

University of Chicago Law School

Chicago Unbound

Journal Articles

Faculty Scholarship

1999

Must Formalism be Defended Empirically ?

Cass R. Sunstein

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles



Part of the [Law Commons](#)

Recommended Citation

Cass R. Sunstein, "Must Formalism be Defended Empirically ?," 66 University of Chicago Law Review 636 (1999).

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

Must Formalism Be Defended Empirically?

Cass R. Sunstein†

INTRODUCTION

A. Puzzles

In the Nazi period, German judges rejected formalism.¹ They did not rely on the ordinary or original meaning of legal texts. On the contrary, they thought that statutes should be construed in accordance with the spirit of the age, defined by reference to the Nazi regime. They thought that courts could carry out their task “only if they do not remain glued to the letter of the law, but rather penetrate its inner core in their interpretations and do their part to see that the aims of the lawmaker are realized.”²

Thus, for example, the German Supreme Court concluded that a law forbidding “sexual intercourse” between Germans and Jews is

not limited to coition. . . . A broad interpretation is . . . appropriate in view of the fact that the provisions of the law are meant to protect not only German blood but also German honor. This requires that in addition to coition, all such sexual manipulations—whether actively performed or passively tolerated—that have as their aim the satisfaction of one partner’s sex drive in a manner other than the completion of

† Karl N. Llewellyn Distinguished Service Professor of Jurisprudence, Law School and Department of Political Science, The University of Chicago. I am grateful to Jack Goldsmith, Saul Levmore, Eric Posner, Richard Posner, and David Strauss for valuable comments on an earlier draft. Participants in the Symposium at the University of Chicago Law School offered a great deal of help; I am especially grateful to Frank Easterbrook and Frank Michelman. Brian Lehman provided superb research assistance and excellent comments as well.

¹ See Ingo Müller, *Hitler’s Justice: The Courts of the Third Reich* 80-81 (Harvard 1991) (Deborah Lucas Schneider, trans) (discussing the adoption of a “teleological method” of interpretation, which “encouraged judges to identify a particular ideological meaning and intent underlying a given law and then use this ‘intent’ to undermine the wording of the law as such”). Compare the intriguing discussion of the formalism of Italian courts under fascism in Guido Calabresi, *The Story of Italy: The Modern Italian Legal System* (Yale/Video nd) (manuscript on file with U Chi L Rev).

² 72 *Entscheidungen des Reichsgerichts in Strafsachen* 9 (1939), translated in Müller, *Hitler’s Justice* at 101 (cited in note 1).

coition, must cease between Jews and citizens of German or related kinds of blood.³

A lower court went so far as to conclude that kissing could take “the place of normal sexual intercourse” and therefore violate the statute, in such a way as to justify a two year jail sentence.⁴

After the war, the Allied forces faced a range of choices about how to reform the German legal system. One of their first steps was to insist on a formalistic, “plain meaning” approach to law. “[S]uch laws enacted in the Hitler period as had not been voided were to be interpreted in accordance with ‘the plain meaning of the text and without regard to objectives or meanings ascribed in preambles or other pronouncements.’”⁵

Now consider the following cases:

1. The Delaney Clause provides that a color additive may not be used if, after appropriate tests, it is “found by the Secretary to induce cancer when ingested by man or animal.”⁶ The Secretary seeks to allow use of a color additive, even if it is carcinogenic, if the lifetime risk of cancer is equal to or less than one in nine million. This is nine times less than the risk of getting cancer from eating, once every two hundred fifty days, a single peanut with the lawfully permitted level of aflatoxins, and also nine times less than the risk of getting cancer from increased air pollution as a result of spending seventeen hours every year in Denver rather than the District of Columbia. Has the Secretary violated the Delaney Clause?⁷

2. The Occupational Safety and Health Act provides that the Secretary of Labor, in issuing standards dealing with toxic materials, “shall set the standard which most adequately assures, to the extent feasible, on the basis of the best available evidence, that no employee will suffer material impairment of health or functional capacity even if such employee has regular exposure to the hazard dealt with by

³ 70 *Entscheidungen des Reichsgerichts in Strafsachen* 375 (1937), translated in Müller, *Hitler's Justice* at 100-01 (cited in note 1).

⁴ See Hans Robinsohn, *Justiz als politische Verfolgung* 35 (Deutsche Verlags-Anstalt 1977), translated in Müller, *Hitler's Justice* at 102 (cited in note 1).

⁵ Müller, *Hitler's Justice* at xvi (cited in note 1).

⁶ 21 USC § 379e(b)(5)(B)(i) (1994).

⁷ For the court's answer, see *Public Citizen v Young*, 831 F2d 1108, 1122 (DC Cir 1987). See also *Les v Reilly*, 968 F2d 985, 989-90 (9th Cir 1992).

such standard for the period of his working life.”⁸ The Secretary issues a regulation controlling benzene at a level of one part benzene per million parts of air, claiming that there is no safe threshold for benzene. Is the regulation lawful?⁹

3. The Food, Drug, and Cosmetics Act defines “drug” to include an article that is “intended to affect the structure or any function of the body.”¹⁰ It is clear that the Congress that enacted the Act did not believe that tobacco counted as a “drug.” Can the FDA regulate tobacco products as such?¹¹

B. Definitions and Distinctions

It is not easy to define the term “formalism,”¹² partly because there is no canonical kind of formalism. My principal concern here is with formalism as an approach to statutory construction, though I will also touch on formalism in the law of contract and constitutional law. The focus throughout will be on formalism as an *interpretive strategy*.

For present purposes, let us understand formalist strategies to entail three commitments: to promoting compliance with all applicable legal formalities (whether or not they make sense in the individual case), to ensuring rule-bound law (even if application of the rule, statutory or contractual, makes little sense in the individual case), and to constraining the discretion of judges in deciding cases. Thus understood, formalism is an attempt to make the law both *autonomous*, in the particular sense that it does not depend on moral or political values of particular judges, and also *deductive*, in the sense that judges decide cases mechanically on the basis of preexisting law and do not exercise dis-

⁸ 29 USC § 655(b)(5) (1994).

⁹ For the Court’s answer, see *Industrial Union Department v American Petroleum Institute*, 448 US 607, 614-15 (1989).

¹⁰ 21 USC § 321(g)(1)(C) (1994).

¹¹ For the court’s answer, see *Brown & Williamson Tobacco Corp v FDA*, 153 F3d 155, 176 (4th Cir 1998).

¹² Consider H.L.A. Hart, *The Concept of Law* 123 (Clarendon 2d ed 1994):

All rules involve recognizing or classifying particular cases as instances of general terms, and in the case of everything which we are prepared to call a rule it is possible to distinguish clear central cases, where it certainly applies and others where there are reasons for both asserting and denying that it applies. . . . This imparts to all rules a fringe of vagueness of “open texture” This aspect of law is often held to show that any elucidation of the concept of law in terms of rules must be misleading. To insist on it in the face of the realities of the situation is often stigmatized as “conceptualism” or “formalism.”

cretion in individual cases. Formalism therefore entails an interpretive method that relies on the text of the relevant law and that excludes or minimizes extratextual sources of law. It tends as well to favor judicial holdings that take the form of wide rules rather than narrow settlements of particular disputes.

Because formalism downplays the role of extratextual sources, it generally denies courts four relevant powers: to make exceptions to the text when those exceptions seem sensible or even necessary; to allow meaning to change over time; to invoke "canons" of construction to push statutes in favored directions; and to invoke the purposes of the legislature to press otherwise unambiguous words in certain directions. Thus understood, formalism stands opposed to "equity," in the form of a willingness to extend or limit the reach of the applicable text, and also to "purposive" approaches to legal terms;¹³ it is opposed to dynamic (as distinguished from static) conceptions of meaning;¹⁴ it is opposed too to the use of canons of construction, at least if they are based on controversial arguments of policy. Frederick Schauer is a prominent academic defender of formalism;¹⁵ on the current Supreme Court, Justice Scalia is its most enthusiastic proponent.¹⁶

There is certainly no canonical form of antiformalism, and those who reject formalism can offer many different competing approaches. But the antiformalist tends to insist that interpretation requires or permits resort to sources other than the text, and the antiformalist tends as well to support judgments that take the form of narrow rather than wide holdings.¹⁷ The antiformalist is less worried about the exercise of discretion in individual cases and is more concerned about avoiding the kinds of rigidity that can lead to injustices and mistakes. Thus the antiformalist might contend that courts legitimately invoke purposes, or background principles of various kinds, to push statutes (or contracts) in what seem to be sensible directions. There is obviously a relationship between the debate over formalism and the debate over rules and standards, a relationship that will inform the analysis through-

¹³ No one denies that terms take their meaning from context. It is possible to favor the ordinary meaning of the text over the law's apparent purposes without denying the dependence of meaning on context, though of course this view raises complexities.

¹⁴ Compare William N. Eskridge, Jr., *Dynamic Statutory Interpretation* 48-80 (Harvard 1994) (arguing the opposite).

¹⁵ See, for example, Frederick Schauer, *Playing By the Rules* (Clarendon 1991). Schauer's version of formalism is highly qualified.

¹⁶ See Antonin Scalia, *A Matter of Interpretation: Federal Courts and the Law* (Princeton 1997).

