

Modeling Law:

Theoretical Implications of Empirical Methods

Charles Cameron & Lewis Kornhauser

Version 2.0

September 15, 2005

Abstract

We examine a long-standing research program in empirical Political Science, fact-pattern analysis (FPA). We connect FPA to definitions of legal rules in jurisprudence and positive political theory. Foundationally, theoretical treatments view rules as functions partitioning case spaces into equivalence classes. Connecting FPA to formal theory has two advantages: first, many elements of the traditional understanding of legal rules become interpretable in terms of FPA, and vice versa. Second, the methodological issues in empirical FPA become much clearer. In particular, the similarity between FPA and related work in artificial intelligence, expert systems, and machine learning becomes obvious. On this basis, following Kastellec 2005, we critique logit, probit, and discriminant analysis-based FPA as less likely to uncover interpretable legal rules than other techniques developed to induce decision trees from data. As noted by Kastellec, classification and regression trees (CART) appear particularly promising in the legal context. We note the possibility of applying CART-based FPA to many areas of the law, with potentially significant applications in legal education and legal practice. We then examine recent attempts to use FPA to uncover changes in “legal regimes.” We suggest that CART could be employed more successfully in this task. We then examine attempts to use FPA in the long-standing “law vs. preferences” debate in Political Science. In our view, this debate as presently framed is unlikely to be productive. We illustrate our points with estimations using simulated data.

Introduction

Fact-pattern analysis (FPA) is a long-standing and on-going research program in empirical Political Science. The object of a fact-pattern analysis is to inductively derive the legal rule used by a judge, court, or courts from their decisions in a series of cases. In that sense, fact-pattern analysis resembles traditional doctrinal analysis. However, FPA supplements the traditional tools of legal reasoning with statistical and mathematical models. Typically, an analyst codes the content of judicial opinions, noting the presence or absence of particular facts and the extent of others, plus the outcome in the cases. Then, using statistical or mathematical methods, the analyst attempts to recover a good mapping from the facts into the legal judgment.

Fact-pattern analysis seems to have been first created by political scientist Fred Kort, whose work spanned the late 1950s to early 1970s. In fact, Kort innovated two distinct methods for fact pattern analysis, the “logical approach” and the “statistical approach.”

The logical approach employed Boolean algebra to uncover the (presumed) underlying logical structure of legal rules (Kort 1963). In this approach, a legal rule might have the form, “If fact A is present and Fact B, or Fact C, then the judgment should be X, but otherwise it should be Y.” On substantive grounds, Kort’s logical approach has great appeal. But in an era before inexpensive computing, actually deriving Boolean rules was an impossibly cumbersome task with data sets of any appreciable magnitude. Moreover, coding errors, randomness, or mistakes in legal judgments have devastating consequences for a method that presumes deterministic, error-free decision-making. Thus, political scientists failed to develop the logical approach beyond Kort’s early papers.

Not so the statistical approach. Kort was perhaps the first to see that discriminant analysis could be applied to fact pattern data and he undertook several remarkable studies, one at book length (1957, Kort and Mars 1957). However, the statistical approach really took off with the publication of Jeffrey Segal's influential 1984 paper, which used probit analysis to undertake a FPA of search-and-seizure data. Segal demonstrated that this area of the law, famously difficult and notoriously chaotic, could be understood in a fairly simple way, using methods readily available in standard statistical packages. Moreover, Segal connected FPA with a long-standing debate in the study of judicial politics, namely, how much influence do the personal ideologies of judges have on their judgments. He and Harold Spaeth subsequently suggested that FPA supported an "attitudinal" (legal realist) view of the law, because Supreme Court doctrine became measurably more conservative with the addition of more conservatives to the high bench. Segal's 1984 work spawned many follow-up studies in a variety of legal areas including ____ (citations).

Subsequent work has used FPA in several ways. Some analysts have used FPA to study power relations in the judicial hierarchy, since FPA affords a crisp way to detect and measure judicial deviations from Supreme Court doctrine. Thus, Cameron, Segal, and Songer embed Segal's FPA in a formal, game theoretic model of judicial auditing (certiorari) (2000). Kritzer and Richards employ FPA to study the stability of legal rules ("legal regimes") (2002, 2003, 2005). On the basis of their statistical analysis, they make strong claims in favor of a "law" rather than "preference" based understanding of judicial decision making.

