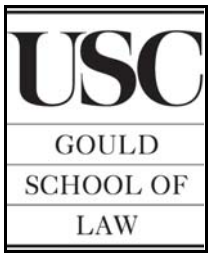
What is Law? A Coordination Model of the Characteristics of Legal Order

Gillian K. Hadfield and Barry R. Weingast

USC Center in Law, Economics and Organization Research Paper No. C10-17
USC Legal Studies Research Paper No. 10-20



CENTER IN LAW, ECONOMICS AND ORGANIZATION RESEARCH PAPER SERIES and LEGAL STUDIES RESEARCH PAPER SERIES

University of Southern California Law School
Los Angeles, CA 90089-0071

What is Law? A Coordination Model of the Characteristics of Legal Order

Gillian K. Hadfield and Barry R. Weingast

University of Southern California and Stanford University

October 2010

Abstract

Legal philosophers have long debated the question, what is law? When can we say that a society is organized as a legal order as opposed to some other type of order such as order based on religious authority, moral principles, emergent social norms, or tyranny? This question is of both theoretical interest and political and economic interest, as countries seek to transition from the rule of power or privilege to the rule of law to build market democracies and generate economic growth. We present a model that seeks to explain the distinctive characteristics of law–such as its generality, abstract reasoning, uniqueness and reliance on open and public processes–on the basis of law's function to coordinate an equilibrium based on decentralized enforcement of rules. We thus depart from the conventional assumption in both law and economics and positive political theory, namely that law is to be defined as a system of coercive enforcement of penalties by the state. We find that the capacity of law–meaning third-party classification of behaviors as wrongful or not–to coordinate enforcement depends on the ability of law to provide unambiguous classifications. Many of the features identified by legal philosophers as characteristic of legal orders, and some that legal theory ignores or deemphasizes, are predicted by the model.

I. Introduction

Legal philosophers have long debated the question, what is law? These theorists are primarily interested in the question of how law is, or is not, distinct from morality: Is an unjust law invalid? Or is any law that meets the socially accepted criteria for validity, which may not include a requirement of justice, properly called a law? In the process of wrestling with the relationship between law and morality, these theorists have developed a significant degree of consensus on the characteristics of legal order. Law, it is generally thought, is a system of governance in which a recognized authority with the power of coercive enforcement establishes rules for behavior that are characteristically *general, stable, public, prospective, clear, not contradictory* and *capable of being followed* and in which enforcement actions are based only on *faithful application* of the rules, not distorted by personal or otherwise extraneous considerations. (This list is from Fuller (1964).)

In this legal philosophical tradition, the elements of law are identified through a combination of moral reasoning and conceptual analysis and in particular are focused on the relationship between the governing authority and the governed. Legal theorists derive their list of characteristics from consideration of what the rules and the legal system that enforces them must look like in order for those subject to them to be reasonably guided by them. For many legal positivists, the requirement that rules be written so as to allow people to plan their conduct around them is necessary as a practical matter (Raz 1977). For modern natural law theorists, the requirement that rules allow people to plan also has moral elements. Luban (2007, 111) interprets Fuller's (1964) requirement of "generality," for example, to derive from the moral obligation of a governor to "treat the citizen as a self-determining agent." More expansively, Finnis (2009, 15-18) declares that "the fundamental equality and dignity of human beings should be defended as part of a rationally sound understanding (concept) of law" because purely positive law is enacted according to socially recognized systems of validity, the "proximate starting point [of which] is the moral need for justice and peace."

This careful normative and conceptual attention to the question of what counts as law

by legal philosophers has not been matched by attention from social scientists. Neither economic nor positive political theory treatments of law offer an account of what law is, why law looks like law or how to distinguish distinctively legal order from other forms of social order such as those based on purely political decisionmaking, spontaneous social norms, or tyranny and organized violence. Social scientists have not thus far sought either to supplement or to contest the legal philosophers' accounts of what constitutes a legal order. We have social science treatments of how more or less general rules might impact efficiency, such as the 'rules versus standards' literature (Kaplow 1992); and explanations about how political circumstances affect doctrine, as in Marks's (1988) explanation for the survival of a narrow reading of Title VII of the Education Act. We now have a relatively large literature on the economic benefits that ostensibly flow from different types of legal regimes, notably in connection with a comparison of the evolution and attributes of common law and civil code regimes (Posner 1977, La Porta et al 1998, Glaeser & Shleifer 2002, Hadfield 2008 & forthcoming). But these literatures largely focus on the economic and political determinants of the substantive content of legal rules. We have no accounts of the economic or political factors that generate an institution that is characterized by general, open and public rules or the rule of law more generally; or by distinctive processes of public, legal reasoning.

Nor do we have systematic treatment of the question of whether the availability of a credible threat of coercive enforcement by a third-party institution is an analytically helpful component of the definition of law. Both law and economics and PPT and the law adopt uncritically the premise that law is what we see today in the developed world: the product of recognized lawmakers (legislatures, judges, voters in a popular referendum) enforced by the coercive power of the state which authorizes a judge to impose penalties. But tying the definition of law to the exercise of government force is a shortcoming of these approaches for both empirical and analytical reasons.

Empirically, many settings exist in which law–in the form of legal reasoning–plays a central role in ordering relationships, despite the absence of a credible threat of coercive

enforcement. The people of Buddhist Tibet prior to 1949 lived under what was clearly a legal system with rules, judges and official processes to resolve disputes, but both the jurisdiction and the judgments of a court had to be consented to by both parties (French 2002, 138). From the tenth to the thirteenth century, Iceland had a developed system consisting of a legislature that enacted rules, a "Law Speaker" who committed customs and rules to memory, recited them publicly, and served as the final word on resolving disputes about the content of customs and rules, and a hierarchical system of courts to resolve disputes under the customs and rules as articulated by the Law Speaker; but it had no centralized system of enforcement (Bryce 1901, Friedman 1979). Similarly, in medieval Europe, there was a wide range of institutions that articulated rules governing trade, many of which relied extensively on non-state enforcement (Mitchell 1904, Milgrom, North & Weingast 1990, Grief 2006) and it is to these beginnings that we trace the development of modern commercial law. International law is characterized by distinctive and recognizable forms of legal reasoning and the use of legal procedures and tribunals, even though there is often no authority capable of enforcing these legal judgments. In many if not most of our daily interactions as citizens and economic agents there is little likelihood of coercive penalties and yet we look to laws and written contracts and legal advice in order to guide our actions: businesses spend significant resources on legal services to draft and interpret contracts even when they recognize that they are extremely unlikely to end up in litigation (Hadfield & Bozovic, in process).

Analytically, by starting with the assumption that law consists of rules coercively enforced by a government, we limit our ability to explain how legal order emerges historically and in settings in which there is no established and legitimate state monopoly on force. We clearly need a model of 'law' that recognizes medieval Iceland's decentralized enforcement of distinctively legal rules as a legal system: it accomplishes everything, and perhaps more, than a system that relies on centralized enforcement can accomplish. In contemporary settings, we need to understand the implications for legal order of the state's limited capacity to enforce compliance through sanction and threat, as the efforts to generate legality inthe

former soviet states demonstrates. To respond to the demand for legal order in a globalizing economy, with no supra-national coercive government, we need models of how law works that explain if and when coercion is necessary. We need a way to distinguish the role and emergence of organic social norms such as reciprocity from the role and emergence of structured, policy-susceptible, legal norms to shape voluntary interactions.

In this paper we initiate the project of filling the gap in law and economics and PPT by developing a positive model of the conditions under which a stable equilbrium can emerge based on an institution with the attributes legal theorists have suggested are necessary for it to be recognized as "legal". To focus the analysis on the structural rather than the third-party enforcement characteristics of law, we assume that the enforcement of a legal rule is exclusively based on community punishment of a rule violation, as in Milgrom, North and Weingast (1990) and Greif (2006). This punishment might, for example, be a boycott or physical retaliation.

Community enforcement presents potential victims of wrongs with a challenge. What actions should be punished? What is the appropriate definition of right and wrong? These questions present a complex coordination problem. Members of the community need to make simultaneous decisions and predictions about how others will act. But heterogeneity among people, including not only their preferences and circumstances but also their knowledge of the particular action, makes coordination problematic. Ambiguities and new circumstances further plague coordination: in the face of new circumstances, for example, multiple interpretations exist about how to extend existing ideas about right and wrong. In the face of these problems, an individual will find it difficult to predict both whether a sufficient number of others will see the action as wrongful and whether they will be willing to participate in a punishment.

