Legal Data Analytics and the Shifting Balance between Caselaw and Statute

David C. Donald*

Abstract

North (1990) draws a distinction between institutions that are 'created', like a constitution or other written rules, and institutions that 'simply evolve,' like the common law. Eisenberg (1988), using a jurisprudential rather than institutional framework, explains in detail how informal institutions ('social propositions') are processed into law through litigation. Deakin and Carvalho (2010) have shown how legal concepts carry information analogous to the instructions of genetic code and how the interaction of enduring concepts with changing circumstances triggers the consistency in variation of legal evolution. Deakin (2015) further argues that disequilibria in the evolution of common law can necessitate legislative intervention, so that both the common and civil law systems are hybrids requiring both statute and case law.

Building on the above theoretical foundation, this paper offers an information-based theory of the changing balance between case and statutory law in common law systems. Growing amounts of real-time information about disputes, solutions and disequilibria increasingly allow lawmakers to self-reflectively engage in the 'second-order observation' (Luhmann, 1993) necessary to form the general and abstract rules composing statutes. Courts can (with a minimal framework of 'writs' and '*stare decisis*') create law spontaneously despite sparse information. Evidence of this can be found in pre-Imperial Great Britain and the 19th century United States. In environments where information rapidly accumulates, 'second order observation' of law expands, accelerating the production cycle from informal institution to abstract statutory rule. Case law remains important (increasingly so in civil law jurisdictions) because constant social and technological change rapidly render legislation incomplete or obsolete. However, there is no reason to think increased information flows will trigger endless rapid change. Therefore, it is likely that increasing information about problems, solutions and disequilibria will continue to diminish the importance of case law for processing informal constraints into the formal institutions of law.

* Professor, Faculty of Law, Chinese University of Hong Kong. The author would like to thank the Hong Kong Research Grants Council for the generous funding of this work under the Theme Based Research Project, "Enhancing Hong Kong's Future as a Leading International Financial Centre" (T31–717/12-R) and Paul Cheuk for providing valuable research support.

CONTENTS

I. INTRODUCTION

"Common law" is a body of decisions reflecting what the recipients of these decisions believe to be just and culturally acceptable rules and principles.[1] Requests for decisions are brought to the courts by litigants, filtered by 'writs' (or 'causes of action') and decided according to this law commonly held. With the passing of time, an ever increasing portion of this "law" has been embedded in decisions previously made, which are accorded a varying degree of respect according to the relevant rules of *stare decisis*.[2] Successful statutory law will also reflect such generally accepted tenets,[3] although the link between generally held notions of justice and written law is made through delegation of authority to appointed representatives rather than organically

[1] For Holmes, it is "what is then understood to be convenient." OLIVER W. HOLMES, THE COMMON LAW p. 21 (2009, Kaplan Classics). For Baker it is the interaction of general beliefs and practices that form an amorphous body of law which accompanies and fills the more definite body evidenced in written cases and statutes. JOHN H. BAKER, THE LAW'S TWO BODIES: SOME EVIDENTIAL PROBLEMS IN ENGLISH LEGAL HISTORY 2 (2001). For Pound, it is "a taught tradition of voluntary subjection of authority and power to reason." Roscoe Pound, *What is the Common Law?* 4 U. CHI. L. REV. 176, 181 (1936-1937).

[2] The Law of England saw a particularly large swing in just how binding judicial precedent was considered to be. In 1852, the House of Lords could explain in Bright v Hutton 3 HLC 343, 388 "You are not bound by any rule of law which you may lay down, if upon a subsequent occasion you should find reason to differ from that rule." at 388. By 1861, the position on precedent was expressed by Beamish v Beamish 9 HL Cas 273: "It is my duty to say that your Lordships are bound by this decision … The law laid down as your ratio decidendi, being clearly binding on all inferior tribunals, and on all the rest of the Queen's subjects, if it were not considered as equally binding upon your Lordships, this House would be arrogating to itself the right of altering the law, and legislating by its own separate authority." *See* the general discussion of strict and relaxed precedent in H. PATRICK GLENN, ON COMMON LAWS 30-31, 50-51 (2005); RONALD DWORKIN, LAW'S EMPIRE 24-26 (1986). The process of looking more to written products of lawmakers and less to what the public holds as just and convenient corresponds to the increasing size of the legal profession. Fewer primary sources of law are found in open sources of culturally embraced norms and more are found in documents produced by legal professionals. Glenn refers to a similar process as "appropriation" of the *jus commune. See* GLENN, *supra,* at 49-53.