¹⁷ See, for example, Cass R. Sunstein, *One Case At A Time: Judicial Minimalism and the Supreme Court* (Harvard 1999).

out. Ronald Dworkin is a prominent antiformalist;¹⁸ on the current Court, Justice Breyer is an outspoken critic of formalism.¹⁹

The real question is "what degree of formalism?" rather than "formalist or not?"²⁰ It is hard to find anyone who believes that canons of construction have no legitimate place in interpretation, or who thinks that literal language should always be followed no matter how absurd and palpably unintended the outcome. No antiformalist thinks that judges interpreting statutes should engage in ad hoc balancing of all relevant considerations. The real division is along a continuum. One pole is represented by those who aspire to textually driven, rule-bound, rule-announcing judgments; the other is represented by those who are quite willing to reject the text when it would produce an unreasonable outcome, or when it is inconsistent with the legislative history, or when it conflicts with policy judgments of certain kinds or substantive canons of construction.

No sensible formalist claims to have offered a complete theory of legal interpretation. Formalist interpretation involves terms that are, or seem, unambiguous. And no formalist should deny that background principles of various sorts are ubiquitous; language has no meaning without such principles, and hence formalism becomes a feasible project only when and because there is agreement on the content of the governing principles. It is another question how the formalist would or should handle terms that are, or seem, unclear, like "equal" or "unreasonable."²¹ Every formalist knows that in many cases it will be necessary to look beyond the text to interpret a statutory term, and to do so self-consciously. The formalist's goal, in such cases, is to develop approaches to interpretation that will push the law in rule-like directions, denying much discretion to judges who will subsequently act on the spot.²²

¹⁸ See his discussion of *Tennessee Valley Authority v Hill*, 437 US 153 (1978), in Ronald Dworkin, *Law's Empire* 20-23 (Harvard 1985). See also James Traub, *Talk of the Town*, New Yorker 35 (June 2, 1997) (reporting Dworkin's discussion of the rule involving basketball players going onto the floor during a fight).

¹⁹ See Stephen Breyer, *On the Uses of Legislative History in Interpreting Statutes*, 65 S Cal L Rev 845 (1992).

²⁰ This suggestion is strongly supported by the comparative analysis in Robert Summers and Neil McCormack, eds, *Interpreting Statutes: A comparative study* (Dartmouth 1991). The analysis shows that while there are significant differences among nations, none is committed to literal language in all cases, and all give text a kind of priority.

²¹ Of course cultural understandings might render these terms unambiguous.

²² This is part of the appeal of formalism in constitutional law. See Scalia, *A Matter of Interpretation*, at 37-47 (cited in note 16). Of course a judge who places a high premium on fidelity to precedent can accomplish the goal of making law more predictable and rule-like, even though such a judge may lack a formalist theory of interpretation. For present pur-

C. Tasks and Claims

In this Article I ask whether a good defense of formalism must be empirical—whether it must depend on factual assumptions and claims. With some qualifications, my answer is “Yes.” Formalism should be defended pragmatically, with close reference to the likely performance of various institutions, and in terms of its consequences; it is not easily defended by reference to quasi-theological claims about the nature of law, legitimacy, or democracy. In this sense a good defender of formalism can and (in my view) had better be a legal realist too.²³ This is not because notions like legitimacy are unimportant—on the contrary, they are extremely important—but because they cannot, without a great deal of empirical supplementation, tell us whether to be formalists or not. And it may indeed be possible to defend formalism in realist terms, as, for example, Lisa Bernstein has done in certain contexts in commercial law.²⁴

More specifically, I claim that formalism, as an approach to statutory interpretation, must be defended by empirical claims about the likely performance and activities of courts, legislatures, administrative agencies, and private parties. To know whether formalism is good, we need to know three principal things:

- whether a formalist or nonformalist judiciary will produce more mistakes and injustices;
- whether the legislature will anticipate possible mistakes or injustices in advance, and whether it will correct them after they occur, and do so at relatively low cost; and
- whether a nonformalist judiciary will greatly increase the costs of decision, for courts, litigants, and those seeking legal advice, in the process increasing the costs associated with unpredictability.

To identify these questions is to say that the ultimate issue is *what interpretive strategy will create lower costs of decision and*

poses I am discussing theories of interpretation and not judges who emphasize precedent as formalists.

²³ In the same general spirit, see William N. Eskridge, Jr., *Textualism: The Unknown Ideal?*, 96 Mich L Rev 1509 (1998) (reviewing Scalia, *A Matter of Interpretation* (cited in note 16), and exploring the practical consequences of textualism); Richard A. Posner, *Bork and Beethoven*, 42 Stan L Rev 1365, 1368 (1990) (criticizing Bork for his failure to supply “convincing reasons” why society should want judges to adopt originalism). See also Calabresi, *Story of Italy* (cited in note 1), which is in the same category insofar as it depicts formalism as a pragmatic approach to concrete institutional problems.

²⁴ See Lisa Bernstein, *The Questionable Empirical Basis of Article 2’s Incorporation Strategy: A Preliminary Study*, 66 U Chi L Rev 710 (1999); Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code’s Search for Immanent Business Norms*, 144 U Pa L Rev 1765 (1996).

costs of error (the reference to “costs” should not imply economic reductionism; the term is a placeholder for the real-world difficulties—in terms of unpredictability of outcomes—that might follow from one or another interpretive strategy). In a legal system in which the legislature is extremely careful before the fact, and highly responsive to judicial interpretations after the fact, formalism might well make sense, especially if a nonformalist judiciary would create confusion and make blunders of its own. By contrast, a nonformalist approach would make sense in a legal system with an excellent judiciary and a legislature that is both careless and inattentive.

The principal qualification to my basic thesis—that formalism must be defended empirically—comes from the fact that without normative claims of some kind, it is impossible to know what counts as a “mistake” or an “injustice” in interpretation, and hence the idea of “error costs” seems dependent on the antecedent theory of interpretation, in which case the theory cannot be chosen on the basis of an (antecedent) inquiry into error costs.²⁵ I discuss this issue in more detail below.²⁶ For now let me suggest that people who disagree about what counts as an “error” should also agree that the case for formalism depends largely on the capacities of relevant institutions, and in particular on the capacities of courts and legislatures. Through agreement on this point, it may be possible to bracket many of the normative debates and to make a great deal of progress by investigating empirical questions. At the very least, it will be possible to see through some prominent but quite implausible arguments for and against formalism (involving, for example, legitimacy, Article I, or separation of powers) and to see the issues on which a choice among interpretive methods really depends.

Indeed, I hypothesize that it is disagreement over the underlying empirical issues—not over large concepts of any kind—that principally separates formalists and nonformalists. Of course, decisive empirical evidence is not easy to find on such questions, partly because of the difficulty of controlling for confounding variables. But by recognizing the empirical issues as such, we would have a better sense of what we are disagreeing *about*, itself a large gain. And with more research, we could learn

²⁵ There is an analogy here with debates over cost-benefit analysis. Many people can agree that cost-benefit analysis makes sense, but they may disagree, sharply, over what should count as a cost or a benefit, and how much various things that count should be allowed to count. By itself, the idea of cost-benefit analysis is empty; the question is whether it can be specified in such a way as to command, and deserve, widespread support.

²⁶ See Section IV.E.

a great deal more about the underlying questions. Of course empirical evidence cannot by itself yield any normative conclusion; the question is whether the normative background is sufficiently shared, so that certain findings would lead apparent antagonists in the same general direction. I believe that this is the case here, and that with imaginable empirical findings, both formalists and antiformalists should be flexible enough to move in the direction of their apparent adversaries.²⁷ This is the great hope of empirical research: to soften legal and political debate by showing that disputes about values are often disputes about facts, at least in the sense that certain answers to factual questions can make it unnecessary to resolve disputes about values.

I should emphasize that I do not actually defend formalism here.²⁸ The important point involves not the conclusion, but the identification of the grounds on which formalists and antiformalists might sensibly disagree. Thus my suggestion to those who are attracted to either Justice Scalia's or Ronald Dworkin's approach to legal interpretation is this: Doesn't the choice between the competing approaches depend on predictive judgments about which a great deal remains to be learned? Don't the relevant disagreements turn, in large part, on those judgments?

I. CONCEPTUAL PRELIMINARIES

Let us distinguish between two phenomena that go under the term "formalism." The first is captured in the pretense that all or most legal terms can resolve hard cases. This pretense is illustrated by (for example) the view that the term "liberty" necessarily includes freedom of contract, or that a requirement of "equal protection" necessarily prohibits (or for that matter requires) affirmative action. The problem with this species of formalism is

²⁷ In this way I suggest that the debate over legal interpretation is not so different from the debate over (for example) the minimum wage: If the minimum wage really does substantially increase unemployment, those who endorse it should be prepared to rethink; if the minimum wage has no adverse effect on employment, those who condemn it should be prepared to rethink.

