

Shedding Light on *Chevron*: An Empirical Study of the *Chevron* Doctrine in the U.S. Courts of Appeals

Orin S. Kerr[†]

The prevalence of the Chevron doctrine in administrative law has prompted widespread scholarly controversy concerning how the doctrine operates in practice. Three descriptive models have emerged. Some commentators have embraced a contextual model and contend that Chevron is a "revolution on paper" that has failed to replace the traditional contextual approach to judicial review of agency action. Others rely on a political model and maintain that the Chevron doctrine is so indeterminate that it serves primarily as cover for judges who decide cases based on their personal political preferences. Other commentators rely on an interpretive model and insist that Chevron is unstable because textualist judges apply the doctrine differently from judges who reject a textualist approach.

This Article presents the results of an empirical study designed to evaluate how accurately the three models describe the Chevron doctrine. By examining the published decisions of the U.S. Courts of Appeals in 1995 and 1996, the author concludes that the three descriptive models largely fail to predict outcomes in actual controversies. The political model mounts the most effective attack, as the study reveals a significant difference in voting patterns between Republican and Democratic judges in ideologically charged cases. However, the results suggest that the contextual model's ability to predict Chevron results is weak at best, and that the interpretive model's predictions are unsupported. The author concludes by offering possible explanations for the divergence between the models' predictions and the reality suggested by the study's results.

Introduction	2
I. Three Models of the <i>Chevron</i> Doctrine.....	5
A. <i>The Contextual Model</i>	6
B. <i>The Political Model</i>	10
C. <i>The Interpretive Model</i>	13
II. Methodology	17
A. <i>Defining a Chevron Application</i>	18
B. <i>Designing the Tests</i>	22

[†] Law clerk to Judge Leonard I. Garth, United States Court of Appeals for the Third Circuit. B.S.E., Princeton University, 1993; M.S., Stanford University, 1994; J.D., Harvard Law School, 1997. I am indebted to Professor Todd Rakoff for providing inspiration, encouragement, and support.

1. <i>The Contextual Model</i>	22
2. <i>The Political Model</i>	24
3. <i>The Interpretive Model</i>	28
III. Study Results	29
A. <i>Overview of the Chevron World</i>	30
B. <i>Results of Testing the Contextual Model</i>	31
C. <i>Results of Testing the Political Model</i>	35
1. <i>Structural Claims</i>	35
2. <i>Individual Claims</i>	37
D. <i>Results of Testing the Interpretive Model</i>	40
IV. Analysis: Do the Three Models Describe <i>Chevron</i> Accurately?.....	43
A. <i>Analysis of the Contextual Model</i>	44
B. <i>Analysis of the Political Model</i>	48
1. <i>Structural Claims</i>	48
2. <i>Individual Claims</i>	49
C. <i>Analysis of the Interpretive Model</i>	53
Conclusion	59

Introduction

In *Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc.*,¹ the United States Supreme Court announced a two-step test for judicial review of agency interpretations of statutes. First, the court must ask if the statute is ambiguous: if the statute is clear, the court should give effect to the expressed intent of Congress. Second, if the statute is unclear, the court must accept the agency's interpretation as long as it is reasonable.² This test, often called the "Chevron two-step," has been an enormously popular subject of realist

1. 467 U.S. 837 (1984).

2. *Chevron* itself expressed the test as follows:

If the intent of Congress is clear, that is the end of the matter; for the court, as well as the agency, must give effect to the unambiguously expressed intent of Congress. If, however, the court determines Congress has not directly addressed the precise question at issue, the court does not simply impose its own construction on the statute, as would be necessary in the absence of administrative interpretation. Rather, if the statute is silent or ambiguous with respect to the specific issue, the question for the court is whether the agency's answer is based on a permissible construction of the statute.

Id. at 842-43.

Thus, if the statute is unambiguous, the *Chevron* analysis is said to be resolved at step one: the court must follow the expressed intent of Congress. If the statute is ambiguous, then the analysis is resolved at step two: the court accepts the agency interpretation as long as it is permissible (i.e., reasonable). See 1 KENNETH CULP DAVIS & RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE §§ 3.2-3.4 (3d ed. 1994).

critique.³ Skeptical that judicial review of agency legal interpretation could operate as simply as the two-step test suggests, commentators have authored a substantial body of work explaining what *really* happens when judges apply the *Chevron* doctrine.⁴

The scholarly critique of *Chevron* offers three distinct models of how courts review agency interpretations of statutory law. The first is a contextual model, which acknowledges that the two-step test exists on paper, but posits that in practice judges continue to adhere to the multi-factored, contextual approach to judicial review that judges are understood to have followed before the *Chevron* case was decided. According to the proponents of the contextual model, the pre-*Chevron* “traditional factors” continue to influence judicial decisions to accept or reject agency interpretations despite the contrary logic of the two-step test.⁵ The second popular critique of the doctrine is a political model, which views *Chevron* through the result-oriented lens of politics. Followers of the political model consider the two-step test to be a justificatory ritual that legitimates results reached to further judges’ political agendas. Accordingly, outcomes are best explained in terms of the political ideologies of individual judges.⁶ The third critique of *Chevron* is an interpretive model.

3. See generally PETER L. STRAUSS ET AL., GELLHORN AND BYSE’S ADMINISTRATIVE LAW 620-36 (9th ed. 1995) (discussing the importance of *Chevron* in the courts, as well as the scholarly reaction to it). The critique of *Chevron* is realist in the sense that it embodies the rule-skepticism often associated with legal realism. See generally JEROME FRANK, LAW AND THE MODERN MIND at vii (Coward-McCann 1949) (1930) (“As these skeptics see it, the trouble is that the formal legal rules enunciated in courts’ opinions—sometimes called the ‘paper rules’—too often prove unreliable as guides to the prediction of decisions. They believe that they can discover, behind the ‘paper rules,’ some ‘real rules’ descriptive of uniformities or regularities in actual judicial behavior, and that those ‘real rules’ will serve as more reliable prediction-instruments”); H.L.A. HART, THE CONCEPT OF LAW 140-41 (2d ed. 1994) (describing a brand of rule-skepticism in which doctrine is viewed as a form of “window dressing” used to explain a judicial result reached intuitively).

4. See *infra* notes 5-7.

5. See Stephen Breyer, *Judicial Review of Questions of Law and Policy*, 38 ADMIN. L. REV. 363, 372-77 (1986); Clark Byse, *Judicial Review of Administrative Interpretations of Statutes: An Analysis of Chevron’s Step Two*, 2 ADMIN. L.J. 255, 266 (1988); Maureen B. Callahan, *Must Federal Courts Defer to Agency Interpretations of Statutes?: A New Doctrinal Basis for Chevron U.S.A. v. Natural Resources Defense Council*, 1991 WIS. L. REV. 1275, 1295-98; Thomas W. Merrill, *Judicial Deference to Executive Precedent*, 101 YALE L.J. 969, 980-90 (1992); Abner J. Mikva, *How Should the Courts Treat Administrative Agencies?*, 36 AM. U. L. REV. 1, 8 (1986); Russell L. Weaver, *Some Realism About Chevron*, 58 MO. L. REV. 129, 129-32, 138-66 (1993); Keith Werhan, *The Neoclassical Revival in Administrative Law*, 44 ADMIN. L. REV. 567, 580 (1992). See discussion of traditional factors *infra* Section I.A.

6. See Linda R. Cohen & Matthew L. Spitzer, *Solving the Chevron Puzzle*, 57 LAW & CONTEMP. PROBS. 65, 68-83 (1994); Christopher F. Edley, Jr., *The Governance Crisis, Legal Theory, and Political Ideology*, 1991 DUKE L.J. 561, 587-88; William S. Jordan, III, *Deference Revisited: Politics as a Determinant of Deference Doctrine and the End of the Apparent Chevron Consensus*, 68 NEB. L. REV. 454, 486-90 (1989); Richard J. Pierce, Jr., *Legislative Reform of Judicial Review of Agency Actions*, 1995 DUKE L.J. 1110, 1123 [hereinafter Pierce, *Legislative Reform*]; Richard J. Pierce, Jr., *The Supreme Court’s New Hypertextualism: An Invitation to Cacophony and Incoherence in the Administrative State*, 95 COLUM. L. REV. 749, 779-80 (1995) [hereinafter Pierce, *Hypertextualism*]; Sidney A. Shapiro & Richard E. Levy, *Judicial Incentives and Indeterminacy in Substantive Review of Administrative Decisions*, 44 DUKE L.J. 1051, 1068-72 (1995); Katherine L. Vaughns, *A Tale of Two Opinions: The*

The interpretive model predicts that deference in individual cases depends upon a judge's approach to statutory interpretation—in particular, the likelihood that text will be considered ambiguous at *Chevron's* step one.⁷ In contrast to the contextual and political models, the interpretive model accepts that judges will attempt to apply the two-step test objectively. Nonetheless, the doctrine is considered inherently unstable because individual judges will defer more or less often depending upon how readily they perceive ambiguity in statutory text.

Despite the debate among proponents of the three models, the evidence to support their often-conflicting claims typically is anecdotal at best. As a result, the scholarship on *Chevron* has reached an impasse, in which three distinct paradigms of the doctrine compete for attention without the substantial empirical evidence needed to advance or impede their claims.⁸

This Article presents the results of an empirical study designed to evaluate the claims of the three competing models of *Chevron*. By examining every application of the *Chevron* doctrine published by the United States Courts of Appeals during 1995 and 1996, the Article hopes to explain how various factors affect the likelihood that a judge will vote to uphold or reject executive interpretations under the doctrine. In other words, what really happens when a judge applies the *Chevron* doctrine? What factors actually influence a judge's decision to accept or reject an agency's interpretation of statutory law?

In particular, the study evaluates the empirical claims of each of the three models of *Chevron* by examining more than 250 applications of the doctrine, culled from over 200 published cases. The contextual model is tested by comparing judicial acceptance rates of agency interpretations when the

Meaning of Statutes and the Nature of Judicial Decision-Making in the Administrative Context, 1995 B.Y.U. L. REV. 139, 149-50.

7. See Bernard W. Bell, *Using Statutory Interpretation to Improve the Legislative Process: Can It Be Done in the Post-Chevron Era?*, 13 J.L. & POL. 105, 132-33 (1997); Gregory G. Garre, *CERCLA, Natural Resource Damage Assessments, and the D.C. Circuit's Review of Agency Statutory Interpretations Under Chevron*, 58 GEO. WASH. L. REV. 932, 951-53 (1990); Michael Herz, *Textualism and Taboo: Interpretation and Deference for Justice Scalia*, 12 CARDOZO L. REV. 1663, 1668-72 (1991); Bradley C. Karkkainen, "Plain Meaning": Justice Scalia's Jurisprudence of Strict Statutory Construction, 17 HARV. J.L. & PUB. POL'Y 401, 459-64 (1994); Bradford C. Mank, *Is a Textualist Approach to Statutory Interpretation Pro-Environmentalist?: Why Pragmatic Agency Decisionmaking Is Better than Judicial Literalism*, 53 WASH. & LEE L. REV. 1231, 1248-50 (1996); Jerry L. Mashaw, *Textualism, Constitutionalism, and the Interpretation of Federal Statutes*, 32 WM. & MARY L. REV. 827, 832-35 (1991); Thomas W. Merrill, *Textualism and the Future of the Chevron Doctrine*, 72 WASH. U. L.Q. 351, 362 (1994); Pierce, *Hypertextualism*, *supra* note 6, at 752; Antonin Scalia, *Judicial Deference to Administrative Interpretations of Law*, 1989 DUKE L.J. 511, 520-21; Bernard Schwartz, "Shooting the Piano Player"?: Justice Scalia and Administrative Law, 47 ADMIN. L. REV. 1, 1 (1995); Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 430 n.91 (1989); Arthur Stock, Note, *Justice Scalia's Use of Sources in Statutory and Constitutional Interpretation: How Congress Always Loses*, 1990 DUKE L.J. 160, 160. See generally WILLIAM N. ESKRIDGE, JR. & PHILIP P. FRICKEY, *CASES AND MATERIALS ON LEGISLATION: STATUTES AND THE CREATION OF PUBLIC POLICY* 876-78 (2d ed. 1995) (summarizing arguments); Gregory E. Maggs, *Reconciling Textualism and the Chevron Doctrine: In Defense of Justice Scalia*, 28 CONN. L. REV. 393, 393-95 (1996) (same).

8. The existing body of empirical work on the *Chevron* doctrine is discussed at length in Part I.

traditional factors are present with acceptance rates when they are absent.⁹ To evaluate the political model, the study compares the voting patterns of conservative and liberal judges in ideologically polarized areas such as immigration and the environment. Finally, the interpretive model is assessed by considering how voting patterns, likelihood of finding ambiguity, and inclination to apply the two-step test vary among judges of different interpretive schools. From these data, I hope to develop a working model of the doctrine that describes “what the courts . . . do in fact”¹⁰ when judges apply the two-step formula.¹¹

The Article contains four parts. Part I explains and develops the three models of *Chevron* that are popular in the literature, paying special attention to the empirical claims made by each. Part II develops the methodology used to evaluate the claims of each model. The results of the study, along with statistical analyses, are presented in Part III. Finally, Part IV presents an analysis of the results, focusing on how well or poorly each of the three critiques of the doctrine predicts actual judicial behavior, and why. Part IV also offers possible explanations for the divergence between the models’ predictions and the reality suggested by the study’s results.

I. Three Models of the *Chevron* Doctrine

The large body of literature on the *Chevron* doctrine draws primarily from three distinct models of how the doctrine functions in practice. Each model presents a jurisprudential paradigm in which a particular set of factors is believed to alter the chances that reviewing courts will uphold agency interpretations of statutory law. Because more than one set of factors can affect the outcomes of *Chevron* cases, these models are not mutually exclusive: Several might be needed to explain patterns of judicial outcomes

9. The term “acceptance rate” is used throughout to represent the percentage rate at which the court accepts the agency’s interpretation. For example, a court that accepts two interpretations and rejects one has an acceptance rate of 67%.

10. Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 461 (1897) (“What the courts will do in fact, and nothing more pretentious, [is] what I mean by the law . . .”). I have been particularly inspired by BENJAMIN N. CARDOZO, *THE NATURE OF THE JUDICIAL PROCESS* 9-19 (1921).

11. Such a working model should be helpful to a broad cross-section of the legal community: scholars will be able to refocus their efforts with a greater knowledge of what it is they are trying to critique; students of administrative law will better understand what the much-discussed *Chevron* doctrine actually does; and administrative law practitioners will be able to predict more accurately which factors help influence the outcomes of their agency-deference cases.

This study also has broad implications for those interested in jurisprudence, particularly hermeneutics and interpretation. Although the three models question the integrity and stability of the *Chevron* doctrine in particular, they implicitly question the integrity and stability of all legal doctrines generally. Each model is the product of a distinct jurisprudential paradigm, which claims its own insights into how legal doctrines function (or malfunction) within judicial decisions. Evaluating how the three models succeed or fail in describing one doctrine provides a case study of how these paradigms succeed in describing the role of doctrine throughout law.

accurately.¹² However, despite their ability to function simultaneously, the three paradigms are conceptually very different. This part discusses the three models in detail, focusing on the theoretical assumptions that inform them and the empirical claims that these assumptions produce.

A. *The Contextual Model*

Scholars generally agree that, before *Chevron* was decided in 1984, judicial review of agency interpretations was governed by a multi-factored, contextual mode of analysis.¹³ Whether a reviewing court chose to defer to an agency's statutory interpretations depended on a number of "traditional factors" that the case law suggested should be considered. These factors included: whether the agency's construction was rendered contemporaneously with the statute's passage; whether the agency's construction was of longstanding application; whether the agency maintained its position consistently; whether the agency's interpretation was based on expertise, or involved a technical and complex subject; whether the agency's construction would expand the agency's jurisdiction; and whether Congress was aware of the agency view and expressly declined to take action to repudiate it.¹⁴ Directed to consider these factors but bound by none, courts applied flexible, functional standards of deference. At its best, this approach preserved judicial power to declare what the law is, but also respected agency judgments where the agency was institutionally better suited to decide the matter, either because the agency had developed specialized expertise, or because the circumstances of the agency's construction suggested an understanding of Congress's true intent that a court might not be able to discern.¹⁵

The central claim of the contextual model is that the traditional factors continue to influence judicial review of agency action in the *Chevron* era.¹⁶ To those who believe in the contextual model, *Chevron* is a "revolution on paper"¹⁷ that has failed to change traditional approaches to judicial review of agency interpretations of statutory law.¹⁸ As then-Judge Stephen Breyer

12. In fact, it is common for commentators on the *Chevron* doctrine to rely on more than one model. Compare Merrill, *supra* note 5 (relying on the contextual model), with Merrill, *supra* note 7 (relying on the interpretative model). Cf. CARDOZO, *supra* note 10, at 10 ("Into that strange compound which is brewed daily in the caldron of the courts, all these ingredients enter in varying proportions.").

13. See Callahan, *supra* note 5, at 1278-79; Colin S. Diver, *Statutory Interpretation in the Administrative State*, 133 U. PA. L. REV. 549, 562-64 (1985); Merrill, *supra* note 5, at 972-75.

14. See Diver, *supra* note 13, at 562 n.95 (citing cases).

15. See Merrill, *supra* note 5, at 973-75. It is worth noting that recent descriptions of the traditional factors regime probably overstate the degree to which the doctrine attained this ideal. For a sophisticated contemporaneous perspective on the traditional factors regime, see LOUIS L. JAFFE, JUDICIAL CONTROL OF ADMINISTRATIVE ACTION 569-89 (1965).

16. See *supra* note 5 for articles that rely on the contextual model.

17. Merrill, *supra* note 5, at 971.

18. See, e.g., Weaver, *supra* note 5, at 137 ("*Chevron* was fully consistent with traditional principles of judicial review.").

explained soon after the case was decided, the rigid two-step formula cannot be followed in every case because “the way in which questions of statutory interpretation may arise are too many and too complex to rely upon a single simple rule to provide an answer.”¹⁹ Because the two-step *Chevron* formula seemed difficult to square with the traditional role of a judge reviewing agency action, Breyer predicted that *Chevron* was unlikely to replace reliance on traditional factors of review.²⁰

Justice Breyer’s initial sense that the traditional factors of judicial review survive in the *Chevron* era has since been shared by many. Professor Thomas W. Merrill of Northwestern argues that the *Chevron* era rejects the traditional factors only “on paper,” and that “in practice the [contextual] approach has lived on.”²¹ Merrill goes so far as to state that he “seriously doubt[s] whether it would ever be possible to decide all deference questions without being drawn into some type of contextual or multivariate inquiry.”²² Former Chief Judge Abner Mikva of the D.C. Circuit has also expressed skepticism about whether the two-step formula could truly replace judicial reliance on contextual traditional factors: “Deference is not an independent judicial technique that can simply be evenhandedly brought to bear on each case. Rather, it expresses a result, the court’s determination in a particular case to leave the agency’s interpretation alone.”²³ Similarly, Professor Clark Byse of Harvard suggests that, in light of the continuing functional importance of the traditional factors, “the *Chevron* model may not be as simple to administer as its literal terms suggest.”²⁴ Other scholars agree that the traditional factors have retained their vitality in the *Chevron* era.²⁵

The jurisprudential underpinning of the contextual model is the belief that courts adopt a managerial stance when reviewing administrative action. According to this Legal Process-inspired vision, agencies are generally trusted to pursue broad Congressional goals through administrative action and policy. The role of reviewing courts is to ensure that agency decisions are “consistent with the purpose properly to be attributed to the statute.”²⁶ Because Congress has created dozens of different agencies and passed thousands of statutes with

19. Breyer, *supra* note 5, at 377.

20. *See id.* at 373.