[3] *See* BAKER, *supra* note [●], at 26-29;

through ideas, history and culture processed in judicial proceedings.[4] In both cases, the generation of law entails a transition from informal "notions"[5] and "practices"[6] of justice to formal expressions of what is enforceable by the state.[7]

Both formal and informal stages of law are found within a society applying the law.[8] This process of circulation between law and society is what Luhmann, borrowing a term from biology, refers as "autopoiesis" – "a general form of system-building using self-referential closure."[9] "Autopoietic systems … not only … produce and eventually change their own *structures* but their self-reference applies to the production of other *components* as well. This is the decisive conceptual innovation."[10] Although Luhmann focuses on systematic process rather than individual actors, the result of the autopoietic feedback loop produces the same, general circular flow of authority that Eisenberg schematizes with the concepts objectivity, support, replicability and responsiveness.[11] Within this feedback loop, information regarding popular notions of justice carried forward in a culture is communicated to government institutions, which apply it according to their own institutional framework and information about preexisting hard law, fixing the received notions into determinate legal concepts and bringing them back to the population as positive law. When performing a comparative analysis of legal traditions, Glenn remarks:

> A given tradition emerges as a loose conglomeration of data, organized around a basic theme or themes, and variously described [with different metaphors] …. In the language of modern information theory, a tradition will always include a great deal of

4 *See* the discussion in MELVIN A. EISENBERG, THE NATURE OF THE COMMON LAW 9-10 (1988).

5 While it may appear too eclectic in a discussion that already borrows concepts from both classical jurisprudence and new institutional economics, the hierarchical development of determinateness from "notion" to "concept" and "idea" found in the work of GWF Hegel is extremely useful for plotting the transition from informal to formal law. GWF HEGEL, SCIENCE OF LOGIC (1816). The transition entails knowledge through self-conscious negation of something by distinction from its opposite, which creates and understanding of a notion's boundaries. This is the of process that takes place when an intuitively grasped aspect of cultural support for a given legal position is taken self-consciously into a judicial decision.

6 Customary practices have been assimilated into the legal system in very significant ways. *See* e.g., van Caenegem's discussion of how the jury combined strands of Nordic community justice present in England pre-1066 with the Frankish practice of the *inquisitio* (use of locals as something resembling witnesses) brought to England by the Normans is a good example of how existing practices are woven into legal institutions. RAOUL C. VAN CAENEGEM, THE BIRTH OF THE ENGLISH COMMON LAW 70-80 (2d. 1988)

7 This most formal stage of law would be its most positivist: "a general habit of obedience" to "bodies of persons giving general orders … and receiving habitual obedience." H.L.A. HART, THE CONCEPT OF LAW 24–25 (3d ed. 2012).

8 History shows that attempts at a reverse process whereby foreign beliefs and social institutions are forced on a population through hard legislation (i.e., a top-down cultural revolution) have little chance of enduring success.

9 Niklas Luhmann, "The autopoiesis of social systems," in SOCIOCYBERNETIC PARADOXES: OBSERVATION, CONTROL AND EVOLUTION OF SELF-STEERING SYSTEMS 172, Geyer & Van Der Zouwen, eds. (1986).

10 *Id.* at 174.

11 EISENBERG, *supra* note [●], at 8-13.

> noise …. The fate of present information, and its effect on the future course of a tradition, will depend on the working and processes of the tradition itself.[12]

The fundamental characteristic of this process is the flow of information and the second is the process by which this information is sifted and applied. As such, information flow will be of central importance for the quality of law.[13]