²⁸ Nor do I mean to suggest, via the title, that any side bears the burden of proof. For clarification, I might add that while formalism captures part of the territory, I believe that it is an inadequate approach to statutory and constitutional interpretation in the United States. In statutory interpretation, I would favor a presumption in favor of that species of formalism that emphasizes the natural meaning of statutory terms, taken in their context; but I would allow the presumption to yield where the result is absurd or where background principles, constitutional or otherwise, argue in favor of pressing the text in particular directions. Cass R. Sunstein, *An Incompletely Theorized Conviction*, 112 Harv L Rev (forthcoming 1999). Constitutional law seems to me to have an even larger nonformalist tendency, and properly so, because it has so many common law characteristics. But a defense of these propositions would go well beyond the current discussion.

that it is dishonorable, a kind of fraud. The relevant legal terms have many possible meanings, and a language lesson is insufficient to justify the choice. But a second possible species of formalism is entirely honorable. On this view, statutory terms should be understood in accordance with their ordinary meaning in the relevant community. When there is such an ordinary meaning, nothing is fraudulent about this approach to legal interpretation. The question is whether it is good, not whether it is fraudulent.

What we might call "the new formalism" is an intriguing blend of realist and formalist arguments.²⁹ It amounts to an embrace of formalism *because of the good effects that formalism has*. Formalism might, for example, increase predictability for all concerned, in the process greatly decreasing the costs of decision. Clear rules in the conflict of laws might be good because they promote planning; a textualist approach to statutes might produce greater clarity in the law. Some of these pragmatic and empirical points appear, for example, in Justice Scalia's argument on behalf of the original meaning of legal texts.³⁰

I think that this is a genuinely new formalism,³¹ sharply distinguished from the old, which was, at least in its self-presentation, not at all empirical and pragmatic in character. At least in public law, the principal defenses of formalism—and the principal challenges to formalism—have been abstract and conceptual, and have not depended on empirical claims at all. These are the arguments that I wish to reject. I believe that the conceptual arguments are best taken as pragmatic arguments in disguise; that claims about "legitimacy" are really consequentialist claims about what system of interpretation is likely to have good effects. Those who wish to defend formalism should move from (unhelpful and question-begging) abstract conceptual arguments to (more promising and tractable) concrete empirical claims.

I will say much more about these issues below. For the moment let us try to locate the less abstract grounds on which for-

²⁹ This is the tendency in, for example, Bernstein, 144 U Pa L Rev 1765 (cited in note 24); Adrian Vermeule, *Legislative History and the Limits of Judicial Competence: The Untold Story of Holy Trinity Church*, 50 Stan L Rev 1833, 1885-96 (1998) (arguing that problems in judicial competence make legislative history an unhelpful source for statutory interpretation). See also Richard Epstein, *Simple Rules for A Complex World* (Harvard 1995), which, when stripped of its libertarian foundations, seems to me a policy-oriented, insistently instrumental argument for rules as the best way of minimizing the sum of decision costs and error costs.

³⁰ See Scalia, *A Matter of Interpretation* at 44-47 (cited in note 16).

³¹ Of course the old formalism was sometimes embraced on more pragmatic terms.

malists and antiformalists might be disagreeing. The law of contract is the place to start.

II. DEFAULT RULES, CONTRACTUAL AND STATUTORY

A major task of contract law is to develop default rules, even in cases in which the contract seems clear. Consider, for example, a case in which an employee complains that he has been fired without any showing of cause. If the contract says nothing about whether employment is “for cause” or “at will,” the court will have to write an implied term, applying a default rule that governs such situations. Indeed, the need for default rules is even more pervasive than it seems. No term, contractual or otherwise, has meaning in isolation; meaning is always a function of both *culture* and *context*. When a contractual term seems clear and self-sufficient, and to require no resort to default rules, it is only because there is general agreement on the relevant culture and the relevant context. The same is true for statutory interpretation.

Good formalists do not deny this point; formalism is a doomed enterprise if it is an effort to give meaning to terms apart from cultural understandings and context. Law can never be autonomous or rule-bound in that sense. Every term requires interpretation. Thus the cases described in the Introduction above³² seem to involve “plain” meaning only to the extent that there is general agreement about the governing background understandings. And when culture and context change, a term that once seemed vague (“equality”) may become clear, and a term that once seemed clear (“carcinogenic,” “psychopathic”) may become vague.

In the law of contract and statutory construction, default rules fall in several categories. Some such rules are *market-mimicking*. Here the goal is to replicate what the parties would have done if they had made specific provision on the point. In the law of contract, whether an employment agreement is for cause or at will might be determined by asking what the parties would have done if they had made an explicit provision on the termination of employment. The key question, “What would the parties have done?” is an empirical one. There is an analogue in statutory construction. If, for example, we do not know whether the Endangered Species Act³³ applies outside the territorial boundaries of the United States, we might construe the statute by asking how Congress would have resolved the issue if it had done pre-

³² See text accompanying notes 6-11.

³³ 16 USC §§ 1536 et seq (1994).

cisely that. The notion of “imaginative reconstruction” in statutory construction embodies this understanding of the interpreter’s task.³⁴

Other default rules are *information-eliciting*. The point of such rules is to put on the person most likely to speak the obligation to obtain a clear statement one way or the other. Hence information-eliciting default rules are sometimes described as “penalties.”³⁵ Such rules attempt to ensure that both parties (and also the courts) will have information at the time the contract is made; they counteract the risk of strategic nondisclosure. Suppose, for example, that a court does not know if the parties would have made an employment contract at will or for cause. The court might want to impose before-the-fact incentives that would encourage the parties to speak unambiguously, both to one another and to the court. It is reasonable to think that if the court wants to do this, it should hold against the employer—that is, it should say that contracts are presumed to be for cause, on the theory that the employer is most likely to make suitable corrections if this provision does not fit the needs of the parties. The key questions here as well are empirical in nature: Who is most likely to make an explicit provision on the subject? At what cost will they do so?

It is also possible to imagine information-eliciting default rules in the area of statutory interpretation. For example, a court might conclude that statutes should be presumed not to apply outside the territorial boundaries of the United States—if it believes that this approach is most likely to elicit a clear statement from Congress—so that all parties (and the courts) are informed, in advance, of the relevant outcomes.³⁶ It should be clear that an information-eliciting default rule can serve formalist goals by decreasing uncertainty and judicial guessing games. The empirical

³⁴ See, for example, Richard A. Posner, *Statutory Interpretation—in the Classroom and the Courtroom*, 50 U Chi L Rev 800, 817 (1983) (“[T]he task for the judge called upon to interpret a statute is best described as one of imaginative reconstruction. The judge should try to think his way as best he can into the minds of the enacting legislators and imagine how they would have wanted the statute applied to the case at bar.”). See also Richard A. Posner, *The Problems of Jurisprudence* 273 (Harvard 1990) (“When confronting unclear statutes, judges . . . have to summon all their powers of imagination and empathy in an effort . . . to place themselves in the position of the legislators who enacted the statute that they are being asked to interpret. They cannot only study plain meanings; they must try to understand the problem that that the legislature faced.”).

³⁵ See Ian Ayres and Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 Yale L J 87 (1989) (introducing the concept of penalty default rules).

³⁶ See, for example, *EEOC v Arabian American Oil Co*, 499 US 244, 248 (1991) (employing such a rule).

questions for statutory default rules are: Which party is in the best position to obtain an explicit provision? What is the cost of doing so?

These points raise the question how a court should choose between a market-mimicking and an information-eliciting default rule. One possibility is to ask *whether the court is confident that it can generate a good rule of either kind*. If a court knows that the parties to a contract would have agreed on some term X, it seems plain that the court should simply adopt that term; there is no need to try to elicit the relevant information. But if a court has no idea what the parties would have done, and if it has a good idea how to allocate the burden so as to elicit a clear statement, an information-eliciting default rule seems better. Thus, for example, an at will background rule would make a good deal of sense if a court was clear that this is the rule that these parties, or most parties, would accept. But if a court is unsure what the parties would do, and if it has confidence about which party would be in a good position to obtain a clear statement, it should adopt an information-eliciting default rule. If the court has no idea what term employers and employees would choose, it might adopt a waivable for cause rule, on the theory that this rule will force employers to come up with a different result if that seems desirable. But if the party in the best position to obtain an explicit provision can do so only at very high cost, perhaps a market-mimicking default rule would be better.

This inquiry into judicial confidence about potential default rules can be understood as a simple way of asking a more systematic question about how to reduce the costs of decision and the costs of errors. (As noted, we should not be too reductionistic here; decision costs and error costs differ from one another, and they include costs of qualitatively diverse kinds.) Decision costs can be understood as the costs of finding out what the law is—a cost faced by courts (attempting to discern the legal rule while deciding the case) and by ordinary citizens (having to invest resources in figuring out the content of the law). Error costs involve both the number of mistakes and the magnitude of mistakes. On some empirical assumptions, a market-mimicking default rule would minimize the relevant costs; on other assumptions, an information-eliciting rule would do so. An independent point involves the *variability* of mistakes. It is possible, for example, that a nonformalist judiciary will produce highly variable, even random errors, while a formalist judiciary will produce the same number and magnitude of mistakes, but in a way that involves no variability and a great deal of predictability.