We argue that law has its distinctive structure in order to serve as an ambiguity-reducing institution that coordinates beliefs among diverse individuals and thus to improve the efficacy of the extra-legal rule enforcement mechanisms that cause behavior to align with rules. We

contend that a designated third-party system of specialized public reasoning–which we call a logic–helps coordinate beliefs. To perform this function, this third-party logic must possess particular structural traits. Some of these traits derive from the need to provide a common knowledge focal point; the logic must for these reasons be a system of authoritative and unique classifications. Other traits derive from the need for the system to be incentive-compatible. In our model, it is key that the system be one in which all of the required participants are willing to participate. This participation requires a degree of convergence between the classifications offered by the public system and those reached on the basis of private idiosyncratic reasoning.

A central piece of our argument thus concerns a question that has arisen in several different strands in the literature: when will an individual punish wrongs against others despite the fact that punishment is costly? Greif (1994) suggests that participation in community punishments is supported by cultural beliefs or preferences for helping others. Milgrom, North & Weingast (1990) posit standard subgame perfect equilibrium strategies where people who fail to punish are themselves punished. Evolutionary game theorists appeal to a biological predisposition to punish wrongs (Boyd, Gintis & Bowles 2010). An important line of experimental work on altruistic punishment looks to the presence of negative emotions such as anger towards wrong-doers (Fehr & Gachter 2002).

We take a different approach in which individuals are concerned about the beliefs of others; specifically, beliefs held by those who might act wrongfully toward them in the future, and those who might participate in punishing a wrongdoer in the future. Suppose there exists a system of reasoning, R, to judge right and wrongful actions. By participating in a punishment based on R, an individual j can affect others' beliefs so as to influence the likelihood that they participate in a punishment in the future; in particular, that they will participate in a punishment following a wrongful action–as judged by R–against j. Participation in the punishment also affects the beliefs of a potential wrongdoer, specifically about the likelihood that an action judged wrongful under R will result in successfully

coordinated punishment.

We show that there exist circumstances where it is rational for j to undertake costly actions to change the beliefs of others (e.g., both potential victims and perpetrators of wrongs) about the likelihood of coordinated punishment of wrongs (against anyone) defined by R. In this way, j's actions help create a coordination equilibrium: acting purely on the basis of self-interest and without any inherent preferences for punishment or cooperation, j will undertake costly actions to change people's beliefs when the benefits to j of coordinating under R exceed j's costs. We claim that this is a central explanation for the emergence of law understood as a distinctive system of reasoning.

This perspective on law – emphasizing law and legal procedure as coordinating diverse people's actions and reactions – provides a new account of the characteristics of law. We demonstrate that several of the features conventionally thought of as distinctive of a legal order can be explained without an appeal to either the practical or moral demands on a third-party enforcer who seeks to exercise control over a subject population. We also identify features that are either missing or de-emphasized in conventional legal theory.

Consider generality. Whereas legal philosophers argue that this is a necessary attribute of law in order to provide adequate guidance to a wide range of people and in a variety of circumstances, we show that a rule must be general enough to cover a wide range of individuals and circumstances in order to attract participants in a coordination equilibrium. That is, generality is necessary to satisfy the incentive constraint on the delivery of effective extra-legal punishment. Rules narrowly focused on a subset of victims are unlikely to gain the support of a wide range of potential victims.

Similarly, we demonstrate that a legal rule must be relatively stable. But in contrast to conventional legal theory, because the need for stability is grounded in the requirements of coordinated punishment among potential victims, our model predicts that law must be stable over a longer time horizon that that strictly needed to allow people to conform their behavior to rules.

Our model also emphasizes distinctive characteristics of law that conventional legal theory has yet to treat as central to the concept of law. We argue, for example, that in order to support coordinated punishment, law must result in a unique classification of actions as right or wrong–even in the face of ambiguity. This predicts that law will be under the stewardship of an identifiable person or entity and with a recognized designation of an ultimately "right" answer to the question of how to classify actions or extend existing classifications to new circumstances. Medieval Iceland 's single Law Speaker, a closed body of professionals authorized to engage in legal analysis and a Supreme Court are examples of institutions that can be understood to have emerged to meet this need.

We highlight another feature of the coordination model, namely, that law must be a system that consists of open, impersonal and public reasoning. In order to be unique in the face of heterogeneity and the need for convergence between idiosynratic and public reasoning, the classifications reached by this coordinating device must allow individuals to present their facts and reasons in the process of elaborating a set of rules and principles to address new or otherwise private circumstances. Coordination also suggests that this process of elaboration needs to be a public process, available to those who must then consult the system and believe that the classifications of the system are common knowledge.

The uniqueness necessary to achieve coordination also implies that legal reasoning must be impersonal reasoning: the classification of conduct must not depend on the identity of the person or entity implementing the reasoning. Unless the ultimate steward of the reasoning system is able to advise everyone who must consult the system to determine their actions with respect to compliance and punishment, others must be able to reliably reach the same classifications as the ultimate arbiter does.

Our paper is organized as follows. In Section II we relate our approach to other coordination odels of law in economics and political science. Section III presents the model and we demonstrate how coordination of extra-legal punishments can lead to an equilibrium organized around a system of reasoning offered by a third-party institution that lacks the

capacity to enforce its judgments about right and wrongful behavior. We then discuss in Section IV what this simple model suggests about the characteristics required of law. Section V concludes.

II. Relationship to the Literature

We are not the first to propose that law and legal institutions can play the role of a coordination device to improve welfare. Our model, however, derives a richer set of implications for the structure of law than do the models proposed by others.

Sugden ([1986] 2005) self-consciously echoing Hume (1739-40), relies on focal point equilibria to explain the spontaneous emergence of self-enforcing conventions (without deliberate design or legal institutions) about coordination, reciprocity and property rights to resolve rival claimants disputes. Binmore (1994, 1998) also approaches the problem of explaining the emergence of conceptions of justice–particularly fairness–as the resolution of a coordination problem in which the equilibrium must be self-enforcing. Dixit (2004) considers multiple settings in which coordination can be achieved by extra-legal conventions, including focal point settings. Cooter (1998)and McAdams (2000) apply these ideas directly to law and the capacity of law as a focal institution to deliberately select an equilibrium in a coordination game where players judge themselves to be better off coordinating than not with their transactional partner. McAdams notes in particular the qualities of law that can confer the attributes required by a focal point, namely publicity and uniqueness.

Myerson (2004) also proposes that a lawgiver can coordinate agents by announcing an equilibrium set of strategies for the agents to pursue. Provided the lawgiver has the qualities of a focal point, he argues, we can expect the agents to coordinate on the equilibrium strategies proposed by the lawgiver. Like Sugden (and Hume), Myerson observes that a rule that deemed the immediate possessor of a piece of a property to be its rightful owner, for example, can coordinate rival claimants so as to avoid wasteful contests over the property. If both claimants expect the other to apply a concept of 'rightful' ownership, then the 'rightful'

owner will rationally claim and the other will rationally recede. Whereas Sugden and Hume look to the spontaneous emergence of this rule, however, Myerson considers the role for an institution such as a legislative assembly which selects generally understood principles to coordinate expectations about who will rightfully claim what. He also, like McAdams, considers the role for an arbitrator who recommends an equilibrium when general principles do not cover the situation or are ambiguous. As McAdams (2000, 1700) puts it, "legal process 'perfects' the convention" by resolving factual disputes about (for example) "possession" in a property rule that makes rightful ownership dependent on "possession."

These approaches locate the coordination problem in the underlying economic game being played by the agents seeking to coordinate. Basu (2000) and Mailath (2001, 2007) focus instead, as we do, on the impact of legal rules on beliefs among not (only) the parties to a transaction but also among those who have a role in enforcing the rules. For these scholars, enforcers are understood to mean officials such as judges and police. Based on this approach, Basu proposes that coordination by law is really no different from coordination by any self-enforcing social norm.[1] Mailath et al use their coordination model to define "authority," particularly legal authority, as the capacity to make "cheap talk" (Farrell & Rabin 1996) announcements that are effective in coordinating behavior. They pay particular attention to the idea that the authority of a legal entity (such as the Supreme Court) can be undermined by attributes of the entity's announcements, such as how far they diverge from public opinion (although they do not model this effect.)

Greif (1994, 2006) is closest to our work. He looks at how cultural beliefs about the punishment of those who defect from expected behavior (such as performance on a contract) can work as focal points to coordinate equilibria: where agents entertain the belief that defectors will be punished by the group, the group can coordinate on an equilibrium without defection; where such beliefs are missing, inducing compliance requires external coercive institutions. He draws on this analysis to explain why the Maghribi who entertained Muslim and Jewish cultural beliefs about the obligation to help others coordinated on an

equilibrium without third-party coercive enforcement whereas the Genoese, operating in a Christian individualist culture, selected a different equilibrium, one that required formal legal institutions and coercive enforcement.