This processing of information expressing notions of justice into formal law is achieved through two general types of institutions: courts and legislatures.[14] Each of these institutions analyze existing data within their own operational parameters to obtain notions of justice and solutions acceptable found in available data resources are sifted and formed into express rules, with the court's rule being in origin specific and the legislature's rule being in origin general. These two institutions receive from different sources their information about both informal notions of justice and existing formal law, and each of them process the information received in different ways, at different levels of generality. Courts hear facts and law as argued by the parties, and have little if any duty or right to explore information not brought to them by these parties in dispute. This is acceptable because, although the court's decision will have retrospective effect on the parties, its rule will have immediate binding effect only on these parties with respect to case in dispute and the exact controlling impact it will have on future cases in controversy will be open to argument as such cases arise. While it is conceivable that a court can set out to create a generally applicable rule rather than settle the dispute at hand, this is a rare exception in judicial decision-making.[15] Legislatures, on the other hand, operate on mandate – usually by popular election – to produce generally applicable rules. It would be highly unusual for there to be procedural restrictions on the information that is admissible to the legislative decision-making process, and the goal of a legislature is to produce workable rules that are general in application. While many acts of legislation are responses to concrete cases of abuse, leading to rules shaped

---

12 H. PATRICK GLENN, LEGAL TRADITIONS OF THE WORLD 15-16 (4d 2010).

13 While the flow of information has always been understood as an important concern for democratic government, it has traditionally been considered from the perspective of freedom of expression and government transparency. Particularly today, it should also be examined from an information theory perspective. This was done when the internet first appeared in the 1990s – *see* e.g., the special issue of the *Journal of the American Society for Information Science*, 45(6), 1994 – but recent changes in the availability of data and its use require serious and continued attention.

14 These can be understood narrowly within the framework of a given legal system. *See* Roscoe Pound, *Common Law and Legislation*, 21 HARV. L. REV. 383 (1908); GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES (1982); BAKER, *supra* note [●]. It can also be understood in a broad form to include civil law systems, including any form of tribunal that hears a dispute between two or more parties and issues a binding order – regardless of whether that order will bind the same court on comparable facts in the future (*stare decisis*) – and a body that following deliberation enacts or issues generally applicable rules with the force of law, without requiring a dispute as prerequisite to the creation of the rule.

15 The "guiding cases" of the Supreme People's Court of China are examples of such general guidelines issued by a court with little regard to the specific, underlying controversy. *See* the Stanford Law School China Guiding Case Project, at https://cgc.law.stanford.edu/.

around such cases,[16] the generally prospective enforceability of legislation means that by definition, rules are created to address a general set of problems.

While sophisticated presentations of information are made in both courts and legislatures, it is the lawmaking activity of legislatures that has the institutional setup best suited to gain from increased availability of data and more powerful tools for data analysis. Working inductively on the basis of evidence, it takes more data to create a rule of true generality than to provide a culturally congruent resolution to a specific case in controversy. In the United States, the process of writing a "restatement" of the law brings the available judicial decisions together in a theoretical framework so as to allow formulation of a general rule, and these restatements make an appropriate and efficient bridge between case law and statute.[17] Indeed, the family of law referred to as "civil law" took its origin from an end product of Roman legal culture– the Code of Justinian – that itself included restating elements, as it followed centuries of judicial decisions and professional commentary by jurists.[18] The interim stage of accumulating rules from judicial decisions is of course not necessary in order to write a statute. A legislative team can consult data in the form of reports and studies from various sources for any legislative project, and the explosion of data on a global scale and increasingly sophisticated analytical tools for processing it has created a very short window from the appearance of a problem to the extraction of its generalized character.

There is considerable evidence that, as in the case of the Roman Empire, history leads from case law to statutory law. At the most basic level, cases accumulate and become information on solutions available to build general models. At a more technical level, from Guttenberg to Zuckerberg the world has experienced a steady (and now torrential) increase in the amount of data available about what people think is just. The course of modern history has brought an increasing amount of law into statutory form, and this is measurable.[19] At the close of the 12th century, an English common law court was fully capable of producing rules on the basis of information provided by the parties in compliance with the requirements of the applicable writ,

---

16 For example, sec. 401 of the Sarbanes-Oxley Act of 2002 specifically targeted a technique used by Enron Corp. chief financial officer Andrew Fastow, by requiring disclosure of transactions previously off-balance sheet and an official report on the extent that special purpose entities are used to facilitate off-balance sheet transactions.

17 *See* e.g., the optimistic verdict in the early days of the American Law Institute: Mitchell Franklin, *The Historic Function of the American Law Institute: Restatement as Transitional to Codification*, 47 HARV. L. REV. 1367 (1934).