21. Merrill, *supra* note 5, at 1032.

22. *Id.* at 1027 (emphasis added).

23. Mikva, *supra* note 5, at 8.

24. Byse, *supra* note 5, at 266.

25. *See Callahan, supra* note 5, at 1299 (arguing that the traditional factors are still relevant); Jordan, *supra* note 6, at 485 (asserting that the traditional factors have retained their vitality); Weaver, *supra* note 5, at 137 (“*Chevron* was fully consistent with traditional principles of judicial review.”).

26. Werhan, *supra* note 5, at 580 (quoting HENRY M. HART, JR. & ALBERT M. SACKS, *THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW* 1417 (tent. ed. 1958)). *See* Richard B. Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1669, 1675 (1975) (describing the “traditional model of administrative law” in which “[t]he court’s function is one of containment; review is directed toward keeping the agency within the directives which Congress has issued”).

varying degrees of textual specificity, this vague but critical task can be accomplished only if the courts exercise substantial discretion to adopt a proper standard of review in each case.²⁷ Courts must employ a contextual, fact-specific approach to judicial review if they are to fulfill their proper constitutional function.²⁸

Contextualists believe that the traditional factors must continue to be relevant because the two-step formula, if taken seriously, would dramatically undermine the courts' ability to fulfill their proper constitutional role.²⁹ Although generalization is somewhat difficult, they typically reason as follows: Because statutes directing administrative action tend to be vague,³⁰ *Chevron* issues usually will be resolved at the deferential second step.³¹ Further, agency interpretations will rarely be so contrary to vague statutory text that they can be labeled unreasonable.³² As a result, an acceptance of *Chevron* would portend the end of real judicial review of agency action. If *Chevron* were truly to eclipse the traditional factors, the presence of ambiguity at the first step would force judges to abdicate their duty to hold agencies legally accountable for their behavior, thereby dramatically shifting power to the executive branch.³³ The flexible, functional traditional factors would no longer give courts the power to ensure that agency action is aligned with statutory purpose. Because courts are unlikely to abdicate their proper function in this manner, the contextualists predict, deference must "operate[] quite differently in practice than it does on paper."³⁴ The traditional factors

27. See Breyer, *supra* note 5, at 379; Werhan, *supra* note 5, at 626; cf. JAMES M. LANDIS, *THE ADMINISTRATIVE PROCESS* 51 (1938) (noting that the precision of statutory directives to administrative agencies will vary with each piece of legislation).

28. See Werhan, *supra* note 5, at 627 (stating that the managerial model "allows reviewing courts to employ a contextual approach to fix their responsibility of reviewing agency decisionmaking").

29. See Cynthia R. Farina, *Statutory Interpretation and the Balance of Power in the Administrative State*, 89 COLUM. L. REV. 452, 456 (1989) (noting that acceptance of *Chevron* at face value would signal "fundamental alterations . . . in our constitutional conception of the administrative state"); Merrill, *supra* note 5, at 994 (claiming that the two-step test threatens "more enduring commitments about the proper role of the courts in a system of separated powers").

30. See LANDIS, *supra* note 27, at 51. This is especially true with regard to New Deal era statutes. See Stewart, *supra* note 26, at 1677-79.

31. See Sanford N. Caust-Ellenbogen, *Blank Checks: Restoring the Balance of Powers in the Post-Chevron Era*, 32 B.C. L. REV. 757, 824 (1991) ("[V]irtually all questions of interpretation logically fall into the second [step] of *Chevron's* two-step approach."); Mark Seidenfeld, *A Syncopated Chevron: Emphasizing Reasoned Decisionmaking in Reviewing Agency Interpretations of Statutes*, 73 TEX. L. REV. 83, 96-97 (1994) (maintaining that resolution at the first step is "relatively rare"); Keith Werhan, *Delegalizing Administrative Law*, 1996 U. ILL. L. REV. 423, 457 (arguing that it is an "unusual case" for a *Chevron* dispute to be resolved at the first step).

32. See Seidenfeld, *supra* note 31, at 96; Werhan, *supra* note 31, at 458 (stating that "post-*Chevron* courts routinely uphold the reasonableness of agency interpretations" and fail "to approach the reasonableness inquiry of step two of the *Chevron* analysis with any discernible vigor").

33. See Caust-Ellenbogen, *supra* note 31, at 824 (arguing that *Chevron* shifts so much power to agencies that it is unconstitutional, and advocating a return to the multi-factored contextual approach); Farina, *supra* note 29, at 456-57; Werhan, *supra* note 31, at 460.

34. Merrill, *supra* note 5, at 1025.

must operate surreptitiously, holding agency action to a higher standard than that which would have resulted if the *Chevron* test recited by the courts were honestly applied.³⁵

The empirical claim of the contextual model can be broken down into two distinct arguments, one of which is beyond the scope of this study, the other of which is not. The claim that is beyond the scope of this Article is that judges sometimes will ignore the two-step test completely, and instead rely exclusively upon the traditional factors for guidance. That is, judges will refuse to apply or even mention *Chevron* when reviewing an agency interpretation of law, and instead analyze the agency action under the traditional-factor test. In an important study published in the *Yale Law Journal* in 1992,³⁶ Professor Merrill provided empirical evidence suggesting that this claim was true at the Supreme Court level in the 1980s.³⁷ It is unclear, however, whether his conclusion holds true today, and whether it was ever true at the circuit court level.³⁸

The contextual model's second claim, which is within the scope of this study, is that the traditional factors continue to influence judicial review even when judges apply the *Chevron* doctrine expressly.³⁹ A recent student comment attempted to evaluate this claim with regard to one factor in particular: the consistency of an agency's interpretation over time.⁴⁰ In the traditional factors era, courts were directed to defer less or not at all to agency interpretations that had been revised.⁴¹ To determine whether courts continue to follow this directive, the author of the comment selected forty-three post-

35. *See id.* at 1027.

36. *See id.*

37. Professor Merrill found that the Court continued to cite to the traditional factors in 37% of its opinions issued from 1984 to 1990 in cases raising the issue of whether to defer to an executive interpretation. *See id.* at 981. Based on the Supreme Court's explicit reliance on the traditional factors during this time period, Professor Merrill concluded that the contextual model survived in the *Chevron* era because the Court was unwilling to apply *Chevron's* framework in every appropriate case. *See id.* at 982-83.

38. As Professor Merrill's own later work shows, the Supreme Court has cited the traditional factors less frequently since 1990. *See Merrill, supra* note 7, at 360-62. It is possible, then, that Professor Merrill's study captures the Court before it had come to an agreement that *Chevron* should replace the traditional factors. Furthermore, the possibility that applications of the two-step test in the circuit courts differ from those in the Supreme Court raises questions about how accurately Professor Merrill's study describes the bulk of the *Chevron* world. *See DAVIS & PIERCE, supra* note 2, at 123 (contrasting the consistency with which circuit courts treat *Chevron* questions with the inconsistency of Supreme Court applications). *Cf.* Peter L. Strauss, *One Hundred Fifty Cases per Year: Some Implications of the Supreme Court's Limited Resources for Judicial Review of Agency Action*, 87 COLUM. L. REV. 1093, 1117-26 (1987) (arguing that the circuit courts are the primary source of resolution of disputes concerning judicial review in the *Chevron* era). In any event, the scope of this study does not include a reinvestigation into whether the traditional factors continue to exist outside of the world of the two-step test.

39. *See, e.g., Weaver, supra* note 5, at 137.

40. *See* David M. Gossett, Comment, *Chevron, Take Two: Deference to Revised Agency Interpretations of Statutes*, 64 U. CHI. L. REV. 681 (1997).

41. *See id.* at 681.

Chevron circuit court cases in which the agency had altered its interpretation of the statute.⁴² Reviewing the cases, the author found that only two of the opinions appeared to refuse to defer to the agency because of the traditional directive that revised interpretations are to be afforded little or no deference.⁴³ The author remarked that the scarcity of opinions resting a refusal to defer on the traditional factor “strongly suggest[ed]” that courts are no longer influenced by the directive.⁴⁴ However, the limitations of the comment suggest that the continuing influence of the traditional factors remains an open question.⁴⁵ Professor Merrill has noted that a few Supreme Court cases in the 1980s applied *Chevron* and simultaneously cited the traditional factors,⁴⁶ but no comprehensive study has attempted to determine whether or how the traditional factors influence outcomes under the *Chevron* doctrine.⁴⁷

B. *The Political Model*

The second model featured in the scholarly commentary on *Chevron* is the political model. According to the political model, the best predictor of deference under the two-step doctrine is the political agenda of the judges deciding each case.⁴⁸

Although the role of politics in legal decisionmaking is a common theme in modern legal scholarship,⁴⁹ the circumstances under which the two-step test arose are particularly conducive to a purely political explanation. Throughout the 1980s, the deregulation-minded Reagan administration openly attempted to weaken the aggressive regulatory efforts of Congress.⁵⁰ In light of the pre-

42. *See id.*

43. *See id.*

44. *See id.* at 701.

45. Several limitations are immediately apparent. First, the study fails to compare deference to revised interpretations with deference to consistent interpretations. *See id.* at 701 n.91. Without this comparison, it is difficult to determine the influence of the agency revision on deference. Second, the study's methodology rests on the author's ability to identify cases in which the traditional factor actually altered the outcome, and to distinguish that set from those in which it did not. This is a difficult, if not impossible, task. Third, the scope of the study is quite narrow: The author considers only 43 cases in which one traditional factor was present.

46. *See Merrill, supra* note 5, at 982 n.54.

47. Note that the largest empirical study of the *Chevron* doctrine, Peter H. Schuck & E. Donald Elliot, *To the Chevron Station: An Empirical Study of Federal Administrative Law*, 1990 DUKE L.J. 984, is not included here. Although Shuck and Elliot found that the *Chevron* doctrine increased the overall rate of affirmance of agency cases in the circuit courts in the months after the case was decided, it did not address whether the traditional factors continued to affect the chances of affirmance in individual cases.

48. *See supra* note 6 for articles that rely on the political model.

49. *See, e.g.*, DUNCAN KENNEDY, *A CRITIQUE OF ADJUDICATION: FIN DE SIÈCLE* (1997); THE POLITICS OF LAW: A PROGRESSIVE CRITIQUE (David Kairys ed., 2d ed. 1990); ROBERTO MANGABEIRA UNGER, *THE CRITICAL LEGAL STUDIES MOVEMENT* (1986).

50. Important cases in administrative law that resulted from these efforts include *Motor Vehicle Mfrs. Ass'n of the United States v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29 (1983); *Farmworker Justice Fund, Inc. v. Brock*, 811 F.2d 613 (D.C. Cir. 1987); and, of course, *Chevron* itself.

Reagan judiciary's eagerness to adopt an active supervisory stance in agency cases,⁵¹ *Chevron's* promise of increased deference to executive interpretation generally served the interests of the Reagan administration.⁵² Greater deference equaled greater agency power, and thus increased the Reagan (and later Bush) administration's power to blunt legislative commands. As a result, throughout the 1980s, proponents of the Reagan administration's efforts tended to approve of *Chevron*, and opponents of Reagan were often harshly critical of the doctrine.⁵³

Responding to this history, proponents of the political model assert two related empirical claims at different levels of generality. First, they suggest that the strength of a judge's inclination to apply and defer under *Chevron* depends on how much the judge agrees with the politics of the administration whose interpretations are being reviewed.⁵⁴ Because the *Chevron* standard is considered particularly deferential to the executive, judges whose ideologies match that of the executive decisionmaker will be more attracted to *Chevron* and more likely to defer than others.⁵⁵ I will call this the structural claim of the political model, because it suggests that the structure and stringency of *Chevron* review depends upon whether there is a match between the politics of the executive and judicial branches.

The second empirical claim of the political model is that judges will reach results in individual cases that further their political ideologies. Liberal judges will reach liberal outcomes, and conservative judges will reach conservative outcomes. According to this model, the *Chevron* doctrine's standards are so indeterminate that judges applying them are free to adopt an outcome-orientation:⁵⁶ the two-step formal shell will support either an acceptance or a

51. See Kenneth W. Starr, *Judicial Review in the Post-Chevron Era*, 3 YALE J. ON REG. 283, 284 (1986).

52. See Pierce, *Hypertextualism*, *supra* note 6, at 779 ("When the Court decided *Chevron*, the Executive Branch was headed by a conservative, Ronald Reagan, and the Judicial Branch included a majority of liberal and moderate judges who had been appointed by Presidents Kennedy, Johnson, and Carter. In that situation, politically conservative Justices could maximize the likelihood of obtaining policies that coincided with their ideological preferences by minimizing the power of federal judges and maximizing the power of agencies. *Chevron* served that purpose well.")

53. See Jordan, *supra* note 6, at 483 (finding "a clear correlation between apparent political ideology and positions favoring or criticizing *Chevron*").

54. See *id.* at 486, 515. Cf. Cohen & Spitzer, *supra* note 6, at 68-83 (developing a model that explains why individual justices would want to adopt *Chevron* to ensure that desired political results become law). See generally William N. Eskridge, Jr. & John Ferejohn, *Making the Deal Stick: Enforcing the Original Constitutional Structure of Lawmaking in the Modern Regulatory State*, 8 J.L. ECON. & ORG. 165 (1992) (developing a similar model); William N. Eskridge, Jr. & John Ferejohn, *The Article I, Section 7 Game*, 80 GEO. L.J. 523 (1992) (same).

55. See Jordan, *supra* note 6, at 515 ("[D]eference doctrine depends in large part upon the politics of the various judges and the relationship between their politics and those of the administrations whose decisions they are reviewing.")

56. See Pierce, *Legislative Reform*, *supra* note 6, at 1123 (arguing that *Chevron* doctrine is imprecise); Shapiro & Levy, *supra* note 6, at 1070-71 (arguing that the tools of statutory interpretation upon which *Chevron* hinges are "inherently imprecise," freeing judges to reach outcomes that they prefer without fearing that their opinions will appear to violate "craft norms"); Vaughns, *supra* note 6, at

rejection of an agency's position equally well in every case.⁵⁷ Thus, for example, Professor Christopher Edley of Harvard suggests that pro-Reagan judges applying the doctrine reach outcomes designed to further pro-Reagan values.⁵⁸ Edley hypothesizes that a pro-Reagan judge's decision whether to choose "deference or intervention would depend upon which posture would, in the particular circumstances, advance the values of anti-federalism, laissez-faire efficiency, and so forth."⁵⁹ Because *Chevron* is an inherently indeterminate doctrine, judges can reach outcomes that conform to their ideological perspectives. I will call this the individual claim of the political model, because it suggests that outcomes depend upon how the effect of individual decisions can further or impede the political agendas of the judges reviewing the administrative action.

Both the structural and individual empirical claims of the political model rely almost entirely upon anecdotal evidence.⁶⁰ Although past studies have found differences in the voting patterns of judges appointed by Republican and Democratic presidents over a wide range of substantive areas,⁶¹ no empirical study has tested whether or how much the ideology of judges alters the outcomes of *Chevron* analyses.⁶² Also, no studies have measured whether

149 (arguing that the *Chevron* two-step test is "elastic enough to permit a judge to build what she will out of the language"); *The Supreme Court, 1994 Term—Leading Cases*, 109 HARV. L. REV. 111, 307 (1995) (arguing that the *Chevron* test is ambiguous and easy to manipulate).

57. See Vaughns, *supra* note 6, at 196 (stating that "appellate courts tend to give *Chevron* either a strong or weak reading which allows them to, in effect, dictate the outcome of a particular case while maintaining the appearance of legitimacy").

58. See Edley, *supra* note 6, at 587-88.

59. *Id.* at 587.

60. See, e.g., Richard J. Pierce, Jr., *Two Problems in Administrative Law: Political Polarity on the District of Columbia Circuit and Judicial Deterrence of Agency Rulemaking*, 1988 DUKE L.J. 300, 302 (looking at several cases in the mid 1980s, and concluding that "democratic D.C. Circuit judges are more likely to reverse agency policies at the behest of individuals, and republican D.C. Circuit judges are more likely to reverse agency policies challenged by business interests"); Vaughns, *supra* note 6, at 139-40 (concluding that judges use *Chevron* to satisfy individual agendas, with the evidence being a majority opinion and a dissent from a single D.C. Circuit case decided in 1986).

61. See, e.g., Timothy B. Tomasi & Jess A. Velona, Note, *All the President's Men?: A Study of Ronald Reagan's Appointments to the U.S. Courts of Appeals*, 87 COLUM. L. REV. 766 (1987). Tomasi and Velona found that in non-unanimous individual benefits cases, Democratic-appointed judges voted in favor of granting benefits in 69% of cases, those judges appointed by a Republican other than Reagan did so 34% of the time, and Reagan appointees voted to extend benefits in only 20% of the cases. See *id.* at 783. Similarly, in regulatory cases, Democratic-appointed judges adopted a pro-regulatory stance in 56% of the non-unanimous cases, judges appointed by a Republican other than Reagan did so in 48% of the cases, and Reagan appointees did so in only 41%. *Id.* This general trend was confirmed by Professor Pierce in his study of the D.C. Circuit in the mid-1980s. See Pierce, *supra* note 6, at 307 n.41.

62. A possible exception is a recent study by Professor Richard Revesz. See Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997). Professor Revesz's study was designed to determine the influence of judicial ideology on the outcomes of challenges to environmental regulation. See *id.* at 1718. After examining approximately 250 of the challenges to EPA action that were litigated in the D.C. Circuit between 1970 and 1994, Professor Revesz concluded that ideology significantly influences judicial decisionmaking in the field of environmental regulation. See *id.* at 1719. Although the scope of Professor Revesz's study was not limited to *Chevron* challenges, the fact that such ideological voting was found after 1984 suggests that

judges are more likely to apply *Chevron* if they approve of the politics of the administration, and none has studied whether they are more likely to defer when they do.

C. *The Interpretive Model*

Unlike the contextual and political models, the interpretive approach does not view the two-step test as a ritual that hides the true nature of judicial review. Instead, the interpretive model accepts that judges take *Chevron* seriously—that they try to apply the two-step formula as objectively as possible. As a result, the model accepts that judicial review under *Chevron* is governed by the language of the two-step test, and nothing but the two-step test.⁶³ The challenge posed by the interpretive model, then, is not to the doctrine's integrity, but rather its stability. The central claim of the interpretive model is that outcomes of judicial review within the new *Chevron* paradigm are influenced strongly by methods of statutory interpretation adopted by individual judges.⁶⁴ If deference kicks in when a statute is found to be ambiguous, then a judge's theory of when ambiguity exists assumes paramount importance.⁶⁵ Judges who find ambiguity more often than others will defer more often, expanding executive power, whereas judges who see less ambiguity will tend to reduce executive power.