These ideas can help guide the selection of default rules. If a court deciding a contract case knows that a certain background rule is what the parties would have chosen, it has (by hypothesis) discovered the right rule and perhaps at low cost. If a court is unaware of the market-mimicking rule, but if it knows what rule will elicit clear statements from the parties, it can impose (at low cost) a rule that will produce the right outcome. When the court lacks confidence in either kind of default rule, it might ask in which rule it has comparatively greater confidence, and choose that one. In the law of contract, an error can be defined as an outcome different from what the parties would have chosen if they had made explicit provision on the point; this simple definition can work for the vast bulk of cases.³⁷ Courts that do not care about what the parties would have done, and that look instead to the objective meaning of contractual terms, should be taken to be saying that this method of interpretation is most likely to minimize decision costs and error costs.

Much the same can be said about statutory interpretation, though there are disanalogies as well as analogies. Begin with the suggestion that when in doubt about the meaning of an ambiguous term, courts should presume in favor of the outcome that Congress would have reached if it had done so expressly.³⁸ Suppose, for example, that one linguistically plausible outcome would produce an extraordinary outcome, one that would risk absurdity. Courts would do well to presume that the absurd outcome was not intended; the presumption could be rebutted by showing something in the context or background to suggest that Congress would have reached that result if it had been explicit on the question at hand. This would be a kind of market-mimicking approach. But if courts have no idea how to understand the term—if the term is ambiguous and if no one can say what resolution Congress would have reached—courts might adopt an interpretation that would impose on those most likely to obtain a clear statement the obligation to do precisely that. Consider, for example, the idea that statutes and treaties will ordinarily be construed favorably to Native Americans.³⁹ This idea can be defended by reference to the fact that such statutes and terms are nearly

³⁷ In some cases, an error may be defined differently, as when some public policy concern suggests that party autonomy is not the foundation for the outcome. This would require a different analysis.

³⁸ This suggestion is controversial, to be sure, and qualifications will be introduced as the discussion proceeds. Note that it does not mean that courts should interpret otherwise clear terms by asking about Congress's subjective intentions.

³⁹ See *Montana v Blackfeet Tribe*, 471 US 759, 766-68 (1985).

dictated by Congress, which is therefore in the better position to get the terms that it wants.

There are, however, some special complexities in the statutory area. First, it can be very hard—even harder than in contract law—to know who is in the best position to obtain a clear statement one way or the other. Partly because of this difficulty, statutory default rules tend, in practice, to be based on considerations of public policy,⁴⁰ not on allocating the burden to the “best eliciter.”

It is also possible to doubt the analogy between the contract law question (“what the parties would have done”) and the statutory question (“what Congress would have done”). In the law of contract, there may be agreement on what counts as a mistake: doing something different from what the parties would have done.⁴¹ This is much less clear in the context of construing statutes; many people think that ambiguous terms should not be understood by asking what Congress would have done.⁴² Perhaps the question is irrelevant or only part of the picture. And perhaps any judgment about what counts as an “error” is an artifact of the theory, and hence it may be impossible to know whether we have an error at all unless we have first selected the appropriate theory. I will return to this question below.⁴³ For now let me suggest that where statutes are genuinely ambiguous, an understanding of what Congress would have done is at least a reasonable place to start the analysis. Those who reject this view (“imaginative reconstruction”) even as a starting point might ask why they do so, and whether empirical questions are relevant to their judgment; I take up this question below.

A final point is that in statutory interpretation, as in contract law, the default rule may have an “endowment effect” and thus tend to stick. The preferences of relevant actors may not be entirely independent of the legal rule. If the background rule is that statutes do not apply outside the territorial boundaries of the United States, those who favor the statute in question may

⁴⁰ See the excellent discussion in Jane S. Schacter, *The Confounding Common-Law Originalism in Recent Supreme Court Statutory Interpretation: Implications for the Legislative History Debate and Beyond*, 51 *Stan L Rev* 1 (1998) (discussing the large role judicially crafted policy norms play in Supreme Court statutory interpretation).

⁴¹ In contract law, however, formalism might be defended not as getting what the parties would have done, but of getting clarity for the parties and the courts, in a way that minimizes the sum of error costs and decision costs.

⁴² See Frank H. Easterbrook, *Statutes' Domain*, 50 *U Chi L Rev* 533, 534 (1983) (arguing that courts should refrain from acting in situations of statutory ambiguity even when they know what Congress would have done).

⁴³ See Section IV.E.

demand a high “price” to give up their presumptive right, whereas if the background rule is against application, those who favor the statute may demand a much lower “price.” Thus background rules can affect the judgments, beliefs, and preferences of the parties to contracts and the parties to law, a phenomenon that much complicates the idea of “market-mimicking.”

Thus far the discussion has involved genuine default rules. But some interpretive rules are *based on public policy and inalienable*; they are, in that sense, far more than mere default rules. Constitutional law amounts to the equivalent of inalienable default rules for statutory interpretation; of course Congress is not permitted to contract around constitutional requirements. These points are pretty obvious, but a less obvious point is in the background: There is a continuum from “other things being equal” default rules to “superstrong” default rules, which require an especially clear statement from the parties or from Congress, to genuinely inalienable default rules.

III. IS FORMALISM A GLOBAL DEFAULT RULE?

A. In General

An understanding of default rules helps to reveal the empirical dimensions of any dispute over formalism. As noted, a central formalist goal is to reduce the burdens of on-the-spot decisions, above all by eliminating the need for the exercise of discretion in particular cases, and by making sure that law is as rule-like as possible, in a way that promotes predictability for parties and lawmakers alike.⁴⁴ Thus we can find formalism in the set of ideas most closely associated with Justice Antonin Scalia. But consider Scalia’s vivid response to the charge of formalism: “[O]f course it’s *formalist*! The rule of law is *about* form.”⁴⁵ This is an emphatically formal response to the charge of formalism; it amounts to a defense of a formal approach not by reference to considerations of democracy, policy, or anything else, but by reference to an “is,” which is taken as an unanalyzed given—here, what the law “is about.” Plainly this way of defending formalism is unsatisfactory. Formalism cannot be justified by an “is”; it must be justified by reference to something outside itself (as Justice Scalia is obvi-

⁴⁴ See Cass R. Sunstein and Edna Ullmann-Margalit, *Second-Order Decisions*, 110 *Ethics* (forthcoming 1999) (evaluating ex ante and ex post burdens).

⁴⁵ Scalia, *A Matter of Interpretation* at 25 (cited in note 16).

ously aware).⁴⁶ The trick is to identify what the justification might be.

The discussion thus far provides several clues. Perhaps formalism can operate as a kind of massive or global default rule, one that operates, on the whole, *either* as a good market-mimicking default rule *or* as a good information-eliciting default rule. More precisely, a formal approach to interpretation might produce simple, clear rules (in the law of contract and the law of statutory interpretation) that reduce the costs of both decision and error. Formalism might reduce decision costs if it is relatively easy for people in a formalist system to know what the law is. Formalism might reduce error costs if it leads to fewer and less damaging mistakes than any alternative (again defining mistakes provisionally and, in cases of ambiguity or genuine doubt, as outcomes that deviate from what the enacting Congress would have done), and if any mistakes are corrected, quickly and at low cost, by the parties or by Congress.

It should be at once apparent that these propositions will depend on empirical issues. Suppose, for example, that a judge who rejects formalism can (at low cost) generate excellent market-mimicking default rules; suppose, that is, that a judge can temper contractual or statutory language in such a way as to discern what the relevant parties (to the contract or in the legislature) would have done. If this is so, the case for formalism is much weakened. But suppose—by contrast—that judges can generate information-eliciting default rules, formalist in character, that work extremely well in creating good *ex ante* incentives for contracting parties and for Congress. If this is so, the case for formalism is greatly strengthened. And to know which of these alternatives holds, we need to know a great deal about likely institutional performance and about underlying facts. Think, for example, of the famous case of *Church of Holy Trinity v United States*,⁴⁷ where the Court appeared to reject a formal approach in favor of one based on legislative “intention.”⁴⁸ If Congress would

⁴⁶ See *id.* at 31-34. This is denied in Ernest J. Weinrib, *Legal Formalism: On the Immanent Rationality of Law*, 97 Yale L.J. 949, 951 (1988) (“Formalism postulates that law is intelligible as an internally coherent phenomenon.”). I believe that Joseph Raz’s response is convincing. See Joseph Raz, *Formalism and the Rule of Law*, in Robert George, ed., *Natural Law Theory* 309 (Clarendon 1992).

⁴⁷ 143 U.S. 457 (1892).

⁴⁸ But see Scalia, *A Matter of Interpretation* at 18-23 (cited in note 16) (referencing *Holy Trinity* as an example of the “danger” of using legislative intent as an interpretive standard); Vermeule, 50 Stan. L. Rev. at 1863-77 (cited in note 29) (arguing that judges are not competent to derive legislative intent from legislative history, as exemplified by *Holy Trinity*). Compare Eskridge, 96 Mich. L. Rev. at 1517-19, 1533-35 (cited in note 23) (arguing

have corrected any problems emerging from a formal approach, and would have done so costlessly, the argument for formalism, in cases of that kind, seems greatly strengthened. (A legislative correction does not of course mean that the court was wrong.)