18 Peter Stein, Roman Law in European History [●] (1999).

19 In 1982, Calabresi pointed out that "the last fifty to eighty years have seen a fundamental change in American law. In this time we have gone from a legal system dominated by common law, divined by courts, to one in which statutes, enacted by legislatures, have become the primary source of law. Calabresi, *supra* note [●] at 1. *See* also Tom Cummins, *Code Words*, 5 JOURNAL OF LEGAL METRICS 89 (2015), noting that the average annual growth rate of the size of the US Code between 1926 and 2012 was 30%.

and the accepted notion of justice might have been clearer to a court at that time than at present. The court had very little tangible information, but enough to announce common law. When Jeremy Bentham pushed for increasingly codified law in 19th century London, information from the Empire was flowing into that city and being processed by a number of central administrations, business headquarters and the House of Lords itself.[20] The tipping point at which available information makes the formulation of a general rule in statute a more feasible source of law than a depending upon specific rules crafted through judicial decision is not reached by all cultures at the same time, and information is certainly not the only reason why statute will be chosen above judicial lawmaking in a given culture.[21] However, it appears both from a theoretical and an historical standpoint that the information view of legal development is extremely significant, and is poised to become even more significant as "big data" and data analytics are now allowing the generation of generalized information at a speed that even legal scholars and social scientists cannot match.[22]

This paper analyzes the role of information as the key factor in the creation of legal concepts from cultural notions of justice, and its importance in distinguishing between case and statutory. The paper is organized as follows. Part II presents the theoretical models that show law moving from informal to formal institutions in an "autopoietic" manner, processing informal understandings of justice into hard rules, which then feed back into the culture affecting their source. Part III looks at the sources of information ordinarily used by courts and legislatures in relation to the level of generality of the rules produced by each. Part IV examines the evolution from case law to statutory law both local level for individual statutes (e.g., via a restatement of the law) and historically for legal systems. Part V offers conclusions and directions for future research on data analytics and law.

---

20 Many law lords were seconded to the Privy Council, which gathered together appeals from the entire British Empire, and included judges from India and other colonies and legal scholars on the civil law formerly governing captured French possessions and the Islamic law formerly governing Ottoman possessions. *See* Ivor Richardson, *The Privy Council as the Final Court for the British Empire*, 43 VUWLR 103 (2012).

21 The popular revolt in Rome leading to the creation of the Twelve Tables to replace the decisions of the Preators as the sole source of law is an example of a political economic cause of statutory law, one repeated to some extent in the creation of the Napoleonic Code and the German Civil Code (whose name is actually Citizen's Law Book – *Bürgerliches Gesetzbuch. See* GLENN, *supra* note [●].

22 Already in 1995, Bernstein remarked that "[t]oday legal information technology, especially electronic data storage and retrieval, offers the single most powerful cure for Corbin's problem of unintended variation in case law …. The United States Supreme Court, for instance, can transmit opinions to one electronic publisher minutes after handing them down, and the opinions are loaded into databases within twenty minutes of receipt." Anita Bernstein, *Book Review of Restatement Redux: Product Liability*, 48 VAND. L. REV. 1663, 1670 (1995). Today, Harvard's Caselaw Access Project (http://lil.law.harvard.edu/projects/caselaw-access-project) is creating an electronic database of all existing US caselaw that could allow high speed machine processing of the law in a way never seen before.

II. LAW AS A SOCIAL FEEDBACK LOOP

A. Jurisprudence and case law on the social content of law

Holmes nicely summarizes the two-part cultural aggregation of institutional machine and processed contents that is law: "The substance of the law at any given time pretty nearly corresponds, so far as it goes, with what is then understood to be convenient; but its form and machinery, and the degree to which it is able to work out desired results, depends very much upon its past."[23] The hard institutions that process informal understandings of "convenience" are also cultural products, but are more enduring. Van Caenegem's history of common law's "birth" shows how the institutions (as Holmes puts it, "machinery") – in particular the "general eyre"[24] – was formed by culture and history in England and then how this machinery processed and applied law that "was essentially autochthonous, based on known rule and familiar practice."[25]