The interpretive method targeted by most proponents of the interpretive model is textualism. Roughly speaking, textualist approaches seek statutory meaning in the language of the statutory text itself, without reference to non-textual sources such as legislative history.⁶⁶ Although not all proponents of the interpretive model agree on how textualism impacts *Chevron* deference,⁶⁷ the majority argue that textualist approaches lead to less deference than intentionalist or dynamic theories of interpretation.⁶⁸ The best-known expression of the majority view appears in a speech delivered by Justice Antonin Scalia at Duke Law School in January, 1989:

we should expect judicial ideology to affect the outcomes of *Chevron* cases. *But see* Jerome Nelson, *The Chevron Deference Rule and Judicial Review of FERC Orders*, 9 ENERGY L.J. 59, 82 (1988) (examining 18 *Chevron* cases involving Federal Energy Regulatory Commission (FERC) orders and suggesting that the ideological orientation of the judges did not affect outcomes).

63. See Werhan, *supra* note 5, at 567-68. See generally Cass R. Sunstein, *Law and Administration After Chevron*, 90 COLUM. L. REV. 2071, 2119-20 (1990) (discussing whether the *Chevron* test has truly replaced other canons of judicial review).

64. See *supra* note 7 for articles that rely on the interpretive model.

65. See, e.g., Merrill, *supra* note 5, at 990 ("In short, under the two-step *Chevron* framework, everything turns on the theory of judicial interpretation adopted at step one.").

66. For a general discussion of textualism and its alternatives, see William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. REV. 621 (1990).

67. See Maggs, *supra* note 7, at 393-95 (documenting the inconsistency of the interpretive model's attacks on textualism).

68. See *id.*

[W]here one stands on [matters of statutory interpretation]—how clear is clear—may have much to do with where one stands on . . . what *Chevron* means and whether *Chevron* is desirable. In my experience, there is a fairly close correlation between the degree to which a person is (for want of a better word) a “strict constructionist” of statutes, and the degree to which that person favors *Chevron* and is willing to give it broad scope. The reason is obvious. One who finds *more* often (as I do) that the meaning of a statute is apparent from its text and from its relationship with other laws, thereby finds *less* often that the triggering requirement for *Chevron* deference exists. . . . Contrariwise, one who abhors a “plain meaning” rule, and is willing to permit the apparent meaning of a statute to be impeached by the legislative history, will more frequently find agency-liberating ambiguity, and will discern a much broader range of “reasonable” interpretation that the agency may adopt and to which the courts must pay deference.⁶⁹

Many followers of the interpretive model agree with Justice Scalia’s conclusion that textualist judges defer to executive interpretations less often than do others.⁷⁰ Most do not consider this to be a good thing, however; they argue that the “strict constructionist” approaches popular among conservative federal judges such as Justice Scalia will lead courts to interfere with agency actions improperly.⁷¹ Instead of respecting agency spheres of expertise, it is believed, textualist judges will go to great lengths to author “intricate and imaginative opinions that . . . support a [false] conclusion that an ambiguous statute has a single plain meaning.”⁷² Finding a plain meaning then allows textualist judges to decide matters for themselves, instead of respecting the agency’s expert judgment.⁷³ By routinely attributing plain meaning to statutory language that most observers would characterize as ambiguous, textualist judges threaten to “virtually emasculat[e] the *Chevron* doctrine.”⁷⁴

A significant minority among the proponents of the interpretive paradigm assert the opposite claim—namely, that textualist methods cause excessive

69. Scalia, *supra* note 7, at 521.

70. See ESKRIDGE & FRICKEY, *supra* note 7, at 876-78 (summarizing arguments); Bell, *supra* note 7, at 132-33; Herz, *supra* note 7, at 1670-72; Karkkainen, *supra* note 7, at 460-61; Mank, *supra* note 7, at 1248-50; Pierce, *Hypertextualism*, *supra* note 6, at 752.

71. See, e.g., Herz, *supra* note 7, at 1670-72 (discussing how the textualist approaches of Reagan-appointed judges such as Justice Scalia and Judge Alex Kozinski lead them to defer only rarely to agency interpretations).

72. Pierce, *Legislative Reform*, *supra* note 6, at 1125. Professor Pierce also makes the dubious claim that textualism exacerbates serious workload problems on the federal bench because the time and effort required of textualists to resolve *Chevron* questions at the first step are overburdening. See *id.*

73. See, e.g., Mank, *supra* note 7, at 1278-84 (maintaining that textualist approaches to *Chevron* lead judges to make decisions that should be left to agencies because the agencies have substantial expertise and judges do not).

74. Pierce, *Hypertextualism*, *supra* note 6, at 752.

deference to agency views.⁷⁵ According to these commentators, textualist judges' reluctance to seek meaning in legislative history leaves them without an important tool to discern whether Congress has spoken to the direct issue at *Chevron's* first step.⁷⁶ Textualist judges are believed to find ambiguity unusually often, leading to excessive deference and an "exalt[ing of] presidential power at the expense of the other branches."⁷⁷

The interpretive model is best understood as a byproduct of the central legal debate of the last several decades: how strictly or loosely to interpret constitutional text.⁷⁸ Within this paradigm, it is axiomatic that a judge's interpretive approach is a telltale indicator of the results that the judge will tend to reach in constitutional cases.⁷⁹ When applied to the statutory realm during the 1980s, this insight triggered a rethinking of the strengths and weaknesses of various approaches to statutory interpretation.⁸⁰ The statutory debate largely traces the constitutional one: Those who advocate a "strict" interpretation of constitutional text are generally attracted to the notion of looking for and adhering to statutory "plain meaning," and those who advocate going beyond the four corners of the Constitution are less likely to look for or feel bound by statutory plain meaning.⁸¹ The interpretive model of *Chevron* applies this framework to judicial review of agency action by assuming that methods of interpreting statutory text exert a measurable effect upon outcomes. Because *Chevron* deference requires a judicial finding of statutory ambiguity, followers of the interpretive model quite naturally look to interpretive methods to determine when that will happen. Followers of different interpretive traditions will approach text differently, which in turn determines whether agency-liberating ambiguity is found, which in turn

75. The inconsistent attacks on textualism's effect upon the *Chevron* doctrine are well-documented in Maggs, *supra* note 7, at 393-95.

76. See Mashaw, *supra* note 7, at 836; Stock, *supra* note 7, at 160; Sunstein, *supra* note 7, at 430 n.91. A related point is made in Mark Burge, Note, *Regulatory Reform and the Chevron Doctrine: Can Congress Force Better Decisionmaking by Courts and Agencies?*, 75 TEX. L. REV. 1085, 1094-103 (1997) (arguing that *Chevron* is unstable because judges use inconsistent canons of statutory construction to determine whether Congress has spoken at step one).

77. Schwartz, *supra* note 7, at 57. Conversely, from a textualist perspective, those judges who rely on nontextual sources such as legislative history to create statutory meaning are guilty of judicial overreaching. See Garre, *supra* note 7, at 970-71.

78. See JOHN HART ELY, *DEMOCRACY AND DISTRUST* 1 (1980) (portraying the debate as being between those who believe that "judges deciding constitutional norms should confine themselves to enforcing norms that are stated or clearly implicit in the written Constitution," and those who believe that "courts should go beyond that set of references and enforce norms that cannot be discovered within the four corners of the document").

79. Compare the results endorsed in RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* 147 (1978) (arguing generally that constitutional text must be interpreted broadly, not as laying down particular conceptions) to those endorsed in ROBERT H. BORK, *THE TEMPTING OF AMERICA* 143 (1990) (arguing generally that interpretations of constitutional text must be strictly limited by the original, particular intent of the Framers).

80. See, e.g., Frank Easterbrook, *Statutes' Domains*, 50 U. CHI. L. REV. 533 (1983); William N. Eskridge, Jr., *Dynamic Statutory Interpretation*, 135 U. PA. L. REV. 1479 (1987).

81. See Scalia, *supra* note 7, at 521.

determines whether agency interpretations are upheld.

Proponents of the interpretive model maintain two related empirical claims. The first and primary claim has already been discussed: It asserts that textualist judges applying *Chevron* will accept executive interpretations at a different rate (whether higher or lower) than will non-textualist judges.⁸² Professor Gregory Maggs recently evaluated this position by studying the voting patterns of one prominent textualist, Justice Antonin Scalia. Professor Maggs found that Justice Scalia accepted agency interpretations in *Chevron* cases about as often as he rejected them.⁸³ Although the study did not directly address how Justice Scalia's acceptance rate compared to those of other Justices, it did find that Justice Scalia usually sided with the majority in *Chevron* cases.⁸⁴ In more than fifty cases, Justice Scalia wrote or joined only five dissents and two partial dissents.⁸⁵ The author concluded that Justice Scalia's textualism did not lead him to defer more or less often than did other Justices.⁸⁶

The second claim of the interpretive model is that textualist judges will be less inclined than will non-textualist judges to apply deference doctrines such as *Chevron*. According to Professor Merrill, the primary proponent of this view, textualism exists in considerable tension with the *Chevron* mindset.⁸⁷ The former relies upon an active and creative judicial role, in which meaning is discovered by a clever judge; the latter is based upon a passive judicial role, in which the courts defer to other institutions. As a result, textualist judges will be less likely to "think *Chevron*," and therefore less prone than other judges to apply its principles.⁸⁸

Professor Merrill attempted to test his theory empirically by studying the usage of legislative history and applications of *Chevron* by the Supreme Court.⁸⁹ The study compared the frequency of judicial reliance on legislative history to the frequency of cases decided under deference principles. The hypothesis was that decreasing use of legislative history over time could be explained by an increasing use of textualism, which would be expected to correspond to fewer applications of *Chevron* and related deference doctrines.⁹⁰ Professor Merrill found that the Court decreasingly relied upon legislative history as the 1980s progressed, which he correlated—albeit very tentatively—with decreasing reliance upon principles of deference to agency

82. See Maggs, *supra* note 7, at 393-95.

83. See *id.* at 409.

84. See *id.* at 413.

85. See *id.*

86. See *id.* at 417.

87. See Merrill, *supra* note 7, at 372.

88. See *id.*; see also Edwin A. Keller, Jr., Comment, *Death by Textualism: The NLRB's "Incidental to Patient Care" Supervisory Status Test for Charge Nurses*, 46 AM. U. L. REV. 575, 613-16 (1996).

89. See Merrill, *supra* note 7.

90. See *id.* at 353.

interpretations of law.⁹¹ From this, Merrill suggested that textualist judges were less inclined to apply *Chevron* and other deference principles than were non-textualist judges.⁹²

Although both Professor Maggs's and Professor Merrill's results are suggestive, neither responds authoritatively to the questions posed by the interpretive model. Professor Maggs' study would have been more persuasive had he compared directly the deference rates of textualist judges such as Justice Scalia to those of non-textualist judges. Without such direct comparison, it is difficult to see whether Justice Scalia's fifty-percent deference rate is unusually high, unusually low, or average.⁹³ Professor Merrill's study would have been more convincing if the data supporting his conclusion had been stronger,⁹⁴ and if he had studied the caseload of a court without entirely discretionary jurisdiction.⁹⁵ In light of these weaknesses, the questions posed by the empirical claims of the interpretive model cannot yet be considered fully answered.

II. Methodology

As with all empirical research, the results of this study are only as useful as the methodology is sound. The purpose of this part is to explore and explain the study's methodology in order to understand both the significance and the limitations of the results it produces.

91. *See id.* at 361-62.

92. *See id.* at 373.

93. Professor Maggs recognized this. *See Maggs, supra* note 7, at 412.

94. Professor Merrill's conclusion that deference doctrines such as *Chevron* were on the wane in the late 1980s and early 1990s rests on fairly weak foundations. Professor Merrill based his conclusion on the fact that during the 1992 Term, only two of the eight cases that mentioned deference actually "turned in any significant degree on" deference to agency interpretations of law. Merrill, *supra* note 7, at 362. By comparison, from 1984 to 1990, 90 cases had mentioned the deference principle, an average of almost 13 cases per Term. *See id.* at 359. Professor Merrill argues that the small number of cases decided upon deference principles during the 1992 Term reveals "a dramatic reduction in the incidence of the use of the [deference] doctrine relative to what we have witnessed throughout most of the modern era." *Id.* at 362.

Professor Merrill's analysis rests upon the assumption that a court's tendency to mention deference without deciding the case on deference principles suggests that the court is oriented away from deference. If anything, one would think that a court's tendency to mention deference when it was not needed to resolve the case before it reveals a tendency *toward* applying deference principles. In any event, Professor Merrill's data show that the number of applications of *Chevron* itself did not drop to an unusually low level in the 1992 Term. *See id.* at 359-60. Finally, Professor Merrill's conclusion does not appear to be supported by data from any Term other than the 1992 Term. *See id.*

95. Whether a court finds that its cases raise issues of deference depends on the court's docket, which in the case of the U.S. Supreme Court is determined by the entirely discretionary certiorari process. *See generally* RICHARD H. FALLON ET AL., HART & WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 37-38 (4th ed. 1996) (describing the certiorari process). As a result, the Supreme Court's failure to rely upon deference principles in a given Term is most easily explained by the Court's decision to grant certiorari in only a few deference cases that Term. *Cf. Strauss, supra* note 38, at 1121 (arguing that the relative consistency of results brought about under *Chevron* allows the post-*Chevron* Supreme Court to review fewer deference cases than before).

To evaluate the challenges to the integrity and stability of the *Chevron* doctrine posed by the contextual, political, and interpretive models, I studied every published *Chevron* opinion decided by the United States Courts of Appeals during the years 1995 and 1996. I chose to study courts of appeals decisions because the bulk of *Chevron* activity is there—the doctrine is applied more often in published courts of appeals decisions than in the published decisions of either federal district courts or the Supreme Court.⁹⁶ My goal was first to locate and then to evaluate every application of the *Chevron* doctrine in order to determine whether the challenges mounted by the three models were valid. In this part I first explain the criteria used to define a bona fide application of *Chevron* for the purposes of the study. Next, I discuss how the data from the *Chevron* applications were used to assess each of the three models.

A. *Defining a Chevron Application*

Because the purpose of this study is to find out what courts really do when they apply the *Chevron* doctrine, the study's methodological foundation is the counting of applications of the two-step doctrine, rather than of the cases that include such applications. Counting applications is superior to counting cases for two important reasons. First, the outcomes of *Chevron* cases are sometimes determined by issues that have nothing to do with the standard of review of agency interpretations of law; a court's willingness to uphold an agency's view does not preclude the court from ruling against the agency on other grounds.⁹⁷ Studying how courts resolve specific *Chevron* disputes, rather than the cases that contain them, is likely to reflect more accurately judicial understandings of how *Chevron* is to be applied.⁹⁸ Second,

96. A Westlaw search reveals that in 1995 and 1996, *Chevron* was cited by courts of appeals in 418 cases, by district courts in 346 cases, and by the Supreme Court in 17 cases. Search of Westlaw, CTA, DCT & SCT Databases (Dec. 11, 1997).

97. For example, a court could accept an agency's interpretation of a statute under *Chevron*, but find that the interpretation was included in a rule that was arbitrary and capricious under the standard of *Motor Vehicle Mfrs. Ass'n of the United States v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29 (1983).

98. Granted, counting *Chevron* applications instead of cases may overestimate the degree to which applications of *Chevron* represent independent applications of the doctrine. Judicial decisions can be the product of compromise, and we may be so cynical as to recognize that sometimes judicial opinions are the contingent products of a result in search of a rationale. See Jerome Frank, *Why Not a Clinical Lawyer-School?*, 81 U. PA. L. REV. 907, 911 (1933) (maintaining that a judicial opinion is "a censored exposition, written by a judge, of what induced him to arrive at a decision which he has already reached"); Laurence H. Silberman, *Chevron—The Intersection of Law and Policy*, 58 GEO. WASH. L. REV. 821, 828 (1990) (noting that judges "too often do not truthfully explain their reasoning" when reviewing agency action). A study counting cases instead of applications is less vulnerable to this charge because who wins and who loses has obvious real-world importance.

Fortunately, the difference between the two was minimized by the fact that I found relatively few cases in which the outcomes of the *Chevron* application did not match the outcome of the case. Furthermore, the critique of judicial opinions as contingent also attacks the case-counting approach. As long as a methodology relies in part on what arguments are put forth in the opinion (such as whether the case involves *Chevron* review at all), this argument applies to some degree.

counting applications is superior to counting cases because many cases include more than one application of the two-step doctrine. Some cases include multiple *Chevron* analyses, in which the agency view is upheld in one or more applications, but rejected in others.⁹⁹ By treating discrete applications of *Chevron* as discrete data points, the chosen methodology tracks judicial efforts to grapple with agency interpretations of law more faithfully than would counting by cases.

The decision to count by applications posed the next methodological problem: How to define a *Chevron* application. The standard was selected with three criteria in mind. First, the standard had to be as objective as possible; second, the standard had to introduce as few systematic biases as possible; and third, the standard had to exclude cases in which the doctrine was mentioned, but was not actually applied. In response to these three criteria, I established the following standard. A section of a judicial opinion counted as a *Chevron* application if and only if:

- 1) the litigation involved either an appeal from an agency adjudication or a direct challenge to an agency regulation concerning a federal administrative agency's interpretation of a statute that the agency was entrusted to administer;
- 2) the majority opinion cited *Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc.*¹⁰⁰ for the standard of review; and
- 3) a discussion of the agency's interpretation appeared in the majority opinion, and the discussion either accepted or rejected the interpretation at *Chevron*'s first step, or else accepted or rejected the agency interpretation under the "reasonable" standard of step two.

Element (1) was necessary in order to filter out uses of *Chevron* in a variety of questionable and controversial circumstances that go beyond the doctrine's original scope. For example, the *Chevron* formula is occasionally used as an interpretive aid in litigation between private parties;¹⁰¹ in attempts to understand the United States Sentencing Guidelines in criminal cases;¹⁰²

99. See, e.g., *Ober v. EPA*, 84 F.3d 304 (9th Cir. 1996) (upholding two EPA interpretations of the Clean Air Act and rejecting one).

100. 467 U.S. 837 (1984)

101. See, e.g., *Phillips Co. v. Denver and Rio Grande W. R.R. Co.*, 97 F.3d 1375, 1377 (10th Cir. 1996) (applying *Chevron* formula to quiet title in private action).

