2002

Book (OUP): On Law, Politics, and Judicialization: Path Dependence, Precedent, and Judicial Power

Alec Stone Sweet, *Yale Law School*

Path Dependence, Precedent, and Judicial Power

Alec Stone Sweet

Given certain conditions, legal institutions will evolve in path dependent ways: that is, the social processes that link litigation and judicial law-making will exhibit increasing returns. Once under way, these processes will build the discursive techniques and modes of decision-making specific to the exercise of judicial power; they will enhance the centrality of judicial rule-making vis-à-vis other processes; and they will, periodically but routinely, reconfigure those sites of governance constituted by rules subject to intensive litigation.

The paper follows in a line of research on how new legal systems emerge, mutate, and mature, and with what political consequences (for example, Stone 1992*a*; Stone Sweet 1997; 2000). The concepts of path dependence and increasing returns have at times been deployed and given empirical content (especially Stone Sweet, Chapters 1 and 4, this volume; Stone Sweet and Caporaso 1998*a*); nonetheless, they were used to complement other theoretical materials and priorities, and were left under-theorized. Here I provide explicit theoretical foundations for the path dependence of legal institutions, and an argument as to why this should matter to social scientists and to lawyers.

A much longer version of this paper, co-authored by Margaret McCown, was presented at the Colloquium on Law, Economics, and Politics, the Law School, New York University, October 2000. The first part, presented here, provided a blueprint for collecting and analysing data on the use of precedent in the European legal system, analysis that comprised the second part of the original paper (Stone Sweet and McCown 2000). Margaret McCown, a graduate student in politics at Oxford, and Stone Sweet continue their collaboration on the second stage of the project. Ms McCown's contributions to this chapter were invaluable, but Stone Sweet is alone responsible for what is written here. He also benefited from discussions with James Caporaso, Paul David, John Ferejohn, Ronald Jepperson, Louis Kornhauser, Paul Pierson, Martin Shapiro, and Mark Thatcher.

Legal institutions are path dependent to the extent that how litigation and judicial rule-making proceeds, in any given area of the law at any given point in time, is fundamentally conditioned by how earlier legal disputes in that area of the law have been sequenced and resolved. The paper elaborates a model of adjudication in which institutional development and decision-making are linked through highly organized discursive choice-contexts, meso structures called 'argumentation frameworks'. Argumentation frameworks are curated by judges as legal precedents. Litigants and judges are assumed to be rational utility-maximizers; but they are also actors who pursue their self-interest in discursive ways, through argumentation and analogic reasoning. Sustained, precedent-based adjudication leads to outcomes that are both indeterminate *and* incremental: that is, they are path dependent. I conclude by addressing various implications of the argument which taken together define an agenda for research.

Theoretical Issues

To invoke the metaphor of path dependence as a shorthand summary of why a particular state of affairs emerges, changes, or persists may be a useful rhetorical device, but it can also be an empty one. Both proponents and critics of path dependence approaches to social institutions are right to insist on the need for clearer theoretical exposition and rigour (Goldstone 1998; Pierson 2000). Path dependent explanations are compelling only to the extent that they elucidate how effects and outcomes that operate and are observable at a systemic—macro—level, are linked, across time, to effects and outcomes that operate and are observable at the domain of the individual decision-maker: the micro level. Such linkages develop through a positive feedback mechanism: a nascent, or maturing, standard of behaviour induces increasingly larger, and better networked, individuals to behave similarly, that is, in ways that adapt to, and thus reinforce, that standard. In the sections that follow, I introduce basic concepts, and then discuss the conditions under which I expect legal institutions, especially judicially-constructed argumentation frameworks—what Americans call 'doctrine'—to develop in path dependent ways. And I specify the kinds of outcomes that mechanisms of positive feedback are likely to generate at both macro and micro levels.

Path Dependence and Feedback

The idea that certain kinds of processes and outcomes are better explained by a logic of path dependence and increasing returns rather than a logic

of path independence and decreasing returns has gained more adherents across the social sciences, while remaining controversial. Long a staple of evolutionary economics (for example, Ayres 1962), the work of Brian Arthur (1994), Paul David (1985; 1994; 1997), and Douglass North (1990) has brought the idea to the forefront, provoking important discussions about its nature, scope, and applicability (for example, Goldstone 1998; Liebowitz and Margolis 1995). The idea is controversial because it is 'fundamentally hostile to a predictable, equilibrium economics' (Goodstein 1995: 1029), wherein the interplay of dichotomous forces—for example, supply and demand—will push systems—for example, markets—to stability through mechanisms triggered by decreasing returns. And, unlike a 'natural selection' approach to, say, the choice of technological standards, the path dependence approach does not assume the superiority of the choice ultimately made relative to other options. Indeed, economic historians have produced a long series of compelling empirical studies to bolster the claim that it is often 'only the sequence of choices—driven by chance and trivial circumstances—that will eventually give one technology the attributes of the fittest' (David and Greenstein 1990; Foray 1997: 735).

Increasing returns approaches are at least presumptively applicable to research on the evolution of political institutions, that is, to the study of how bodies of law—policy—are chosen through the activities of governmental organizations. I understand Pierson's (2000) arguments on this point to be more or less definitive, and will not rehearse them here. Instead, I take up two standard reference points for any general discussion of the topic: Arthur's (1994: Chs 2, 4) and David's (1994; 1997) summaries of their own ideas and findings.

Arthur identified the basic characteristics of social processes driven by increasing returns dynamics. First, initial conditions do not determine outcomes. Second, precisely because, at the *ex ante* moment, end points are unpredictable, the analyst will be able to explain any subsequently constituted state of affairs only in relation to a particular sequence of 'choices or outcomes of intermediate events [that have taken place] between initial conditions and the endpoint' (Goldstone 1998: 834). Third, to the extent that any intermediate choice determines the sequence and content of subsequent choices, the observed process will exhibit *non-ergodic*[1] properties. 'Small historical events can become durable effects' (Baumann *et al.* 1996: 160) through positive feedback: 'The micro behaviour of the system'—for example, the

[1] 'A path dependent process is "non-ergodic": systems possessing this property cannot shake off the effects of past events, and do not have a limiting, invariable probability distribution that is continuous over the entire state space' (David 1993: 29).

relevant decisions taken by individuals in the second instance—will amplify the 'distribution of choices produced during the first period' (Foray 1997: 741). Fourth, in so far as outcomes are embedded as aggregate social choices, or investments, the cost of transition away from the new standard will be high, even pre-emptively so. Last, these processes possess a common structure. At a beginning point, a range of choices, formats, or templates for a particular form of behaviour are available; at one or several 'critical junctures'[2] one of these choices gains an advantage, however slight, and this advantage is continuously reinforced through positive feedback. Ultimately, the choice becomes dominant, or 'locked in', as a relatively taken-for-granted state of affairs.

