelling* (New York, NY: NASA Goddard Institute for Space Studies, 1986).

Zweigert, Konrad, and Hein Kötz, *Introduction to Comparative Law*, 3rd ed. (Oxford-New York: Oxford University Press, 1998).

Tables

Table I: Variables – Description and Source

Genetic_Distance =	The genetic distance between the transplanted country and the code/precedents drafter has been obtained as follows. First each country has been matched to the principal population set in the relative territory in year 1500 according to the historic information reported in CIA (2008) and in Encyclopedia Britannica (2008). Next, following the methodology illustrated in Spolaore and Wacziarg (2009), I matched each population to one of the macro groups for which the coancestry coefficient has been calculated by Cavalli-Sforza et al. (1994) in Table 2.3.1.A, page 75. The variable is normalized in order to range between 0 and 1.
Av_Elf =	Index of ethnolinguistic fractionalization, approximating the level of lack of ethnic and linguistic cohesion within a country, ranging from 0 (homogeneous) to 1 (strongly fractionalized) and averaging 5 different indexes for the 1960. Sources: La Porta et al. (1999) and, for Central and Eastern Europe, Persson and Tabellini (2003).
Constraint_Executive =	Data for 1850-2000 in Polity IV. The variable is calculated as the average of constraint on the executive over the period between the year of independence and 2000 (ignoring missing data). The original variable is normalized in order to range between 0 and 1. Sources: Polity IV project, Marshall and Jaggers (2004).
Time_Independence =	Number of years from independence to 2000, ranging from 0 to 250 (the latter value is used for all non-colonized countries). Sources: CIA (2008) and Encyclopedia Britannica (2008).
GDP =	Log of mean of real GDP per capita over the 1950-2000 period (Chain Method in 2000 prices). Source: Penn World Tables, Heston et al. (2006).
Population =	Mean of total population in thousands over the 1950-2000 period. Source: World Bank (2002), CIA (2008).
Soil_Diversity** =	The variable is calculated as 1 - (number of different types listed for each country/number of grid cells for each country).* Source: G-Econ dataset, Nordhaus et al. (2006).
Second_Person =	Dummy variable which equals 1 if the number of second person pronouns that might be used in spoken language varies according to the social proximity between speakers and 0 otherwise. Source: Kashima and Kashima (1998).
Altitude_Diff =	Maximum difference in mean grid-cell altitude.* Source: G-Econ dataset, Nordhaus et al. (2006).
Economic_Power =	The index of distribution of economic power resources is obtained adding the product of the value of the degree of decentralization of non-agricultural economic resources and of the share of non-agricultural population to the product of the family farm area and of the share of agricultural population. The variable is the mean over the 1918-1998 period and higher values implies a more egalitarian distribution of economic power. Source: Vanhanen (2003).

Notes   * In the G-Econ dataset variables are reported for each grid-cell within countries. Each grid-cell corresponds to an area of 1 degree latitude times 1 degree longitude, which is approximately 100km by 100km.
** The dominant types among 27 possible choices are listed for each grid-cell.

Table II: Variables – Means and Standard Deviations

	Cultural Heterogeneity and Democracy			Instruments and Other Controls	
	Law_Making Sample (94)	Tradition Sample (130)		Law_Making Sample (94)	Tradition Sample (130)
Genetic_Distance	0.504	0.442	Soil_Diversity	0.787	0.783
	(0.288)	(0.286)		(0.195)	(0.215)
Av_Elf	0.358	0.320	Second_Person	0.723	0.692
	(0.309)	(0.295)		(0.450)	(0.463)
Constraint_Executive	0.420	0.488	Altitude_Diff	1.276	1.496
	(0.315)	(0.318)		(1.203)	(1.345)
			Economic_Power	36.846	37.551
				(15.738)	(17.409)

Table III: Statute Law versus Case Law - Multinomial Logit

PANEL A

	(1)		(2)		(3)	
	The dependent variable is the likelihood of a reform toward:					
	Case law	Statute law	Case law	Statute law	Case law	Statute law
Genetic_Distance	0.153	- 0.185			0.186	- 0.153
	(0.052)***	(0.079)**			(0.078)**	(0.075)**
Av_Elf			0.109	- 0.197	0.051	- 0.167
			(0.087)	(0.090)**	(0.072)	(0.083)**
Time_Independence	0.0004	0.0003	0.0007	0.0003	0.0006	0.0003
	(0.0002)**	(0.0004)	(0.0003)**	(0.0004)	(0.0003)**	(0.0004)
Estimation	Multinomial Logit					
Other Controls	None		GDP, Population			
Number of Obs.	156		132		132	
Pseudo R^2	0.09		0.10		0.13	
Log Pseudo-Likelihood	-161.868		-134.926		-130.801	