Eisenberg presents something resembling a "phenomenology" of common law, in that he shows how the judicial activity brings social elements to bear on a given dispute in order to create a decision that "congruently" balances key elements taken from the legal system, social beliefs, science and accepted routines, and institutional restrictions on the court's action.[26] Eisenberg begins with the question of common law's (unelected) legitimacy, explains that it arises because courts apply "social propositions" already known to the litigants, and argues that each respected judicial decision will present a congruent balance of precedent, morals, judicial policies, and the fruits of science.[27] This theory departs from some leading jurisprudence in that it does not assert the law to be "gappy", but rather to have access to the notions of justice necessary to "answer every question."[28] Change occurs as one element of the complementary congruence that makes the law swells to cause imbalance, so that a change in morality or science can cause the court to adjust its interpretation of precedent.[29]

Eisenberg's theory of common law development accords very well with that offered by Baker.

[To be concluded]

---

23 OLIVER W. HOLMES, THE COMMON LAW p. 21 (2009, Kaplan Classics).

24 RAOUL C. VAN CAENEGEM, THE BIRTH OF THE ENGLISH COMMON LAW 20-28 (2nd ed., 1988).

25 *Id.* at 91.

26 Eisenberg, *supra* note [●].

27 *Id.* at 8-43.

28 *Id.* at 159.

29 *Id.* at 104-145.

B. Law as assimilation of informal institutions into formal rules and principles

As explained above, legal scholars have long known that in both common and civil law, the ideas, concepts and rules formulated as law are derived from notions existing in the environment of social institutions. The discipline of institutional economics provided a valuable analysis demonstrating the coexistence of such informal constraints and formal ordering systems. North explains how the two levels of order work together:

> In the modern Western world, we think of life and the economy as being ordered by formal laws and property rights. Yet formal rules, in even the most developed economy, make up a small (although very important) part of the sum of constraints that shape choices; a moment's reflection should suggest to us the pervasiveness of informal constraints. In our daily interaction with others, whether within the family, in external social relations, or in business activities, the governing structure is overwhelmingly defined by codes of conduct, norms of behavior, and conventions.[30]

However, North does not address the manner in which informal constraints "migrate" from cultural notions into formal law. When expressly examining common law, he appears to understand its development as one between older and newer judicial decisions: "Common law is precedent based …. Past decisions become embedded in the structure of law, which changes marginally as new cases arise involving new, or at least in terms of past cases unforeseen, issues; when decided these become, in turn, a part of the legal framework. The judicial decisions reflect the subjective processing of information in the context of the historical construction of the legal framework."[31] North mentions a role of informal constraints in the evolution of common law only as those determining the ideological preferences of judges ("their subjective and ideologically conditioned views of how the world ought to be"[32] This view persists 15 years later, in 2005, when North rhetorically queries, what is the relationship between evolution of informal constraints and "changes in the formal rules?"[33]

Examining a second generation of institutional scholars, we still find that the relationship between informal constraints and formal rules is still not understood dynamically. K and P offer a theory of institutional change that supplements that proposing exogenous shocks, but do so by returning to the argument that the rules of the institution must be "gappy" so that this ambiguity can be exploited: "[W]here we expect incremental change to emerge is precisely in the "gaps" or

---

30 DOUGLASS C. NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE (POLITICAL ECONOMY OF INSTITUTIONS AND DECISIONS) 36 (1990).

31 *Id.* at 96-97.

32 *Id.* at 97.

33 Douglass C. North, Understanding the Process of Economic Change, Kindle Location 1326 (2005).

"soft spots" between the rule and its interpretation or the rule and its enforcement. This is an analytic space that other conceptions of institutions (as behaviors in equilibrium, or as scripts) essentially rule out by definition, but as a practical matter this is exactly the space in which contests over – and at the same time within – institutions take place."[34]

[To be concluded]

B. Law as autopoiesis

Luhmann extensively explores the interaction between the social and the legal, what the institutionalists call informal constraints and formal rules. The interaction is described as a form of communication: "Social systems use communication as their particular form of autopoietic reproduction."[35] This communication can take place at a number of levels of perception.