B. Disputes About What?

To get hold of the resulting debates, and to understand more concretely what formalists and nonformalists are disputing, it is necessary to make some further distinctions. First consider the problem of *statutory ambiguity*, a problem that arises when Congress uses a term that is reasonably understood in several different ways. Every sensible formalist is aware that when a statute is ambiguous, it is necessary to consult something other than the text. Here there is no dispute between formalists and nonformalists. A consistent formalist might appeal to background principles of various kinds, like the rule of lenity (counseling courts to construe criminal statutes narrowly) or the nature of the federal system (often said to require courts not to allow national law to preempt state law). An analogy can be found in Justice Scalia's suggestion that substantive due process should be understood by reference to tradition, understood at the lowest level of generality.⁴⁹ The advantage of an approach of this kind is that it promises to contain judicial discretion on the spot, by requiring courts to use rules where text is unclear. Nonformalists are likely to be skeptical of the overall value of these second-best efforts to constrain decisional discretion.

A second problem involves the problem of *statutory generality*, which arises when Congress uses a broad term that appears to cover the case at hand, but that may produce unreasonableness or absurdity in doing so. Here the formalist wants to understand the term in the general sense suggested by its literal meaning; the nonformalist tends to disagree, finding in excessive generality a genuine interpretive puzzle. Consider Wittgenstein's famous example: "Someone says to me: 'Shew the children a game.' I teach them gaming with dice, and the other says, 'I didn't mean that sort of game.' Must the exclusion of the game with dice have come before his mind when he gave me the order?"⁵⁰ The cases described in the Introduction⁵¹ are possible cases

that formalism is indeterminate in *Holy Trinity* because the statutory terms are ambiguous).

⁴⁹ *Michael H. v Gerald D.*, 491 US 110, 127 n 6 (1989).

⁵⁰ Ludwig Wittgenstein, *Philosophical Investigations* 33 (Blackwell 1963) (G.E.M. Anscombe, trans).

⁵¹ See text accompanying notes 6-11.

of *excessive generality*—problems in which the text appears to run in one direction, but in which the collective speaker is not likely to have had the particular application in mind, and would likely have been aghast at that application.

In their modern incarnation, formalist approaches to interpretation tend to share a number of positive and negative features:

- insistence on using the text as the exclusive or principal source of interpretation;
- a belief that the original or ordinary understanding of the text is what governs (acknowledging that in some cases the text will be ambiguous);
- a refusal to make exceptions to the ordinary meaning of terms, even where the outcomes were probably not intended and indeed seem far afield from the concerns of the enacting legislature;
- where text is unclear, a search for second-best approaches to constraining judicial discretion;
- an insistence that when the text is ambiguous, courts should follow the understanding of the agency entrusted with its administration;
- a refusal to consult legislative history;
- caution about the ideas of legislative “purpose” and legislative “intention”;
- unwillingness to invoke policy arguments of various kinds, involving, for example, the consequences of one or another approach;⁵² and
- attention to canons of construction that help in limiting judicial discretion and in uncovering meaning, and also to canons that reflect a distinct constitutional commitment or otherwise give clear signals to Congress, but not to canons of interpretation that are not time-honored or that embody controversial judgments about public policy.

C. Might Formalism Mimic the Political Market?

One way to approach the resulting debate is to ask whether formalism can be justified in the same terms as any of the various kinds of default rules. Perhaps formalism, as an approach to statutes or contracts, can be understood as a kind of global market-mimicking approach, one that accurately divines Congress’s instructions in particular cases (understood, controversially, as

⁵² But see Schacter, 51 Stan L Rev 1 (cited in note 40) (collecting striking evidence of the use of policy arguments by the Supreme Court).

“what Congress would have done if it had decided the question”), or that does this at least as well as any alternative, while at the same time minimizing decision costs. This would be a strong argument for formalism.

But it is not clear that this is true; whether it is depends on underlying facts. It should be clear by now that on certain assumptions, formalism should ensure low decision costs and relatively low error costs. Perhaps formalist judges can reach decisions quickly, and perhaps they will do as well as anyone else, or better than anyone else, in reaching the right result (if the criteria of rightness are as stated, a point taken up below). But are the assumptions accurate? In general, formalism should indeed reduce the costs of decision, but this is not inevitable (text may cause a lot of confusion too,⁵³ and in some cases study of legislative history might, at least in theory, reduce the costs of decision), but it might also increase the number and seriousness of errors. If nonformalist judges are perfect—if they could, without error or high cost, examine all of the legal materials to uncover what Congress really meant by a statutory term—they would do much better than formalist judges.

Consider, for example, the Delaney Clause problem presented in the Introduction.⁵⁴ If literal interpretation of the Delaney Clause was a blunder—because Congress would not have wanted courts to ban the agency’s action⁵⁵—nonformalist judges will probably do better. Now it is possible that as a guide to actual congressional judgments, formalism is superior to any other approach; if formalism is fairly accurate though crude, and if alternative approaches produce numerous misreadings, then formalism might be best. And even if nonformalist judges can arrive at good answers, they are likely to sacrifice predictability. Perfect nonformalist judges will not be able to let people know the law in advance. Where predictability is especially important—such as in areas involving commercial and criminal law—formalism might be favored.

⁵³ See, for example, Eskridge, 96 Mich L Rev at 1516-22 (cited in note 23) (examining the problems of textual coherence by reference to *Holy Trinity*).

⁵⁴ See text accompanying notes 6-7.

⁵⁵ There is of course a question whether this is the appropriate test. To say that the text binds, but the unenacted intention does not, requires an argument, likely with large empirical dimensions. Article I is insufficient; the question is what the text means, and Article I does not tell us that. To say that the text is binding, but the unenacted intention is not, is probably to say that the consequences will be better if we proceed accordingly—that is, predictability will be increased, judicial discretion will be reduced, mistakes will be more readily corrected ex post or ex ante, and so forth. These points are discussed below. See Section IV.D.

But these issues are hard to resolve in the abstract. An important empirical question involves the relative competence of formalist and nonformalist judges and likely legislative reactions. Excellent nonformalist judges would be best of all. On the other hand, moderately competent formalist judges would be better than incompetent nonformalist judges, and may even be better than moderately competent nonformalist judges. At least we are now in a position to see the empirical questions that might separate those who accept and those who reject formalism as a market-mimicking default rule.

D. Might Formalism Result in the Disclosure of More Information?

It seems most straightforward to defend formalism as a massive or global information-eliciting default rule. Perhaps formalism is likely to produce greater clarity from Congress, precisely because it ensures that statutory language will be understood by reference to its terms. Thus the notion that statutes will be taken in their “plain meaning” might be understood as a way of encouraging Congress to speak unambiguously. Knowing that courts will not correct mistaken applications, Congress (and interested groups hoping to win in court) should take greater care before the fact. Indeed, formalism might even be seen as a close cousin of the nondelegation doctrine, a doctrine grounded explicitly on the notion that courts will require Congress to speak with some degree of clarity.⁵⁶ (The nondelegation doctrine is in turn related to the idea that some statutes and some contracts are void for vagueness; this is also an information-eliciting idea.) If legislators know in advance that courts will not ask “what Congress would have done,” and will instead take legislation at its word, then Congress will be penalized by careless language, with the result that Congress will speak more clearly, and no one will have to guess what it “would have done.”

On this view, formalism is an information-eliciting default rule, in a way that has two salutary consequences. First, it requires the most knowledgeable parties to disclose relevant information to all participants in the political debate, thus making judgments informed rather than ignorant. Second, it discloses information to courts, which might otherwise be forced to guess about what Congress would have done if it had decided the is-

⁵⁶ For the classic, and lonely, applications of the nondelegation doctrine, see *Panama Refining Co v Ryan*, 293 US 388 (1935); *Schechter Poultry Corp v United States*, 295 US 495 (1935).

sue—a guess that might be wrong, and a guess whose very nature requires an increase in the costs of decision.

Or consider the debate over the use of legislative history. One defense of ignoring the history is that it puts pressure on the legislature to focus on the text of the statute, and not to rely, in case of contest, on nuggets placed in a committee report. The result is a kind of information-eliciting rule, one that—if it works well—will put appropriate burdens on Congress to write law, and will create a strong incentive to do exactly that. The *Chevron* principle,⁵⁷ counseling courts to defer to administrative interpretations of ambiguous statutes, can be understood in similar terms. A hoped-for consequence of the *Chevron* approach is to force Congress to speak with clarity; if Congress knows that the executive (often its institutional rival) will be empowered to resolve ambiguities, perhaps there will be more clarity where clarity matters. In this sense, *Chevron* itself can be seen as a kind of nondelegation doctrine. It is therefore possible to imagine an argument that would defend formalism on the ground that it serves the same good functions that information-eliciting default rules serve in the law of contract.

For the reasons suggested above, this would be a largely empirical argument. The question is whether Congress can respond, *ex ante* or *ex post*, to a formalist regime, and whether it can do so at relatively low cost. If a formalist judiciary does not in fact elicit information from Congress—if Congress is relatively unresponsive to the formalist signals—the case for formalism, as a global information-eliciting default rule, is much weakened.

IV. IS FORMALISM GOOD?

A. In General

From the discussion thus far, the question whether it is right to approach statutes in formalist fashion seems to depend on three principal questions:

1. How much, if any, additional *inaccuracy* (however defined) would be introduced by a nonformal approach? If courts would blunder a great deal, or introduce irrelevant considerations of policy or principle, the case for formalism is fortified.