A recent paper by Kastellec (2005) advances the methodology of FPA research program in a significant way, by reviving the logical approach and grounding it in feasible and readily implemented statistical methods. Kastellec notes that a modern statistical innovation, classification and regression trees (CART), can be used to uncover Boolean-like legal rules, within a statistical framework that accommodates randomness, error, and missing data. Kastellec estimates CARTs for several noted FPA data sets, showing that the rules uncovered by CART are readily interpretable and have a considerable degree of plausibility.

By any reasonable standard, FPA is a successful research program. Statistical methods for modeling law are well advanced. But what are FPA's *theoretical* underpinnings? What theories of the law do these methods implicitly assume? Do the implicit theoretical foundations make sense? In light of theory, how should we interpret a FPA? Conversely, what does a theoretically informed understanding suggest about appropriate empirical methods? And, what does theory tell us about the use of FPA to uncover legal regimes and adjudicate the hoary law vs. preferences debate?

In this brief paper we can only begin to sketch some answers. However, we suggest that, on theoretical grounds, CART and similar methods are probably preferable to logit, probit, or discriminant analysis. We suggest that CART can be used to study structural breaks in legal regimes, although some methodological issues remain. Finally, we argue that naïve applications of statistical methods to fact pattern data cannot yield other than crude evidence in the law vs. preference debate. To make stronger claims requires the recovery of the underlying structural parameters in judicial utility functions. However, we are somewhat skeptical about the empirical feasibility of this project, due to

identification problems. Rather than pursue an analytic will-o-the-wisp, it may be more productive to use FPA to gain deeper insights into the evolution and performance of legal systems, including the stability of rules and the sources of doctrinal change, the dynamics of statutory interpretation, relations among courts in the judicial hierarchy, among many other topics. The application of theoretically grounded FPA in legal education and legal practice remains largely untapped. Pursuing these matters seem worthwhile goals.

What is a Legal Rule?

A legal rule partitions a case space Ψ . We may describe a case c in Ψ as a vector of characteristics $c = (c_1, c_2, \dots, c_n)$. Doctrine partitions this case space in a complex way. (For more detail see Kornhauser 1992) The partition arises from a series of dichotomous decisions that we might describe as a decision for plaintiff (designated by 1) and a decision for defendant (designated by 0).

For our purposes, we may restrict attention to a particular *cause of action*. A cause of action identifies a set of facts on the basis of which a plaintiff is entitled to prevail; a cause of action partitions the case space into two sets: those cases in which the plaintiff prevails and those cases in which he loses. A cause of action generally can be decomposed into a set of *issues*, on each of which the plaintiff must prevail in order to prevail on a case. We may model an issue as a function from a subvector of facts on a case into the set $\{0,1\}$ where 0 indicates that plaintiff loses and 1 that he prevails.

As an example, consider a simple contract action. To prevail on a contract claim, plaintiff must establish that a contract was formed – an issue that itself can be decomposed into a set of sub-issues of offer, acceptance, and the absence of duress—and

that defendant did not meet her obligation to perform. Plaintiff prevails only if she prevails on each issue.

Now consider a simple case before a court. The case may be clear in the sense that the functions specifying the resolution of each issue in the case are clearly specified and known. The court may simply follow the rule; or it may attempt to distinguish the new case from prior cases. To do this, it refers to some facts not considered in a prior case and argues that some legal issue should be understood to be conditional on these new facts as well as on the facts previously considered. Alternatively, the court might argue that the criterion for deciding a particular issue, on given facts, should be amended. A decision that modifies prior doctrine in either way develops or changes the law.

To connect these ideas with FPA, consider a hypothetical fourth Amendment doctrine involving the law of search-and-seizure. The hypothetical rule has the following structure: If a search is too intrusive, the acquired evidence should be excluded. However, the acceptable level of intrusiveness depends on a reasonable expectation of privacy, e.g., one has a higher degree of protection from intrusive searches in one's home than on the street.

To be more specific, the hypothetical legal is:

$$\left\{ \begin{array}{l} \textit{exclude} \text{ if } \left\{ \begin{array}{l} \text{intrusive} \geq 9, \text{ or} \\ \text{private} \geq 7 \text{ and } \text{intrusive} \geq 4 \end{array} \right. \\ \textit{admit} \text{ otherwise} \end{array} \right.$$

(The indicated cut-off values are arbitrary and purely illustrative).

This legal rule has a natural implementation in a decision tree in which the judge asks herself two questions. First, was the search very intrusive? If so, the evidence should be excluded. If not, did the search occur in a protected private space (e.g., a home)? If

not, the search should be admitted. If so, was the level of intrusive moderate or higher? If so, the evidence should be excluded; otherwise, it should be admitted. The decision tree is shown in Figure 0.