102. See, e.g., *United States v. LaBonte*, 70 F.3d 1396, 1403-08 (1st Cir. 1995) (interpreting the United States Sentencing Guidelines), *rev'd* 117 S. Ct. 1673 (1997). The Supreme Court's majority opinion did not reach the question of whether *Chevron* should be applied to administrative interpretations of the Sentencing Guidelines. See *LaBonte*, 117 S. Ct. at 1679.

and as a tool to review the legal interpretations of state agencies.¹⁰³ Depending on one's view of *Chevron's* optimal scope, these applications may or may not be proper.¹⁰⁴ However, the fact that these applications are not widespread suggests that to include them in a study of the *Chevron* doctrine would introduce a systematic bias that is better avoided. For example, in litigation between two private parties, a judge might invoke *Chevron* if the judge happened to agree with the agency interpretation, but not if the judge disagreed with the agency. The result would be an artificially high acceptance rate. Of course, it is possible that judges could do the same with *Chevron* in less controversial areas, but *Chevron's* wide acceptance in the courts for review of agency action minimizes this concern.¹⁰⁵ Element (1) thus excludes marginal cases and limits the study to appeals from adverse agency adjudications and direct challenges to agency regulations.¹⁰⁶

Element (2) requires the case in which the application appears to cite *Chevron* specifically. This requirement arises out of the need to establish simple and reasonably objective criteria for inclusion in the study. Of course, whether a judge applying the two-step formula cites *Chevron* or a subsequent case is entirely arbitrary, and disregarding those cases that do not cite *Chevron* excludes cases that apply the test in substance, if not in name.¹⁰⁷ However, excluding all cases that fail to cite *Chevron* is helpful because it dramatically simplifies the search for applications. A standard that includes all applications of two-step tests that resemble *Chevron* would be difficult to monitor: It would require plowing through thousands of individual cases for those that replicate the two-step test. Because there is no reason to believe that the cases that cite *Chevron* are systematically different from those that apply an identical standard without citing the case, Element (2) should present an accurate picture of the *Chevron* world which can be extrapolated to cases that do not cite the case itself.¹⁰⁸ The added requirement that applications cite the

103. See, e.g., *Perry v. Dowling*, 95 F.3d 231, 236-37 (2d Cir. 1996) (applying *Chevron*-type analysis to a state agency's interpretation of the Social Security Act).

104. *Chevron's* apparent capacity to take on new life in a variety of contexts has been discussed in recent scholarship such as Dan M. Kahan, *Is Chevron Relevant to Federal Criminal Law?*, 110 HARV. L. REV. 469 (1996) (arguing that the Department of Justice (DOJ) should be the primary source of interpretations of federal criminal law, and that courts should review DOJ interpretations under the *Chevron* standard).

105. Cf. DAVIS & PIERCE, *supra* note 2, at § 3.6 (stating that circuit courts almost invariably apply *Chevron's* two-step framework).

106. As a result, in most cases the government agency, its chief administrator, or the United States is a named party in the appellate litigation.

107. Indeed, studies of judicial deference in Supreme Court opinions generally have attempted to include cases that apply a *Chevron*-like framework even if the cases do not cite *Chevron* itself. See, e.g., Maggs, *supra* note 7, at 408; Merrill, *supra* note 5, at 981 n.53. Such an approach is feasible in a study of Supreme Court cases, given the Court's limited docket, but it would be impracticable in a study of circuit court cases.

108. Of course, it is possible that some judges will cite *Chevron* merely as lip service when they vote to uphold an agency interpretation. See Silberman, *supra* note 98, at 827. However, it is also possible that they will cite *Chevron* when they vote to reject an agency interpretation. Furthermore,

case specifically *for the standard of review* is meant to be relatively weak. The opinion need not state explicitly that the two-step formula provides the standard of review; rather, the requirement is merely designed to exclude cases in which *Chevron* is clearly being cited for an unrelated proposition.

The purpose of Element (3) is to define the depth of judicial review necessary for the citation to constitute a *Chevron* application. I deliberately chose a relatively liberal standard. For example, this standard includes the common judicial practice of condensing *Chevron*'s two-step formula into a single inquiry into "reasonability" or "permissibility."¹⁰⁹ Because this test is functionally identical to the two-step formulation,¹¹⁰ the study includes them as *Chevron* applications. One complication arising from the liberal construction of Element (3) was the question of how to apply it to cases that included several potential *Chevron* applications. In particular, I was concerned that applying the standard indiscriminately would overcount and thus overemphasize the importance of a few heavily litigated cases in which many agency interpretations were challenged at once. Heavily litigated cases could include frivolous claims, to which a court might respond by raising and quickly dismissing a host of *Chevron* issues. The result would be an artificially high rate of acceptance of agency interpretations. In order to counteract this effect, multiple *Chevron* applications were counted within a single case only when it seemed clear from the opinion that each application was analyzed separately, and that the citation to *Chevron* was not simply a passing comment.¹¹¹

Although every attempt was made to delineate carefully the required parameters for inclusion in this study, it must be emphasized that defining a *Chevron* application proved to be a subjective and difficult task.¹¹² Of particular concern is the possibility of bias. It is a central tenet of empirical research that one who conducts an experiment should not be aware of the hypothesis that the study seeks to evaluate if the criteria for selecting the

absent reliance on the court's own statement of the basis of its opinion, it would be impossible to determine when the court is applying the *Chevron* framework at all. Possible judicial reluctance to explain honestly the basis of decisions is a limitation inherent in all empirical research of this type. See *id.* at 828 (stating that judges "too often do not truthfully explain their reasoning" when reviewing agency action).

109. In fact, my study found that 28% of *Chevron* applications are resolved using this condensed test. See *infra* Part III.

110. See Byse, *supra* note 5, at 256.

111. This subjective "ratcheting-up" of the standard was subsequently validated by the fact that the acceptance rate of agency interpretations in cases with multiple applications matched that in cases with only a single application. The overall acceptance rate for applications in both categories was 73%. The agency's interpretation was accepted in 36 of 49 applications in multiple-application cases, and 149 of 204 applications in single-application cases.

112. This difficulty seems to be part and parcel of this type of empirical work. Cf. Merrill, *supra* note 5, at 981-82 n.53 (acknowledging that Professor Merrill's definition of a *Chevron* application is "highly subjective," involving "determin[ing] whether the author of the controlling opinion was thinking about *Chevron* in setting forth the analysis of deference").

sample are subjective; knowledge of the hypothesis creates biases in the study itself in favor of finding the hypothesis to be true.¹¹³ Unfortunately, like most others who have published empirical studies in this field, I was unable to comply with this tenet.¹¹⁴ However, it is hoped that any bias inherent in the endeavor was minimized by the study's attempt to test several hypotheses at once. Because the outcomes predicted by the different models were often in conflict, it was significantly less likely that bias in favor of one or all of the models would distort the results of the study than would have been the case had only one hypothesis been tested.¹¹⁵

B. *Designing the Tests*

With the framework for analyzing *Chevron* applications in place, the next challenge was to devise tests to evaluate the claims of the contextual, political, and interpretive models. As the scarcity of conclusive empirical work on *Chevron* might suggest, this proved a difficult task. Accordingly, I placed a high priority on developing a methodology able to test each model's claims as faithfully as possible.

1. *The Contextual Model*

The goal of testing the contextual model was to see whether and how judicial acceptance rates changed when traditional factors were present or absent. The study focused upon three traditional factors: whether the agency's construction was rendered contemporaneously with the statute's passage; whether the agency's construction was of longstanding application; and whether the agency maintained its position consistently over time. According

113. See R. CLAY SPROWLS, *ELEMENTARY STATISTICS* 117 (1955) ("A researcher with a vested interest in a problem cannot be trusted to select by judgment the sample to be used in the test.").

114. Data selection in previous empirical studies appears to have been conducted by the authors of the articles in which the studies appeared. These include Maggs, *supra* note 7; Merrill, *supra* note 5; and Tomasi & Velona, *supra* note 61. Having spent hundreds of hours reading and analyzing cases for this study, I promise that I would have gladly passed off the job to another had I been given the opportunity to do so.

115. The bias here relates almost entirely to the question of whether to count a reported decision as a *Chevron* application when it is unclear whether the court has applied *Chevron*. For example, imagine that the question was whether to count as a *Chevron* application a case in which a politically conservative, textualist Reagan appointee authored an opinion upholding an agency's denial of welfare benefits. Imagine further that the court's reported decision suggests indirectly that the court's rationale was an application of *Chevron*; although the opinion is not entirely clear, it is reasonable to conclude that the court considered the statute ambiguous and the agency construction reasonable. A subconscious bias in favor of validating the political model would influence the individual conducting the study to count the application because it would tend to validate the political model—a conservative judge had reached a conservative result. However, a subconscious bias in favor of the interpretive model would influence the author not to count the case, as the interpretive model generally predicts that textualist judges will not find statutes to be ambiguous. The effect of testing several empirical claims at once was to make it substantially less likely that a subconscious desire to prove any of the models correct would distort the results.

to the contextual model, contemporaneous, longstanding, and consistent rules should be upheld more frequently than others.¹¹⁶ These three factors were selected because they are particularly conducive to objective measurement. It would have been ideal to measure many more, if not all, of the traditional factors; however, many of the traditional factors are sufficiently elastic so as to defy objective analysis. For example, it proved impossible to determine whether an agency interpretation was based on expertise or involved a technical and complex subject. By contrast, it was possible to formulate objective tests that tracked relatively closely the substantive concerns of the three selected traditional factors.

To study whether contemporaneous agency interpretations received more deference than later interpretations, I recorded the year in which each challenged regulation was published as a final rule and compared it to the year in which the relevant portion of the associated statutory text was passed.¹¹⁷ For the purposes of this study, it was assumed that a contemporaneous interpretation is one that was published as a final rule very soon after the relevant bill became law; a non-contemporaneous interpretation would be one that appeared as a final rule many years later. Granted, appearance in the *Federal Register* as a final rule is not a perfect indicator of the true timing of an agency's adoption of a particular construction, because an agency informally can adopt an interpretation long before it appears in the form of a final rule. However, federal law mandates that no regulation can become effective until thirty days after it is published in the *Federal Register*.¹¹⁸ Consequently, using the date of publication as the "date of birth" for each rule will be accurate in most cases. Using this date allows a comparison of the acceptance rates of contemporaneous interpretations with those of later ones by comparing the acceptance rates of those interpretations promulgated soon

116. See Sunstein, *supra* note 63, at 2101. The rationale for this heightened deference is somewhat unclear, but the most common explanation is that longstanding and consistent rules deserve more respect because they are more likely to have prompted reliance interests, and that contemporaneous interpretations deserve more respect because they are more likely to reflect Congressional will, since they were adopted in the same political milieu as was the statute. See *id.*

117. I did not do the same for challenged adjudications because there was no simple way to determine the date of the initial adoption of an agency's view adopted by adjudication.

I identified the date of passage of the statutory text by studying the United States Code Annotated, the notes in which describe how the relevant sections of the statute changed over time. By using the notes, it is relatively straightforward to identify the date of passage of the particular word or phrase of statutory text that a regulation interprets. Similarly, I determined the year of promulgation of each regulation by looking up each in the *Code of Federal Regulations* (C.F.R.). I considered the date of promulgation of each rule to be the date on which the rule was published in the *Federal Register*. The C.F.R. cites to the *Federal Register*, setting forth the date of the promulgation of the final rule of each regulation. When a regulation had been amended, I consulted the *Federal Register* itself to determine the year in which the particular interpretation being challenged first appeared in the form of a final rule.

118. See 5 U.S.C. § 553(d) (1994). See generally JAMES T. O'REILLY, ADMINISTRATIVE RULEMAKING: STRUCTURING, OPPOSING, AND DEFENDING FEDERAL AGENCY REGULATIONS 144-47 (1983) (discussing the relationship between the effective dates of final rules and their publication in the *Federal Register*).

after statutory passage with those of interpretations promulgated many years later.

I assessed whether longstanding agency interpretations were upheld more frequently than recently-adopted interpretations by determining how many years had passed between the publication of each challenged final rule and the publication of the appellate opinion evaluating the challenge. That is, the study looks at whether the regulation, not the interpretation, was longstanding. The date of the final rule may not always correspond to the date that the interpretation was adopted, but using the date of the final rule permits a simple test of whether longstanding interpretations receive greater deference under *Chevron*—namely, a comparison of the acceptance rates of older rules with those of more recent ones.

Assessing whether consistent interpretations were upheld more frequently than inconsistent ones required scanning each opinion for mention of changes in the agency's interpretation over time. If the opinion did not mention any changes, it was assumed that the agency's interpretation was consistent; otherwise, the agency's position was deemed inconsistent. I then compared the acceptance rates of the applications in the two groups. Limiting the investigation to the text of the opinion itself has an obvious flaw: A court will not always mention whether the agency's interpretation was inconsistent. In fact, because one would expect a court to be more likely to mention that an interpretation was inconsistent when the court decides to reject an agency's interpretation, reliance on the opinion's text should tend to underestimate the acceptance rate of inconsistent interpretations. Nevertheless, I felt forced to adopt this approach because it would have been impractical to conduct independent investigations into the histories of 253 agency interpretations. Thus, this test is used with the understanding that it at best can disprove the hypothesis that inconsistent interpretations are upheld less often than consistent ones.¹¹⁹

2. *The Political Model*

Assessing the political model's structural claims required examining whether judges were more likely to apply and defer under *Chevron* when reviewing interpretations produced by same-party executives. This in turn required formulating ways to determine each judge's party; the administration that produced each interpretation; and how likely a judge was to apply and defer under *Chevron*.

I assigned a party affiliation to the author of each opinion by assuming

119. That is, if the test itself tends to underestimate the acceptance rate of inconsistent interpretations, then a finding that the acceptance rate of inconsistent interpretations is the same as those of consistent ones would seem to disprove the hypothesis that inconsistently held interpretations receive less deference.

that judges were members of the same party as the Presidents who appointed them.¹²⁰ This is true in all but a few cases, as only seven of the 189 circuit judges appointed from 1977 to 1995 did not come from the party of the appointing President.¹²¹

It proved more difficult to assign each interpretation to a particular administration. Unable to research the history of each agency action, I opted for two simple conventions. First, I assumed that all agency adjudications reviewed in 1995 and 1996 were the products of Clinton-era agencies. There may have been a few Bush-era adjudications that were not reviewed by the circuit courts until 1995, but there were probably only a handful. Second, I assumed that regulations were promulgated by the administration in office during the majority of the year in which the final rule was published. For example, I assumed that all regulations promulgated from 1977 through 1980 were Carter-era rules; that all final rules issued from 1981 through 1992 were Reagan and Bush-era regulations; and that all regulations promulgated in 1993 or later were the products of the Clinton administration. This convention is imperfect because it fails to account for rules promulgated in the last few weeks of each administration, and also for those created by one administration but put into effect by another. However, in most cases it should properly identify the administration responsible for each regulation.¹²²

I determined whether Republican or Democratic judges were more likely to apply *Chevron* by comparing the proportion of applications authored by active judges appointed by a particular President to the proportion of active circuit judges appointed by that President.¹²³ By collecting the sums for both parties, I could then see whether Democratic or GOP judges were more or less likely to apply *Chevron*. The theory is that judges who are more likely to apply the doctrine will author more opinions applying it than their

120. This proxy has been used before. See e.g., Revesz, *supra* note 62, at 1718 n.6.

121. See Sheldon Goldman, *Judicial Selection under Clinton: A Midterm Examination*, 78 JUDICATURE 276, 287 (1995). According to Goldman's statistics, Clinton appointed one Republican; Bush two Democrats; and Carter four Republicans to the courts of appeals. See *id.*

122. See generally O'REILLY, *supra* note 118, at 144-47 (describing the relationship between the creation and implementation of a rule and the date of its publication as a final rule in the *Federal Register*).

123. If an opinion was signed, I used 2 ALMANAC OF THE FEDERAL JUDICIARY (1996) to determine which President appointed its author. In the case of per curiam opinions, I adopted the following convention. If all judges on a panel were appointed by the same President, I would record the opinion as having been authored by an appointee of that President. See, e.g., *Texas Mun. Power Agency v. EPA*, 89 F.3d 858 (D.C. Cir. 1996) (per curiam) (a panel consisting of three Reagan appointees). If all judges on a panel were appointed by a President of the same party, and the majority of the judges were appointed by a single President, I would record the opinion as having been authored by an appointee of the President who appointed the majority. See, e.g., *Environmental Defense Fund, Inc. v. EPA*, 82 F.3d 451 (D.C. Cir. 1996) (per curiam) (a panel consisting of two Reagan appointees and one Bush appointee). In all other cases, I excluded the opinion from the relevant section of the study. See, e.g., *Kennecott Utah Copper Corp. v. United States Dep't of Interior*, 88 F.3d 1191 (D.C. Cir. 1996) (per curiam) (a panel consisting of one Bush appointee, one Reagan appointee, and one Clinton appointee).

representative proportion of the active judge pool would suggest.¹²⁴ For example, a group that comprises 20% of the active judges, but authors 40% of the applications of the doctrine in the study, can be said to be particularly *Chevron*-prone.

To calculate the proportion of all active circuit judges appointed by each President, I went one step beyond simply tabulating the distribution of active judges during the period of the study.¹²⁵ The problem with simple tabulation is that it fails to acknowledge that the distribution of judges in some circuits is more important than in others. For example, the distribution of judges in the D.C. Circuit is unusually important.¹²⁶ To correct for the unequal *Chevron* caseload in the different circuits, I calculated a weighted proportion by circuit for each set of judges. Thus, the distribution of judges on a circuit with a *Chevron* caseload twice as heavy as another is considered twice as important.¹²⁷ Using the weighted proportion allows a more direct comparison between the actual proportion of *Chevron* opinions each group authored and the proportions we would expect them to author if all groups were equally likely to apply the doctrine.

The individual hypothesis of the political model asserts that liberal judges will use the doctrine to reach liberal results, and that conservative jurists will reach conservative results. Following my earlier convention, I assumed that Democrat-appointed judges are political liberals and that GOP appointees are political conservatives. Although this is an oversimplification, it will be

124. Authorship of majority *Chevron* opinions is of course not a perfect indicator of whether a given judicial group was *Chevron*-prone; the views of judges who often write in dissent, or who choose not to write opinions on agency questions generally, will be underrepresented. However, the very low frequency of dissents in circuit court cases, combined with the absence of a reason to think that judges appointed by any particular President would be more or less attracted to agency cases, suggests that the proportion of authorship of the study's *Chevron* opinions accurately reflects each group's overall tendency to apply the doctrine.

125. I calculated the flat proportion of active circuit judges appointed by each President by using a *Federal Reporter* volume covering cases at the end of 1995 and the beginning of 1996. I selected 72 F.3d at vii-xv, which revealed that, in the middle of the study period, there were 30 active Clinton circuit judges, 40 active Bush judges, 65 active Reagan judges, 25 active Carter judges, 2 active Ford judges, 3 active Nixon judges, and 1 active judge from each of the Johnson, Kennedy, and Eisenhower administrations. See 72 F.3d at vii-xv.

I did not count senior circuit judges, because they generally take a lighter caseload than active judges. Without knowing the caseload of each senior judge, it would have been difficult to figure out how many opinions their numbers should have produced if the null hypothesis (that all judges apply *Chevron* equally often) were true.

126. The D.C. Circuit decides an inordinately large share of *Chevron* cases. It also, for example, has no active judges appointed before the Carter administration. Because about 5% of the active appellate judiciary is pre-Carter, the heavy D.C. Circuit caseload could create a false impression that pre-Carter judges are not as willing as later appointed judges to apply *Chevron*.