Positive political theory and the law has long recognized the importance of coordination in one aspect of the law, namely, constitutional law with a focus on constitutional stability. Most new constitutions fail (Elkins, Ginsburgh and Melton 2008), so why do those few survive? The literature provides some insight into this question, and it concerns the central issue of this paper, coordination. Hardin (1989, 2006), following Hume (1739-40), argues that the central feature of constitutions is to provide coordination for citizens around various rules (see also Ordeshook 1992). Constitutions, in this view, create focal solutions that allow citizens to create order. In a model closely paralleling that in this paper, Weingast (1997) argues that constitutional stability requires that citizens have the ability to coordinate against governments that seek to transgress constitutional provisions. To do this, citizens must create focal solutions to the problem of what features of the constitution are worth defending. Constitutions that become focal points (typically in moments of crisis) have greater ability to survive then ones that do not. Similarly, Fearon (2006) argues for the coordination effect of elections in democratic (and hence democratic constitutional) stability.

These approaches in the literature bear many similarities to our own but our approach significantly expands the explanatory scope and power of a coordination model of law. It does so because 1) it focuses on the coordination of distributed enforcement of rules, thereby expanding the relevance of the theory beyond underlying transactions that are themselves coordination games; 2) it focuses on explaining law not in terms of the substantive content of rules (recommended equilibrium strategies) but rather in terms of the attributes of law that make "law" a distinct and identifiable form of social ordering, apart from the reliance on coercive enforcement by an organized state; and 3) it pays express attention to the incentives of agents to participate in an equilibrium potentially coordinated by legal classification, without attributing the incentive to participate to exogenous preferences or beliefs

(culture), biological or emotional dispositions to punish wrongdoers, or meta-level punishment strategies (punishing non-punishers.) In doing so, we link the attributes of law not only to their capacity to serve as focal points but also to their capacity to attract participation in distributed enforcement.

III. Model

Assume there are two infinitely-lived buyers, A & B who in each period $t = 1, ..., \infty$ purchase a good from an infinitely-lived seller. Future profits are discounted with a common discount factor, δ. Buyers value the good at V and contract with the seller to pay a price $P < V$ prior to delivery. The seller incurs a cost, c, to perform on the contract and deliver the good as promised. Let the seller's performance be characterized by an n-tuple $X = (x_1, x_{2,}...x_n)$ of factors with $x_1 \in \{A, B\}$ indicating the identity of the buyer. For each buyer j, let $\mathfrak{X}^j$ represent the set of n-tuples X in which the identity of the buyer $x_1 = j$.

The elements of X capture a wide variety of considerations relevant to the buyer's and the seller's assessment of the value of the deal: attributes of the seller and the buyer, the buyer's use for the good, discussions and correspondence at the time of contracting, promised delivery date, the history and nature of the relationship, the type and quality of good, the location of delivery, location of production, delivery method, insurance terms, risks of loss or damage including "force majeure" type risks, etc.

In each period and for each buyer, with probability θ, the seller has an opportunity to engage in a performance that is "wrongful" in some way. A key feature of this model is that we pay close attention to the inherent ambiguity of what it means for the seller's performance to be wrongful. A performance is deemed wrongful by a system of reasoning–a set of principles–or what we will call a *logic*. Formally a logic maps the potentially very large set of all possible X vectors into a binary $\{0, 1\}$classification of "wrongful" and "not wrongful."

Each buyer possesses an *idiosyncratic logic,* $I^j : \mathfrak{X}^j \rightarrow \{0, 1\}$, to assess whether the

seller's performance of a contract with that buyer is wrongful, that is, whether the buyer believes that the seller's contract obligated a delivery different in some way from the one performed by the seller. A performance might be judged by the buyer to be wrongful, for example, if nothing is delivered, if the wrong number or quality of the good is delivered, or if it is delivered late or at the wrong place or subject to conditions not included in the contract. This logic represents the buyer's assessment of the content of its own deals struck with the seller and the terms on which the buyer understands the seller to be obligated to deliver.

Each buyer's idiosyncratic logic is not accessible to others: others cannot reproduce the buyer's logic to predict how the buyer will analyze and hence categorize a given delivery failure. By designating this analysis as "idiosyncratic" we emphasize that the assessment is conditional on the buyer's particular circumstances and that the buyer's decision to buy is based on its own evaluation of the circumstances in which the deal is valuable. We posit idiosyncrasy as a source of value-generating diversity in an economy (Hong and Page 2001). [2] Idiosyncrasy may in part derive from private information and experience with factors that differs from others, but the buyer's logic is not 'information' in the sense of statements that can be unambiguously conveyed to others in a common language via a report or signal. The buyer's logic is a system of private reasoning to organize and analyze information and make judgments. Because people's situations and experiences differ, so too will their systems of judgment in ways that are not apparent to others. This is what makes idiosyncratic logic *ex ante* inaccessible to others.[3]

The seller's performances are observable to all. The seller keeps the payment P regardless of whether performance is judged wrongful or not. Buyers, after observing a potential delivery failure, to any buyer, can choose to boycott the seller and obtain a payoff of 0 for each period in which they boycott. The 2-period profit for the seller who sells to and avoids

wrongful delivery to both buyers is

$$2(P-c)+2\delta(P-c).$$

The 2 period profit for the seller who sells to both buyers, makes a wrongful delivery to one buyer and expects to lose a next-period sale to that buyer only is

$$P+(P-c)+\delta(P-c).$$

The 2 period profit for the seller who makes a wrongful delivery to a buyer and expects to lose 2 future sales is

$$P+(P-c).$$

The seller will not be deterred by a lone boycotter if

$$c>\frac{\delta}{(1+\delta)}P$$

and will be deterred by two boycotters if

$$c<\frac{2\delta}{(1+2\delta)}P.$$

We call a boycott that deters the seller from taking the opportunity to engage in a wrongful performance an *effective boycott* and assume that a boycott is only effective if both buyers participate in the boycott.

Condition 1. *Effective boycotting requires both buyers to simultaneously boycott:*

$$\frac{\delta}{1+\delta}P<c<\frac{2\delta}{1+2\delta}P. \tag{1}$$

We assume that both buyers will purchase the good in period t even if they anticipate

that a wrongful delivery will never result in an effective boycott. Expected profits in this case are

$$(1-\theta)V - P > 0$$

and both buyers will continue to purchase despite the lack of coordination if

$$\theta < \frac{V-P}{V}. \tag{2}$$

(This assumption makes the exposition simpler and eliminates the need to analyze the incentive for the seller to coordinate boycotts–something to be considered in future work.)

We assume an institutional environment as follows. No third-party institution exists that is capable of enforcing a penalty against the seller for a wrongful performance. There are, however, a variety of third-party institutions capable of classifying performances as wrongful or not and articulating reasons for their classifications–that is, implementing a logic. Examples of such institutions could include "English common law" "the Law Merchant" "the customs of this village as articulated by the elders" "rabbinical teachings" "the Dutch merchants' guild" "the Archibishop of Hamburg" and "the United States Supreme Court." Our goal is to identify the characteristics such an institution must have in order to coordinate the boycotting decisions of the buyers and effectively deter wrongful deliveries.

Assume that at $t = 0$ a seller has made a delivery to a buyer that would be classified as wrongful by at least one of the available third-party logics and that this has been observed by both buyers. In period 1, then, both buyers face a decision about whether to boycott the seller or not. Each buyer knows that the boycott will be effective only if the other buyer also boycotts. We assume that the buyers cannot get together to agree on a boycott; each decides independently whether to boycott.

As with any repeated coordination game, a subgame perfect Nash equilibrium to boycott can be supported by strategies under which buyers punish buyers who fail to participate in the boycott, who fail to punish buyers who fail to punish buyers who fail to participate in

the boycott, and so on. (This is the type of strategy that supports the coordinated boycott equilibrium posited in Milgrom, North & Weingast (1990).) We exclude such strategies both on principle (it is hard to provide reasons external to the equilibrium for engaging in the strategy) and because of our assumption that each buyer possesses an idiosyncratic logic for classifying non-deliveries as wrongful or not: there is no *a priori* unique common logic to determine when punishments should be delivered and when not. (Most game theory models of reputation and coordinated punishment, assume a unique common classification of actions as "cheating" or not.) We instead develop a model in which we can make it clear why an individual buyer might be willing to boycott based on the private benefit of a successful boycott, and we address explicitly how this incentive depends on the availability of a common logic with particular characteristics.