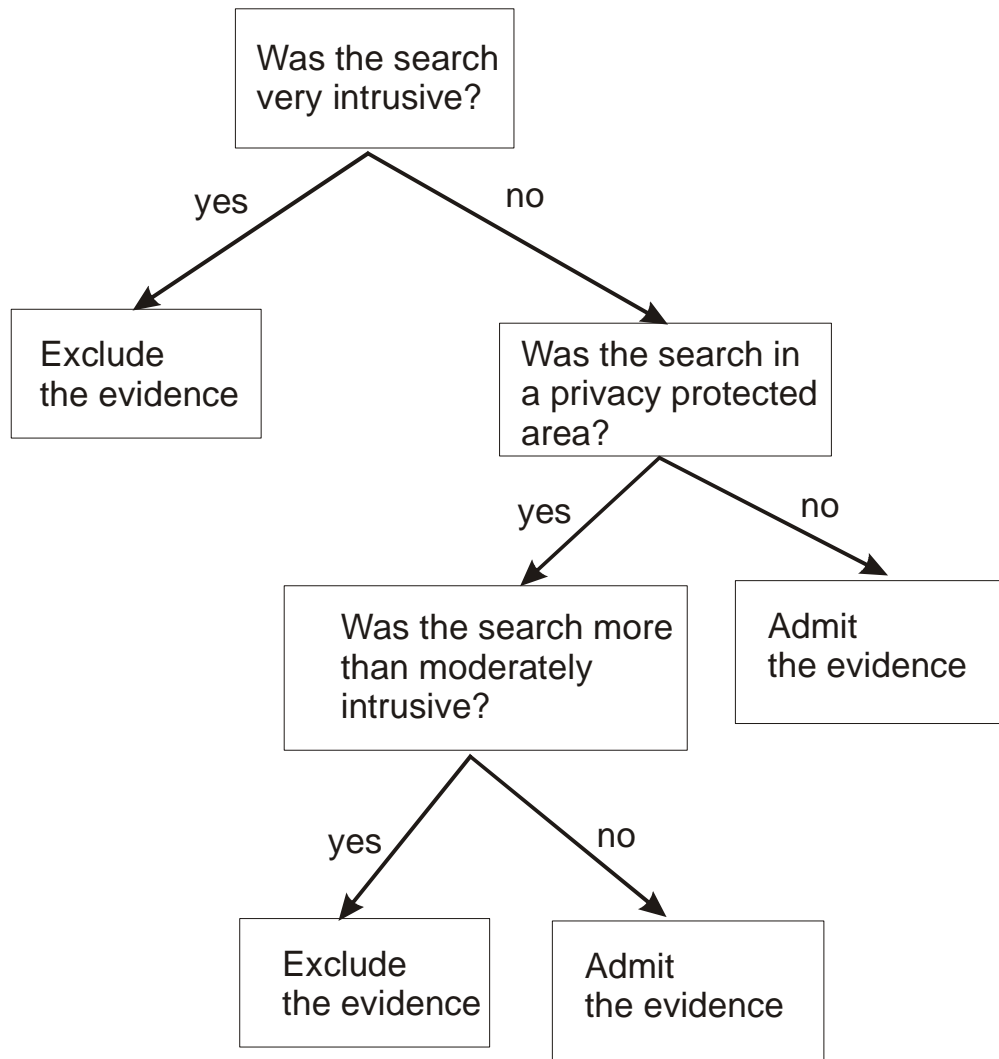


Figure 0. A Decision Tree for the Hypothetical Search & Seizure Rule

The case space Ψ is composed of the Cartesian product of “intrusiveness” and “privateness” (assumed to be compact intervals on the real line). The hypothetical legal rule partitions the case space in the fashion shown in Figure 1.