127. In theory, there is a shade of circularity in this test; if different judges approach *Chevron* very differently, the mix of judges in a circuit will affect the circuit's load. In practice, however, the effect is infinitesimal, as the flat and weighted proportions proved to be within 1% of each other in each case. From note 125, *supra*, we can see that 17.8% of active judges are Clinton appointees; 23.8%, Bush appointees; 38.7%, Reagan; and 14.9%, Carter. These figures are nearly identical to the weighted proportions calculated and shown in Chart 9, p. 41, *infra*.

accurate in most cases: circuit judges typically are former party activists of the appointing president's party,¹²⁸ and assertions that circuit judges are appointed to further political agendas are all too common.¹²⁹

Sorting outcomes into liberal and conservative results required adopting a simple convention. Although the political spectrum in the United States does not often divide neatly along conservative/liberal lines, I was able to classify outcomes as either liberal or conservative in several substantive areas that altogether included about 75% of the applications in the study. Specifically, it was assumed for the purposes of this study that liberal outcomes were those that furthered the causes of environmental protection, immigration, entitlement benefits, government regulation of business, and employees' rights against employers. Decisions that impeded these causes were classified as conservative.¹³⁰ Six types of cases emerged in which rejection of an agency's view generally led to a liberal result: individual entitlement benefits cases (typically involving an agency's denial of benefits); immigration cases (typically reviewing an agency's denial of asylum to an individual); environmental challenges by public interest groups (such as a suit by the Sierra Club to block lax environmental regulations); suits by employees, unions, and public interest groups alleging health, safety, or labor law violations, often in the workplace; suits by hospitals seeking reimbursement for Medicare/Medicaid expenses; and suits involving law enforcement or parole questions. Rejection of an agency's interpretation led to conservative results in four categories. The most common was the umbrella category of economic matters, which included commerce, trade issues, tariffs, taxes, finance, and customs issues. These cases typically were appeals by corporations seeking to avoid having to pay fees, taxes, or tariffs levied by an agency. The other three categories were corporate challenges to environmental rulings; corporate challenges to health, safety, and labor rulings; and challenges by states against federal regulations.

128. See Goldman, *supra* note 121, at 280 tbl.2 (showing that about 60% of Carter, Reagan, and Bush appointees, and about 50% of Clinton appointees, have histories of party activism).

129. See, e.g., Tomasi & Velona, *supra* note 61, at 766-70 (reviewing the Reagan administration's judicial philosophy and the reactions to it); see also RICHARD A. POSNER, *THE FEDERAL COURTS: CHALLENGE AND REFORM* 18 (1996) (noting the Reagan administration's attempt to shift the judiciary to the right following the appointment of "notably liberal" Carter judges). Although this might suggest that Reagan and Bush judges are more conservative than the appointees of previous Republican administrations, Tomasi and Velona's data suggest that the voting behavior of all Republican appointees was fairly similar, and that that of all Democratic appointees was also fairly consistent. See Tomasi & Velona, *supra* note 61, at 779. My own data also supported this conclusion. Consequently, I have compared the voting patterns of Republican and Democratic judges instead of the voting patterns of judges of particular Presidential appointments.

130. See generally NEWT GINGRICH, *TO RENEW AMERICA* (1995) (advocating a brand of political conservatism that trumpets free market economics, the end of the welfare state, and limited immigration).

3. *The Interpretive Model*

According to the interpretive model, judges from different interpretive schools apply *Chevron* differently. To ascertain whether this is true, it was necessary to assign each judge to an interpretive school. Making an individual determination for each judge proved a daunting task; each judge is unique, and few judges have written opinions or articles consistently endorsing particular approaches to statutory interpretation. Also, because the sample size of the study was too small to use results based on only a very limited subset of judges with known views, I needed a convention to assign each judge to an interpretive school.

The critical insight to the solution of this problem was that there should be a significant correlation between judges' interpretive philosophies and the Presidents who appointed them. Specifically, textualism is strongly associated with Reagan and Bush judges,¹³¹ and more dynamic forms of statutory interpretation are known to be more popular among appointees of Democratic Presidents.¹³² Of course, this does not mean that all Reagan and Bush judges are strict textualists, but it does mean that enough are influenced by the textualist mindset that the voting patterns of the judges *en masse* should reflect its impact. As a result, a comparison of the voting patterns of Reagan/Bush judges as a whole compared to Carter/Clinton judges as a whole should reflect the difference between textualist and non-textualist judges.

Admittedly, the accuracy of this proxy depends upon the range of approaches placed under the umbrella of "textualism." Relatively few judges adhere to the most strict form of textualism, marked by a strident refusal to employ legislative history and a frequent reliance on dictionary definitions.¹³³ If this strict form is considered the only incarnation of textualism, then there may be too few textualist judges for the proxy of presidential appointment to produce meaningful results. The reason this strict form should not be considered the only form of textualism, however, is that the crucial characteristic of textualism in the interpretive model is its claim to allow

131. See, e.g., Steven R. Greenberger, *Civil Rights and the Politics of Statutory Interpretation*, 62 U. COLO. L. REV. 37, 51-52 (1991) (noting the interest in textualism among influential Reagan appointees to the federal bench); Jeffrey W. Stempel, *The Rehnquist Court, Statutory Interpretation, Inertial Burdens, and a Misleading Version of Democracy*, 22 U. TOL. L. REV. 583, 591 (1991) (same); Nicholas S. Zeppos, *Justice Scalia's Textualism: The "New" New Legal Process*, 12 CARDOZO L. REV. 1597, 1599 (1991) (same).

132. See generally Stempel, *supra* note 131, at 590-98.

133. For an example of this approach see *MCI Telecomms. Corp. v. AT&T Co.*, 114 S. Ct. 2223, 2229-31 (1994) (Scalia, J.) (relying on several dictionary definitions instead of legislative history to determine that the word "modify" is not ambiguous at step one). Circuit judges that might be considered strict textualists are Frank Easterbrook of the Seventh Circuit and Alex Kozinski of the Ninth Circuit. See, e.g., Frank H. Easterbrook, *Text, History, and Structure in Statutory Interpretation*, 17 HARV. J.L. & PUB. POL.'Y 61 (1994) (advocating a textualist mode of statutory interpretation); Alex Kozinski, *Hunt for Laws' "True" Meaning Subverts Justice*, WALL ST. J., Jan. 31, 1989, at A18 (endorsing the use of "plain meaning" in statutory interpretation).

judges to extract an unusually high degree of meaning from text.¹³⁴ This claim is not limited to strict textualists; rather, it is a relatively common theme in the jurisprudence of the Reagan/Bush judiciary.¹³⁵ Because Reagan and Bush judges on the whole claim a commitment to finding statutory meaning in text, the interpretive model's claims about how "textualists" approach *Chevron* also applies more broadly to the mainstream of the Reagan/Bush judiciary. If the interpretive model is understood to assert that judges who find more meaning in text apply *Chevron* differently from those who find less meaning, then a comparison of the voting patterns of Reagan and Carter judges (to use an example) should reflect that difference. Therefore, this study adopts a relatively broad definition of textualism; its focus is the interpretive model's emphasis on ambiguity-production, not the technicalities of whether a judge uses dictionaries often or consistently rejects reliance on legislative history.

The test of whether judges from different interpretive schools were more or less likely than others to apply *Chevron* was a comparison of the proportion of applications authored by active circuit judges appointed by a particular President to the proportion of all active circuit judges appointed by that President.¹³⁶ If textualist judges are less inclined to apply *Chevron*, then the proportion of *Chevron* applications authored by active Reagan and Bush judges should be lower than their representative proportion of the pool of active circuit judges. To assess whether textualist and non-textualist judges deferred at different rates, I compared the overall acceptance rates of Clinton, Bush, Reagan, Carter, and pre-Carter judges. Similarly, to test the claim that textualist judges will perceive less or more statutory ambiguity, I compared the proportions of cases in which judges of each group found that a statute was ambiguous under *Chevron*'s first step.

III. Study Results

This part presents the quantitative results of the study. It begins with an overview of the *Chevron* world, which provides a context for subsequent findings relating to the three models. This overview includes figures on how often courts deferred, how often the application was resolved at step one or

134. For the purpose of this analysis, I focus on the interpretive model's claim that textualism leads to less deference than would be optimal. This position is the majority view among proponents of the interpretive model. See Maggs, *supra* note 7, at 393-95.

135. See Lino A. Graglia et al., *The Goldwater Institute and the Federalist Society: Federalism and Judicial Mandates*, 28 ARIZ. ST. L.J. 17, 107 (1996) (noting that a central claim of judicial conservatives is that "by engaging in statutory construction that ignores the plain meaning of the statutes, the courts substitute their own view of what the law should be for what the legislature has enacted"); Roger Pilon, *Rethinking Judicial Restraint*, WALL ST. J., Feb. 1, 1991, at A10 (noting that Reagan and Bush judicial nominees professed their beliefs in literal interpretation of statutory text during their confirmation hearings). Similar approaches, such as "strict construction," have been rallying cries for judicial conservatives since at least the Nixon era. See Stempel, *supra* note 131, at 590-98.

136. See *supra* notes 120-130 (discussing how this information was collected).

step two, and the *Chevron* caseload of each circuit. Next, the data on the three models are presented, beginning with the contextual model, moving to the political model, and finishing with the interpretive model.

A. *Overview of the Chevron World*

In the years 1995 and 1996, the U.S. Courts of Appeals applied the *Chevron* doctrine in 223 published cases that fit the criteria of this study. In these 223 cases, the courts applied the doctrine 253 times. The agency interpretation was accepted in 73% of these applications.¹³⁷ The number of adjudications that were challenged was roughly equal to the number of regulations that were challenged; both regulations and adjudications were upheld at essentially the same rate.¹³⁸ Overall, 76% of the interpretations were products of Democratic administrations, the great majority from the Clinton administration; only 24% were products of Republican administrations, all but one from the Reagan and Bush administrations. The most common substantive regulatory areas were the environment, trade and commerce, immigration, labor, and entitlement benefits.¹³⁹ As might be expected, the D.C. Circuit authored a disproportionate share of *Chevron* applications: Fully 30% percent of the applications of the *Chevron* doctrine originated in the D.C. Circuit. The Ninth Circuit was responsible for another 15% of the applications. Other circuits shared smaller portions of the *Chevron* caseload. The lightest output was found in the Seventh Circuit, where the doctrine was applied in only six cases—about 2% of the total.¹⁴⁰

Courts applying *Chevron* condensed the two-step test into a single question of whether the interpretation was “reasonable” in 28% of the applications.¹⁴¹ Of those cases, agency views were rejected as unreasonable 16 times and upheld as reasonable 56 times, for an acceptance rate of 78%. When the courts applied the full two-step framework, they resolved the *Chevron*

137. One hundred eighty-five interpretations were accepted, and 68 were rejected. A copy of the data is on file with *Yale Journal on Regulation*.

138. The 129 adjudications that were challenged were upheld 72% of the time. The 124 regulations that were challenged were upheld at a 74% rate.

139. The substantive breakdown was as follows: 39 applications involved the environment; 39 involved trade and commerce (including tariffs, taxes, and finance); 35 applications concerned immigration; and 18 applications involved individual entitlement. These subject areas covered about 65% of the total *Chevron* applications.

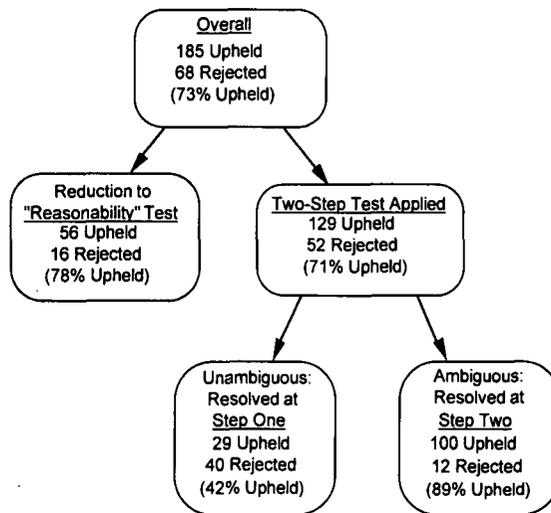
140. By circuits, the *Chevron* caseload was distributed as follows: The D.C. Circuit was responsible for 77 applications; First Circuit, 7 applications; Second Circuit, 14 applications; Third Circuit, 17 applications; Fourth Circuit, 17 applications; Fifth Circuit, 12 applications; Sixth Circuit, 9 applications; Seventh Circuit, 6 applications; Eighth Circuit, 9 applications; Ninth Circuit, 37 applications; Tenth Circuit, 14 applications; Eleventh Circuit, 9 applications; Federal Circuit, 25 applications.

141. The equivalence between the full two-step test and a single question of whether the agency interpretation is “reasonable” is suggested in Byse, *supra* note 5, at 256; see also Stephen F. Williams, *Developments in Judicial Review with Emphasis on the Concepts of Standing and Deference to the Agency*, 4 ADMIN. L.J. 113, 123-26 (1990) (statements of Judge Stephen F. Williams).

question at step one 38% of the time and at step two 62% of the time. When the analysis was resolved at step one, agency views were upheld 29 times and rejected 40 times. When the statute was declared ambiguous and the court moved on to step two, the agency constructions were accepted in 100 cases and rejected in 12 cases. Thus, courts resolving applications at step one upheld the agency interpretations only 42% of the time (compared to 73% overall), and those resolving applications at step two upheld the agency view in 89% of the applications.

These results are shown in Figure A.

FIGURE A
Acceptance Rate at Each Stage



B. Results of Testing the Contextual Model

With a general overview of *Chevron* in place, we now turn to the results for the contextual model, presented in Charts 1 through 3. Chart 1 compares the acceptance rate for consistent interpretations to that for inconsistent interpretations; Chart 2 explores whether longstanding interpretations receive greater deference; and Chart 3 examines the relative deference paid to contemporaneous interpretations.

If the contextual model is accurate, then we would expect that consistent, longstanding, and contemporaneous interpretations will be upheld more often than inconsistent, newly enacted, and noncontemporaneous interpretations.¹⁴² Looking at Charts 1, 2, and 3, the key issue is whether the results reveal a

142. See Sunstein, *supra* note 63, at 2101 (noting that consistent, contemporaneous, and longstanding interpretations receive greater deference in the contextual model).

statistically significant difference among the acceptance rates in each chart, as only the finding of a significant difference will support rejection of the null hypothesis that the contextual model is inaccurate and that traditional factors have no effect on the outcomes.

CHART 1
Consistent and Inconsistent Interpretations

Acceptance Rate for Consistent Interpretations	74% (166/225)
Acceptance Rate for Inconsistent Interpretations	68% (19/28)
Statistically Significant at 95%/90% Confidence?	No/No

CHART 2
Longstanding Interpretations

Years Between Publication of Final Rule and Court of Appeals Opinion	Percentage of Regulations Upheld Under <i>Chevron</i>
0-1	58% (14/24)
2-4	82% (37/45)
5-10	77% (23/30)
>10	72% (18/25)
Statistically Significant at 75% Confidence? ¹⁴³	No

143. $\chi^2 = 1.2, d.f. = 3.$

CHART 3
Contemporaneous Interpretations

Years Between Passage of Statute and Publication of Final Rule	Percentage of Regulations Upheld Under <i>Chevron</i>
0-4	80% (51/64)
>4	68% (41/60)
Statistically Significant at 90%/80% Confidence?	No/Yes

Chart 1 reveals that inconsistent interpretations were upheld 68% of the time, and consistent ones were upheld slightly more often, or 74% of the time. An independent means t-test for statistical significance reveals that this difference is insignificant at both the 95% and 90% confidence levels.¹⁴⁴ In other words, the difference is small enough that we cannot confidently reject the hypothesis that it is due to chance.

Chart 2 presents data showing the acceptance rate of interpretations as a function of how long regulation embodying the interpretation has been binding. It shows that very recent interpretations (a year old or less) are, in fact, upheld at a lower rate than more longstanding interpretations. However, there does not appear to be a systematic preference for more longstanding interpretations: Agency interpretations that were at least a decade old were accepted at the same rate (72%) as the overall set of regulations (74%).¹⁴⁵ Furthermore, the chi-square test of statistical significance indicates that we cannot reject the hypothesis that the variation in the acceptance rates was due to chance, even at a mere 75% confidence level.¹⁴⁶

Turning to the contemporaneous interpretations hypothesis tested in Chart 3, we see that interpretations promulgated within four years of the passage of a statute are indeed accepted at the higher rate (80%) than those that came later (68%). Although this difference is not statistically significant at a 95% or 90% confidence level, it is significant at a more forgiving 80% confidence level. In other words, we can be 80% confident that the higher acceptance rate

144. See generally THOMAS H. WONNACOTT & RONALD J. WONNACOTT, *INTRODUCTORY STATISTICS FOR BUSINESS AND ECONOMICS* 265-67 (4th ed. 1990) (explaining the basic theory behind independent means sampling for statistical significance).

145. See *supra* note 138 and accompanying text (discussing overall acceptance rate of regulations).

146. See generally WONNACOTT & WONNACOTT, *supra* note 144, at 549-55 (explaining the chi-square test). The chi-square test is a popular non-parametric test for determining statistical significance. See *id.*

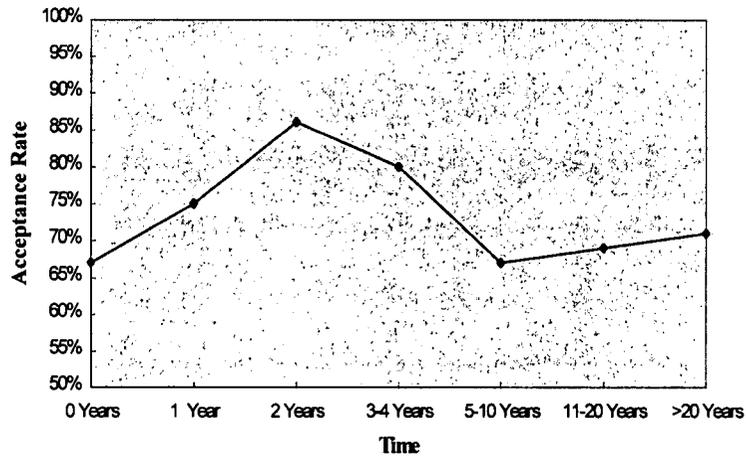
for more contemporaneous interpretations is not due to random chance. As Chart 4 and Figure B show, however, the difference shown by Chart 3 is not quite as simple when we break down the data further. Final rules that were adopted in the same year as the statute's passage were actually accepted at a relatively low rate; that rate increases and reaches its peak at the two-year mark, and then it begins to drop. Furthermore, when the data is broken down into the detail of Chart 4 and Figure B, the chi-square test fails to find the differences in the acceptance rates to be statistically significant, even at a 75% confidence level.