⁵⁷ *Chevron USA, Inc v Natural Resources Defense Council, Inc*, 467 US 837, 866 (1984).

2. What kind of ex ante *incentives* are created by formal and nonformal approaches, and how, if at all, would Congress *respond* to those incentives? Would judicial formalism encourage Congress to legislate more clearly? Would Congress correct the problems of excessive generality, at least after the fact?

3. How much additional *uncertainty* would be introduced by a nonformal approach, taking uncertainty to include the need for litigation and the difficulty of planning? How bad would such uncertainty be? What would be its cost? Here the inquiry is into the size of decision costs associated with formal and nonformal approaches.

From these questions, we have a research agenda that could keep a lot of people busy. It is possible, for example, to imagine state-by-state comparisons (what has happened with more, or less, formalist state courts?), area-by-area comparisons (do some areas of law, for example tax, see a high or low degree of formalism, and with what consequences?), and general national studies across time (how often does Congress overrule formalist, or non-formalist decisions? Is Congress responsive to formalist signals?).

B. Constitutional Formalism: A Note

We can now obtain a better understanding of debates over formalism in constitutional law. Here, too, decision costs and error costs are highly relevant. Justice Scalia, for example, thinks that use of the original understanding of the text will simplify constitutional inquiry (a point about reducing decision costs); he also believes that the fabric of rights enjoyed by Americans is more, not less, likely to be at risk from nonformalist approaches (a point about error costs).⁵⁸ In his view, serious “gaps” in rights protection will be filled by democratic arenas. A very different constitutional formalist, Akhil Amar, believes that the constitutional text, rightly understood, will protect an ample category of rights, and hence that those who believe in these rights have nothing to fear from his species of formalism (a point about error costs).⁵⁹ Antiformalists suggest, by contrast, that the constitutional text leaves many indeterminacies (a point about decision costs), and also that constitutional formalism would result in an

⁵⁸ Scalia, *A Matter of Interpretation* at 37-47 (cited in note 16).

⁵⁹ See generally Akhil Reed Amar, *The Bill of Rights: Creation and Reconstruction* (Yale 1998).

unacceptably narrow category of constitutional rights (a point about error costs).⁶⁰

Now it is unlikely that empirical inquiry will resolve all of the relevant debates in the constitutional domain, partly because there is so much disagreement in that domain about what counts as an error, and partly because any judgment to that effect may well be theory-dependent. (Does it count against an interpretive method that it leads to a rejection of *Roe v Wade*?⁶¹ *Griswold v Connecticut*?⁶² *Craig v Boren*?⁶³) But there are important empirical issues here too. For example, the argument for formalism in constitutional law would be strengthened if it could be shown that if courts using that method would not protect certain rights, the democratic process would do so instead. The argument against formalism would be strengthened if the opposite could be shown. Many debates over *Roe v Wade*, in fact, turn on that question.⁶⁴

C. England, America, and the Regulatory State

From the discussion thus far, I speculate that other things being equal, formalism is likely to hold sway in democracies that have a competent, highly responsive legislature, and that less formal approaches will be found in democracies in which the legislature is incompetent, inattentive, or both.⁶⁵ The speculation is no more than that, but it is supported by some evidence from the corporate field, where formalism is more likely when legislatures are attentive, and also by comparing statutory interpretation in England with that in the United States.⁶⁶

⁶⁰ See generally Ronald Dworkin, *Bork's Jurisprudence*, 57 U Chi L Rev 657 (1990).

⁶¹ 410 US 113 (1973).

⁶² 381 US 479 (1965).

⁶³ 429 US 190 (1976).

⁶⁴ See, for example, Gerald Rosenberg, *The Hollow Hope: Can Courts Bring About Social Change?* 173-265 (Chicago 1991) (arguing in part that political and legislative movements prior to *Roe* may have been at least as significant a factor in realizing the right to an abortion as *Roe* itself).

⁶⁵ These comments are restricted to democracies. In the case of Nazi Germany, the nonformalism came from a judiciary that did not insist that anything be codified in actual law, and that simply implemented the Nazi program, whether or not it had legislative support. The Allies' insistence on formalism should be taken as an effort to inculcate rule-of-law values. The nonformalist approaches I am discussing here are entirely consistent with the ideal of the rule of law; they operate only in cases of ambiguity or excessive generality.

⁶⁶ I draw in this Section on the extremely illuminating discussion in P.S. Atiyah and Robert S. Summers, *Form and Substance in Anglo-American Law: A Comparative Study of Legal Reasoning, Legal Theory and Legal Institutions* (Clarendon 1987).

English law is far more rule-bound and formalistic than American law. The British Parliament is less likely to delegate discretionary authority to judges. For their part, English judges tend to treat statutes as rules, generally refusing to investigate whether the particular application of the rule makes sense as a matter of policy or principle. In England, lawmaking and law interpretation are far more rigid than in the United States, where judges tend to avoid formalism, and to allow themselves to depart from general language in the process of confronting particular cases.

Institutional differences between England and America help explain the different methods. Laws in England are drafted by an Office of Parliamentary Counsel, a highly professional body that consists of skilled authors of laws. The Parliamentary Counsel is concerned to ensure a uniform style of drafting. The Counsel is also closely attuned to the methods of English judges. The judges' practice is itself uniform and relatively simple. In a parliamentary system, the government and the legislature are allied, and the high degree of party control means that there is a high level of homogeneity in England at the lawmaking stage. Moreover, and critically, Parliament revisits statutes with some frequency, and it fixes mistakes that are shown as such when particular cases arise.⁶⁷

The situation in the United States is very different. There is less centralized drafting and hence less uniformity in terminology. Although in America the Office of Legislative Counsel does help with much drafting, there is less professionalization in the production of statutes. In America, the drafters of legislation are multiple and irregularly coordinated. Except in rare cases, the party system no longer provides a great deal of coherence, and the executive and legislature are hardly aligned. Congress appears only intermittently aware of the judges' interpretive practices, which are themselves not easy to describe in light of the sheer size of the federal judiciary and the existence of sharp splits, on just this point, within the Supreme Court. It would be wrong to say that Congress is entirely oblivious to judicial decisions interpreting statutes.⁶⁸ But Congress is not in the business of responding rapidly and regularly to particular cases in which interpretations, literal or otherwise, tend to misfire. Hence both

⁶⁷ There have been interesting recent changes in English practice, with legislative history now playing a role. See Eskridge, 96 Mich L Rev at 1542 (cited in note 23).

⁶⁸ See Eskridge, *Dynamic Statutory Interpretation* at 151-53, 259, 262, 295-96, 316-22 (cited in note 14) (citing instances where Congress overrode Supreme Court interpretations of statutes).

lawmaking and law-interpreting practice are very different from what they are in England. None of this suggests that England or America has the optimal system of interpretation in light of its own institutional characteristics. But it does suggest that the two legal systems are highly responsive to their distinctive institutional features.

Compare in this regard the interpretation of law by administrative agencies, which are the first-line interpreters of a wide range of regulatory statutes. Of course agencies must decide whether or not to be formalists. Should courts allow them to be?⁶⁹ An implication of the analysis thus far is that the answer is "Yes." For the most part, courts should allow agencies to choose whether to be formalists, as a way of reducing overall costs of decision and overall costs of error.

The place to start is with *Chevron USA, Inc v Natural Resources Defense Council*,⁷⁰ which has emerged as the most important case about legal interpretation in the last thirty years. *Chevron* holds that where statutes are ambiguous, courts should accept any reasonable interpretation by the agency charged with their implementation.⁷¹ *Chevron* appears to rest on the suggestion, central to legal realism, that the decision how to read ambiguities in law involves no brooding omnipresence in the sky but is an emphatically human judgment about policy or principle. *Chevron* concludes that, where underlying statutes are ambiguous, Congress should be taken to have decided that agencies are in a better position to make that judgment than courts. Agencies are in that better position because, *Chevron* emphasizes, the President is generally in charge of their policy judgments, and hence agencies have a kind of democratic pedigree, certainly a better one than the courts.

Seen through the lens of *Chevron*, the debate over formalism might be understood very differently in the context of the twenty-first century, whose public law would pose as a central question: What are the views of any agency charged with implementation of this law? As against modern formalists, we might urge that administrative agencies should be authorized to reject the "text" in situations where common law judges would not be so authorized, at least when there is no evidence of a considered legislative judgment against the agency's interpretation. An especially im-

⁶⁹ I borrow here and adapt an earlier discussion. See Cass R. Sunstein, *Justice Scalia's Democratic Formalism*, 107 Yale L J 529, 550-55 (1997), reviewing Scalia, *A Matter of Interpretation* (cited in note 16).

⁷⁰ 467 US 837 (1984).

⁷¹ Id at 842-45.

portant point here involves the agency's hierarchical and centralized character. The hierarchical structure eliminates much of the unpredictability and variability that could come from nontextual courts.