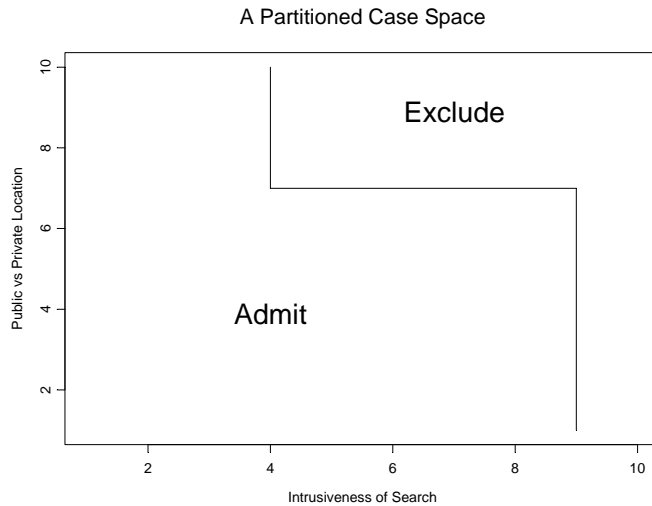


Figure 1. A Partitioned Case Space – Hypothetical Search and Seizure Doctrine.

Empirical FPA and Partitioned Case Spaces

The object of a fact pattern analysis is to recover the legal rule used by judges by observing a series of decisions. In this sense, a FPA is not so different from a traditional doctrinal analysis, except that statistical inference supplements the traditional tools of content analysis and legal reasoning.

To illustrate a FPA, we employ simulated data generated by the hypothetical search-and-seizure rule. The cases come from a ten-by-ten grid superimposed on the case space, so there are observed outcomes for 100 cases, each of whose fact pattern is different from the others. The outcome in each case is assumed to be decided according to the hypothetical doctrine. One could make the exercise more realistic by repeatedly sampling from the space and adding a degree of noise to the data. But we will proceed most directly by estimating models using all the noise-free data.

First consider the consequences of applying a logit analysis to the simulated data. The results are shown in Table 1. By statistical standards, the logit appears to function very well. Results like these would be considered a significant success in a FPA.

Intrusiveness	.1.38*** (.31)
Privateness	1.03*** (.26)
Constant	-.14.66*** (3.33)
Degrees of Freedom	97
Residual Deviance	44.4

Standard error in parentheses. *** = significant at the .01 level.

Table 1. Logit Analysis of Artificial Search & Seizure Data (Regime 1)

Figure 2 displays the results of the analysis, that is, the estimated probability of excluding for any value of intrusiveness and privateness. Then, in Figure 3, we impose the implied partition of the case space, using probability = .5 as the dividing line between the two partitions.

As shown in Figure 3, the logit analysis seriously mis-represents the legal rule that actually generated the data.

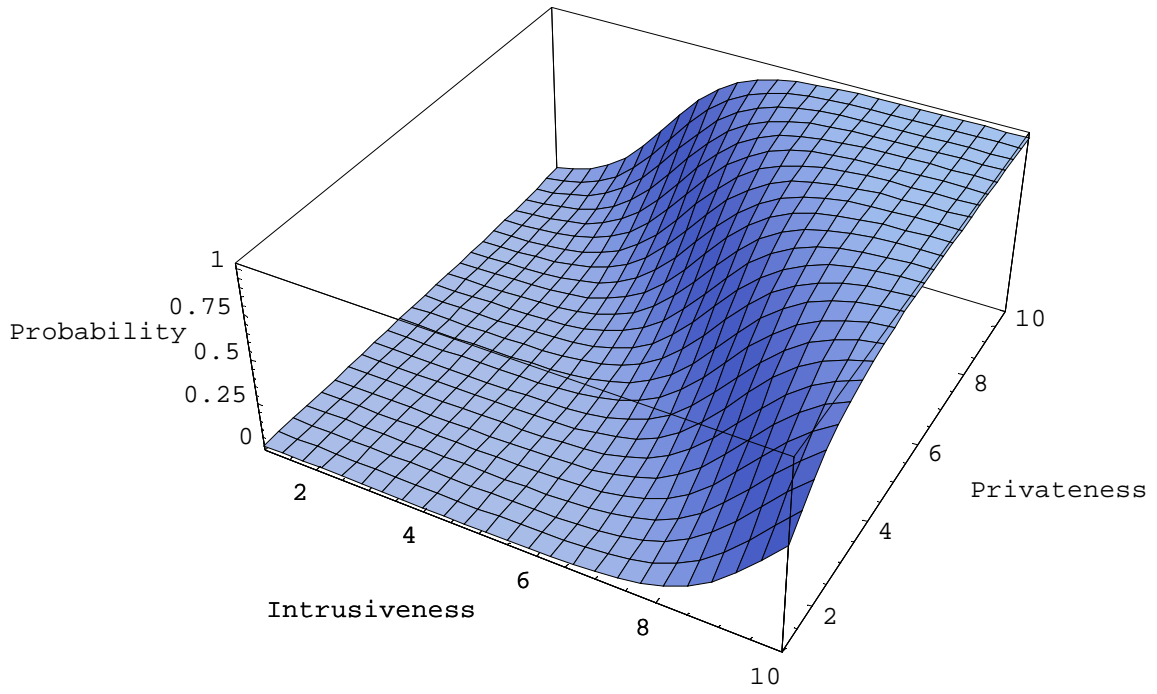


Figure 2. Estimated response surface in the search and seizure example.