We allow a buyer when boycotting to make an announcement about the basis for the boycott. In particular, a boycotting buyer designates the logic under which it classifies the seller's performance as wrongful. Suppose, for example, that at $t = 1$ buyer A engages in a boycott and announces that it does so under a logic R. (The mnemonic is R for "reasoning".) A's announcement can be interpreted as an endorsement of R although we do not ascribe to A any inherent preferences (such as cultural beliefs) directly over R: A does not derive utility directly from the adoption or endorsement of R by itself or others, but only from the consequences of choices made from following (or not following) R. For ease of exposition we will refer to a performance classified as wrongful by R as "R-wrongful."

Clearly the only reason for A to engage in this behavior is if by doing so A changes the likelihood that the seller makes a wrongful delivery to A in the future. A solitary boycott, however, does not change the seller's expected payoff sufficiently to deter a wrongful delivery. Thus A's decision (including the announcement of R) must be premised on the impact of its actions on the seller's expectation of a two-buyer, effective, boycott in the future. This, in turn, depends on the impact of A's choice on B's beliefs and actions.

We start by examining B's decision in Period 2 following the observation at $t = 1$ that

A has boycotted and designated R as the logic under which A has done so. Suppose that A announces an R with the attribute that $R(X) = I^B(X)$ for all $X \in \mathfrak{X}^B$. That is, suppose A designated the logic underlying its boycott as the equivalent of B's idiosyncratic logic for cases involving B. And suppose that B believes that A will continue in the future to boycott any R-wrongful performance if A observes B to join in the boycott. What is B's incentive to join in the boycott in Period 2? Provided the seller believes that buyers who participate in an effective boycott will continue to boycott R-wrongful performances in the future, by joining A's boycott, B's choice to boycott under R changes the seller's beliefs about the likelihood of an effective boycott in the future. This expectation induces the seller to avoid performances that can trigger a boycott, that is R-wrongful performances. Because, in this example, R coincides with B's idiosyncratic assessment of when the seller has violated the contract with B, this implies that B secures the elimination of performances it judges to be wrongful, starting in Period 3. Formally, joining the boycott in Period 2 will then be worth it to B if

$$\sum_{t=3}^{t=\infty} \delta^{t-2} (V - P) > \sum_{t=2}^{t=\infty} \delta^{t-2} ((1-\theta)V - P). \tag{3}$$

Rewriting 3 we have the condition that, assuming that both A and the seller believe that once A and B have coordinated on R they will continue to coordinate on R, B will join the boycott if

$$\theta > (1-\delta)\frac{V-P}{V}. \tag{4}$$

That is, B will join the boycott based on its idiosyncratic logic provided θ, the risk of wrongful performances under that logic, is sufficiently great or the discount factor sufficiently close to 1. This result does not require that B be the victim of the original wrongful performance. B's incentive to boycott in response to a wrong to A arises because under the logic designated by A and with the assumption that when they have coordinated once they will continue to

coordinate in boycotting future wrongs under R, all R-wrongful performances are deterred. For B, this includes all potential wrongs done to B. This analysis demonstrates the potential for A to designate a logic and demonstrate a willingness to boycott so as to induce the expectation of a coordinated boycott in response to wrongs in the future.

We have no reason, however, to believe that A will in general want to coordinate boycotts for I^B-wrongful performances unless the logic that does so also coincides sufficiently with I^A on $\mathfrak{X}^A$. Moreover, by assumption, B's logic is presumed to be idiosyncratic and inaccessible to others. Thus even if A wanted to, it is not possible for A to deliberately designate a logic that replicates I^B on $\mathfrak{X}^B$. Nor is it possible for the seller to condition future delivery decisions on whether or not they will trigger a boycott coordinated by judgments reached using I^B.

What, then, must the characteristics of a logic be such that it can support an equilibrium in which the buyers are coordinated in effective boycotting and the seller is able to predict the effect of alternative performances on the likelihood of a boycott?

First, it must be the case that R is *general* in the sense that R is defined over $\mathfrak{X}^A \cup \mathfrak{X}^B$. It must address the potential circumstances of both A and B. Second, all players–the buyers and the seller–must be able to access or query the logic in order to determine how it would classify particular performances. This means that the logic must be both common knowledge and *publicly accessible* (with "public" defined as the buyers and the seller). This means, at a minimum that all the players must be able to implement the logic, perhaps with the aid of experts, to analyze possible configurations of circumstances: the set of possible X vectors. For this to be possible the logic must be relatively *clear* (again, perhaps with the use of expert assistance) and *not contradictory.* Moreover, because decisions by B to adopt R reflect payoffs over an inifinite time horizon, R must be *stable.* B's decision to join a boycott, for example, will depend on B's comparison of the costs of the boycott today as against the future benefit of the deterrence of performances judged wrongful under the announced logic R; this means R must be expected to be the relevant logic used by the

buyers and the seller in the future.

We will develop the implications of the model for the characteristics of R as an institution more fully in Section IV, below. For now we proceed to develop the formal model by assuming that A can designate an $R : \mathfrak{X}^A \cup \mathfrak{X}^B \rightarrow \{0, 1\}$ that possesses the minimal qualities of being publicly accessible and stable: all players can access the logic to determine how it would classify a performance vector X and all players expect R's classifications to remain unchanged in all future periods. We will call this a stable *common logic*.

Let r^j be a measure of the *convergence* between j's idiosyncratic logic and the common logic R. Formally, r^j is the likelihood that given an opportunity for wrongful performance, R classifies the performance in the same way that I^j does. [4] Evaluated as of Period 1, the expected payoff to buyer j if R, as of Period 3, coordinates the expectations of both buyers and the seller that an R-wrongful performance will be met with an effective boycott is then given by

$$\sum_{t=3}^{t=\infty} \delta^{t-1}(1 - (1 - r^j)\theta)V - P). \tag{5}$$

The expression in brackets preceding V in (5) reflects the probability that, with the threat of an effective boycott in response to R-wrongful performances, buyer j enjoys performance with value V. It does so in all cases except where the seller has an opportunity to enage in a performance that j's idiosyncratic logic judges to be wrongful but R does not. (Note that if $r^j = 1$, meaning that R is fully convergent with I^j, then the buyer always receives performance with value V.)

It is clear from (5) that A would prefer to designate a common logic R that is as convergent as possible with I^A, that is, with r^A close to 1. But A's effort to coordinate the market on R will only be successful if B chooses to join the boycott under R. This requires r^B sufficiently high to satisfy B. But A cannot assess r^B: only B can query the common logic to determine the convergence between R and I^B. We cannot, therefore, solve

A's problem as a strategic matter. A will have to test available common logics. We will therefore focus on the characteristics of an equilibrium solution.

In Period 1 A has to decide whether it is willing to boycott and announce that the boycott is based on a particular common logic R in the hope of getting B to join the boycott and inducing in the seller the expectation that R-wrongful performances will result in an effective boycott. Both A and the seller lack information about B's idiosyncratic logic and thus can only learn if R is sufficiently convergent with I^B if B joins the boycott in Period 2. A must plan, therefore, on sustaining a boycott of the seller for two periods before it can become common knowledge that R will coordinate an effective boycott in response to wrongful performance. We presume that A puts the probability that any given R is sufficiently convergent with I^B to induce B to endorse R and join the boycott at $\frac{1}{2}$.[5] If A is willing to boycott in Period 1, then, it must be that

$$(6)\quad \frac{1}{2}\sum_{t=3}^{t=\infty}\delta^{t-1}\left[(1-(1-r^A)\theta)V-P\right]+\frac{1}{2}\sum_{t=3}^{t=\infty}\delta^{t-1}\left[(1-\theta)V-P\right]>\sum_{t=1}^{t=\infty}\delta^{t-1}((1-\theta)V-P).$$

Rewriting, this gives us the following necessary condition for A's participation in the early stages of a sequence leading to an equilibrium coordinated by R:

Condition 2. *A necessary condition for buyer A to be willing to begin a boycott in Period 1 and designate a logic R as the basis for the boycott is that R is sufficiently convergent with* I^A*:*

$$\begin{aligned} r^A &\geq \frac{(1-\delta^2)}{\delta^2}\frac{(1-\theta)}{\theta}\frac{(V-P)}{V} \\ &\equiv r_1. \end{aligned}$$

Meeting Condition 2 requires that $0 \leq r_1 \leq 1$. The existence of such an r_1 can be seen by noting first that if $\delta^2+\theta=1$,

$$r_1 = 1-\frac{P}{V}.$$

Then, since $0 < \frac{P}{V} < \frac{(1-\theta)V}{V}$, we have $\theta < r_1 < 1$.