Systems become path dependent through positive feedback, essentially adaptation and network effects that are gradually institutionalized as stable practices. David[3] and Arthur showed that some kinds of situations would be more conducive to producing feedback than would others, especially those featuring positive network externalities. Because the adoption of a technical standard or certain kinds of social norms reduces uncertainty and enables large numbers of individuals to construct productive relationships with one another, the marginal benefits of adoption will rise with each decision by other actors to adopt. Snowballing ensues, which institutionalizes, or 'locks in', the choice. David (for example, David 1992) and Arthur also argued that increasing returns dynamics would be particularly prevalent in knowledge-based industries: localized learning mixes with network effects to produce path dependent dynamics and systemic inflexibility. Where certain fixed costs are associated with adoption, such as investments in infrastructure or training, the costs of reversing course or adjusting to a different or presumptively better standard is commensurably higher.

Some elements of David's work (see especially David 1994) connect to more sociological views of institutions. Institutions—from conventions to formal rules—structure social settings, 'by aligning individual's expectations' and priorities in ways that enable coordination, given cognitive limitations and the existence of multiple solutions to any given cooperation game (David 1994: 209–10), a standard formulation. Pushing further, David

[2] Critical junctures are not necessarily events that actors understood as 'big' or 'important' when they took place. Their importance may become clear only further downstream. Consider *Marbury v. Madison* (USSC 1803) or the European Court of Justice's decisions on direct effect and supremacy (see Chapters 4, 6).

[3] David contributed to the development of, and largely accepts, Arthur's account of path dependent dynamical systems. Nonetheless, he rarely invokes the concept of increasing returns, focusing instead on the consequences of non-ergodicity.

suggests that it is the dynamic, and path dependent, interplay between institutions, organizations, and agency that generates complex institutional arrangements (see also North 1990). For our purposes, the critical issues are how such arrangements are selected and how domains, once organized through rule structures, reproduce themselves. As sociologists have long understood, organizations not only become possible on the basis of rule structures but also serve to reinforce and to evolve new institutions, allowing actors to respond incrementally and more or less predictably to changes in the environment.

Simplifying, David argues that the more any organization—or linked clusters of institutions and organizations, an *organizational field*—possesses effective procedures for processing information, adapting rules to situations, and monitoring the activities and performance of individuals, the more likely it is that the domain of action governed by that organization will evolve in path dependent ways. Organizations develop and instil 'codes' that socialize actors into relatively coherent, 'ideal-type' roles (following Arrow 1974; cited in David 1994: 212–13). Socialization, of course, is a ubiquitous form of positive feedback, and therefore is easily understood in terms familiar to students of path dependence more generally. The codes—heuristics, ideal-type roles, or 'performance scripts' (Jepperson 1992)—that guide action in any field typically exhibit some measure of non-ergodicity and irreversibility. Whether borrowed and adapted wholesale or gradually developed through trial and error, such codes:

> ... are not instantly created and learned, and so constitute a form of 'durable capital' whose costs are 'sunk'; they become more refined and ... ingrained through repeated use, rather than eventually wearing out, [in contrast to] most tangible forms of productive capital. (David 1994: 212–13)

With use, institutions supporting the domain typically become more differentiated and the codes more increasingly articulated. In so far as they do, actors will, in order to access and exploit the organization's resources, specialize 'in the information capable of being transmitted by the codes'. Although specialization enables productive activity, it also imposes costs, to the extent that actors, over time, 'become less efficient in acquiring and transmitting information not easily fitted into the code' (Arrow 1974: 57). Specialization provokes branching within the codes and vice versa; branching is a prototypically path dependent mode of institutionalization; and institutionalization proceeds through feedback mechanisms (see also Stone Sweet, Chapter 1, this volume).

The similarity between these ideas and those of the 'new institutionalists' in sociology (for example, Powell and Dimaggio 1992), and particularly the John Meyer group at Stanford (research surveyed in Jepperson 2001) should be obvious.[4] In myriad ways, organizations function as gatekeepers to discrete domains or fields. They recruit, instruct, and credential actors; and they authorize certain kinds of political activities while discouraging others.

The Path Dependence of Legal Institutions

Theoretical interest in the process through which judicial authority emerges and evolves over time seems never to have been widespread or sustained in the social sciences.[5] Law is commonly presumed to be relatively stable feature of the social environment; and the sources and consequences of judicial rule-making are often ignored or explained away. Game theorists assume fixed rules in order, among other things, to assess the likely impact of different legal regimes (for example, Baird, Gertner, and Picker 1994), and they have not resolved problems associated with modelling the dynamics of judicial rule-making (see the exchange between Vanberg 1998*b* and Stone Sweet 1998*a*). Constitutional political economy focuses almost exclusively on institutional design, as if constitutions do not evolve meaningfully once they enter into force and as if constitutional adjudication and rule-making did not exist (for example, Brennan and Buchanan 1985; Voigt 1999). The field of law and economics generally seeks 'to explain and to evaluate the content of legal rules rather than the process by which they are created' (Kornhauser 1992*a*: 169). And in much of the world where a formalist legal positivism still reigns as orthodoxy, scholars typically depict the law as a self-contained system that predetermines outcomes (Merryman 1985; Schlink 1993).

There are some obvious reasons why the processes through which legal institutions evolve ought to be subject to the dynamics that typify the evolution of non-ergodic dynamical systems more generally. Legal rules are, after all, standards of behaviour that judges select, enforce, and revise. Judicial rule-making, being more or less authoritative, should function to reduce uncertainty about the nature and scope of the standard, and also to provoke and reinforce feedback effects: there is rarely an analog to judicial authority

[4] David himself does not make these connections explicitly.