PANEL B

	(1)		(2)		(3)		(4)	
	The dependent variable is the likelihood of a reform toward:							
	Case law	Statute law	Case law	Statute law	Case law	Statute law	Case law	Statute law
Genetic_Distance			0.222	- 0.263			0.195	- 0.246
			(0.090)**	(0.113)**			(0.089)**	(0.108)**
Av_Elf					0.145	- 0.165	0.105	- 0.098
					(0.112)	(0.096)*	(0.100)	(0.088)
Constraint_Executive	- 0.253	- 0.371	- 0.157	- 0.377	- 0.321	- 0.197	- 0.227	- 0.240
	(0.110)**	(0.114)***	(0.087)*	(0.122)***	(0.128)**	(0.174)	(0.112)**	(0.166)
Estimation	Multinomial Logit							
Controls	*Time_Independence, GDP, Population*							
Number of Obs.	127		127		115		115	
Pseudo R^2	0.21		0.26		0.21		0.25	
Log Pseudo-Likelihood	- 116.556		- 109.082		- 104.775		- 100.049	

Notes:
1. Robust standard errors–z distribution–in parentheses;
2. *** denotes significant at the 1% confidence level; **, 5%; *, 10%.

Table IV: Statute Law versus Case Law - IV Probit

	(1)	(2)	(3)	(4)	(5)
	The dependent variable is the likelihood of a reform toward:				
	Case law	*Case law*	*Case law*	*Case law*	*Case law*
Genetic_Distance	0.064 (0.175)	- 0.425 (0.182)**	0.148 (0.012)***	0.126 (0.046)***	0.124 (0.0004)***
Av_Elf	0.844 (0.794)	0.160 (0.189)	0.092 (0.008)***	0.071 (0.026)***	0.087 (0.0000)***
Constraint_Executive	- 0.313 (0.186)*	- 1.644 (0.143)***	- 0.166 (0.014)***	- 0.172 (0.063)***	- 0.155 (0.0006)***
Altitude_Diff				0.034 (0.021)*	
Economic_Power					- 0.001 (0.0000)***
First Stage for *Av_Elf*					
Soil_Diversity	0.282 (0.141)**		0.291 (0.142)**	0.280 (0.149)*	0.298 (0.144)**
Second_Person			- 0.087 (0.061)	- 0.083 (0.062)	- 0.091 (0.063)
R^2 *in the First Stage*	0.44		0.45	0.45	0.45
First Stage for *Constraint_Executive*					
Soil_Diversity			0.128 (0.147)	0.143 (0.152)	0.106 (0.143)
Second_Person		- 0.275 (0.062)***	- 0.280 (0.060)***	- 0.285 (0.061)***	- 0.267 (0.062)***
R^2 *in the First Stage*		0.50	0.50	0.50	0.52
Estimation	IV Probit				
Other Controls	*Time_Independence, GDP, Population*				
Endogenous	*Av_Elf*	*Constraint_Executive*	*Av_Elf, Constraint_Executive*		
Instruments	*Soil_Diversity*	*Second_Person*	*Soil_Diversity, Second_Person*		
Num. of Obs.	94	94	94	94	94
Exogeneity Test	0.21	0.00	0.98	0.89	0.99
Log Pseudo-Lik.	- 20.700	- 13.755	- 12.377	- 11.004	- 10.045

Notes:
1. Robust standard errors in parentheses;
2. *** denotes significant at the 1% confidence level; **, 5%; *, 10%;
3. Control variables used in the second stage are also included in the first stage but not reported to save space;
4. In the exogeneity test row the p-value of a joint test of a zero covariance between the two reduced form equations' errors and the structural equation's error is reported.