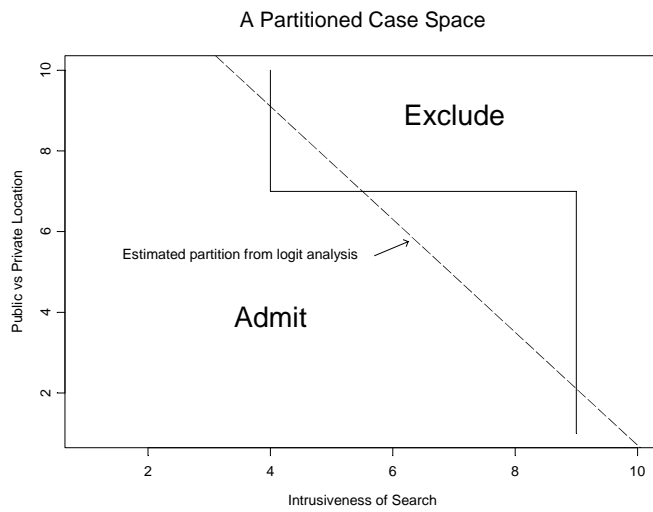


Figure 3. Logit-based FPA misrepresents the actual legal rule.

We now estimate the rule using CART. Specifically, we employ the `tree()` function available in S-plus. Omitting statistical details, we uncover the following decision tree:

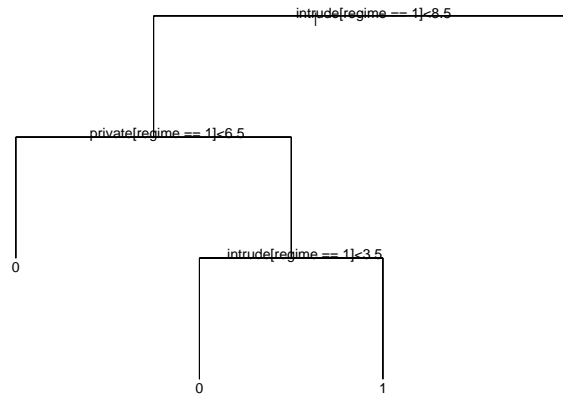


Figure 4. A decision tree induced from the artificial search and seizure data using `tree()`.

Modulo on the integer values used in the grid of cases, the CART recovers the doctrine exactly correctly. The decision tree induced from the data corresponds exactly to that shown in Figure 0. The induced decision tree implies the partitioning of the case space shown in Figure 5. Again in this representation, it will be seen that the statistical procedure virtually perfectly reproduced the legal rule that generated the data.

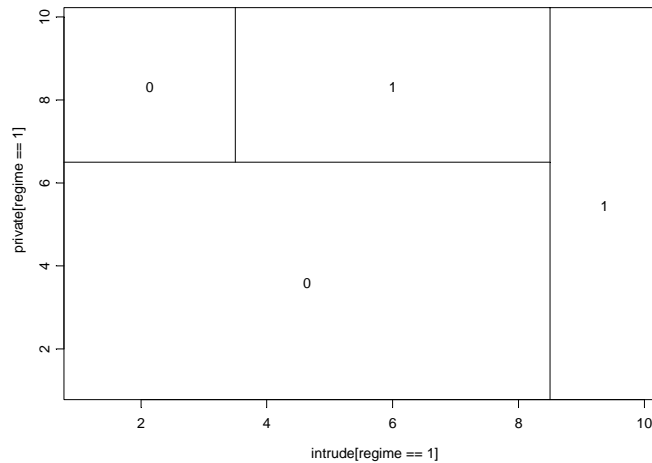


Figure 5. Partitioned case space implied by the induced decision tree: the legal rule is recovered (virtually) perfectly.