CHART 4
Contemporaneous Interpretations

Years Between Passage of Statute and Publication of Final Rule	Percentage of Regulations Upheld Under Chevron
0	67% (4/6)
1	75% (12/16)
2	86% (19/22)
3-4	80% (16/20)
5-10	67% (20/30)
11-20	69% (11/16)
> 20	71% (10/14)
Statistically Significant at 75% Confidence?¹⁴⁷	No

147. $\chi^2 = 0.89$, d.f. = 5.

FIGURE B
 Contemporaneous Interpretations: Acceptance Rate as a Function of Time
 Between Passage of Statute and Publication of Final Rule



To summarize, few significant differences were found among the data produced by tests of the contextual model. Consistent and inconsistent interpretations were upheld at very similar rates. Although very recent interpretations were rejected slightly more often than average, there was no general tendency for older interpretations to be upheld more frequently than younger ones. Finally, those agency regulations promulgated within four years of statutory passage were upheld at a slightly higher rate than less contemporaneous interpretations.

C. Results of Testing the Political Model

1. Structural Claims

The structural claim of the political model asserts that judges will be more likely to apply the deferential *Chevron* standard, and then more likely to accept executive interpretations, if their politics coincide with those of the administration whose interpretations are reviewed.¹⁴⁸ Because more than 75% of the interpretations reviewed in 1995 and 1996 were products of Democratic administrations, the first part of this claim amounts to a prediction that Democratic judges are more likely than Republican judges to apply the *Chevron* framework. Chart 5 examines this proposition by comparing the percentage of the applications authored by active judges of each party to the

148. See *supra* notes 54-55 and accompanying text (discussing structural claim).

percentage of active circuit judges affiliated with each party at the time of the study.¹⁴⁹ If the political model is an accurate predictor, then we would expect the percentage of *Chevron* applications authored by active Democratic judges to exceed the percentage of the active circuit court judges appointed by Democratic Presidents.

CHART 5
Tendency to Apply *Chevron*, by Party

Party Affiliation	Proportion of <i>Chevron</i> Applications	Weighted Proportion of Circuit Court Judges	Statistically Significant at 95%/90% Confidence?
Republican	68.2% (137)	66.5%	No/No
Democrat	31.8% (64)	33.5%	No/No

As Chart 5 shows, the percentage of *Chevron* applications authored by active judges of each party matches very closely each party's respective proportion of the pool of active judges. Democratic judges comprised 33.5% of the active bench and authored 31.8% of the *Chevron* applications, even though most of the executive interpretations reviewed were products of Democratic administration.

Chart 6 considers the second structural claim of the political model, that judges applying *Chevron* will tend to accept interpretations authored by their own party and to reject those of the opposing party. The chart shows that no such trend exists: judges of both parties upheld Democratic interpretations at an identical 71% rate and Republican interpretations at very similar rates (78% and 73%). These differences were very small given the sample size; neither of them is statistically significant at either a 95% or 90% confidence level.

149. The method for calculating the weighted proportion of active judges is discussed in notes 125-127 and accompanying text.

CHART 6
Tendency to Accept Interpretations by Party

Party of Executive Branch	Acceptance Rate, Republican Judge	Acceptance Rate, Democratic Judge	Statistically Significant at 95%/90% Confidence?
Republican Executive Interpretations	78% (32/41)	73% (11/15)	No/No
Democratic Executive Interpretations	71% (82/115)	71% (48/68)	No/No

Thus, neither of the political model's structural claims seems to be supported by the data. Contrary to expectations, judges aligned with the politics of the administration under review were no more likely to apply *Chevron*, and no more likely to defer when they did, than judges of the opposite party.

2. Individual Claims

Charts 7 and 8 consider the individual claims of the political model. If the individual political model correctly describes the *Chevron* world, then Republican judges should tend to use *Chevron* to reach conservative results, and Democratic judges to reach liberal results.¹⁵⁰ Charts 7 and 8 test this hypothesis by comparing the acceptance rate between Republican and Democratic judges in a variety of liberal/conservative contexts.

150. See *supra* notes 128-130 and accompanying text (discussing individual claims).

CHART 7
Cases in Which Acceptance Leads to “Conservative” Results

Subject Matter	Acceptance Rate, Republican Judge	Acceptance Rate, Democratic Judge	Statistically Significant at 95%/90% Confidence?
Individual Entitlement Benefits	100% (13/13)	40% (2/5)	Yes/Yes
Immigration Appeals	71% (17/24)	42% (5/12)	No/Yes
Environment—Public Interest Challenge	87% (13/15)	67% (2/3)	No/No
Hospitals Seeking Medicare/Medicaid Reimbursements	57% (4/7)	71% (5/7)	No/No
Health, Safety, Labor—Employee, Union, Public Interest Challenge	100% (7/7)	100% (5/5)	No/No
Law Enforcement, Prisons, Parole	66% (2/3)	0% (0/2)	No/No
TOTAL	81% (56/69)	56% (19/34)	YES/YES

CHART 8
Cases in Which Acceptance Leads to "Liberal" Results

Subject Matter	Acceptance Rate, Republican Judge	Acceptance Rate, Democratic Judge	Statistically Significant at 95%/90% Confidence?
Economic Matters (e.g., Commerce, Trade Tariffs, Taxes, Finance)	68% (19/28)	92% (12/13)	No/Yes
Health, Safety, Labor—Corporate Challenge	54% (7/13)	80% (8/10)	No/No
Environment—Corporate Challenge	78% (7/9)	75% (9/12)	No/No
Challenges by States	67% (2/3)	100% (2/2)	No/No
TOTAL	66% (35/53)	84% (31/37)	NO/YES

These data reveal that, indeed, there is a tendency for Republican and Democratic judges to reach results consistent with their political ideologies in certain areas. For example, Chart 7 shows that when individuals appealed denials of entitlement benefits, Republican judges upheld the decision to deny benefits 100% of the time (thirteen out of thirteen cases). By contrast, Democratic judges upheld the government's denial only 40% of the time (two out of five cases). Similarly, individual immigration appeals were received very differently by Republican and Democratic judges. Republican judges sided with the government against the individual 71% of the time, while Democratic judges did so only 42% of the time. These results were statistically significant, the former at a 95% confidence level, the latter at a 90% level. Chart 8 shows that this trend continued when deferral to the executive branch would further liberal policies. For example, economic regulations dealing with commerce, trade, and taxes were upheld significantly more often by Democratic judges than by Republican judges—92% versus

68%.¹⁵¹ Overall, looking at the total difference between Democratic and Republican judges in cases with relatively clear ideological significance, there is a gap of about 20% between the acceptance rates of the two groups.¹⁵²

D. *Results of Testing the Interpretive Model*

The interpretive model makes three central claims: That textualist judges will be less likely to apply *Chevron* than judges who approach text more loosely; that the two groups of judges will find statutory text to be ambiguous at significantly different rates; and that the two groups will uphold interpretations at significantly different rates.¹⁵³ The results of testing these three hypotheses are presented, respectively, in Charts 9, 10, and 11.

Chart 9 considers whether judges appointed by particular Presidents are more or less likely than others to apply *Chevron*.¹⁵⁴ Because textualist and strict constructionist approaches are closely linked with Reagan and Bush judges, the interpretive model predicts that Reagan and Bush judges will author fewer *Chevron* opinions relative to their number than will Carter and Clinton judges.¹⁵⁵ Chart 9 compares the proportion of the applications authored by active judges appointed by each President to the weighted proportion of all active judges appointed by that President.

151. In some areas, such as law enforcement (Chart 7), and challenges by states (Chart 8), differences did exist but were not statistically significant due to small sample sizes.

152. See the last lines of Charts 7 and 8. It should be noted that Charts 7 and 8 include 193 applications, more than 75% of the applications in the study. The remaining 25% covered substantive areas such as energy and communications, in which there was no clear ideological significance. The acceptance rates among Republican and Democratic judges were essentially identical in these other areas.

153. See *supra* Section I.C. (discussing interpretive model).

154. This table is a close cousin of Chart 5; it presents by President what Chart 5 presented by party. The extra detail is needed in Chart 9 because not all Republican administrations have been equally supportive of textualism, and not all Democratic executives have been equally opposed to it. In particular, Reagan and Bush judges are generally believed to be most attracted to textualism, while Carter judges, as the most liberal set of Democratic judges, are believed to be most repelled by it. See *supra* notes 131-133, 135 and accompanying text. Cf. Ronald Stidham et al., *The Voting Behavior of President Clinton's Judicial Appointments*, 80 JUDICATURE 16, 20 (1996) (comparing the voting patterns of Nixon, Ford, Carter, Reagan, Bush, and Clinton judges and finding that Carter judges vote for the liberal position significantly more often than do Clinton judges).

155. See *supra* notes 87-88 and accompanying text.

CHART 9
Tendency to Apply *Chevron*, by Presidential Appointment

Appointing President	Proportion of <i>Chevron</i> Applications	Weighted Proportion of Circuit Court Judges
Clinton	17.9% (36)	17.3%
Bush	19.9% (40)	23.5%
Reagan	44.8% (90)	39.7%
Carter	13.9% (28)	15.3%
Ford	1.0% (2)	0.8%
Nixon	2.5% (5)	1.6%
Other	0% (0)	1.7%
Statistically Significant at 75% Confidence? ¹⁵⁶		No

Chart 9 fails to support the claim that textualist judges will be less likely than others to apply the two-step test. The data show that Reagan judges are slightly more likely to apply *Chevron* than their numbers would suggest, while Bush and Carter judges are slightly less likely. Furthermore, these differences are very small and fail even a 75% confidence level chi-square test for statistical significance. In other words, the data fail to reject the hypothesis that appointees of all Presidents are equally likely to apply *Chevron*, regardless of their interpretive approaches.

The data in Chart 10 consider whether judges appointed by particular administrations found significantly more or less ambiguity than other judges. According to the interpretive model, judges who have adopted interpretive approaches that tend to find meaning in text will perceive less statutory ambiguity at step one, and therefore reach *Chevron*'s step two less frequently than other judges.¹⁵⁷ Most proponents of the interpretive model believe that textualist-influenced judges see more meaning in text, and should see less ambiguity; a minority believe that they will find less meaning, and thus more ambiguity.¹⁵⁸ Chart 10 evaluates both hypotheses by comparing the percentages of applications that reached step two authored by judges appointed by different Presidents, which corresponds roughly to different interpretive philosophies.

156. $\chi^2 = 7.2$, d.f. = 6.

157. See *supra* Section 1.C. (discussing interpretive model).

158. See *supra* notes 63-77 and accompanying text.

CHART 10
Tendency to Find Ambiguity, by Presidential Appointment

Appointing President	Percent of Cases in Which Ambiguity Was Found at Step One
Clinton	70% (21/30)
Bush	59% (17/29)
Reagan	63% (44/70)
Carter	55% (16/29)
Ford/Nixon/Eisenhower	57% (8/14)
Johnson/Kennedy	67% (2/3)
Republican	61% (69/113)
Democratic	63% (39/62)
Statistically Significant at 75% Confidence?¹⁵⁹	No

Chart 10 reveals no clear correlation among presidential appointments and the rates at which judges found ambiguity under *Chevron's* step one. Overall, Republican and Democratic judges found ambiguity at almost exactly the same rate. Notably, Carter judges found the least ambiguity; they resolved the *Chevron* analysis at step two only 55% of the time. Clinton judges found the most ambiguity, reaching step two 70% of the time. Reagan and Bush judges were sandwiched between Carter and Clinton judges, finding ambiguity in 63% and 59% of the cases, respectively. These differences were small, however, failing to achieve statistical significance.

Chart 11 compares overall acceptance rates by Presidential appointment. The interpretive model predicts that Reagan and Bush judges, being more inclined to be textualists, will have acceptance rates different from Carter and Clinton judges.¹⁶⁰ As Chart 11 illustrates, however, there is a remarkably uniform acceptance rate among judges.

159. $\chi^2 = 0.7$, d.f. = 5.

160. See *supra* Section I.C. (discussing interpretive model).

CHART 11
Overall Acceptance, by Presidential Appointment

Appointing President	Overall Acceptance
Clinton	71% (27/38)
Bush	82% (33/40)
Reagan	69% (65/94)
Carter	71% (29/41)
Ford/Nixon/Eisenhower	71% (15/21)
Johnson/Kennedy	80% (4/5)
Republican	73% (113/155)
Democratic	71% (60/84)
Statistically Significant at 75% Confidence?¹⁶¹	No

With the exception of Bush appointees, appointees of every President since Nixon accepted agency interpretations at a rate between 69% and 71%, inclusive. Bush appointees accepted agency interpretations at a slightly higher rate (82%), but this difference is not statistically significant.

Chart 11 confirms the result suggested by Charts 9 and 10. In all three, there is no clear relationship between the interpretive philosophy embraced by a set of judges and their approach to *Chevron*. All sets of judges were roughly as likely to apply *Chevron*, perceived the same degree of ambiguity, and were roughly as likely to accept agency interpretations. In short, the trends predicted by the interpretive model have not been found by this study.

IV. Analysis: Do the Three Models Describe *Chevron* Accurately?

This part examines how the results either support or contradict the predictions of the contextual, political, and interpretive models of *Chevron*. For each model, three issues are considered. The first issue is how well the model predicts patterns of judicial decisionmaking. Second, to the extent that the model does describe the patterns, I consider whether the patterns can also be explained by alternative theories. Third, to the extent that the model fails to

161. $\chi^2 = 0.7688$, d.f. = 5.

describe the patterns accurately, I examine possible misunderstandings that might have led proponents of the model to adopt a mistaken view of how judges apply the doctrine.

A. *Analysis of the Contextual Model*

The data suggest that the contextual model is a poor predictor of judicial behavior under *Chevron*. Contrary to the predictions of the contextual model, inconsistently held interpretations were accepted at essentially the same rate as those that were consistent over time.¹⁶² Although freshly enacted regulations were invalidated unusually often, there was no systematic preference in favor of upholding regulations that were longstanding: regulations that had been in effect for a decade or more were upheld at the same rate as newer ones, and the regulations that were upheld most often were those that had been enacted only two to four years prior to judicial resolution in the courts of appeals.¹⁶³ These results indicate that, at best, the traditional factors exert a weak influence on *Chevron* outcomes.

The likelihood that the traditional factors continue to influence post-*Chevron* courts appears strongest when we consider the acceptance of contemporaneous interpretations. Executive interpretations that appeared in final rules within four years of a statute's passage were upheld at a higher rate (80%) than those that appeared later (68%).¹⁶⁴ Although the difference is significant only at an 80% confidence level, the difference is precisely what the contemporaneous adoption factor would predict—that courts should defer more frequently to interpretations that followed soon after a statute was passed.

Does this mean that the contemporaneous-interpretation traditional factor is still vital? Not necessarily. Courts might accept nearly contemporaneous interpretations more frequently than later ones because contemporaneous interpretations actually reflect statutory text more accurately. That is, it is

162. Although a six percent difference exists, it is statistically insignificant and likely the result of chance. *See supra* p. 32 cht.1. Further, because the test itself was biased in favor of finding a difference, the absence of a larger difference in acceptance rates suggests that we can be fairly confident that consistency does not impact *Chevron* deference. *See supra* note 119.

163. *See supra* p. 32 cht.2. The fact that regulations that were promulgated less than a year prior to judicial resolution were invalidated unusually frequently cannot convincingly be explained by the traditional preference for longstanding interpretations, because this preference generally refers to interpretations that have existed for decades. *See, e.g., Andrus v. Shell Oil Co.*, 446 U.S. 657, 667-68 (1980) (recognizing increased deference for interpretation that had been in effect for several decades). Perhaps it can be explained by judicial willingness to invalidate a rule with regards to which few reliance interests have developed, and which may never have been put into effect. Alternatively, the high invalidation rate for very young regulations may be explained by the Clinton administration's promulgation of several weak regulations during 1994 and 1995. For example, several amendments to the Clean Air Act were not received well by the courts. *See, e.g., Davis County Solid Waste Mgmt. v. EPA*, 101 F.3d 1395 (D.C. Cir. 1996) (invalidating emissions standards regulation); *American Petroleum Inst. v. EPA*, 52 F.3d 1113 (D.C. Cir. 1995) (invalidating reformulated gasoline regulations).

164. *See supra* p. 33 cht.3.

unclear whether the higher acceptance rates are due to the courts' *sense* that contemporaneous interpretations better reflect congressional intent (a traditional explanation of the contemporaneous factor)¹⁶⁵ or, alternatively, the *reality* that they do. If the former is true, then the traditional factor is still vital; if not, then the theory inspiring the traditional factor is validated, but the factor itself is dead.

There are two reasons to think that the latter understanding better explains the slightly higher acceptance rate of contemporaneous interpretations. First, the failure of most judicial opinions to report the dates of both statutory passage and the appearance of the final rule suggests that many courts were unconcerned or even unaware of whether the interpretations at issue were contemporaneous.¹⁶⁶ If judges paid attention to the timing of final rules and used that timing either to relax or ratchet-up standards of deference, as the contextual model predicts, we would expect that these dates would usually be noted in judicial opinions.

A second reason to accept the latter theory is the possibility that it can explain the peculiar shape of the curve in Figure B. As Figure B shows, truly contemporaneous interpretations—those promulgated less than two years after the passage of a statute—are actually accepted at a relatively low rate. The rate climbs as time passes, reaches a peak at two years after the statute was passed, and then declines. Although these patterns could be attributable to chance fluctuations,¹⁶⁷ they might also be explained by a functional understanding of agency behavior. Perhaps final rules promulgated within months of the passage of a statute are less faithful interpretations of statutory text because they are products of hurried efforts to publish a rule as soon as possible. By comparison, rules promulgated two years after the statute was passed might be expected to be more careful and faithful products of notice and comment procedures. Finally, regulations promulgated many years later could be the product of an agency's attempts to do more with the statute than its text and original purpose might allow, as the contemporaneous interpretation theory suggests. The result would be a pattern of acceptance rates resembling the curve of Figure B.

Ultimately, it is unclear whether the results in Chart 3 are products of the contextual model's influence, the reality of how well agencies interpret statutes at different times, or chance fluctuations. In any event, the data suggest that, at best, the contextual model exerts only a very weak influence on the outcomes of applications of *Chevron*.

Although the data suggest that traditional factors have not retained their significance in the *Chevron* era, two important limitations must be taken into

165. See Sunstein, *supra* note 63, at 2102.

166. I would estimate that both dates were included in about 30% of the applications I analyzed.

167. Recall that a chi-square test of significance found that the results of Figure B could not reject (even at a 75% confidence level) the hypothesis that the fluctuations were due to chance.

account. First, traditional factors appear to remain important to particular judges, who continue to discuss them in their published opinions.¹⁶⁸ If judges assigned to a particular appellate panel believe that the traditional factors are still relevant, then, *a priori*, they are. Of course, judicial citation alone does not necessarily connote importance; it is possible that judges only discuss traditional factors when the factors support the conclusion that they would otherwise reach.¹⁶⁹ In any event, discussions of traditional factors were limited to about 5% of the cases studied.