[5] Interest has grown substantially in recent years, for example: Dezelay and Garth (1996); Greif (1989); Milgrom, North, and Weingast (1990); Stone Sweet (1999*b*); Shapiro (1988).

in accounts of how technological standards are selected or discarded in a decentralized market. If judges begin to construe contracts in ways that impose one type of risk allocation on parties rather than another, then future contractants will have an interest in adapting to this rule-making to the extent that risk allocation is basic to commercial exchange and given that the parties know that a legal dispute between them is always possible. Judges help actors resolve basic assurance dilemmas: their activities increase the confidence, among those contemplating adoption, that a sufficient number of others will make the same choice (see Gillette 1998: 819–20). Further, if the judicially curated standard becomes the actual standard, then it is all the more likely that disputes that arise under it will be resolved through subsequent adjudication.

Some institutionalists, concerned more with identifying basic mechanisms of institutional change rather than with the work of law and courts per se, give judicial rule-making pride of place. North (1990), for example, argues that courts, by enforcing property rights and adapting the law to changing circumstances, reduce the transaction costs associated with impersonal exchange, helping societies expand markets and wealth (see also Stone Sweet and Brunell 1998*a*). For March and Olson (1989), the incremental techniques of common law adjudication comprise a paradigmatic mechanism for generating, and then continuously revising, the norms and 'logics of appropriateness' that organize human community. In the legal literature, a handful of important articles explicitly take, or argue for, a path dependence perspective on the development of legal institutions (for example, Gillette 1998; Kornhauser 1992*a*, *b*; Roe 1996; Stearns 1995), if for varied purposes.

In the rest of the paper, I seek to push this agenda forward by specifying the theoretical foundations of a path dependent model of adjudication.

Organizational Factors

Legal institutions—for example, substantive law, rules of standing and jurisdiction, doctrinal principles, cannons of construction, and so on—partly evolve through adjudication. I argue here that certain common features of adjudication favour the path dependence of these institutions *if some minimally robust conception of precedent exists*. I do not claim that these features are essential characteristics of all legal systems, or that they exert the same effect everywhere. Indeed, where these features are not present, or where the effects of the factors discussed are mitigated by other

unspecified factors, I expect the path dependence of legal institutions to be commensurably weaker.

First, courts typically do not control the temporal order of cases that come before them. In any legal domain, judicial rulings on cases that come first will exert influence on subsequent litigation and judicial decision-making. This will be so to the extent that courts frame their decisions in light of past decisions and litigants adapt their claims to the courts' rule-making (see below). Where a judicial decision itself provokes a stream of cases, such as when a ruling expands the opportunities for, or enhances the benefits of litigating for some class of individuals, increasing returns properties are most obvious. In any event, in the overall process of selecting a legal rule, or in clarifying its meaning over time, there are likely to be significant 'first mover advantages' not unlike those identified by economists (see Mueller 1997). In negative terms, the claim is that, if one could run a sequence of cases differently, one would find that judicial outcomes too would be different, and they would trace a different path.

Second, adjudication constitutes a dichotomous discursive field in which specific legal questions are typically asked in a 'yes-no' format (Kornhauser 1992*a*: 171). Each party seeks different answers from the judge, and each makes counter-arguments in light of the other's claims. In the next section, I draw out some of the consequences of this structure for how legal actors are likely to behave. For now, it is enough to note that judges not only answer at least some of the questions posed to them but also give reasons for why they have chosen one result over another.[6] Dispositive answers given to yes-no questions possess the inherent capacity to block one path of development while encouraging another. Further, when judges justify their rulings with reasons, they at least implicitly announce prospective rules which subsequent litigants will understand and use as templates for future argumentation in cases that are similar in some salient way. If this is so, the *path independence* of legal institutions will be unlikely or exceptional.

Third, adaptation and network effects, to be registered on the behaviour of judges, lawyers, and professors, are facilitated by the fact that records of judicial activity are heavily documented. Published decisions typically include a description of the relevant or settled fact context of the case, a survey of the main arguments and counter-arguments, the dispositive decision, and the justification. The lawyers' briefs and other summaries of the parties' arguments are also usually available. Lawyers, who through training and practice

[6] In some polities, some courts are legally required to provide written legal justifications for rulings made.

learn to take their cues from these materials, are commonly organized into relatively autonomous groups that roughly map onto relatively autonomous domains of the law. If domains of law develop along separate path dependent lines, then the processes that serve to construct various networks of legal professionals will also be path dependent.

Fourth, judicial rule-making is typically embedded in dense matrices of legal rules, some of the properties of which will favour, again, increasing returns logics. As noted, the pre-existing records of relevant judicial activity will heavily condition the work of lawyers and judges, given almost any conception of precedent. Further, at any given point in time, the existing law relevant to any specific dispute may itself be relatively inflexible and path dependent in the sense of *being more or less immune to change except through adjudication*. This point applies to bodies of substantive law—particular distributions of rights and obligations governing interactions in a specific social context—but also to procedure and to principles of standing and jurisdiction. Last, but not least, modern legal systems typically organize legal acts hierarchically, and this fact has consequences for system inflexibility, that is, for the capacity of other political actors to alter or reverse judicial outcomes.

To take a concrete example, the German Basic Law confers upon the Federal Constitutional Court the authority to defend constitutional rights, and declares the primacy of those rights over any other conflicting legal provision. Any German statute can be attacked for violating rights provisions by almost any actor in society, public or private; if the German Court agrees with the petitioner and strikes downs a statute, reversal is virtually impossible, except through a subsequent decision of the Court. This is because the Basic Law protects basic rights from substantial change through constitutional amendment. More generally, wherever constitutional judicial review operates in a routine and minimally effective manner, the path dependence of constitutional law, because of its normative supremacy vis-à-vis other legal norms, will be assured. Given network effects, the relative inflexibility of any body of law will be largely determined by two factors: (1) the density of judicial rule-making in that area, and (2) the relative ease or difficulty of overturning the law-making effects of judicial decisions through non-judicial means (see Stone Sweet 2000: Chs 2, 3).