Some readers may suspect that we have stacked the deck with this example – and to some extent they would be right. There are legal rules (partitions of case spaces) that logit or discriminant analysis will recover virtually perfectly, which CART will struggle with. An example is the “balancing rule”:

$$\begin{cases} \text{action 1 if } x \geq y \\ \text{action 2 otherwise} \end{cases}$$

The implied partitioning is a 45-degree line in the plane. Given reasonably clean data, a logit or discriminate analysis should recover this rule. Standard ways of implementing CART can approximate the rule, but only with a succession of steps resulting in a lengthy and somewhat artificial decision tree.

In our view, however, balancing rules are a relative rarity. The reason is, balancing rules are hard to implement and often result in inequitable enforcement across cases. In contrast, bright line, dimension-by-dimension standards are far easier to apply and result in greater consistency across cases. Judges strongly favor rules that are easy to

understand and easy to implement. These kinds of rules are friendly to CART – and less friendly to logit, probit, and discriminant analysis.

Kastellec’s innovative use of CARTs to recover legal doctrines highlights a hitherto unnoticed similarity between fact-pattern analysis in legal studies and efforts in other fields, notably the machine learning subfield in artificial intelligence, pattern recognition in engineering, decision table programming in decision theory, and neural networks in neurosciences.¹ As an example, CARTs have been used extensively to study medical diagnoses (citations). In this application as in many others, relatively simple CART-derived decision trees often perform as well as human experts.²

In our view, Kastellec’s innovation opens up new directions for legal research, education, and practice. CART-based decision trees, derived with the assistance of legal experts, might provide quick and simple guides to current doctrine. More than that, they could prove an invaluable tool in doctrinal analysis, but quickly uncovering the logical issues in play when considering alternative doctrinal structures. For example, what are the logical implications of over-turning a given case? If one over-turns Case A, then exactly what other cases must be considered over-turned as well? Simply as a device for teaching legal reasoning, CART-based decision trees may have a valuable role to play.

¹ Murthy 1997 provides a cross-disciplinary review.

² It is interesting that Martin employed CARTS in his recent “bakeoff” between legal experts and computer models, to predict Supreme Court decisions. But the CARTs employed there used arbitrary predictors and made no effort to recover actual legal rules. Citations.

Structural Breaks in Legal Regimes

Legal rules change and evolve over time, often with momentous consequence. How, when, and why do structural breaks occur? This question is of interest to legal historians, political scientists, and legal issue specialists.

Consider a structural change in a legal rule. For example, suppose the hypothetical search-and-seizure rule changes to:

$$\left\{ \begin{array}{l} \textit{exclude} \text{ if } \left\{ \begin{array}{l} \textit{intrusive} \geq 10, \text{ or} \\ \textit{private} \geq 9 \text{ and } \textit{intrusive} \geq 6 \end{array} \right. \\ \textit{admit} \text{ otherwise} \end{array} \right.$$

We show the new partitioning in the right hand panel of Figure 6; the left hand side shows the earlier rule, for ease of comparison. The new doctrine is much more conservative, in that only the most intrusive searches outside the home are disallowed, the threshold for the privacy exception has been narrowed, and even there the allowable level of intrusion is higher. We also estimate a logit FPA from the new doctrine, employing exactly the same methods used earlier. The results are shown in Table 2, and by the dashed line in the right-hand panel of Figure 6.