For B to be willing to join the boycott in Period 2, it must be that

$$\sum_{t=3}^{t=\infty} \delta^{t-2}((1-\theta+r^C\theta)V^C - P) > \sum_{t=2}^{t=\infty} \delta^{t-2}((1-\theta)V^C - P). \tag{7}$$

Rewriting, this gives us the following necessary condition for B's willingness to join a boycott, on the assumption that this will result in an equilibrium coordinated by R:

Condition 3. *A necessary condition for buyer B to be willing to join a boycott in Period 2 given the designation by A of R as the basis for the boycott is that R is sufficiently convergent with I^B:*

$$\begin{aligned} r^B &\geq \frac{(1-\delta)}{\delta}\frac{(1-\theta)}{\theta}\frac{(V-P)}{V} \\ &\equiv r_{2.} \end{aligned}$$

Again, to check the existence of $r_2 < 1$, note first that

$$\underset{\delta \to 1}{Lim}\ r_2 = 0$$

and that if $\delta^2 + \theta = 1$, then $\delta + 1 > 0$ and

$$\begin{aligned} r_2 &< 1 - \frac{P}{V} \\ &< 1. \end{aligned}$$

Conditions 2 and 3 give us a constraint on the degree to which an equilibrium R must converge with the idiosyncratic logics of the buyer who moves first to establish R and the buyer who waits for another to propose R by starting a boycott, respectively. First, these

conditions establish[6] that

$$r_1 > r_2. \tag{8}$$

This is intuitive: the first mover who bears the cost of a longer boycott in order to test the unknown acceptability of R to B must enjoy a higher minimum return than B from the implementation of R. Second, B must bear the cost of a one-period boycott to signal the acceptability of R. Although this is a lower cost than A bears, it still imposes a constraint on the nature of R in equilibrium: A cannot establish an equilibrium with R if R shows too little convergence with B's idiosyncratic logic; by moving first A cannot pull the equilibrium too close to its own idiosyncratic logic if that pulls it too far away from B's.

Both r_1 and r_2 are increasing in V.[7] This implies that for an equilibrium to emerge, the higher is the margin enjoyed by the buyers in this market, *ceteris paribus*, the greater must be the convergence between R and the buyers' idiosyncratic logic. This is somewhat suprising, and it suggests that a more competitive market can sustain an equilibrium around a 'less good' logic (from the perspective of the individual buyers' idiosyncratic reasoning) than a market with substantial rents.

The fact that r_1 and r_2 are increasing in V also indicates that if A and B enjoy different margins ($V^A \neq V^B$), then we are more likely to see an equilibrium established by the lower margin buyer than the higher margin buyer. To see this, note that if $r^A = r^B = r$ and $r < r_1(V^B)$, then B is unwilling to bear the cost of being the first mover. Then equilibrium requires $V^A < V^B$ in order to ensure that $r_1(V^A) < r < r_1(V^B)$, that is that A is willing to move first.

Conditions 2 and 3 gives us necessary conditions for the emergence of an equilibrium coordinated by R, but not sufficient conditions. In order to establish an equilibrium we have to look at the beliefs that underlie the incentives of A and B to start or join a boycott organized on the basis of R. We turn to beliefs now.

Conditions 2 and 3 capture the decision problems for A and B respectively, only if all players–both buyers and the seller–believe that if an effective boycott based on R is coordinated in Period 2, then both buyers can be expected to boycott in any future period in which an R-wrongful performance occurs. This is what produces the payoff starting in period 3 and continuing indefinitely into the future in which the buyer enjoys a lower rate of wrongful performances: sellers anticipate that an R-wrongful performance will lead to a 2-buyer boycott and will therefore (by Condition 1) be deterred from R-wrongful performances. Additionally, both buyers must believe that if either A fails to boycott in Period 1 or B fails to boycott in Period 2, or if either buyer fails to boycott an R-wrongful performance in the future, then no buyer will ever boycott R-wrongful performances in the future. This is what produces the payoff resulting from a decision not to initiate (A) or join (B) a boycott.

Consider first the belief that once a two-buyer boycott has been observed in Period 2, both buyers will choose to boycott an R-wrongful performance in the future. We can justify these beliefs on the basis of the presumption that once the buyers have coordinated their boycott once, and recognizing that coordination on a two-buyer boycott produces a higher payoff for each than coordination on not boycotting, they will continue to coordinate the two-buyer boycott. (See Crawford & Haller (1990, 575) for this approach.)

But we may be able to ground this presumption more firmly in the structure of this game and what the parties are presumed to know about the game they are playing. The structure of the interaction between the buyers and the sellers in this repeated game is that (as in Crawford & Haller 1990) the initial strategies of the buyers are deliberately focused around testing (A) and confirming (B) the acceptability of R as a coordinating device. These buyers are self-consciously learning about whether they can coordinate.

The essential attributes of a focal point emphasized by Schelling ([1960] 1981) are thus generated endogenously by the sequence of A's and B's repeated interaction: A selects a *single* device (R) and makes it *salient* by taking an action (boycotting) that is only rational if A expects to continue coordinating with that device in the event that B indicates that it

too would be better off coordinating with that device. B then takes an action in Period 2 that is only rational if B judges coordination under R to produce a higher payoff and, again, only if B expects to continue to coordinate under R whenever an R-wrongful performance is delivered.

We can take this logic a step further by considering the seller. It is common knowledge that both A and B care about the seller's beliefs about whether the buyers can coordinate and that the seller, too, is observing the early-stage play in this repeated game to update its beliefs about the prospect for coordination. The seller begins the game with no information about the capacity for these two (idiosyncratic) buyers to coordinate. It then observes the proposal of R by A, backed up by a costly action (the boycott) that indicates that A is willing to invest in changing the seller's beliefs about whether the buyers can coordinate. In Period 2 the seller observes an action by B that indicates that it too is willing to invest in changing the seller's beliefs, that is, to demonstrate coordination. If the seller believes that the buyers have taken these costly actions because they believe that by doing so they will cause the seller to update its beliefs and expect a two-buyer boycott in response to an R-wrongful performance, then in fact it is rational for the seller to update in this way.

Suppose the seller chose to ignore the Period 2 demonstration and to exploit an opportunity to engage in a wrongful performance to one of the buyers in the future. Clearly this is only rational for the seller if the seller entertains the belief that the buyers will not be able to coordinate. But then both A and B can act in such a way as to demonstrate that they are coordinated–and to thus cause the seller to update this belief. In effect, each buyer can be thought to recognize that if the other buyer boycotts as expected, then a failure to join that boycott will cause the seller to update in the wrong direction: to attach a lower probability to the likelihood of a boycott. This makes a deviation by the seller in the future isomorphic with A's proposal in Period 1: the seller is testing whether R is a coordinating device. And we have shown that both buyers are willing to incur a cost to change the information structure of the game in such a way as to confirm R as a focal point for coordination on

higher-value payoffs. The players communicate through actions dictated by their strategies. As Crawford and Haller (1990, p. 574) put it: "They cannot communicate [about the game] except by playing it."

We can therefore state our main result.

Proposition 1. *Given that Conditions 1, 2 and 3 are satisfied, the logic R and the following strategies and beliefs support a perfect Bayesian equilibrium in the repeated game such that beginning in Period 3 all players will expect a coordinated boycott in response to R-wrongful performances and the seller will be deterred from R-wrongful performances: (1) Following an R-wrongful performance in Period 0, A boycotts in Periods 1 and 2 and announces R; (2) B boycotts in Period 2 if A boycotted in Period 1; (3) thereafter, if and only if an effective boycott was achieved in Period 2, A and B boycott whenever an R-wrongful performance occurs; (4) the seller exploits the opportunity to make a wrongful delivery in every periods unless and until an effective boycott in achieved in Period 2; (5) all agents entertain the belief that for either buyer, r^j is sufficiently convergent with R if and only if the buyer boycotts as prescribed by the equilibrium strategies.*

The rationality of the beliefs in this equilibrium follow from the definition of sufficient convergence: A is better off participating in a boycott in Period 1 and 2 than not if R is sufficiently convergent with its idiosyncratic logic, given the common knowledge prior that there is a 50-50 chance that R is sufficiently convergent for B. B is better off participating in the boycott in Period 2 having observed A's participation and knowing the structure of A's problem and hence A would only have boycotted under the R proposal in Period 1 if R was sufficiently convergent with A's idiosyncratic logic. It is common knowledge for both the buyers and the seller that A and B seek to coordinate their boycotting strategy because they each achieve a higher payoff if coordinated. It is therefore rational for all agents to update their beliefs once an effective boycott organized around A's proposed focal point–the logic R–has been observed to expect that with probability 1 both buyers will boycott in the future in the event of an R-wrongful performance. Given satisfaction of the

effective boycott condition, this deters the seller from exploiting the opportunity to engage in R-wrongful performances. No buyer has an incentive to deviate from the equilibrium strategy at any point, that is, to fail to participate in an expected boycott (provided R is sufficiently convergent for both) because doing so leads the other buyer to believe that R is not sufficiently convergent for the defecting buyer and therefore to conclude that this buyer will not boycott R-wrongful performances in the future. These beliefs defeat coordination, and the deviating buyer is worse off.