In summary, how courts typically operate and how legal actors typically behave are likely to provoke and then sustain the path dependent development of litigation and judicial rule-making. Given some underlying notion of precedent, these processes can be expected to exhibit some significant degree of randomness—through the vagaries of sequencing—and non-ergodicity—through the survival of rules announced in past rulings;

and judicial rule-making can be expected to provoke positive feedback effects—more litigation and the construction of litigation networks—and to move the law along paths that are relatively inflexible—that is, costly or impossible to reverse.

Judging and Precedent

Precedent is basic to the path dependence of the legal system. It is at the heart of the organizational code that judges curate and lawyers learn. If courts did not justify their decisions with reasons, or if legal communities did not consider the records of these justifications to be relevant to the conduct of future litigation, then there would be no increasing returns to this code. In this section, I stipulate the micro-foundations of precedent, and thus elaborate a very specific view of the phenomenon. As with any such stipulation, I will surely leave out elements that others consider to be crucial.

I note beforehand the existence of a dense, if disjointed, literature on precedent and the doctrine of *stare decisis*. The orthodox view, continuously reiterated in traditional academic discourse, has it that precedent and *stare decisis* perform a set of linked functions. They help judges discover and to apply rules of law presumed to be discoverable; they enhance legal certainty for the law's subjects; and they help judicial law-making—an intrinsically decentralized mode of governance—achieve a semblance of centralization and systemic coherence. Others take perspectives external to the law per se. For Stone Sweet (1999*b*) and Chapman (1994), precedent follows naturally from giving reasons for decisions, which is one of the basic techniques judges have to counter the permanent 'crisis of legitimacy' in which they find themselves (see also Shapiro 1980 and Chapter 4). Shapiro (this Chapter) argues that it enables courts to receive and to emit clearer signals in the face of an 'acute noise problem', the sheer quantity and richness of litigation that confront judges.[7] Rasmussen (1994) argues that the doctrine that precedents are binding serves to enhance each judge's power with respect to future judges, thus accounting for its popularity. For Stearns (1995: 1357), *stare decisis* allows courts to escape 'cycling' problems—majority voting situations in which there is no Nash equilibrium—being the 'social choice equivalent of the proscription on reconsideration of a rejected motion' that one finds in legislatures (see also Kornhauser 1992*a*). Each of these logics has merit, and I draw connections with some of them in the discussion that follows.

[7] Heiner (1986) re-specifies the problem as one of uncertainty and imperfect information.

Further, since the advent of legal realism there has been a long-running debate on the question of whether, why, and how precedents are binding (recent expressions include: Caldeira 1994; Knight and Epstein 1996; Segal and Spaeth 1996). Empirically, we know that legal actors routinely behave as if the legal materials contained in past court decisions were directly relevant to their purposes in the present; but we also know that judicial outcomes regularly escape parameters that pre-existing law would fix. Although by no means my central objective, this paper responds to this controversy. Most important, I reconceptualize the relationship between precedent and judicial rule-making as one in which the former enables the latter.

Precedent: Analogic Reasoning and Argumentation

Two basic claims about law and judges underlie my conception of precedent. First, judicial outcomes are fundamentally indeterminate. The generic source of the law's indeterminacy lies in the essential tension between the abstract nature of the social norm on the one hand, and the concrete nature of human experience on the other. Any particular social situation is in some meaningful sense unique, whereas norms are specified in light of an existing or evolving typology of fact contexts, an abstraction that deprives situations of some of their richness. One could imagine a world in which this tension did not exist. In a community governed by what Neil MacCormick (1989: 1) has called the 'special providence model of law', for example, a perfect normative system clearly specifies, for every individual in every possible situation, the rights and duties necessary for the community to achieve optimal collective good. In such a community, the law would constitute a 'complete' social contract (see Milgrom and Roberts 1992: 127–32). In the world we know, all bodies of law are imperfect and incomplete, hence one of the basic social functions of courts: to adapt norms of behaviour to the needs of those who live under them, over time, given changing circumstances.[8]

Second, judging, far from being a *sui generis* activity, is instead a manifestation of a deeply rooted human penchant for using analogic reasoning to make sense of, and to manage, the complexity of the environment. For cognitive psychologists, analogical reasoning is the process through which people 'reason and learn about a new situation (the target analog) by relating it to a more familiar situation (the source analog) that can be viewed as structurally parallel' (Holyoak and Thagard 1997). The ability to construct analogies is widely considered to be an innate part of thinking (Keane 1988; Vosniadou and Ortony 1989; Holyoak and Thagard 1995; Mayer 1992). Unfamiliar situations, those that individuals cannot understand through their generalized knowledge, stimulate the formation of analogies which are used to conceptualize *and* to find solutions to problems (Keane 1988: 103). The set of potential source analogues is defined jointly by (1) the specific, immediate problem to be resolved, or situation to be conceptualized, and (2) the past experiences of the individuals constructing the analogy. Legal argumentation constitutes a species of analogical reasoning: actors reason through past decisions of the court—the equivalents of source analogues—to characterize the interplay of new fact contexts and legal interests raised by a dispute—the target analogue—and to find an appropriate solution to it.[9]

Psychologists have also engaged the question of what constitutes an appropriate or effective analogy. Research has shown that the most successful analogies—those that best enable people to conceptualize situations and solve problems—maximize certain values, what the literature refers to as 'constraints' (Spellman and Holyoak 1992). The greater the conceptual similarity between source and target, the more the internal relationships between their core elements are structurally parallel, and the better able to offer solutions to the problems posed, the more effective is the analogy. Holyoak and Thagard (1997) have shown that problem solvers generally do seek to maximize each of these values, and in doing so they enhance the overall coherence of analogic thinking. Moreover, even when the choice of source analogue is guided by the conscious, goal-oriented purposes of the individual, the constraints guiding analogy formation nonetheless impose a degree of consistency on the sorts of mappings that will be considered effective and legitimate.

It is an obvious, but not a trivial, thing to note that analogy formation is a perfectly path dependent process, in that each transfer or adaptation of a source analog to a target analog is possible only by virtue of a prior outcome of analogic reasoning which, in turn, depends upon the sequence of situations and problems that individuals have *already* confronted and resolved.