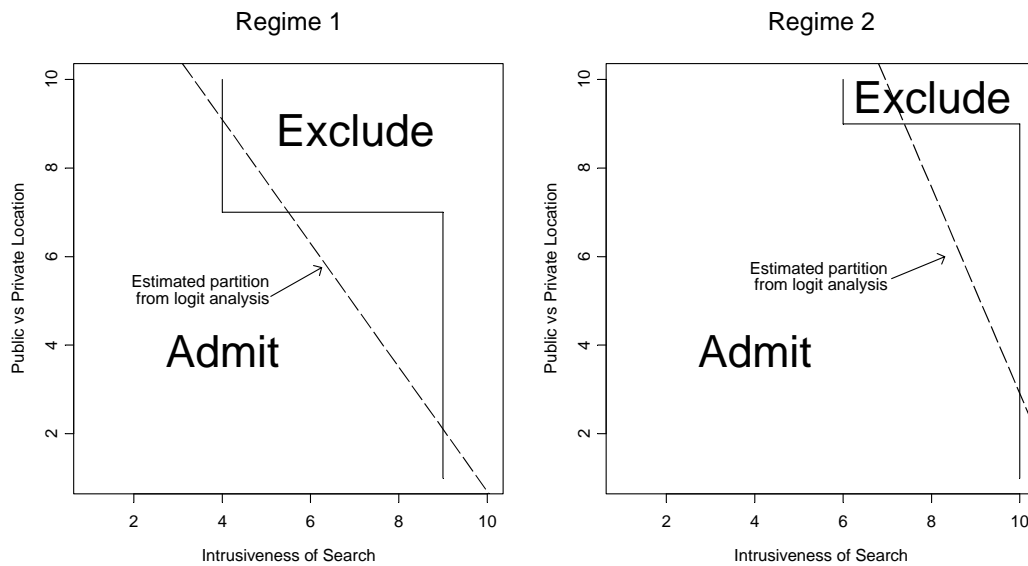


Figure 6. A doctrinal shift.

Suppose an analyst suspected a structural break in a legal regime, that is, a fundamental change in the legal rule at a specific time period. How could one detect it? Kritzer and Richards suggest the following approach. First, undertake logit-based FPAs, before and after the suspected break. Then perform a Chow test, to see if any of the estimated coefficients display statistically significant differences before and after the hypothesized break. In terms of Figure 6, one would estimate the logit shown by the dashed line on the left; estimate that shown by the dashed line on the right, and see if the slope or intercept differs dramatically between the two.

//For brevity we omit doing this in this version of the paper.// Though the logit-based FPA analysis fundamentally misrepresents the actual legal doctrines, the Kritzer-Richards method does flag the structural break in doctrine.

Is it possible to do a similar analysis using CART? In fact, one can do so very simply by adopting methods employed in, for example, Capelli and Reale 2005. Simply add an indicator variable to the data set, the fact pattern “case decided before vs. after the hypothesized structural break”. If the new variable enters the estimated decision tree, so that doctrine varies by regime, and the indicated changes survive standard “pruning” procedures, one may take this as evidence of a structural change. Then, having uncovered the structural break, one can estimate separate decision trees for the two regimes to confirm the differences.

Intrusiveness	.1.37*** (.37)
Privateness	0.63*** (.20)
Constant	-.15.71*** (3.97)
Degrees of Freedom	97
Residual Deviance	37.5

Standard error in parentheses. *** = significant at the .01 level.

Table 2. Logit Analysis of Artificial Search & Seizure Data (Regime 2)

//For brevity, we omit the analysis but this procedure works very well with the simulated data. We do not provide a formal statistical test across CARTs, but it appears feasible to create one.//

Law Versus Preferences

Students of judicial politics will be quite familiar with the hoary “attitudinalist” (legal realist) vs. “legal model” (legal positivist, “neoinstitutional”) debate, which dates back almost three-quarters of century: how much in judicial decision-making is due to the political preferences of judges and how much due is due to the independent influence or authority of the law and legal materials? Perhaps not surprisingly, both sides claim FPA supports their view.

How FPA could support both sides is no mystery. On the one hand, it is easy to show with FPA that Supreme Court doctrine changes when large numbers of conservatives or liberals enter the Court, and that liberal and conservative lower court judges decide cases somewhat differently (employ different doctrines) when they can do so. On the other hand, it is also easy to show with FPA that legal rules have a considerable degree of sticking power, so liberal and conservatives seem to keep rules they might not like, even for extended periods. And, the same analysis showing deviations by lower court judges also shows a great deal of conformity and consistency in decision making, even across judges with different party affiliations.

In our view, naïve applications of statistical methods to fact pattern data cannot possibly yield other than crude evidence in the law vs. preferences debate. For example, when putative liberals and conservatives employ the same doctrine, it may indicate a strong adherence to “stare decisis,” trumping ideological differences. But it may simply reflect power relations in the judicial hierarchy or between the Supreme Court and Congress. Or it may indicate that liberals and conservative share views in a particular area of the law – preferences are still at work, all the way down, but in this area of the

law all judges are “moderates.” If putative liberals and conservatives employ different doctrines, the natural inference may well be that their ideological differences explain their doctrinal preferences. But uncovering evidence of “preferences” in an area of the law where the social costs of legal change are low does not guarantee ideological behavior in other areas where the social costs of change may be very high.

A convincing analysis will need to go beyond simple observation of difference or similarities in doctrine across putative liberals and conservative to uncover structural parameters in judicial utility functions, including the weight placed on the costs of legal change (stare decisis values). One could then make meaningful statements about when ideological differences are apt to weigh large, when not, and with what consequences for doctrine.

Unfortunately, the identification problem in recovering structural preference parameters will be severe. It is an open question whether such a research program could ever succeed.