IV. Discussion

We noted in the development of the model that for R to coordinate the buyers on an equilibrium in which the seller is deterred from engaging in R-wrongful performances, it must be the case, minimally, that R is publicly-accessible and stable. This is necessary if the buyers are to be able to consult R to determine how it classifies different configurations (X) constituting performances and for B to make decisions about current boycotting behavior on the basis of expected benefits from future classifications and hence equilibrium boycotting choices.

In this section we delve more deeply into what R's attributes must be for it to perform the two functions required of it in our model: to coordinate expectations about how performances will be classified and to support the incentives of both buyers to participate in boycotts of performances that the logic deems 'wrongful', even when those are wrongs suffered by the other buyer. Our aim is to demonstrate that the model predicts, on positive grounds, several of the features of law that legal theorists have posited based on a conceptual or moral analysis of what it means to be governed by the rule of law. We also seek to provide an account for the role of public, distinctively legal, reasoning.

Generality. — Our model makes it plain that the logic R must be general in the sense that it addresses itself to the interests and situations of both buyers. Mathematically, this seems unremarkable: R must be defined over the set of circumstances that both A and

B consider relevant. Nothing in our model, of course, prevents R from being specifically conditioned on the identity of the buyer; R could be elaborated in statements of the form "when the buyer is A" or "when the buyer is a duke rather than a peasant." But the rationale for generality in this model suggests that R will not be particularized. Generality in our model derives from the equilibrium requirement that A designate a logic that attracts B to participate in a coordinated boycott triggered by the application of that logic. If, instead, A proposes a logic that is too personalized, too focused on supporting A alone, B will refuse to participate in the boycott. The third-party (but not coercive) institutions offering the alternative logics from which A can choose (we consider below whether A can itself be the provider of the logic) must therefore develop their offerings without access to the idiosyncratic logic, and perhaps even the identity, of any potential buyer. It is precisely the idiosyncrasy of the circumstances that make up the deal each buyer strikes individually with the seller that generates the ambiguity of wrongfulness in this model and thus the demand for a coordinating common logic. The common logic must, therefore, be addressed to abstract circumstances ("the document was not signed" "the buyer has a history of accepting delivery without objection up to a week after the delivery date") rather than the particularized circumstances A and B encounter. Relatedly, in an easy extension of the model, if there are a large number of both existing and unknown future buyers and endorsement of R in equilibrium means that an individual buyer participates in a boycott of R-wrongful deliveries to any such buyer, it must be that R is capable of addressing itself to the as-yet unidentified buyers.

Consider, more closely, B's assessment in Period 2 of its willingness to endorse R by boycotting. We can see that our notion of generality includes more than the abstract terms in which R's elements are expressed. R must not only be general, it must be *generalizable*. By this we mean that R must be capable of sustaining a inquiry by B into the classification that would be made of hypothetical circumstances different from the circumstances that have been observed to date. Coordination in our model drives generality, which is fundamentally

tied to the feasibility of predicting the classification of wrongful behavior in order to assess one's willingness to signal participation in a proposed logic. (For an evolutionary game model that focuses on the incentive of a potential participant in community punishment to signal an–exogenous–willingness to punish, see Boyd, Gintis & Bowles 2010.)

Generality in our account is not merely the tautological content of what it means to govern by "rules" (as Fuller (1964)[8] asserts.) Nor is it simply an economizing feature of lawmaking (as Myerson (2004) seems to suggest[9])–although the transaction–cost minimizing quality of generality will almost certainly play a role in a more elaborate coordination model that accounts for the costs of analyzing and predicting the classifications of a proposed logic. Our account in this model is closer to Raz (1977). Raz suggests that generality is necessary for the "rule of law" in order to control the unpredictability of particular orders which makes it difficult for people to guide their behavior on the basis of law. He also argues that generality is necessary because a legal system must minimally consist of judicial institutions and "there could not be institutions of any kind unless there are general rules setting them up." (Raz 1977, 206). Our model clearly requires that the classifications reached by R be predictable. (Generality might, as Raz asserts, promote predictability–although this does not seem necessary and, as a conceptual matter, is certainly not obvious.) But in our model generality performs a function beyond simple predictability: it addresses an incentive constraint. In order to coordinate participation in the punishment of wrongs against others–the only form of enforcement in our model–it must be that the logic addresses the circumstances of all those who might be required for effective enforcement. The idiosyncrasy of individual circumstances then generates the requirement that the law be written in general terms, general enough to encompass the particularities of the other players being asked to endorse the law through their actions.

Ours is also not a story about the preferences of agents for fair or equal treatment. Our buyers derive utility only from the transactions they engage in with the seller. They do not enjoy community benefits or good feelings about themselves or the goods associated

with conformity to norms *per se*. Similar to Binmore's (1994, 1998) effort to ground the Rawlsian "justice as fairness" principles in game theory, our analysis grounds the emergence of "general" rules on the interaction of self-interested agents who do not possess an inherent set of values over their relative treatment by the rules.

Uniqueness.—Our model clearly requires that R be unique: both the buyers and the seller must be consulting the same logic to classify performances, and it must be common knowledge that they all do so. This is the case even if the buyers never interact with one another other than through the enforcement equilibrium. Think about this in geographical terms. Suppose buyer A is in Genoa and buyer B is in Flanders and that both are purchasing from the same seller. In conventional accounts of law, it would be perfectly acceptable for A and B to be governed by different legal rules. All that matters in these accounts is that the law meet minimal conditions so that people can act so as to comply with a legal command and so avoid punishment. As Raz (1977, 208) puts it, "The law to be law must be capable of guiding behaviour, however inefficiently." The seller will have to know the rules in Genoa and the rules in Flanders, and this may be inefficient, but the lack of uniqueness is not a problem so long as the law in each country is enforced by a third-party coercive institution. In our model, however, the buyer in Genoa and the buyer in Flanders need to share the same logic, the same set of rules, for determining when the seller is in breach of its obligations, in order for coordinated enforcement to be an equilibrium.

This requirement of uniqueness is a deep one, going beyond a mere identification of the same body of law by all agents. In our simple model, each agent must not only consult the same logic. It must also be able reasonably to reach–and it must be common knowledge that it is reasonably able to reach–the same classification of a particular performance as the other agents. This requirement, in turn, implies that application of the logic by diverse agents must be expected to result in the same *unique classifications*. This imposes a constraint on the structure of the reasoning employed by the logic itself when accurately applied: There must be, at least in theory, a "right" answer to the question of whether a particular performance

is wrongful or not. The logic must be coherent and not contradictory. (Coherence is considered an especially important attribute of law by Continental jurists in the civil law tradition (Merryman and Perez-Perdomo 2007).)

Importantly, this requirement of unique classification does not imply that the rules and principles that make up the logic produce an *obvious* classification. The set of rules and principles that comprise the logic could be complex and ambiguous and capable of producing multiple answers, although this would make it more costly. (Our simple model assumes all logics are costless to use.) Agents may make errors in applying the logic. What is important is that there be a recognized process for determining a unique answer among a set of possible answers implied by the rules and principles. This observation gives content to our original definition of a logic not merely as a set of rules or principles but rather as the product of a third party institution. The problems of ambiguity seems to necessitate that there be *authoritative stewardship* of the classifications reached by the logic, to resolve complexities, ambiguities and gaps. Maintaining a coordination equilibrium seems to require that the steward (a person, a profession, a machine) be unique. This sheds light on why we generally find that in an established legal system in a complex environment there is usually a single Supreme Court, for example. And why tenth century Icelanders designated a single person to serve as the Law Speaker.

Impersonal Reasoning .—The requirement that all agents in the model reach the same unique, common knowledge classifications in order to coordinate also implies that the logic consist of *impersonal reasoning* in a community of any significant size or duration. That is, performing the operations of the logic on a set of facts regarding a performance must be invariant to the identity of the person or entity engaged in the operation. Buyer B for example, in order to decide whether to join the boycott in Period 2, must be able, either personally or through the retention of an expert, to assess the classification that R will make of the transactions to which B is a party. If the community were very small, it might be possible for B (as well as A and the seller) to consult a single person (the "institution"

providing R) to make these predictions. But in general, the agents will need access to more "legal" advice than a single person can provide. Such a system can only be provided, with unique classification, if the logic is invariant to the identity of the person applying the logic.