The second limitation on the conclusion that the contextual model is a poor predictor of post-*Chevron* deference is the fact that this study examined only three factors that were themselves interrelated.¹⁷⁰ It is possible that some untested factors remain strong predictors of deference, so that the contextual model remains viable in some areas, but not in others. It would be odd if *Chevron* wiped out some factors while leaving others in place, but it would not be inexplicable. Commentators have noted that the functional rationale for deferring more often to consistent, contemporaneous, and longstanding interpretations has never been entirely clear.¹⁷¹ It is possible that these traditional factors have been preempted by *Chevron* whereas other, more functional factors have survived. For example, courts might ignore whether interpretations were longstanding, consistent, and contemporaneous, but nonetheless defer at a higher rate to interpretations involving scientific or technical questions in tacit recognition of agency expertise.¹⁷²

Despite these two limitations, this study's results reject the stronger claims of the contextual model, that the traditional factors are alive and well in the *Chevron* era. The traditional factors may impact some cases in small ways,

168. See, e.g., *Visiting Nurse Ass'n v. Bullen*, 93 F.3d 997, 1007 (1st Cir. 1996) (Cyr, J.) (asserting that longstanding interpretations receive more deference); *Credit Union Ins. Corp. v. United States*, 86 F.3d 1326, 1332 (4th Cir. 1996) (Russell, J.) (arguing that inconsistent interpretations receive less deference than consistent interpretations); *Hanauer v. Reich*, 82 F.3d 1304, 1311 n.4 (4th Cir. 1996) (Hamilton, J.) (same); *Sarasota Mem'l Hosp. v. Shalala*, 60 F.3d 1507, 1511 (11th Cir. 1995) (Dubina, J.) (maintaining that interpretations involving agency expertise receive more deference); *Florida Mfrd. Hous. Ass'n v. Cisneros*, 53 F.3d 1565, 1573 (11th Cir. 1995) (Carnes, J.) (arguing that inconsistent interpretations receive less deference than consistent interpretations); *Barrera-Echavarria v. Rison*, 44 F.3d 1441, 1444-45 (9th Cir. 1995) (Hall, J.) (claiming that congressional acquiescence validates earlier agency interpretation).

169. If this is the case, then the traditional factors may be cited even though they are not actually influencing judicial outcomes. Compare *Florida Mfrd. Hous. Ass'n v. Cisneros*, 53 F.3d 1565, 1574 (11th Cir. 1995) (rejecting inconsistent interpretation (citing *INS v. Cardoza-Fonseca*, 480 U.S. 421, 446 n.30 (1987) for the proposition that inconsistent interpretations receive less deference)), with *Massachusetts v. FDIC*, 102 F.3d 615, 621 (1st Cir. 1996) (upholding inconsistent interpretation (citing *Rust v. Sullivan*, 500 U.S. 173, 186 (1991) for the position that inconsistent interpretations do not receive less deference)).

170. The factors are related in that they all concern the timing of the administrative interpretation, as compared to the substance of the regulation or the agency involved.

171. See, e.g., Sunstein, *supra* note 63, at 2101-02.

172. Note, however, that it is difficult to distinguish between interpretations that involve technical and scientific questions and those that do not, making measurement difficult. See *supra* Subsection II.A.1.

but the assertion that *Chevron* is a “revolution on paper”¹⁷³ that has left the traditional factor regime completely intact seems difficult to maintain. The traditional factors may once have governed judicial review of agency action, but the evidence suggests that, at the court of appeals level, they have been largely eclipsed by *Chevron*.

The contextual model’s apparent failure to predict accurately post-*Chevron* results prompts the question: Where have followers of the contextual model erred? One possibility is the contextualists’ misunderstanding of how often courts reject agency interpretations. Recall that the contextualists’ faith in the traditional factors typically is inspired by an understanding that taking the two-step model seriously would shift powers dramatically to the agencies and lead to an abdication of the proper judicial function.¹⁷⁴ This concern follows from an understanding that the *Chevron* test nearly always will force courts to accede to agency interpretations: Most statutes will be ambiguous, and nearly all ambiguous statutes will be upheld. As a result, the test is seen as a “blank check” to the agencies which must be tempered by the traditional factors regime.

The fear that *Chevron* is a blank check appears to be unjustified, however. As Figure A illustrates, fully 38% of the applications of *Chevron* were resolved at step one, and 58% of these resolved applications were decided *against* the agency. Even when the courts acknowledged statutory ambiguity and moved on to the second step, the agency’s interpretation was still rejected 11% of the time. Furthermore, when the court collapsed the *Chevron* doctrine into a single test of reasonableness, the court rejected the agency view as “unreasonable” 22% of the time.

Judicial willingness to declare text unambiguous, and even to invalidate interpretations of admittedly ambiguous statutes, suggests that *Chevron* is probably not the blank check that many contextualists have feared it to be. Even at step two, judges have refused to abdicate their task of reviewing agency action. Granted, some may question whether this willingness to reject agency positions reflects a true reading of the language of *Chevron*. Perhaps judges have adopted readings of “ambiguous” and “reasonable” that are more or less strict than the two-step’s text or history would support in order to attain the rough degree of deference that most judges deem proper. Regardless, the unexpected and previously unmeasured “bite” of both the first and second steps explains how the traditional factors could be eclipsed to a large extent. Judicial standards of what is “ambiguous” at step one and “reasonable” at step two have been set at levels that rein in agencies reasonably effectively. Perhaps the traditional factors are not needed to temper *Chevron*, because judges have interpreted *Chevron* such that it tempers itself.

173. Merrill, *supra*, note 5, at 971.

174. See generally *supra* notes 28-30. ↘

B. *Analysis of the Political Model*

The political model predicts *Chevron* outcomes with only mixed success. The model's structural claims, which predict that the strength of a judge's inclination to apply and defer under *Chevron* depends on how much the judge agrees with the politics of the administration reviewed, appear to be incorrect. However, the study generally validates the individual claim—namely, that judges will sometimes vote with their politics in individual cases. This finding is made with the caveat that the tendency appears to be no more pronounced in the *Chevron* era than under the traditional factors regime.

1. *Structural Claims*

The results of the study fail to support the structural claims of the political model. Democratic judges were not more likely to apply *Chevron* than Republican judges, even though Democratic judges were presumably more “in tune” with the politics of the Democratic administrations the courts spent most of their time reviewing. Furthermore, when judges did apply the two-step *Chevron* test, they did not agree with the interpretations of like-minded administrations more frequently than with those of the opposing party. These results suggest that judges do not use *Chevron* as a weapon to support the administrative efforts of political allies and to undermine those of political enemies. Judges who are politically aligned with a particular administration do not go out of their way to invoke *Chevron* in an effort to make sure that its administrative judgments are upheld, and they are no more likely to accept the interpretations of agencies controlled by political allies than those controlled by enemies.¹⁷⁵

The likely error made by proponents of the structural claims of the political model is acceptance of an oversimplified view of agency output. Proponents generally assume that Democratic and Republican administrations produce systematically “liberal” and “conservative” interpretations, respectively. This is misleading. Although Republican agencies may in general produce new interpretations that are more conservative than those of Democratic agencies, executive interpretations are almost always the products of some ideological compromise.¹⁷⁶ As a result, Democratic executive interpretations will often be challenged by groups seeking more liberal results, and GOP executive constructions will be challenged by parties desiring more

175. This conclusion is supported by a recent empirical study of Supreme Court review of agency action. See John C. Kilwein & Richard A. Brisbin, Jr., *Supreme Court Review of Federal Administrative Agencies*, 80 JUDICATURE 130, 135 (1996) (“Apparently party control of political institutions does not have a significant influence on the Court’s support of agencies. The Court seems to be at best only slightly affected by the partisan context of national politics when it decides agency cases.”).

176. See generally O’REILLY, *supra* note 118, at 62.

conservative results. Furthermore, parties often challenge agency actions that resist ideological categorization, such as the details of tariff calculation.¹⁷⁷ Given the full panoply of agency action, it is wrong to assume that accepting the interpretations of Republican executives will further conservative views, and that upholding Democratic executive constructions furthers liberal policies. The correlation between the politics of an administration and the consequences of upholding its interpretations is too tenuous to allow political judges to approach *Chevron* differently when reviewing the interpretations of different administrations.

2. Individual Claims

In contrast with the structural claims of the political model, the individual claim accurately predicted voting trends in several types of cases. Just as the political model asserts, judges who applied *Chevron* in cases with clear political ramifications tended to reach results consistent with their political views. The acceptance rate among Democratic and Republican judges in these cases differed by about 20%, with Democratic judges more likely to author opinions reaching liberal results and Republican judges more likely to author opinions reaching conservative results.¹⁷⁸

Admittedly, these differences are somewhat smaller than the numbers suggest due to judicial self-selection. Self-selection may occur when the presiding judge on an appellate panel assigns opinions among the judges on the panel: Judges will tend to be more attracted to authoring opinions aligned with their political philosophy than to authoring those that oppose it. For example, imagine that a panel of two Carter appointees and one Reagan appointee decides an immigration case, and that all three judges agree on the merits that the government reasonably interpreted the law in denying asylum

177. See, e.g., *Zenith Elecs. Corp. v. United States*, 77 F.3d 426 (Fed. Cir. 1995) (challenging the methodology of how an agency calculated the fair market value of products to arrive at the proper tariff); see also Ronald M. Levin, *Judicial Review and the Uncertain Appeal of Certainty on Appeal*, 44 DUKE L.J. 1081, 1084 (1995) (“[M]any cases do not involve such high stakes and public passions. In those cases, it is much easier to believe that judges are honestly willing to be guided by [doctrine].”).

178. See *supra* pp. 38-39 chs.7-8. This difference is particularly dramatic in certain areas. In Chart 7, for example, we see that Republican judges affirmed the government’s decision to deny entitlement benefits 13 out of 13 times, whereas Democratic judges affirmed such decisions only two out of five times. Similarly, in Chart 8, Democratic judges upheld government economic regulations at a 92% rate, whereas Republican judges upheld the economic regulations at a 68% rate. In both cases, Republican judges were more likely to reach results that furthered the agenda of the Republican party, and Democratic judges were more likely to reach results that furthered the Democratic party’s agenda.

The executive interpretations examined in Charts 7 and 8 represent 193 out of the 239 *Chevron* applications that were identifiable as having been authored by a judge of a particular party. (Although there were 253 *Chevron* applications, the authorship of 14 applications could not be identified by party because they appeared in *per curiam* opinions authored by a panel of both Republican and Democratic judges.) The remainder of the applications were primarily in the energy, communications, and transportation fields, in which there were no obvious ideological stakes and no difference in the acceptance rates between Democratic and Republican judges.

to the appellant. If judges are attracted to writing opinions that are roughly aligned with their political leanings, we would expect that the Reagan appointee would be more likely than either other judge to author the opinion. If this dynamic is repeated, then liberal judges will tend to author more liberal opinions, and conservative judges will tend to author more conservative opinions, even if judges of both parties are in complete agreement with each other on the merits of every case.

The existence of a self-selection problem was confirmed by studying the panels involved in deciding each case in two particularly divergent areas: immigration and individual entitlement benefits. In particular, I was interested in unanimous panels comprised of both Republican and Democratic appointees. If there is a self-selection problem, then the minority party judge will be less than 33% likely to author the opinion if the opinion resonates with the politics of the majority party, and more than 33% likely to author the opinion if it instead aligns with the agenda of the minority party.¹⁷⁹ In fact, this is exactly what happened. When the outcome was aligned with the politics of the majority party, the minority judge authored only 13% of the opinions (2 out of 15); when it was aligned with the politics of the minority party, the minority judge authored 46% of the opinions (6 out of 13). The presence of self-selection means that the different rates in Charts 7 and 8 overstate the actual impact of political differences on the outcomes of *Chevron* cases. Part of the difference is due to judges reaching *Chevron* outcomes that are consistent with their politics, but part of it is due to the judges' eagerness to write opinions when their legal conclusions happen to match their political preferences.

Even acknowledging the self-selection problem, there are two ways to interpret the divergent voting patterns of liberal and conservative judges in politically charged subject areas. On an absolute level, it is clear that judges applying *Chevron* are sometimes influenced by the political stakes of the litigation. The more interesting question, however, is the relative one: Accepting the realist insight that politics can influence outcomes, how does the impact of politics in the *Chevron* world compare to its importance *outside* of *Chevron*? This is the true test of the political model—whether the *Chevron* framework reduces or expands the discretion of those judges who may be inclined to use judicial review as an opportunity to enact their personal policy preferences into law.¹⁸⁰

179. Because the cases were heard before a three-judge panel, each case was heard either by two GOP appointees and a Democratic appointee, or vice-versa. Two judges can be termed majority party judges on such a panel; the third is the minority party judge. The 33% figure represents the random likelihood that any one of the three judges would be assigned authorship of the opinion in each case.

180. See Shapiro & Levy, *supra* note 6, at 1070 (arguing that the tools of statutory interpretation upon which *Chevron* hinges are “inherently imprecise,” freeing judges to reach outcomes that they prefer); Vaughns, *supra* note 6, at 149-50 (arguing that the *Chevron* two-step test is “elastic enough to permit a judge to build what she will out of the language”).

Comparing the results of this study with those of an earlier study by Professor Jon Gottschall suggests, albeit tentatively, that *Chevron* has neither significantly increased nor significantly decreased the degree of outcome-orientation that existed during the pre-*Chevron* era. Professor Gottschall analyzed the voting records of Republican-appointed and Democrat-appointed circuit judges starting just over a year before *Chevron* was decided.¹⁸¹ Although there is some overlap between the period of Professor Gottschall's study and the beginning of the *Chevron* era,¹⁸² the impact of *Chevron* on Gottschall's results is negligible.¹⁸³ Therefore, we can use his study to draw a rough comparison between how judges voted in the traditional factors era and how they vote in the *Chevron* era.

Gottschall analyzed the voting patterns of Republican and Democratic appointees in two areas that are similar to those used in this study: labor and welfare.¹⁸⁴ Instead of examining opinion authorship, Gottschall tallied the votes of all judges on each panel, assigning to each vote a liberal or conservative orientation.¹⁸⁵ Because this approach corrects for any self-selection that may overstate the degree of political difference between the two sets of judges, small differences in the voting patterns between Republican and Democratic judges observed by Gottschall should correspond to larger differences in our own study. Examining labor cases, Gottschall found a 20% gap between how often Republican and Democratic judges voted for liberal outcomes.¹⁸⁶ By way of contrast, in this study's category of health, safety, and labor cases, there was a gap of about 25% between the rates at which Republican and Democratic judges authored such opinions.¹⁸⁷ In the welfare area, Gottschall found an 8% difference between GOP and Democrat appointees' votes.¹⁸⁸ In my own study, the 18 welfare cases yielded a large gap

181. See Jon Gottschall, *Reagan's Appointments to the U.S. Courts of Appeals: The Continuation of a Judicial Revolution*, 70 JUDICATURE 48, 51 (1986).

182. *Chevron* was handed down on July 25, 1984. *Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc.*, 467 U.S. 837 (1984). Gottschall's study analyzed cases from July 1, 1983 to December 31, 1984. See Gottschall, *supra* note 181, at 51.

183. A Westlaw search in the court of appeals database in the year 1984 reveals that *Chevron* was cited in only seven majority opinions by the courts of appeals during the five month period of overlap. Search of Westlaw, CTA Database (Dec. 11, 1997). Because Gottschall's study encompasses almost four thousand cases, see Gottschall, *supra* note 181, at 51, the effect of *Chevron* was insignificant.

184. See Gottschall, *supra* note 181, at 51.

185. See *id.*

186. See *id.* at 54 tbl.6. Overall, Republican judges voted for the liberal result about 50% of the time, and Democratic judges about 69% of the time. To reach these figures, I have taken the liberty of combining Gottschall's results for the Kennedy/Johnson and Carter judges to arrive at the 69% figure for Democratic appointees and the Nixon/Ford and Reagan judges to arrive at the 50% figure for GOP appointees.

187. See *supra* p. 39 cht.8.

188. See Gottschall, *supra* note 181, at 54 tbl.6. Again, I have taken the liberty of combining Gottschall's results for Kennedy/Johnson and Carter appointees to arrive at a Democratic figure, and those for Nixon/Ford and Reagan appointees to arrive at an overall Republican figure.

of 60%.¹⁸⁹ Combining all the categories Gottschall studied, he found a gap of about 10 or 20%;¹⁹⁰ the last lines of Charts 7 and 8 reveal an overall gap in this study of about 20 or 25%.

Comparisons between the two studies are complicated by different methodologies and time periods and the relatively small number of cases I studied. Most glaringly, Gottschall's study is not limited to judicial review of agency interpretations of statutory law, which makes a direct comparison between the *Chevron* regime and the traditional factors regime difficult.¹⁹¹ Despite this caveat, one can draw the very tentative conclusion that, overall, *Chevron* has probably not dramatically altered the pre-*Chevron* level of judicial outcome-orientation in politically charged cases. Keeping in mind the self-selection biases, it appears that *Chevron* has reduced the zone of discretion in the labor area. However, it apparently has expanded the zone in the welfare area, although probably not as profoundly as one might infer from the 60% difference, which was likely due in large part to the extremely small sample size.¹⁹²

Looking at the aggregate data, the difference between the overall gap shown by Gottschall and that in this study appears well within the order of bias caused by self-selection. At the very least, the comparison fails to support the conclusion that *Chevron* outcomes are unusually indeterminate in a way that frees judges to pursue politically desirable outcomes more than they may under other legal regimes. Although more evidence would have to be collected before a more definite conclusion could be reached, this admittedly rough comparison fails to support the view that the *Chevron* regime is either dramatically more or less determinate than the traditional factors regime.¹⁹³

189. See *supra* p. 38 cht.7.

190. See Gottschall, *supra* note 181, at 54 tbl.6.

191. However, appeals in the welfare and labor areas will be almost entirely reviews of agency action, making the comparison between Gottschall's numbers and my own at least plausible.

192. The 40% acceptance rate for Democratic judges was based on only 5 cases; the 100% rate for Republican judges was based on 13 cases.

193. A review of the literature on pre-*Chevron* judicial review of agency interpretations of law suggests that *Chevron*'s failure to be less determinate than the traditional factor test should not be altogether surprising. Before *Chevron* was decided, there was a fairly broad scholarly consensus that judicial review doctrine was chaotic and unpredictable, if not nonexistent. See Stephen M. Lynch, Note, *A Framework for Judicial Review of an Agency's Statutory Interpretation: Chevron, U.S.A., Inc. v. Natural Resources Defense Council*, 1985 DUKE L.J. 469. It was generally understood that the Supreme Court had developed two independent lines of authority on the issue of whether a reviewing court should defer to agency interpretations of law: One line of cases granted deference to agency interpretations, and the other did not. It was also understood that there was little or no way to determine when the Court would choose one line of cases over the other. See, e.g., 4 KENNETH CULP DAVIS, ADMINISTRATIVE LAW TREATISE § 30.07 (1958) (noting that the Court had never attempted to explain why it applied a given line of cases); Louis L. Jaffe, *Judicial Review: Question of Law*, 69 HARV. L. REV. 239, 258-62 (1955) (identifying the two lines of cases); Henry P. Monaghan, *Marbury and the Administrative State*, 83 COLUM. L. REV. 1, 4 (1983) (terming the case law "apparently erratic"). Thus, what today is sometimes considered the elegant "traditional factors regime" was then considered essentially unpredictable and result-oriented. However indeterminate the *Chevron* doctrine may be, it is difficult to see how it could be less determinate than the doctrinal vacuum it replaced. See Richard J. Pierce, Jr., *Chevron and Its*

C. *Analysis of the Interpretive Model*

The most conclusive results produced by this study relate to the interpretive model. The data uniformly fail to support the model's claims, finding no correlation between a judge's interpretive method and his or her approach to *Chevron*. This unexpected result suggests the need for a reexamination of the connection between *Chevron* and textualism in order to gain a better understanding of both.