[8] If judges are expected to resolve legal disputes in light of the dictates of the law, then they will have little choice but to 'complete' the law on an incremental basis. That is, judges will become law-makers.

[9] I have quite consciously not referenced the scholarly literature in law and philosophy on the uses and abuses of analogy in reasoning. My point is that analogical reasoning is part of the deep structure of legal reasoning, regardless of how legal actors actually use it. In legal scholarship, analogical reasoning is more or less taken for granted as one of the mechanisms through which case law develops. Murray (1982: 833) calls it 'a vital tool in legal reasoning' and Sunstein (1993: 741) states that 'reasoning by analogy is the most familiar form of legal reasoning', but both lament the lack of serious attention given to its actual role in producing legal outcomes. I am aware of only two studies that relate the cognitive psychology of analogical reasoning to precedent-based legal reasoning (Marchant *et al.* 1991; 1993); both report experiments using subjects untrained in the law, and neither examines litigation or judicial outcomes.

Research has shown that, at any given moment in time, people have some discretion, or exercise some degree of conscious choice, over which analogue gets mapped onto unfamiliar situations and how; but the set of available analogues is finite, having been determined by prior events and choices. In adjudication, analogic reasoning has been institutionalized as a set of relatively stable practices. Lawyers and judges use it self-consciously, and the results of their deliberations are, I repeat, heavily documented, which facilitates the identification and mapping of appropriate analogies. Of course, in crafting their arguments and decisions, litigators and judges also work to achieve logical or systemic coherence through other methods, such as deduction and abduction.

Although analogic reasoning infuses the organizational codes that structures the domain of law and adjudication, it is also a functional instrument of decision-making. Litigation raises questions concerning which and how legal norms are to be applied to concrete situations, and with what social consequences. Judges will answer these questions in light of how they have answered questions raised with respect to prior, analogous situations rather than willy-nilly, because skilled analogic reasoning is basic not only to skilled action per se but to the judges' own social legitimacy. By formalizing the results of analogic reasoning into precedents, say through abduction, judges give the legal system a measure of 'relative determinacy' (see Sartor 1994: 200–2). In providing analogues to other actors, they enhance the coherence of the litigation environment, thus reducing randomness and noise. And by grounding judicial rule-making in discourses present in existing case law, the judges mitigate the problem that the precise content of the law governing a particular social situation could not have been known to the parties at the time their dispute erupted (see Stone Sweet, Chapter 1, this volume). None of these points is incompatible with the view that those who litigate and those who judge are goal-oriented. Indeed, the argument is that legal actors will sharpen their analogic reasoning skills through training and use, precisely because they are rational, strategic actors operating within an organizational code that requires such skills of them.

I now turn to precedent or, more precisely, to *argumentation frameworks*. Prior legal decisions constitute the materials that enable the construction of such frameworks. Legal systems are webs or clusters of relatively autonomous argumentation frameworks.

Argumentation frameworks are discursive structures that organize (1) how parties to a legal dispute ask questions of judges and engage one another's respective arguments, and (2) how courts frame their decisions.[10] Following

[10] Due to space limitations, only the main features of such frameworks are discussed here.

Sartor (1994), these can be analysed as a series of inference steps represented by a statement justified by reasons or inference rules that lead to a conclusion. Such frameworks typically embody inconsistency to the extent that they offer, for each inference step, both a defensible argument and a counter-argument from which contradictory—but defensible—conclusions can be reached. In resolving disputes within these structures, judges choose from a menu of those conclusions.

In practice, legal argumentation, unlike purer forms of logical deduction, tends to be 'non-monotonic'. In monotonic reasoning, 'no logical consequence gets lost by extending the premises from which it has been deducted' (Sartor 1994: 191). The primary source of non-monotonicity in legal discourse is the dialectic relationship between new facts or information about the context of the dispute, and the existing premises that underlie normative conclusions. Particularly susceptible to this dialectic are those argumentation frameworks that contain normative statements that (1) announce general rules and then justified exceptions to those rules, and (2) govern how conflicting legal interests ought to be balanced, or how the conflict between two otherwise applicable norms is to be resolved (for example, when two rights provisions conflict). In balancing situations, which are ubiquitous in constitutional and administrative law, 'only the particular circumstances of a case allow a choice to be made', while any new situation 'may render new arguments possible, which may defeat previously valid ones' (Sartor 1994: 191, 197).

We have returned to the point made earlier : that legal systems and judicial outcomes are indeterminate. Nonetheless, argumentation frameworks provide a measure of at least short-term systemic stability to the extent that they condition how litigants and judges pursue their self interest, social justice, or other values through adjudication. To be effective, legal actors have to be able to identify the type of dispute in which they are involved, reason through the range of legal norms that are potentially applicable, and assess available remedies and their consequences. Such frameworks, being formalized analogues, help actors do all of these things and more. They require actors to engage not only in analogic reasoning but in argumentation. Considered in more sociological terms, they are highly developed, meso-level structures that connect institutions—the law—to the domain of individual agency by sustaining deliberation about the nature, scope, and application of norms. In culturalist terms, they enable specifically placed social actors to adjust existing if abstract 'guides to action' to 'the relentless particularity of experience' (Eckstein 1988: 795–6) on a continuous basis. Precedent is basic to judicial governance, allowing it to proceed incrementally from pre-existing institutional materials.

An Example

One can expect to find argumentation frameworks wherever one finds sustained litigation, irrespective of standard distinctions made in comparative law.[11] The basic elements of argumentation frameworks are legal norms—normative statements[12]—as these norms have been interpreted and applied by judges in past disputes. Such structures amalgamate normative materials for the purposes of solving particular legal problems. In tort litigation, for example, judges resolve the problem 'is the defendant responsible for damages allegedly suffered by the plaintiff?' on the basis of how a set of lower-level questions has been argued and answered. These include questions concerning the definition and measurement of negligence, risk assumption, and so on, as well as how chains of causality are to be constructed from the facts. Clusters of rules assembled from different streams of case law typically organize how these different questions are posed and answered. The precise mix of questions and the amalgamated clusters of normative statements that enable legal actors to respond to any given set of questions constitute the relevant argumentation framework.