On this view, the choice between formalist and nonformalist interpretation should generally be made by the relevant agency, which would be permitted to balance the relevant, largely empirical issues as it sees fit. The agency is in a good position to know whether the additional uncertainty introduced by purposive or dynamic interpretation is worth the candle. Thus courts should defer to nonformalist administrative agencies. It follows that the easiest case of those referred to above⁷² involves the Delaney Clause. Courts should have allowed the FDA to make exemptions for de minimis risks (as they did not⁷³); a nonformalist FDA should have been permitted to ensure that the statute did not generate an unreasonable result, one that could not be characterized as an actual judgment of the enacting Congress.

Why might anyone disagree with these judgments? The best answer lies in the empirical judgments that underlie any embrace of formalism. It would be possible to believe that armed with the right incentives, Congress would correct any problems produced by literalism, and also that agencies are subject to parochial pressures, so as to weaken the suggestion that they have a comparative advantage by virtue of their democratic pedigree and technical ability. Perhaps agencies are worse than courts, simply because they are vulnerable to factional influences. To the extent that this is true, the argument for judicial use of the text against nonformalist agencies is strengthened. And to the extent that Congress will provide correctives to the problems created by formalism, there is much less need to allow agencies to proceed nonformally. I believe that in general nonformalist agencies are unlikely to abuse their discretion and that legislative responses to the problems introduced by formalism would be too little, too late.⁷⁴ But these issues require further empirical study.

⁷² See text accompanying notes 6-11.

⁷³ See, for example, *Public Citizen v Young*, 831 F2d 1108, 1122 (DC Cir 1987); *Les v Reilly*, 968 F2d 985, 989 (9th Cir 1992).

⁷⁴ The Delaney Clause, not amended until 1997, is itself an example. See The Food Quality Protection Act of 1996, Pub L No 104-170, codified at 21 USCA §§ 346 et seq (1997).

D. Defending Formalism Unempirically? Of Concepts and Legitimacy

It would be possible for some formalists (and some antiformalists) to respond to the arguments I have made thus far in the following way: The analysis has not addressed the most fundamental concerns. No good formalist should be persuaded by it. The real case for formalism depends on constitutional considerations, or legitimacy, or democracy. Contingent institutional judgments, with an empirical cast, are entirely beside the point. In a metaphor: Bentham could not persuade Kant of the correctness of some course of action by showing that utilitarian arguments favor it. That is not what Kant is concerned about.⁷⁵ So too for the suggestion that empirical arguments can settle the choice among interpretive methods.

This problem raises the question whether it is possible to defend formalism without reference to empirical issues—whether there is an analogy to a deontological position on formalism's behalf. I do not believe that there is any such analogy. There are important questions about legitimacy and democracy, but they cannot be brought to bear on the choice among interpretive methods without resort to empirical claims. But these questions require discussion.

1. The Constitution.

Formalism might be defended by reference to the Constitution. On this view, the Constitution sets up certain procedures for the making of law, and certain things—canons, policy concerns, “equity,” legislative intentions—are not the law. Under Article I, the text is the law.

The problem with this suggestion is not that it is wrong but that it is irrelevant. To be sure, the text is the law, and we can agree that policy judgments, and intentions standing alone, do not represent the law. Something does not become the law because Congress intended it to be the law. But the question here is how to decide what the text means. The fact that the text is the law does not establish whether its meaning is to be assessed by looking at subjective intentions (at what most members of Congress intended it to mean), by examining canons of construction (as in the idea that text should not lightly be understood to raise constitutional doubts), or by understanding text in light of traditional equitable concerns. If meaning is to be established by

⁷⁵ I am grateful to Frank Easterbrook for pressing this question at the Symposium.

looking at the objective meaning of the words, rather than anything within the minds of the legislators, this is not because the Constitution says so—it does not—but because that approach leads to the most sensible system of interpretation, for reasons akin to those I have discussed.⁷⁶

2. Law's ontology.

A second nonempirical defense of formalism might rely on a judgment about the nature of law, or in other words, on a claim about law's ontology. Do we not—it might be asked—have to decide what counts as “law” first, before we start to ask empirical questions? After all, a presidential seizure of the steel mills does not count as lawful, or as law, even if it would be good, all things considered, to allow the President to seize the steel mills. Perhaps a formalist approach stems from a (nonempirical, pre-empirical) conception of what allows something to count as law, and perhaps that conception precedes anything of the kind I have suggested here.

For present purposes the simplest answer to this objection is that the debate between formalists and antiformalists cannot be resolved by asking about the nature of law, because any answer is unhelpful on this particular debate. Consider the three cases with which this Article begins.⁷⁷ In each of them, people could go either way without making a contestable claim about the nature of law. Those who think that courts should not understand literal language to produce highly unreasonable outcomes need not venture a theory of what makes something “law.” Nor need their adversaries. Those who think that legislative history is relevant do not need to say anything controversial about the nature of law. Those who emphasize “purpose,” or traditional canons of construction, offer no disputed account of “law.” Both formalists and antiformalists agree that courts should not understand a statute to mean whatever they think that a good statute would say, and this agreement may indeed have something to do with a theory of

⁷⁶ It would be possible to argue that the constitutional text, read in light of its history, argues in favor of one or another approach to statutory interpretation. For the formalist, this view raises two problems. First, the history seems to argue against formalism; the framing generation did not embrace formalism. See Eskridge, 96 Mich L Rev at 1523-26, 1529-31 (cited in note 23). Second, resort to the history, as an authoritative guide to the text, requires an argument of some kind, and that argument must be, roughly, the sort of empirical argument I have discussed here. All I mean to suggest is that the Constitution by itself does not support, or undermine, formalism; the real issues lie elsewhere.

⁷⁷ See text accompanying notes 6-11.

law's ontology.⁷⁸ But the disputes I am discussing have a different character.

3. Legitimacy and democracy.

Some people defend formalism on emphatically nonempirical grounds, claiming that formalist approaches stem from the right theory of political legitimacy. Here, I think, are the most serious questions about the argument I have been making.

Frank Easterbrook, for example, claims that for "the textualist a theory of political legitimacy comes first, followed by a theory of interpretation that is appropriate to the theory of obligation."⁷⁹ He offers what he takes to be "two major strands" of the relevant account. According to the first, the "fundamental theory of political legitimacy in the United States is contractarian," and without some kind of formalist approach, "a pack of lawyers is changing the terms of the deal, reneging on behalf of a society that did not appoint them for that purpose."⁸⁰ The second kind of account is based on an answer to the question why judges should be obeyed. On Easterbrook's view, judges deserve obedience only because they are adhering to the Constitution. "When form comes from evolving institutions and ideas, then the living must decide by elections Nothing beats textualism *in court*, because nothing else is capable of supporting a judicial veto."⁸¹

But these arguments fail to show that the idea of legitimacy supports textualism. Those who seek to invoke intentions, or equitable considerations, or canons of construction, or evolving, widely held social understandings can also claim the mantle of political legitimacy;⁸² we cannot know if they are wrong without

⁷⁸ Actually I do not believe that this is the case. I think that the empirical and institutional issues I have discussed are crucial here too—that these issues precede ontology all the way down. This is a pragmatic claim, suggesting that what makes something qualify as law turns on the consequences, or the cash-value, of allowing it to qualify as law. Thus, for example, the Constitution is binding for a set of (loosely speaking) pragmatic reasons, not because of anything about the nature of law. The Constitution is binding because it is a good constitution and because we are much better off if we take it to be binding. But for present purposes it is not necessary to accept these more controversial claims.

⁷⁹ Frank H. Easterbrook, *Textualism and the Dead Hand*, 66 Geo Wash L Rev 1119, 1119 (1998). Easterbrook is speaking of textual interpretation of the Constitution, but I believe that what he says is the most lucid account of how textualism might be supported nonempirically for statutes as well.

⁸⁰ *Id.* at 1121.

⁸¹ *Id.* at 1125.

⁸² The most famous effort to justify a form of constitutional review that is not textual in Easterbrook's sense is self-consciously based on the ideal of democracy. See John Hart Ely, *Democracy and Distrust: A Theory of Judicial Review* (Harvard 1980). In statutory construction, it is possible to imagine analogies. See, for example, Cass R. Sunstein, *After the Rights Revolution: Reconceiving the Regulatory State* (Harvard 1990) (offering argu-

knowing a great deal more about institutional performance. Nor does judicial review become “legitimate” by virtue of the fact that it is based on the text of a document two centuries old. The “legitimacy” of judicial review, like the legitimacy of the Constitution itself, is a complicated question, one for which the bare fact of the “text” is neither a necessary nor a sufficient condition. Nor is it clear how the notion of a social contract can show that formalism is a uniquely legitimate approach to interpretation. The signatories to the relevant text are all long dead, and to say that nonformalist judges are reneging on some “deal” seems to me no more than a metaphor. In what sense did any of us agree to that particular “deal”? If we are to be held to “it”—and the real issue is how we establish its terms—this is because it is good for us to be held to it, not because we agreed.

What underlies Easterbrook’s argument, I think, is not a notion of contract, but one of democratic self-government. He thinks that the Constitution can claim support in popular sovereignty, and so too for what emerges from “elections,” whereas judicial “reneging” lacks anything like popular support. And to the extent that Easterbrook is connecting the idea of legitimacy with democracy, and seeing these as linked to debates over interpretation, he is certainly on firm ground. Some methods of interpretation would indeed be illegitimate (though to see why, some empirical questions would probably have to be answered as well).