Conclusion

Connecting fact-pattern analysis to the theory of legal rules has strong implications for empirical work. The new directions for FPA, suggested in Kastellec’s paper and further explored here, open up many possibilities for fresh research into the evolution and performance of legal systems. Worthwhile topics include the stability and duration of legal rules and the sources of doctrinal change, the dynamics of statutory interpretation (which has yet to be considered within a FPA framework), and the relations among courts in the judicial hierarchy, among many other topics. The application of theoretically grounded FPA to legal education and legal practice remains largely untapped, and could

be an exciting area. Pursuing these matters seem worthwhile goals.

References

- Breiman, Leo, Jerome H. Friedman, Richard A. Olshen and Charles L. Stone. 1985. *Classification and Regression Trees*. Belmont, California: Wadsworth.
- Cameron, Charles, Jeffrey Segal and Donald Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions," *American Political Science Review* 94(1):101-116.
- Cappelli, Carmela and Marco Reale. 2005. "Detecting Multiple Structural Breaks in the Mean with atheoretical Regression Trees," Working paper, Department of Mathematics and Statistics, University of Canterbury, New Zealand.
- Clark, Linda A. and Darly Pregibon. 1992. "Tree-Based Models," pp. 377-419 in John Chambers and Trevor Hastie (eds) *Statistical Models in S*. New York: Chapman and Hall.
- George, Tracy and Lee Epstein. 1992. "On the Nature of Supreme Court Decision Making," *American Political Science Review* 86(2): 323-37.
- Hagle, Timothy. 1991. "but Do They Have to See It to Know It: The Supreme Court's Obscenity and Pornography Cases," *American Politics Quarterly* 24:237-61.
- Ignani, Joseph. 1994. "Explaining and Predicting Supreme Court Decision Making: The Burger Court's Establishment Clause Decisions," *Journal of Church and State* 36:301-21.
- Kastellec, Jonathan P. 2005. "Legal Rules and the Classification of Supreme Court Decisions," paper presented at the 2005 annual meeting of the American Political Science Association.

- Kornhauser, Lewis A. 1992. "Modeling collegial Courts I: Legal Doctrine." *Journal of Law, Economics, and Organization* 8(3):441-70.
- Kort, Fred. 1957. "Predicting Supreme Court Decisions Mathematically: A Quantitative Analysis of the 'Right to Counsel' Cases," *American Political Science Review* 51(1): 1-12.
- . 1963. "Simultaneous Equations and Boolean Algebra in the Analysis of Judicial Decisions," *Law and Contemporary Problems* 28(1): 143-163.
- . 1968. "A Nonlinear Model for the Analysis of Judicial Decisions," *American Political Science Review* 62(2): 546-555.
- . 1973. "Regression Analysis and Discriminant Analysis: An Application of R.A. Fisher's Theorem in Data in Political Science," *American Political Science Review* 67(2): 555-559.
- Kort, Fred and David Mars. 1957. *The Administration of Justice in Connecticut*. Storrs, Connecticut: Institute of Public Service, University of Connecticut.
- Kritzer, Herbert M., J. Mitchell Pickerell and Mark Richards. 1998. "Bringing the Law Back In: Finding a Role for Law in Models of Supreme Court Decision-making," paper prepared for the Midwest Political Science
- Kritzer, Herbert M. and Mark J. Richards. 2002. "Jurisprudential Regimes in Supreme Court Decision Making," *American Political Science Review* 96(2): 305-320.
- . 2003. "Jurisprudential Regimes and Supreme Court Decisionmaking: The Lemon Regime and Establishmen Clause Cases," *Law and Society Review* 37(4): 827-840.

- . 2005. "The Influence of Law in the Supreme Court's Search and Seizure Jurisprudence," *American Politics Quarterly* 33(1): 33-55.
- Martin, Andrew D., Kevin Quinn, Theodore Ruger and Pauline Kim. 2004. "Competing Approaches to Predicting Supreme Court Decision Making," *Perspectives on Politics* 2(4):761-68.
- McQuire, Kevin. 1990. "Obscenity, Libertarian Values, and Decision Making in the Supreme Court," *American Politics Quarterly* 18:47-67.
- Murthy, Sreerama K. 1998. "Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey," *Data Mining and Knowledge Discovery* 2, 345-190.
- Segal, Jeffrey. 1984. "Predicting Supreme Court Cases Probabalistically: The Search and Seizure Cases, 1962-1981," *American Political Science Review*