It may be useful at this point to provide an example of the development of one such argumentation framework, that governing the question 'is flag burning a form of symbolic speech that is protected by the first amendment?' Simplifying, the framework is designed to handle two kinds of lower-order questions. First, what constitutes symbolic speech? Second, to what extent does government have a legitimate, constitutionally recognized interest in limiting rights to symbolic speech?

The answer to the second question came first. In *O'Brien*, the US Supreme Court (USSC 1968*a*) recognized that certain communicative acts, although wordless, might constitute forms of speech protected by the Constitution. Nonetheless, the Court ruled that the government could regulate such acts if certain criteria had been met. *O'Brien* had burned his draft card, in violation of a federal statute, in protest against the Vietnam War and military conscription. The Court let stand his conviction on the grounds that the statute's purposes went substantially beyond the mere stifling of speech; the

[11] The best systematic, comparative research has shown that 'there are no great differences in the use [of precedent] between the so-called common law and civil law systems' (Summers and Tartuffo 1991: 481). Sartor (1994) gives two detailed examples of argumentation frameworks, the first drawn from Italian torts, the second from German constitutional law.

[12] 'All statements (obligations, permissions, authorizations, applicability rules, interpretation canons, preference evaluations, definitions, etc.) susceptible to being used in justifying legal conclusions are fully entitled legal norms', not least because they 'interact reciprocally in the game of arguments and counter arguments' (Sartor 1994: 200).

draft card fulfilled a variety of administrative functions, such as communicating the rights and duties of the bearer. In deciding the case, the Court was led to formulate the following 'test':

> a government regulation is justified if it is within the constitutional power of the government; if it furthers an important or substantial government interest; if the government's interest is unrelated to the suppression of free expression; and if the incidental restriction on alleged First Amendment freedoms in no greater than is essential to the furtherance of that interest.

Announced in this general and prospective way, this rule could be expected to induce argumentation and deliberation of a specific kind, in light of specific fact contexts.

In subsequent cases, the Court gave the concept of 'symbolic speech' a name and conferred on it the protection of the first amendment (for example, USSC 1969). In *Spence* (USSC 1974*a*), the Court defined symbolic speech as those wordless acts in which 'an intent to convey a particularized message [is] present' and that 'in the surrounding circumstances the likelihood [is] great that the message would be understood' by those who witness it. Henceforth, those who would claim first amendment protections for symbolic speech will have to convince the Court that their acts 'pass' the Spence test. They do so in full knowledge that the government will argue the contrary.

By the time of the Texas flag burning case (USSC 1989), the argumentation framework governing symbolic speech was sufficiently developed to structure, more or less completely, the proceedings. The Court, working through the Spence test, agreed with Mr Johnson that his act of burning the flag, on public grounds during the 1984 National Republican Convention, constituted symbolic speech. Proceeding to balancing, the Court rejected Texas' assertion that the State's interest in protecting the flag—because Texans 'venerate' it—outweighed Johnson's speech rights. After all, to the extent that the flag actually did function as a symbolic repository of collective meaning, Johnson's act could not but be collectively understood (Spence). Given its declared interest, the State's prohibition of flag burning could not be dissociated from the direct suppression of speech, and it therefore failed the third prong of the *O'Brien* test.

In *Texas v. Johnson*, the law governing symbolic speech was extended to flag burning, with spectacular political effects. But the decision, from the point of view of the argumentation framework, broke no new ground. Instead it reinforced the centrality of a specific cluster of normative statements, found in *O'Brien* and *Spence*, to these types of situations. Further,

the outcome in Texas is a path dependent one, proceeding from a particular sequence of accumulated judicial refinements of answers to questions that serve to specify the legal interests of those who would litigate symbolic speech cases. It bears mention in this regard that the law of symbolic speech is relatively inflexible, as defined above,[13] and will likely remain so as long as the framework remains stable. At the same time, like any balancing framework, this structure provides courts with a great deal of flexibility in that it allows them to justify a ruling for either the speaker or the government: the fact context, not the law, will vary.

Precedent: Rationality and Adaptation

I assume that all legal actors are rational in the sense of being utility-maximizers. I assume that judges seek to maximize, in addition to private interests, at least two corporate values (see also Stone Sweet 2000: 139–50). First, they work to enhance their legitimacy vis-à-vis all potential disputants by portraying their own rule-making as meaningfully constrained by, and reflecting the current state of, the law. Second, they work to enhance the salience of judicial modes of reasoning vis-à-vis disputes that may arise in the future. Propagating argumentation frameworks allows them to pursue both interests simultaneously. Judges may also seek to enact their own policy preferences through their decisions. Yet the more they do so, or the more that other actors understand them to be doing so, the more likely judges will be to attempt to hide their policy behaviour in legal doctrine. Once policy is packaged as doctrine, it will operate as a constraint on future judicial law-making to the extent that doctrine narrows the range of arguments and justifications that are available to litigators and judges, and to the extent that the law is path dependent.

I assume that litigators are seeking to shape the law in ways that will most benefit their clients, at the moment and in the future. Given the costs of litigation, non-judicial actors will pursue their interests through adjudication only so long as expected returns exceed those costs. In public law litigation, this calculation will be partly conditioned by perceptions of the relative cost and likelihood of achieving the same policy change through other means, such as lobbying for legislative reform. They will do so in full knowledge that there will be others who will work to block legal change. In any given domain of law, rationality means deploying those litigation strategies best adapted to achieving desired legal outcomes, given the current state of the law, as

[13] In response to the decision, the US Congress adopted a statute to protect the flag, a law later annulled by the federal courts. A constitutional amendment designed for the same purpose was proposed but never adopted.

that law is constituted by argumentation frameworks. Those who litigate more frequently in any given domain will invest more heavily in mastering the intricacies of relevant frameworks, the organizational field's codes, and in charting their evolution over time. This is another way of saying that the more the law is path dependent the more we can expect it to branch, and the more incentives legal elites will have to specialize, not least in order to participate in the development of the law.

Implications

If the law is path dependent for the reasons identified, then a range of complex issues are raised. It may also be that, at some times and in some areas, legal institutions are relatively *path independent*. This possibility, too, raises a set of important and unanswered questions.