The problem here is that the ideal of democratic self-government does not by itself justify formalism. Many antiformalists are enthusiastic democrats too. They might invoke legislative history on the ground that judges should consult the will of the representatives of the people, rather than dictionaries or their own judgments, to determine what vague words mean. They might think that canons of construction should defeat literal language, on the ground that those canons have support within the norms and traditions of the public as revealed over time, or otherwise have a good democratic pedigree—as loose, general language from an occasionally inattentive Congress in, say, 1992, does not. This does not mean that the antiformalists are right. It means only that the ideal of democracy, or political legitimacy, cannot by itself support formalism. That ideal must be supplemented by a set of institutional and empirical claims. With certain findings about institutional capacities and performance,

ments in this vein); Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (MIT 1996) (William Rehg, trans).

formalism would be more legitimate than nonformalism; with contrary findings, the opposite would be true.

In his remarks at the Symposium, Frank Michelman went further. He criticized formalism itself on grounds of legitimacy. He claimed that formalism is illegitimate. In his view, formalism could be legitimate only if certain democratic preconditions were met—only if law had, in its genesis and susceptibility to revision, a kind of strong democratic pedigree because (for example) it was made by free and equal citizens. Because law does not, in our society, have this pedigree, formalism is illegitimate. I agree with much of Michelman's argument, and I think I accept his conception of political legitimacy, but I also think that he neglects some important institutional questions. To be sure, formalism has, under his assumptions, a serious legitimacy deficit. But from the standpoint of legitimacy, things may be worse, not better, if courts go in nonformalist directions. Judges may, for example, push statutes in directions that increase the legitimacy problem (if, for example, they conclude that statutes should generally be construed narrowly and in conformity with the common law). Michelman's effort to show that formalism is illegitimate disregards comparative institutional questions that, I have suggested, are crucial to the choice.

For example, suppose that if people do not like the outcomes they get from a formalist system of interpretation, they will elect legislators who will give them better outcomes. If this is true—and it is an empirical proposition—then there is an argument for formalism: Perhaps formalism will create a set of dynamic incentives for genuine self-government, and, in the process, for better law. Perhaps it will lead citizens to elect representatives who do what they want them to do. If this is an empirical possibility, there is a good argument for formalism. In sum, defenses of formalism that speak to the nature of law, or to questions of legitimacy, are tacitly based on some empirical projections, and they stand or fall on those projections.

E. Defining Errors

A possible response would underline a serious problem mentioned but not resolved to this point: How do we know if there has been an interpretive error at all? It seems reasonable to say that any identification of an "error" depends on an interpretive account of some kind, and thus that any assessment of whether there are errors, and error costs, is fatally circular. Perhaps errors cannot be identified before the theory of interpretation is already in place. If this is so, then we cannot choose among alterna-

tive approaches by examining errors and their “costs”; whether there is an error depends on what theory has been chosen. Textualism offers a theory by which to tell whether there have been errors; so too do certain nonformalist accounts. We cannot decide which account is right by asking which account produces more errors. By suggesting that we assess errors by asking “what Congress would have done,” I have loaded the dice from the start.

There is an important point and much truth here. The simplest reply is that even if an inquiry into errors is theory-dependent, people with different theories might be brought to agree on the right interpretive method, if the empirical findings turn out one way rather than another. If legislatures are highly responsive and if antiformalist courts produce unpredictability and arbitrariness, the case for formalism is greatly strengthened. Suppose, for example, that in the area of taxation, the legislature will anticipate possible problems produced by general language, and that it will promptly correct problems that emerge after the fact; suppose too that in that area, nonformalist courts will make the rules of taxation hard to predict and also arbitrary on any view. By contrast, the case for formalism would be undermined, in tax law, if legislation were quite crude and if the legislature was unlikely to correct unanticipated unreasonable outcomes, and if courts were excellent at adapting the text to unforeseen problems. If the use of legislative history greatly increases the cost of decision without helping to make better decisions, the argument against using legislative history is quite strong. In this way people who have abstract theories about interpretation, or political legitimacy, should be flexible enough to move in one or another direction with imaginable empirical findings.

There is a deeper point here. A textualist cannot simply stipulate that judges make mistakes if they reject textualism; textualists have to come up with some criteria to give some grounds for their commitment to textualism. The fact that Article I sets up certain mechanisms for the production of law does not offer any such criteria; the question is what any “law” means. As we have seen, the notion of market-mimicking and information-eliciting default rules provides a start. Perhaps it can be added that where the outcome of a particular interpretive approach would be absurd or highly unreasonable by Congress’s own lights, the interpretation produces errors and error costs. Perhaps it can be agreed—indeed, it generally is agreed—that excessive generality can produce absurdity, and the question is whether the costs (in terms of decisions and mistakes) are higher if courts reject the absurd application. Approaches of this kind would permit us to

make a great deal of progress in choosing between formalist and nonformalist approaches. If formalists (or nonformalists) do not wish to claim support on this kind of ground, on what ground do they wish to defend their position? And recall that we are dealing here with a continuum of approaches, not with any rigid dichotomy.

Some space for disagreement would of course remain, and it is here that we might find disputes that are not empirical in nature. Some people think that courts err if they attempt to decide statutory cases the way that Congress would have wanted them to do, even if the statute is ambiguous or excessively general. Moreover, different kinds of debates can be imagined among those who disagree about the role of appropriate background norms or canons of statutory construction. One person strongly favors the use of clear statement principles; another is sharply critical of them. And people can disagree about the content of such principles. One person claims that statutes in derogation of the common law should be narrowly construed; another urges that statutes should be interpreted favorably to traditionally disadvantaged groups, or generously to promote environmental protection. The dispute here depends largely on judgments about the nature of a just or good legal order. But even here there are empirical questions. If courts indulge one or another interpretive principle, what will legislatures do? Will they respond at all? Respond in a way that defeats the purposes of the relevant canon of construction? Or internalize the judgments that underlie the canon? It would be possible to know a lot more about these questions than we now do.

In sum: My narrowest claim here is that if we agree that a basic task of statutory interpretation is, in cases of doubt, to discern what Congress would have done if it had made provision on the point ("imaginative reconstruction"), the debate between formalists and nonformalists is highly empirical, and we can understand the grounds on which reasonable people might disagree. A somewhat broader claim is that empirical debates also lie at the heart of competing views about the role of legislative history, "purpose," and the canons of construction; at least such debates would be much informed by judgments about the likely performance of courts and legislatures. The broadest claim is that there is no plausible basis for formalism, and textualism, that does not depend on pragmatic and empirical considerations; that, for example, imaginative reconstruction cannot be rejected as a criterion (in favor of objective "meaning") without resort to those con-

siderations. The question that remains to be answered is this: If this claim is wrong, on what basis might formalism be defended?

CONCLUSION

I have argued in this Article that any defense of formalism must have a large empirical dimension. Formalism, as an approach to statutory interpretation, is best defended on the view that

- under a formalist regime, legislatures will be responsive to the incentives, created by formalism, to do pretty well before the fact, and to correct mistakes after the fact;
- legislatures will respond to a nonformalist regime by legislating sloppily and by delegating power to the judiciary;
- a nonformalist judiciary will create blunders and injustices; and
- a nonformalist judiciary will create considerable uncertainty, with high costs within the legal system (as cases increase in complexity) and in ordinary practice (as people do not know how to conduct their affairs).

These are empirical claims that formalist interpretation will decrease error costs and decision costs. The nonformalist denies these claims, suggesting that in the relevant jurisdiction, courts are not likely to blunder, or to create significant increases in unpredictability, and that legislatures are not likely to be responsive to formalist methods. Here there are many questions for empirical testing. The answer may well vary from nation to nation and across diverse subject matters.

For modern public law and for formalism, a significant complication is introduced by the rise of administrative agencies, which do much of the day-to-day work of statutory interpretation. Agencies must decide in the first instance whether to be formalists. Formalism may be good for the judiciary but bad for the administrative state, and the judiciary would do well to recognize that possibility. Thus I have argued that the greater accountability and specialization of agencies should permit them to choose between formalist and nonformalist statutory construction, so long as both are reasonable.

The broadest lesson has to do with the relevance of empirical claims to many topics in legal theory, and the great difficulty of doing the latter without attending to the former. Of course it is not possible to evaluate facts without normative criteria; but in many contexts, there is sufficient agreement on the criteria to proceed, if the facts are clearly one way or the other. At the very least, an understanding of facts will clarify where we disagree

and exactly why, and in the process it may well lower the volume of legal and even political disputes, in a way that should produce more light and less heat. Antiformalists should agree that they would change their view if it were clear that nonformalist courts would blunder, or that legislatures, when confronted with a formalist judiciary, would make ex post corrections or respond to formalism by legislating with greater clarity before the fact. For their part, formalists ought to be shaken if nonformalist judges would proceed wisely and without imposing high costs of decision, or if legislatures would be entirely unresponsive to formalist interpretation and if palpable errors would go uncorrected. Mutual acknowledgement of these points should make one of the great, quasi-theological debates in legal theory at once more tractable, less theological, and less puzzling. If formalism must be defended empirically, future progress will depend on isolating the factual disagreements and attempting to make an investigation of what, in fact, is true.