[To be concluded]

III. INFORMATION PROCESSING IN COURTS AND LEGISLATURES

A. Law is information-dependent

Whether the process be undertaken in deciding a judicial decision or in formulating a rule in a statute, the lawmaker brings its abstract understanding of justice into collision with facts (whether from a case, from historical information or from hypothetical possibilities) for a given situation and creates a rule. This dialectic between abstract justice and concrete fact always requires information about both the notion of justice to be applied and the facts of the case or cases at hand. Whether the resulting rule is immediately applicable (as in common law) or part of a general framework to be applied to future cases (as in civil law), it is a synthesis of the information that is available as part of the law-making process. Congruence of the resulting law with the balance perceived by relevant social actors on the basis of knowledge available and in fact weighed will determine social perception of the law's legitimacy.

[To be concluded]

B. The information content of pleadings

Case law is specific, and need address only a single instance. A case is decided on the basis of pleadings and evidence introduced by the parties. Statutory law is general, and should be

---

[34] James Mahoney & Kathleen Thelen, *A Theory of Gradual Institutional Change*, in EXPLAINING INSTITUTIONAL CHANGE 1, 12-14 (James Mahoney & Kathleen Thelen eds. 2010).

[35] Luhmann, *supra* note [●], at 174.

able to address many individual instances of a given problem. While statutory law could be written by a single legislature plumbing the depths of humanity like William Shakespeare penning a play, statutes are usually assembled by large teams of legislators and their assistants, drawing on information presented by an even larger body of people.

[To be concluded]

C. Social science in lawmaking

General rules require much more information than do specific rules in order to be formulated. An analysis of the informational input into law-making will show how the two types of law at the core of the common and civil law traditions require and use very different sources of information. The best proxies to measure information flows are (i) persons involved in the formulation of a rule, (ii) average pages of material used in formulating a rule, (iii) average number of persons consulted directly and indirectly in formulating a rule, and (iv) operation of bodies (such as UNICTRAL or IOSCO) that support the creation of rules frameworks.

[To be concluded]

IV. INFORMATION IN THE PATH FROM CASE LAW TO STATUTORY LAW

A. The form of law in agrarian England and America

In the UK and the US, the periods of relatively sparse population were the strongest periods of common law. A judge does not need significant infrastructure in order to create rules. One or more judges can travel through a sparsely populated area (like Medieval England or the Western Territories of the US) and provide legal decisions as necessary.

With increasing connectivity of population centers through communication and transportation links, the mass of information available about disputes arising and solutions formulated multiplied. This information accelerated steadily since the late 19th century and began to expand dramatically in the late 20th century.

[To be concluded]

B. The critical mass of information necessary for codification

As large urban masses developed and spawned broad information about general problems (such as in Victorian England and New Deal US), statutes start to replace case law as the primary source of law. In other countries, such as France and Germany, the adoption of civil codes was a project of centralization undertaken in Paris and Berlin by a central government (in Germany, following a 25 year study by 10s of leading professors and jurists). Today, statutes far outnumber

case law as the source for legal rules in all common law countries, and this increase has accompanied an exponential growth in the amount of information available about both problems and possible solutions (facilitated by many national and international bodies dedicated to providing model or contractual solutions to a given problem). In the US, a process of gathering together cases on a given area of law and generalizing the common principles (restatements of the law) is a precursor to the writing of statutes formulating general rules about an area previously governed by case law. Cases also often precede statutes in new areas technological development where little information is available about the process (such as in 1990s e-commerce or 2010s social impact of big data and search engines), and a critical mass of cases followed by a restatement fills the knowledge deficit necessary for legislation.

[To be concluded]

C. Globalization, "big data" and the development of law through machine learning

The convergence of common and civil law has largely paralleled the convergence of other elements of society during the 19th and 20th centuries. Since the 19th century, people in developed countries have had an increasingly accurate knowledge of how people in other countries solve scientific, technical, social and political problems. This is often referred to as 'globalization', and one result is that solutions reached in one country (whether on how to dress or eat or communicate) are copied and transplanted into another. This has been a direct result of increases in communication and knowledge transfer. Science, as the idea that a positively achievable truth or best solution can be distilled from data about base circumstances and partially successful or failed attempts, has also contributed to converging forces. The utility of case law in situations of limited knowledge has led such case decisions to gain increased influence in civil law jurisdictions and the general increase in knowledge about problems and possible solutions has rapidly increased the body of statutory rules in all jurisdictions. The result has been a convergence of common and civil law.

[To be concluded]

V. CONCLUSION: THE FUTURE OF LAW IN DATA ANALYTICS

[To be concluded]