The Indeterminate Norm and Judicial Discretion

Figure 2.1 helps to simplify and summarize the argument made thus far, namely, that adjudication functions to reduce the indeterminacy of legal norms through (1) use, that is, argumentation, interpretation, application, and (2) the propagation of argumentation frameworks. The line between

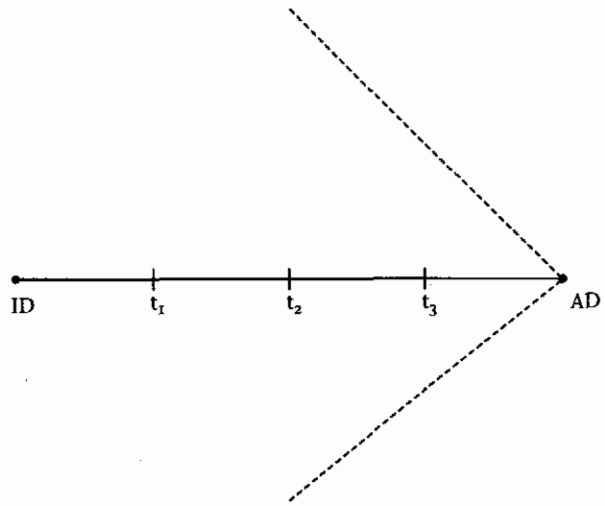

Fig. 2.1. The indeterminate norm and judicial discretion

point ID—absolute indeterminacy—and point AD—absolute determinacy—defines the extent to which any given rule can vary along one dimension, that of determinacy. Point ID represents a theoretical pole at which there exist no stable, collective understandings of the meaning and scope of application of the rule. Point AD represents the opposite theoretical pole, at which the collective understandings of the meaning and scope of application of the norm are perfect; neither ID nor AD is realistic. For any given norm, indeterminacy is a relative condition: it will vary in context and over time. If the law is path dependent, for the reasons given, the position of the norm on this continuum will move from left to right over time, as the rule is adjudicated. The broken lines spreading from point AD define the range of defensible arguments available for how a rule must be applied to resolve disputes arising within that rule's purview. The space constitutes the judge's *zone of discretion* within which the judge chooses among available options and justifications. If the law is path dependent, this zone will narrow over time, as the rule is adjudicated. That is, argumentation frameworks that appear will move from left to right: choice narrows; and frameworks become more articulated. As this space narrows, so does the discretion of the judge.

The argument made here bears directly on certain debates within contemporary legal theory. In so far as legal institutions are in fact path dependent, certain strains of positive legal theory must be considered at least presumptively right or supportive of the theory presented here, and other strains of legal theory must be presumptively wrong. In the former category we find the positivist legal theory of H. L. A Hart and his heirs. Hart (1994: especially 124–41), understood judicial discretion to be inversely proportional to normative indeterminacy so long as judges did their jobs in an 'adequate' or 'reasonably defensible' rather than in an 'arbitrary' or 'irrational' manner. Judicial law-making—the use of discretion—is defensible rather than arbitrary to the extent that it proceeds through analogic reasoning, in light of precedent, and to the extent that it 'renders' legal rules 'more determinate'. The underlying criteria governing a justified extension of the law into new areas is the 'relevance and closeness' of the new situation to a pre-existing, norm-governed situation.

For Neil MacCormick (1978), a close student of Hart's, the primary objective of legal theory is the development of standards for evaluating judicial decisions as 'good or bad', 'acceptable or not acceptable', 'rational or arbitrary'. Bad decisions are those that cannot, ultimately, be packaged as a deduction, that is, where the justification given does not proceed logically from explicated norms and principles. Bad decisions also fail to adhere to the 'principle of formal justice' defined as *like cases shall be decided in like fashion*. For MacCormick, judges can produce good decisions only through a combination of *good* analogical reasoning and *good* justification:[14] precisely the mix of ingredients that generates argumentation frameworks and makes them path dependent. Further, good decisions definitively structure lawyers' strategies (positive feedback): *good* strategies will reinforce existing frameworks and will plead for incremental adjustments to the law from judges (see the related comments in the Introduction and Shapiro's contribution to this chapter).[15] Last, even where consequentialist—policy—arguments trump purely legal ones, judges are likely to choose from a menu comprised of defensible rulings, not all rulings that are imaginable. These points, because they link the production of doctrine to the dynamics of prospective judicial law-making,[16] are generally consistent with arguments made in this chapter and elsewhere in this book.

But they also contrast with what the law and economics movement typically asserts. Law and economics imagines a market for legal institutions. Stable institutions that do emerge are assumed to be Pareto-efficient since if they were not they would be abandoned and replaced. In the world imagined, institutions are plastic—both malleable and disposable—and new ones can be generated with little or no cost. In such a world, the analyst needs to know a great deal about the needs of the economy and about how to calculate transaction costs and efficiencies, but little or nothing about how the law has developed through use. We agree that there is a market for legal rules that litigation and adjudication help to construct. But we see

[14] 'In short, rules can be ambiguous in given contexts, and can be applied in one way or the other only after the ambiguity is resolved. But resolving the ambiguity in effect involves choosing between rival versions of the rule ... once that choice is made, a simple deductive justification of a particular decision follows. But a complete justification of that decision must hinge then on how the choice between the competing versions of the rule is justified ...' (MacCormick 1978: 67–8).

[15] '[T]he requirement upon a litigant's lawyers to frame his case as a legal claim or defense imposes ... two limits on the formulation of the case: first, it must be so formulated as to avoid conflict with existing rules—here the possibility of 'explaining' and 'distinguishing' unfavourable precedents ... must be borne in mind; and secondly, it must be formulated in such a way that it can be shown to be supported by analogies from existing ... law ... preferably authoritatively stated by judges ... or by ... respectable legal writers ... or *faute de mieux* newly minted by counsel as explaining and rationalizing some relevant group of acknowledged rules' (MacCormick 1978: 121).

[16] '[I]t seems that appreciation of the necessary universality of justifying reasons for the decision of particular cases can enable us clearly to explain otherwise puzzling features of the doctrine of precedent. It does so by focusing on the way in which, quite apart from any doctrine of [*stare decisis*] in any official or binding sense, the constraints of formal justice obligate a court to attend to the need for generic rulings on points of law, and their acceptability as generic rulings, as essential to the justification of particular decisions' (MacCormick 1978: 86).

no reason to assume, a priori, that the world constructed is an optimally efficient one.