Table V: Building the *Convergence* Measure

	Law_Making	*Appeal*	*Equity*	*Procedure*	*Evidence*
Factor Loadings	1 (fixed)	0.746 (0.403)*	0.548 (0.262)**	0.691 (0.310)**	0.109 (0.085)
Likelihood Ratio Test (4 degrees of freedom)			38.53 (0.00)		
Number of Units		980 (Level 1)		196 (Level 2)	
Factor Loadings	1 (fixed)	0.758 (0.409)*	0.543 (0.254)**	0.701 (0.317)**	
Likelihood Ratio Test (3 degrees of freedom)			3.23 (0.36)		
Number of Units		784 (Level 1)		196 (Level 2)	
Estimation	Item Response Model With Multinomial Logit Linking Function–Adaptive Quadrature				

Notes
1. Standard errors clustered at the country level in parentheses;
2. *** denotes significant at the 1% confidence level; **, 5%; *, 10%.
3. The null hypothesis of the likelihood ratio test is that all the factor loadings do not differ from 1. The relative p-value is reported in the parentheses.

Table VI: Civil Law versus Common Law - Multinomial Logit

PANEL A

	(1)		(2)		(3)	
	The dependent variable is the likelihood of a reform toward:					
	Common law	Civil law	Common law	Civil law	Common law	Civil law
Genetic_Distance	0.160	0.430			0.122	0.313
	(0.161)	(0.212)**			(0.196)	(0.235)
Av_Elf			0.217	0.159	0.183	0.048
			(0.227)	(0.226)	(0.256)	(0.241)
Estimation	Multinomial Logit					
Controls	*Time_Independence, GDP, Population*					
Number of Obs.	196		168		168	
Pseudo R^2	0.19		0.15		0.18	
Log Pseudo-Likelihood	- 173.422		- 155.381		- 150.313	

PANEL B

	(1)		(2)		(3)		(4)	
	The dependent variable is the likelihood of a reform toward:							
	Common law	Civil law	Common law	Civil law	Common law	Civil law	Common law	Civil law
Genetic_Distance			0.027	0.303			- 0.086	- 0.360
			(0.196)	(0.214)			(0.242)	(0.263)
Av_Elf					0.025	0.244	0.059	0.135
					(0.309)	(0.291)	(0.350)	(0.321)
Constraint_Executive	- 0.478	0.551	- 0.466	0.567	- 0.736	0.967	- 0.729	1.016
	(0.222)**	(0.286)*	(0.220)**	(0.285)**	(0.292)**	(0.350)***	(0.288)**	(0.338)***
Estimation	Multinomial Logit							
Controls	*Time_Independence, GDP, Population*							
Number of Obs.	164		164		148		148	
Pseudo R^2	0.21		0.22		0.25		0.26	
Log Pseudo-Likelihood	- 142.938		- 140.274		- 121.920		- 119.547	

Notes: 1.   Robust standard errors (z distribution) in parentheses;
2.   *** denotes significant at the 1% confidence level; **, 5%; *, 10%.

Table VII: Civil Law Versus Common Law - IV Probit

	(1)	(2)	(3)	(4)	(5)
	The dependent variable is the likelihood of a reform toward:				
	Common law	Common law	Common law	Common law	Common law
Genetic_Distance	- 0.056	- 0.353	- 0.056	- 0.088	- 0.172
	(0.267)	(0.233)	(0.272)	(0.013)***	(0.211)
Av_Elf	- 0.010	0.205	- 0.011	- 0.050	0.012
	(0.369)	(0.342)	(0.339)	(0.007)***	(0.001)***
Constraint_Executive	- 0.589	- 1.689	- 0.589	- 0.608	- 0.525
	(0.241)**	(0.212)***	(0.271)**	(0.090)***	(0.065)***
Altitude_Diff				0.084	
				(0.012)***	
Economic_Power					- 0.006
					(0.0007)***
Estimation	IV Probit				
Other Controls	*Time_Independence, GDP, Population*				
Endogenous	*Av_Elf*	*Constraint_Executive*	*Av_Elf, Constraint_Executive*		
Instruments	*Soil_Diversity*	*Second_Person*	*Soil_Diversity, Second_Person*		
Number of Obs.	130	130	130	130	130
Exogeneity Test	0.97	0.00	0.99	0.62	0.99
Log Pseudo-Lik.	- 42.427	- 44.317	- 26.614	- 23.403	- 24.753

Notes: 1.   Robust standard errors in parentheses;
2.   *** denotes significant at the 1% confidence level; **, 5%; *, 10%;
3.   Control variables used in the second stage are also included in the first stage but not reported to save space;
4.   In the exogeneity test row the p-value of a joint test of a zero covariance between the two reduced form equations' errors and the structural equation's error is reported.