This discussion implies another key feature of the logic. The institution providing the logic must be neutral and independent: it must have no interest in the classifications reached by the logic. This also helps to explain why it is probably the case that–even presuming that A could communicate the content of its idiosyncratic logic and, further, even presuming that I^A and I^B are sufficiently convergent–A cannot just propose I^A –a logic it controls–as the basis for its boycott.

Public Reasoning and Open Process.— We have shown that the logic R must be *publicly accessible*: both the buyers and the seller must have access to the logic to consult it in making their decisions about boycotting and performance. More subtly, our model implies that the requirement of publicness goes beyond mere publication of the rules, as most legal theory presumes.[10] Our model suggests that the common logic must also be a form of *public reasoning* elaborated in an *open process* to which an interested party might introduce their private information and reasoning. (See Waldron 2008 for a proposal similar to this from a legal philosophical perspective.) Both of these ideas find their root in the heterogeneity and idiosyncrasy that generates the problem of ambiguity and the need for a common logic in the first place. Put differently, in a homogeneous world with shared and unambiguous classifications of all performances as "cheating" or not, there is no need at all for an external institution to provide a common logic; in such a world, we will find it easier to predict, as do Hume (1740) and Sugden (1986, 2005), the spontaneous emergence of norms to coordinate behavior. Milgrom, North & Weingast (1990) find that all that is needed in such a world is an institution that serves to share information across traders separated in time.

The requirement of open and public reasoning appears to follow from our model because the assumption of idiosyncrasy suggests that the classifications reached by the logic must

be *immanent*, not fully articulated in any form that can be consulted *ex ante* by all agents. (The idea of immanence will be recognizable to those schooled in the traditional legal concept that the common law is "found" not "made": it contains all of its principles even if they are not articulated until a specific case is adjudicated.[11]) Recall that we have defined each buyer's idiosyncratic logic as an inaccessible reasoning process that maps (potentially private) information into an assessment of the value of a potentially complex set obligations on the seller. (In a world where delivery in 10 days is generally considered acceptable, for example, buyer B may be an innovative manufacturer which has discovered how to employ just-in-time delivery or variations in wholesale packaging to improve the allocation of inventory.) Having no access to the idiosyncratic reasoning of individual buyers when it offers its logic as a candidate coordination device in Period 1, a third-party institution must provide a logic that is capable of integrating, coherently, the information and reasoning from individual buyers through an infinite horizon. The logic, therefore, cannot be (just) a dataset collecting classifications already reached by the logic; it must contain the placeholders for dealing with as-yet-unimagined circumstances. Nor can it be a complete prescription of how all possible circumstances would be classified by the logic. To do this would require access to the idiosyncratic logic of (possibly as yet unknown) buyers who are uniquely able to assess the value and intended content of their transactions with sellers. As we have seen, the logic R must be sufficiently convergent with each buyer's idiosyncratic (*ex ante* inaccessible) logic in order to attract the buyer's participation in the coordination equilibrium.

Because the classifications must be immanent, the reasoning in R must be one that requires elaboration in particular circumstances. Those particular circumstances are initially private information. In order to elect to participate in the boycott equilibrium, each buyer must be able to elaborate the logic privately as it applies to these privately known circumstances and considerations. We have already discussed the requirement that this elaboration ultimately produce a unique classification. Ultimately, when this set of circumstances becomes relevant (the seller is contemplating a potentially wrongful performance

or the buyers are determining whether to engage in a boycott in response to a potentially wrongful performance), this classification must be capable of becoming public.

Thus the elaboration of the reasoning–its application to particular circumstances–must be conducted in public and it must be open to a presentation from the initially privately-informed buyer (more generally, also the seller) of how its idiosyncratic reasoning plays out in the common logic. Buyer A does not care about buyer B's idiosyncracies unless and until B is the potential victim of a wrongful performance and A has to decide whetehr to boycott or not. At that point, R must include an open and public reasoning process to determine, uniquely, whether the performance is R-wrongful or not. More generally, although we are not modeling the selection of A as a strategic choice in our simple model, we would predict that A would be more likely to propose an R that is open to hearing from B and sufficiently public. Open process and public reasoning are likely to give B greater confidence R will, in practice, converge sufficiently with I^B.

Stability.—All legal theorists emphasize that law must consist of relatively stable rules. Our model also requires this. But conventional accounts of law, focused on the need to provide individuals with sufficient guidance that they can conform their conduct to law and so avoid punishment, imply a different timeframe for stability. In conventional accounts, a rule needs to be stable between the time an agent chooses an action and the time at which there is the potential for having that action judged and penalized under the rule. This is the timeframe the seller in our model cares about: rules need to be stable during a period, but from the seller's perspective they could change period to period.

Stability in our model, however, is also required to meet the requirements of the buyers. They require stability over a much longer horizon than the seller does. The buyers must be able to anticipate in the early stages of the game (Periods 1 and 2) that the logic they are evaluating as a potential coordinating device will retain its classifications over an infinite horizon. To meet this stability requirement, R cannot change from period to period.

Prospectivity.—Compare the departure between our theory and conventional legal

theory with respect to stability to the implications for prospectivity. Conventional legal theory, again on the basis of what a person requires in order to conform and avoid punishment, asserts that law must be prospective: the seller cannot condition behavior at the start of Period N on a rule that is not available until the end of Period N. But rules could change from period to period. In our model, buyers do not care about prospectivity except to the extent that they can predict that if they coordinate on R, the seller will be able to effectively condition its behavior on R. Thus our prospectivity requirement is addressed to the same need as that proposed in conventional theory.

V. Conclusion

We began with the question, what is law? Our answer is that law is, at least in part, a system of distinctive reasoning used to classify conduct as right or wrong that serves to coordinate distributed agents in delivering punishments to deter wrongdoing. Our model predicts the characteristics that such a system must have in order to meet two basic constraints: the contraint that it be common knowledge that all are coordinating on the same system and the constraint that participation in the system must be incentive-compatible. We rest community punishment of wrongs on incentive-compatibility, rather than meta-levels of punishment, biology or cultural preferences. This distinguishes our approach significantly from the existing literature. Our approach to solving the common knowledge problem departs from the existing literature in that we do not look for the exogenous environment to supply focal points that are salient and unique. Rather we look to how the deliberate effort to coordinate punishment of defectors can lead to the designation of a third-party institution–perhaps from many on offer–as an equilibrium coordination device. Although we do not in this paper model the selection or supply process, our approach paves the way for a deeper analysis of the endogenous emergence and indeed construction of institutions that serve to coordinate.

Our approach allows us to identify key characteristics that an institution that can sup-

port an equilibrium in distributed enforcement must possess. The institution must supply a logic for classifying conduct that is both general and stable in order to meet the incentive-compatibility constraint. This provides a novel explanation of the generality and stability requirements that many legal philosophers believe is a *sine qua non* of a legal, as opposed to tyrannical or merely mangerial, order. In order to meet the common knowledge constraint, the institution must supply a logic that possess many of the features legal philosphers emphasize on other practical or moral grounds: it must be publicly accessible, open to the presentation of facts and arguments, and impersonal.

Our model allows us to suggest the importance of features that are missing or de-emphasized in conventional legal theory. Also to meet the common knowledge constraint, the institution providing the logic likely must have a mechanism for reaching a unique classification. This sheds light on a debate in legal theory between Dworkin (1986), who maintains that at least the ideal of a "right answer" is essential to law, and other legal positivists who emphasize the practical existence of deep disagreements among legal experts in difficult and contested cases. Our analysis suggests that even in the presence of ambiguity–indeed especially because of ambiguity due to heterogeneity among individuals and circumstances–in order to coordinate enforcement behavior, a logic-providing institution must assume ultimate stewardship of an ultimately final classification of behavior as right or wrong.

The need to create a unique, common knowledge means to coordinate provides a basis for understanding another distinctive feature of legal institutions, namely that they assume the role of final arbiter of the meaning of a constitution or a contract. Modern legal systems, for example, generally possess a clear hierarchical structure, a system for resolving overlapping jurisdiction into a unique jurisdiction, and a 'supreme' court of last resort in each jurisdiction. The link we expose between heterogeneity and these features directed at uniqueness also holds out the possibilty of distinguishing between the features we should expect to see in different environments. In homogeneous environments with low risks of violence we expect less organized systems for reaching ultimately unique classifications. This

might explain the pre-1949 Tibetan system governing a subsistence economy, for example, which offered multiple overlapping jurisdictions with no clear system of hierarchy and resisted the announcement of a definitive conclusion other than the one ultimately agreed to by all parties.