Efficiency and Rigidity

The allocational efficiency of any institution is not determined by the degree to which that institution was the outcome of a path dependent process. As David (1997: 21) puts it, 'Path dependence is not a sufficient condition for market failure'; the two questions are analytically distinct. The account of legal institutions offered here does not need the assumption that the law is at least as 'efficient' as all other possible institutional arrangements, only that such institutions are relatively functional for actors and make them better off than they would be in a world without such arrangements. Put differently, between some theorized optimality point or curve and a situation in which there is no rule structure at all, there exists a huge range of possible outcomes that may become institutionalized through sequencing and feedback of judicial rulings.

The issue of institutional rigidity or inflexibility is potentially more embarrassing. The concepts of path dependence, positive feedback, and increasing returns are all variables, yet this variance is rarely operationalized, let alone explained. Economic historians have produced persuasive, perhaps definitive, explanations of specific episodes of the path dependent development of technology, production standards, and networked systems. Taken as a whole, these explanations add up to something suggestive of a common framework of analysis. But little attention has been paid to path dependence as a dependent variable. If the argument of this paper is right, than ongoing, intensive adjudication will produce relatively discrete, path dependent lines of case law. There would seem to be no good reason to think that all argumentation frameworks, let alone the legal rules supported by case law, should be equally path dependent in the sense of being inflexible or 'locked in'. After all, courts may abandon precedent and start over; they may borrow doctrinal materials from other lines of case law considered more successful in some way; and the rule of incrementalism may be violated by dramatic new rulings. Likewise, there would also seem to be no good reason to assumed that legal elites should be networked similarly across time and domain. These are empirical issues that deserve sustained empirical attention.

Increasing Returns: To What, for Whom?

In this paper, the phrase 'increasing returns to legal institutions' is suggestive shorthand for the ways in which positive feedback, as managed by

judges, generates network effects , helps to spread and entrench organizational codes, and produces path dependent legal outcomes. Yet increasing returns is technically a function, a measurable relation among variables: for example, the relationship of costs, or investments, to profits. It may turn out to be quite difficult for political scientists to produce compelling increasing returns explanations of political institutions if problems associated with operationalization and measurement or variables are not overcome. Where the law directly imposes transaction costs and conditions the exercise of property rights—for example, economic regulation, product liability—calculating the marginal benefits of adjustment to a nascent standard may be relatively straightforward.[17] Yet, in many salient areas of law and politics, calculating marginal costs and benefits is not straightforward, including in the example given above on symbolic speech.

I notice the problem but do not pretend to have resolved it. It is worth emphasizing that the argument developed here is that *argumentation frameworks are subject to increasing returns*. I do not deny that specific, dispositive legal outcomes can be produced through increasing returns dynamics, and may even become 'locked in' as any norm or standard of behaviour might be locked in. But my focus is on the development of argumentation frameworks, *not on the specific rule governing the situation*; the orientation flows in part from the assumption that legal rules are indeterminate. The production of argumentation frameworks is best modelled as a coordination game in so far as all legal elites—litigators, law professors, and judges—need such codes and are better off with than they would be without such frameworks. Such codes may have distributive consequences in and of themselves, but they also help to define and network the field of adjudication. They are thus basic to the social power of legal elites. The struggle, through litigation, to resolve specific disputes and to make rules to govern specific situations falls outside of cooperative game theory. Argumentation frameworks are the meta-rules that govern this struggle, and meta-rules are probably more easily produced and sustained than are outcomes pursuant to litigation, which is 'play within the meta-rules'. That said, the stability of judicially curated rules heavily depends upon the stability of doctrinal structures.

[17] Yet even here there are problems, which I thank James Caporaso for pointing out. In invoking increasing returns, do we mean to say that 'the marginal cost of making the *n*th decision is less than the average cost of all decisions?' Is there 'a supply curve for judicial decisions, in which the costs decrease and benefits increase at the margins?'

Conclusion

I have argued that certain characteristics of judging and of courts will favour the path dependent development of argumentation frameworks, which are the basic codes that define adjudication as an organizational field. At any given point in time, existing argumentation frameworks, and the substantive legal outcomes they help to sustain, will constitute the necessary causal conditions for the emergence of new frameworks, and thus for new extensions of the law.

In principle, the analyst should be able to chart the development of frameworks which should, in principle, chart the evolution of the legal system more broadly. The development of legal institutions will provoke the development of networks of legal actors specializing in that area of the law. For these actors, existing argumentation frameworks establish the basic parameters for action.

The basic unit of observation is the doctrinal or argumentation framework. If a legal system exhibits path dependent qualities, then the following expectations will be met *and* will be causally related to one another.

First, argumentation frameworks will link, through feedback loops, the activities of litigators and judges over time. Litigation will provide opportunities for judges to construct such frameworks, and subsequent litigants will take their cues from judicial rule-making. If such frameworks are path dependent, the sequence *litigation –> judicial rule-making –> subsequent litigation –> subsequent rule-making* will tend to reproduce itself as a self-reinforcing process sustained by a mix of analogic reasoning and the self-interest of legal actors, that is, potential litigants and judges.

Second, legal outcomes will be both indeterminate and incremental. By 'indeterminate' I mean that, the further downstream we are from any given prior point in the process of selecting or refining legal institutions through adjudication, the more unpredictable will be the content of the rule selected. By 'incremental' I mean that the substantive law and argumentation frameworks produced through this process will be predicated on outcomes and frameworks that had emerged upstream. Stated differently, how cases are sequenced and decided will organize the future.

Third, legal institutions will produce network effects that embed them in wider social practices and interests. Substantive legal outcomes and argumentation frameworks will become parameters for, focal points of, social action. They will be referenced and used by an increasing numbers of individuals and groups operating in an increasing numbers of arenas, including those not otherwise directly associated with adjudication.

Fourth, legal institutions will be relatively resistant to reversal or de-institutionalization. This outcome will be governed partly by the decision-making rules that govern reversal of the judicial law-making and partly by the relative embeddedness, through positive feedback, of outcomes over time.