Our aim in this paper has been to initiate the project of using the formal tools of law and economics and PPT and the law to move these fields beyond a focus on law as simply another mechanism for determining economic or political outcomes, approaches that miss most of what is distinctive about legal order. We need positive models of law in order both to understand and to develop policy to support the emergence, stabilization and dynamic evolution of legal orders.

REFERENCES

Basu, K. (2000). *Prelude to Political Economy: A Study of the Social and Political Foundations of Economics*. New York: Oxford University Press.

Binmore, K. G. (1994). *Game Theory and the Social Contract: Playing Fair*. Cambridge, MA: MIT Press.

Binmore, K. G. (1998). *Game Theory and the Social Contract: Just playing*. Cambridge, MA: MIT Press.

Boyd, R., H. Gintis, and S. Bowles (2010). Coordinated punishment of defectors sustains cooperation and can profilerate when rare. *Science 328*, 617–620.

Bryce, J. (1901). *Studies in History and Jurisprudence*. New York: Oxford University Press.

Cooter, R. (1998). Expressive law and economics. *Journal of Legal Studies 27*, 585–608.

Crawford, V. P. and H. Haller (1990). Learning how to cooperate: Optimal play in repeated coordination games. *Econometrica 58*, 571–595.

Dixit, A. K. (2006). *Lawlessness and Economics: Alternative Modes of Governance*. New York: Oxford University Press.

Elkins, Z., T. Ginsburg, and J. Melton (2009). *The Endurance of National Constitutions*. Cambridge: Cambridge University Press.

Eskridge, W. N. (1991). Overriding supreme court statutory interpretation decisions. *Yale Law Journal 101*, 331–455.

Farrell, J. and M. Rabin (1996). Cheap talk. *Journal of Economic Perspectives 10*, 103–118.

Fearon, J. D. (2006). Why use elections to allocate power. Paper presented at the annual meetings of the American Political Science Association.

Fehr, E. and S. Gachter (2002). Altruistic punishment in humans. *Nature 415*, 137–140.

Finnis, J. (2009). Natural law theories. *Stanford Encyclopedia of Philosophy*.

French, R. R. (2002). *The Golden Yoke: The Legal Cosmology of Buddhist Tibet*. Ithaca: Snow Lion Publications.

Friedman, D. (1979). Private creation and enforcement of law: A historical case. *Journal of Legal Studies 8*, 399–415.

Fuller, L. (1964). *The Morality of Law*. New Haven: Yale University Press.

Glaeser, E. L. and A. Shleifer (2002). Legal origins. *Quarterly Journal of Economics 117*, 393–406.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy 102*, 912–950.

Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.

Hadfield, G. K. (2008). The levers of legal design: Institutional determinants of the quality of law. *Journal of Comparative Economics 36*, 43–74.

Hadfield, G. K. (forthcoming). The dynamic quality of law: The role of judicial incentives and legal human capital in the adaptation of law. *Journal of Economic Behavior and Organization*.

Hardin, R. (1989). Why a constitution? In B. Grofman and D. Wittman (Eds.), *The Federalist Papers and the New Institutionalism*. New York: Agathon Press.

Hardin, R. (2006). Constitutionalism. In B. R. Weingast and D. Wittman (Eds.), *Oxford Handbook of Political Economy*. New York: Oxford University Press.

Hong, L. and S. E. Page (2001). Problem solving by heterogeneous agents. *Journal of Economic Theory 97*, 123–163.

Hume, D. (1739-40). *A Treatise of Human Nature.* New York: Oxford University Press.

Kaplow, L. (1992). Rules versus standards: An economic analysis. *Duke Law Journal 42*, 557–623.

Kramarz, F. (1996). Dynamic focal points in n-person coordination games. *Theory and Decision 40*, 277–313.

Luban, D. (2007). *Legal Ethics and Human Dignity.* Cambridge: Cambridge University Press.

Mailath, G. J., S. Morris, and A. Postlewaite (2001). Laws and authority. Manuscript, University of Pennsylvania.

Mailath, G. J., S. Morris, and A. Postlewaite (2007). Maintaining authority. Manuscript, University of Pennsylvania.

Marks, B. A. (1988). A model of judicial influence on congressional policymaking: Grove city college v. bell. Working Papers in Political Science P-88-7, Hoover Institution.

McAdams, R. H. (2000). A focal point theory of expressive law. *Virginia Law Review 86*, 1649–1729.

Merryman, J. and R. Perez-Perdomo (2007). *The Civil Law Tradition, 3rd Edition: An Introduction to the Legal Systems of Europe and Latin America* (3 ed.). Stanford: Stanford University Press.

Milgrom, P. R., D. C. North, and B. R. Weingast (1990). The role of institutions in the revival of trade: The medieval law merchant, private judges, and the champagne fairs. *Economics and Politics 2*, 1–23.

Mitchell, W. (1904). *An Essay on the Early History of the Law Merchant*. London: University Press.

Myerson, R. B. (2004). Justice, institutions, and multiple equilibria. *Chicago Journal of International Law 5*, 91–107.

Ordeshoook, P. (1992). Constitutional stability. *Constitutional Political Economy 3*, 137–175.

Porta, R. L., F. L. de Silanes, A. Shleifer, and R. Vishny (1998). Law and finance. *Journal of Political Economy 106*, 1113–1155.

Posner, R. A. (1977). *Economic Analysis of Law*. Boston: Little Brown.

Raz, J. (1977). The rule of law and its virtue. *Law Quarterly Review 93*, 195.

Schelling, T. C. (1981). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Sugden, R. (2005). *The Economics of Rights, Cooperation and Welfare* (2 ed.). New York: Palgrave Macmillan.

Waldron, J. (2008). The concept and the rule of law. *Georgia Law Review 43*, 1–61.

Weingast, B. R. (1997). The political foundations of democracy and the rule of law. *American Political Science Review 91*, 245–263.

Notes

[1]Basu's "Core Theorem" is " Whatever behavior and outcomes in society are legally enforceable are also enforceable through social norms." He articulates two corollaries: "What can be achieved through the law can, in principle, also be achieved without the law" and "If a certain outcome is not an equilibrium of the economy, then no law can implement it." (2000, 117).

[2]Hong and Page (2001) present a model in which "collections of agents outperform individuals partially because people see and think about the problems differently" (p. 130) diversity is captured by characterizing individuals in terms of their individual *internal language* (used to represent objects), *perspective* (a mapping from objects into the internal language) and a *heuristic* (a set of rules for moving around the space of objects in his or her internal language, a logic)

[3]Crawford & Haller (1990) present the idea that agents may lack a common language for representing the structure of a game and thus cannot reproduce the reasoning of others (for purposes of coordination) except on the basis of observed outcomes that can be uniquely associated with a particular action. See also Kramarz (1996) solving an N-player coordination game in the absence of a pre-existing common language. Both Crawford & Haller and Kramarz analyze the dynamic process of reaching coordination through the generation of a common language based on the evolving history of a game.

[4]We abstract here from the implications of an R that makes type 2 errors: finding a performance wrongful that is not judged by the buyer to be wrongful. Error of this type will impact the price of the good inefficiently and the buyer will not want to induce boycotts to deter such performances. This raises a set of strategic issues that we do not take on here (including whether the buyer can avoid triggering a boycott by not revealing the performance or pressing a claim with the other buyer.)

[5]See Crawford & Haller (1990) for a similar modeling strategy in a game with strategic uncertainty based on the lack of a common language to describe a game.

[6]This can be confirmed by noting that the derivative of r_2 with respect to δ is negative and $\delta^2 < \delta$.

[7]The derivative of r_1 with respect to V is $\frac{P}{V^2\theta\delta^2}(1-\theta)\left(1-\delta^2\right) > 0$. The derivative of r_2 is $\frac{P}{V^2\theta\delta^2}(1-\theta)(1-\delta) > 0$.

[8]"The first desideratum of a system for subjecting human conduct to the governance of rules [Fuller's definition of "law"] is an obvious one: there must be rules. This may be stated as the requirement of generality." Fuller (1964, 46).

[9]Myerson does not deal extensively with the concept of generality, but notes that "general" principles might be established by an assembly in order to resolve a larger set of cases. Myerson (2004, 100).

[10]"The law must be open and adequately publicised. If it is to guide people, they must be able to find out what it is." Raz (1977, 198-199)

[11]Blackstone held that it was not the judicial function to "pronounce a new law, but to maintain and expound the old one". 1 William Blackstone, Commentaries 69.