The results directly contradict the most common and important claim made by followers of the interpretive model, which is that textualist judges will find a different level of ambiguity at step one than will nontextualist judges.¹⁹⁴ In fact, judges whom we know to be particularly influenced by textualism were no more likely than other judges to find that a statute had a plain meaning at step one.¹⁹⁵ The data also fail to support the related assertions that textualist judges will either interfere with agency actions improperly or be too deferential. For example, the overall rate at which textualism-influenced Reagan judges upheld agency interpretations of law was nearly identical to the rate at which Carter judges did so.¹⁹⁶ Finally, the data contradict Professor Merrill's claim that textualist judges will be less *Chevron*-prone than nontextualist judges by demonstrating no correlation between a given group's

Aftermath: Judicial Review of Agency Interpretations of Statutory Provisions, 41 VAND. L. REV. 301, 313 (1988) (concluding that under *Chevron*, "judges will have less room to infuse their personal political philosophies in the Nation's policy making process").

194. See *supra* note 7 for articles that rely on the interpretive model.

195. Chart 10 on page 41 shows no correlation between a judge's interpretive philosophy and his or her tendency to find ambiguity. Clinton and Carter judges found the most and the least amount of ambiguity, respectively; Bush and Reagan judges were in the middle. These differences were slight, statistically insignificant, and almost certainly explained by chance, suggesting that the judges in each group were equally prone to find ambiguity in statutory text.

196. See *supra* p. 43 cmt.11. The acceptance rate among all judges was remarkably uniform, with the vast majority between 69% and 71%. See *id.*

There are several potential explanations for the higher acceptance rate (82%) of Bush judges. One explanation is that Bush judges took the bench at a time when *Chevron* was well-established, Congress was and had been Democratic for some time, and the presidency had been held by Republicans since 1980. Looking back at the structural claims of the political model, one could imagine that judges trained in this atmosphere would associate deference with helping the administration that appointed them. These judges thus would have been "trained" to accept agency interpretations. This explanation is undercut by the apparent inability of the structural claims of the political model to predict patterns of *Chevron* cases accurately.

A second and related explanation is the Bush administration's public search for judicial candidates who favored judicial restraint. See, e.g., Nadine Cohodas, *Bush on Judiciary: The Signals Are Mixed*, CONG. Q. WKLY. REP., Nov. 26, 1988, at 3394 (discussing the Bush administration's focus on judicial restraint). Because a judge might see rejection of an administrative interpretation as "activism," it is possible that Bush judges deferred more often because they were restraint-oriented.

However, the fact that Reagan judges did not share the high acceptance rate of Bush judges suggests that the unusually high acceptance rate among Bush judges is best explained by chance. See Sheldon Goldman, *The Bush Imprint on the Judiciary: Carrying on a Tradition*, 74 JUDICATURE 294, 306 (1991) (arguing that Bush judges share the jurisprudential moorings of Reagan judges). In any event, it is important to note that despite the higher acceptance rate among Bush appointees, Bush judges found statutes to be ambiguous a tad *less* often than average (59% as compared to 62%).

opinion of textualism and its tendency to apply *Chevron*.¹⁹⁷ Judges from different interpretive schools appear equally likely to apply *Chevron*'s two-step framework, equally likely to find that a statute is ambiguous at step one, and equally likely to uphold agency interpretations of law.

Because textualism is almost uniquely associated with Reagan and Bush appointees,¹⁹⁸ the fact that the Reagan/Bush judiciary in the courts of appeals applied *Chevron*, found ambiguity, and deferred to agency interpretations at the same rate as other judges directly challenges the validity of the interpretive model. Textualism may or may not lead judges to defer in the correct set of cases (assuming there is such a set), but it seems fairly clear that, at least in the *Chevron* context in the courts of appeals, the textualism associated with Reagan and Bush judges has not introduced "cacophony and incoherence throughout the administrative state."¹⁹⁹

Proponents of the interpretive model might make two responses to my conclusion that the interpretive model is inaccurate, both of which present alternative visions of what it means to be a textualist. The first response questions the accuracy of the fit between presidential appointment and a judge's interpretive school. The argument is as follows: If we adopt a stricter notion of what it means to be a textualist, then we find that there are really only a handful of textualist circuit judges, such as Frank Easterbrook²⁰⁰ and Alex Kozinski.²⁰¹ Even if all of those textualist judges are Reagan and Bush appointees, their numbers are too small to affect the overall acceptance rates of the 105 active Reagan and Bush judges included in the study. Therefore, this study shows only that mainstream Reagan and Bush judges are not textualists—not that textualists do not defer differently from non-textualists.

This argument falters primarily because it fails to account for the centrality of the textualist orientation to the brand of conservative legal thought embraced by mainstream Reagan and Bush appointees. The judicial orientation of the Reagan and Bush administrations was defined by its opposition to the broad use of non-textual sources in constitutional and statutory interpretation; a central theme of the legal writings of the proponents of these administrations is that meaning *can* be found in text, whether the text is statutory or constitutional.²⁰² From this perspective, the law

197. Reagan judges and Clinton judges applied *Chevron* slightly more frequently than their numbers would predict; Bush and Carter judges applied it slightly less frequently. See *supra* p. 41 ch.9. These differences were small and almost certainly due to chance.

198. See *supra* notes 131-133 and accompanying text.

199. Pierce, *Hypertextualism*, *supra* note 6, at 752. Note that this result is consistent with those found by Professor Maggs. See Maggs, *supra* note 7, at 416-17.

200. See Easterbrook, *supra* note 133 (advocating a textualist mode of statutory interpretation).

201. See Kozinski, *supra* note 133 (endorsing the use of "plain meaning" in statutory interpretation).

202. See, e.g., CHARLES FRIED, ORDER AND LAW: ARGUING THE REAGAN REVOLUTION—A FIRSTHAND ACCOUNT 61-65 (1991) (noting that interest in originalism among legal conservatives in the Reagan era was fueled by originalism's promise to find definiteness in text); U.S. DEPT OF JUSTICE,

is found in the text, and a judge who looks beyond the text for meaning is an activist, because the judge is no longer interpreting the law but making it.²⁰³ Given the centralized screening process employed by the Reagan and Bush administrations to ensure that all judicial selections shared their judicial philosophy,²⁰⁴ it is unlikely that many Reagan or Bush appointees would fundamentally disagree with this viewpoint.²⁰⁵ Indeed, contemporary reports of the confirmation proceedings of Reagan and Bush judges noted that “literal interpretation of statutory . . . texts” was among the “bedrock conservative doctrine[s]” articulated by the nominees.²⁰⁶

Although this does not mean that all Reagan and Bush circuit judges are textualists in the mold of Justice Scalia, it does mean that Justice Scalia’s textualism is one particular instantiation of a belief in text that is shared by most Reagan and Bush circuit judges.²⁰⁷ If the interpretive model is correct in positing that judges who profess to find meaning in text will find ambiguity less frequently at *Chevron*’s step one, then Reagan and Bush judges should

USING AND MISUSING LEGISLATIVE HISTORY: A REEVALUATION OF THE STATUS OF LEGISLATIVE HISTORY IN STATUTORY INTERPRETATION (1989); E. Spencer Abraham, *Introduction: Twentieth Anniversary Volume*, 20 HARV. J.L. & PUB. POL’Y 1, 5 (1996) (“If we are to confine courts to their proper role, . . . [we] must revive the notion that there are intelligible legal principles that courts can follow, distinct from the judge’s will.”); Richard A. Epstein, *Some Doubts on Constitutional Indeterminacy*, 19 HARV. J.L. & PUB. POL’Y. 363, 365 (1996) (suggesting that indeterminacy is often used as a ploy “to avoid . . . embarrassing silence” when a judge or advocate “ha[s] nothing to say on the merits” as to why a certain result is legally correct); A. Raymond Randolph, *Dictionaries, Plain Meaning, and Context in Statutory Interpretation*, 17 HARV. J. L. & PUB. POL’Y. 71, 77 (1994) (D.C. circuit judge appointed by Bush notes that he finds “rigorous textual analysis” to be “essential” to statutory interpretation).

203. See Abraham, *supra* note 202, at 5; Graglia, *supra* note 135, at 107 (noting that a central claim of judicial conservatives is that “by engaging in statutory construction that ignores the plain meaning of the statutes, the courts substitute their own view of what the law should be for what the legislature has enacted”). This orientation is also revealed by President Reagan’s oft-repeated statement that “the proper role of the courts is to interpret the law, not make it.” Stephen Wermiel & Gerald F. Seib, *U.S. Judge Ginsburg Nominated by Reagan for Supreme Court Post*, WALL ST. J., Oct. 30, 1987, at A22.

204. See generally HERMAN SCHWARTZ, *PACKING THE COURTS: THE CONSERVATIVE CAMPAIGN TO REWRITE THE CONSTITUTION* (1988).

205. Chief Judge Richard A. Posner of the Seventh Circuit is a prominent exception. See, e.g., Richard A. Posner, *Legislation and Its Interpretation: A Primer*, 68 NEB. L. REV. 431 (1989) (advocating a “pragmatist” approach to statutory interpretation).

206. Pilon, *supra* note 135, at A10 (quoting Terrence Moran, *In His Own Image; Bush Judiciary: White, Male, and Cautious*, LEGAL TIMES, Dec. 31, 1990, at 1, available in LEXIS, News Library, *Arcnws File*).

207. The likely reason that this fact is under-appreciated is that the majority of circuit judges do not maintain a high profile in academic circles. Of those Reagan and Bush judges who do publish regularly in law reviews, however, most are identified as textualists. See Easterbrook, *supra* note 133; Kozinski, *supra* note 133. Furthermore, most of the Reagan and Bush appointees to the Supreme Court have at one time or another been labeled textualists. See, e.g., Pierce, *Hypertextualism*, *supra* note 6, at 751-52 (maintaining that Justices Scalia, Kennedy, and Thomas and Chief Justice Rehnquist are “hypertextualists”); Karkkainen, *supra* note 7, at 401 (arguing that Justices Scalia, Kennedy, and Thomas are adherents of textualism and that Justice O’Connor and Chief Justice Rehnquist frequently join textualist opinions). Most of these Justices are former Republican-appointed circuit judges. It would be remarkable if Presidents Reagan and Bush had selected several textualist judges for seats on the Supreme Court if in fact only a handful of their circuit court appointees could be construed as textualists.

find ambiguity less frequently than do other judges. No such trend exists, however, making it difficult to maintain the belief that textualist judges defer differently under *Chevron* from judges who are less committed to finding meaning in text.

The second argument that proponents of the interpretive model could make is that all circuit court judges will tend to see the same ambiguity in statutes because they take their marching orders from the Supreme Court. This view assumes that the approach to *Chevron* adopted by the Supreme Court will be dutifully followed by the circuit courts. For example, if the Supreme Court adopts a textualist interpretation of *Chevron*, then the circuit courts will do the same.²⁰⁸ Accordingly, the only school of interpretation that matters is that which is embraced by the Supreme Court, and the uniformity among different judges' approaches to *Chevron* shows only that the circuit court judges are faithfully following the Supreme Court's interpretive lead.

The problem with the second argument is that the circuit court judges included in this study found statutory meaning using a variety of interpretive methods. Some judges used legislative history to look for meaning;²⁰⁹ others did not.²¹⁰ Some judges approached step one as if it were a search for plain meaning,²¹¹ while others did not.²¹² This range of interpretive techniques could be explained by the fact that the *Chevron* case itself fails to specify what interpretive method to use at step one: The case punts, telling judges simply to use "traditional tools of statutory construction."²¹³ In any event, the absence of a dominant interpretive method makes it difficult to accept the proposition that the uniformity of results is explained by circuit court judges' blind obedience to the latest interpretive orders of the Supreme Court. In direct contravention of the central claim of the interpretive model, judges applied different interpretive methods but perceived the same overall degree of ambiguity at *Chevron*'s first step.

208. See *Pierce*, *Hypertextualism*, *supra* note 6, at 752 ("If the Court persists in its use of hypertextualism, lower courts will have no choice but to adopt that method of statutory interpretation.").

209. See, e.g., *Strickland v. Commissioner*, 96 F.3d 542, 547 (1st Cir. 1996) (Selya, J.) (examining legislative history to see whether Congress had spoken to the precise issue); *United Servs. Auto. Ass'n v. Perry*, 92 F.3d 295, 299 n.4 (5th Cir. 1996) (per curiam) (considering whether legislative history transformed a statute that was "ambiguous on its face, [into one that was] unambiguous in fact").

210. See, e.g., *Ethyl Corp. v. EPA*, 51 F.3d 1053, 1058 (D.C. Cir. 1995) (Edwards, C.J.) (examining the plain language of the statute at step one); *American Legion v. Derwinski*, 54 F.3d 789, 795 (D.C. Cir. 1995) (Rogers, J.) (examining only the plain language and the structure of the statute at *Chevron*'s first step).

211. See, e.g., *Cabell Huntington Hosp. Inc. v. Shalala*, 101 F.3d 984, 988 (4th Cir. 1996) (Wilkinson, C.J.) (finding that the "plain meaning" of the statute resolves the *Chevron* issue at step one); *Sidwell v. Express Container Servs.*, 71 F.3d 1134, 1138 (4th Cir. 1995) (Luttig, J.) (same) (citing WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 27 (1993)).

212. See, e.g., *Ramsey v. Kantor*, 96 F.3d 434, 442 (9th Cir. 1996) (Reinhardt, J.) (concluding that the *Chevron* analysis was resolved at step one because the administrative action was "clearly contemplated" by the statute).

213. *Chevron*, 467 U.S. at 843 n.9.

In light of the substantial body of scholarship that accepts the claims of the interpretive paradigm, it is surprising that its claims are unsupported by the data. The question arises: Where did proponents of the interpretive model go wrong?²¹⁴ I propose that the model errs by understating the degree to which theories of statutory interpretation are normative, rather than descriptive. I submit that jurists internalize interpretative norms based on their largely intuitive understandings of the proper role of the judiciary in a constitutional democracy, not on their personal answer to the hermeneutic question of how much meaning can be extracted from text. Roughly speaking, those judges who follow text more closely tend to profess a belief in a more limited, rule-following judiciary,²¹⁵ while those who endorse a more dynamic interpretative method tend to appreciate judicial rule-making power.²¹⁶ Whether a judge advocates or rejects textualism does not reflect the judge's capacity to find more or less meaning in text. Instead, it means that the judge believes that the body politic is better served by judges who try more or less hard to find what meaning may be there.

The reason that the interpretive model fails in the *Chevron* context, then, is that *Chevron* asks judges an interpretive question in a context that disrupts the usual relationship between the outcomes served and the political theories that typically inform judges' interpretive methods. *Chevron* upsets the usual

214. I focus here on the interpretive model's claim that textualist judges will more frequently see plain meaning at step one and defer less frequently than other judges. This is the claim most often made by the interpretive model's proponents, and it is also the most intuitively persuasive. As for the belief that textualist judges will too frequently see ambiguity and defer, one senses that it arises largely from the fact that super-textualist Justice Scalia was a longtime advocate of executive power. See, e.g., Schwartz, *supra* note 7 (maintaining that Justice Scalia will find ambiguity at step one because he is committed to agency power); Stock, *supra* note 7 (same). In addition, Professor Merrill's suggestion that textualist judges might be less inclined than other judges to apply *Chevron* is somewhat puzzling. Even assuming that textualist judges see plain meaning particularly often, why would textualist judges be less likely to apply *Chevron* at all? Would they not simply be more likely than other judges to resolve the analysis at step one, rather than at step two? In that case, Professor Merrill's claim collapses into the more common prediction (addressed in depth in the main text) that textualist judges will defer less frequently than other judges.

215. See, e.g., ANTONIN SCALIA, A MATTER OF INTERPRETATION 23 (1997) (asserting that the central belief underlying textualism is that judges have no authority to act in a legislative capacity); Abraham, *supra* note 202, at 5 (asserting that judges must feel bound by intelligible legal principles if they are to accept a more limited role); Frank H. Easterbrook, *Legal Interpretation and the Power of the Judiciary*, 7 HARV. J.L. & PUB. POL'Y 87 (1984) (arguing that judicial perception of ambiguity in statutes gives judges opportunities to enact their policy preferences into law); Easterbrook, *supra* note 133, at 63-64 (defending textualism on the grounds that it leads to clear rules and confines judges).

216. See, e.g., GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES 2 (1982) (suggesting that judges should have the authority to overrule obsolete statutes); RONALD DWORKIN, LAW'S EMPIRE 313 (1986) (asserting that a judge's view of statutory meaning should "depend on the [judge's perception of what is the] best answer to political questions"); WILLIAM N. ESKRIDGE, JR., DYNAMIC STATUTORY INTERPRETATION 50 (1994) (suggesting that a judge interpreting statutes should consider not only the text, but also the judge's belief as to which rules will best fulfill the needs and goals of present day society); Ronald Dworkin, *Law as Interpretation*, 60 TEX. L. REV. 527 (1982) (describing a mode of interpretation in which the judge is a "partner" with the legislature in creating a new and evolving meaning from text).

relationship between interpretation and the judicial role in two ways. First, statutory ambiguity no longer expands judicial power; it constricts it, limiting the judicial role to deferential review for unreasonableness. Conversely, finding meaning in the text no longer limits judicial power; it expands it by granting to the courts plenary review of administrative action. Second, *Chevron* transforms a judge's degree of commitment to the text from a means of allocating power between the legislature and the judiciary (its usual function) into a means of allocating power between the judiciary and the executive. Finding meaning in the text no longer enhances the power of the legislature over the judiciary; instead, it emphasizes the power of the judiciary over the executive. I propose that these disruptions of the typical association between interpretative method and the judicial role explains why judges do not approach *Chevron*'s first step with their usual interpretive associations intact. *Chevron*'s atypical interpretive context in effect suspends judges' normative associations between their approaches to text and political theory.

Consider the case of a judge who adopts an expansive view of the judicial function and believes that the proper judicial role is to ensure that the broad policy concerns of Congress are carried out in a fair and just way. Because textualism requires a judge to adhere to text instead of purpose and justice, the judge would likely eschew textualism and instead find that most texts were ambiguous enough to allow the judge to fashion a just remedy.²¹⁷ In the *Chevron* context, however, the ambiguity that would normally allow the judge to fashion a just remedy backfires. A finding of ambiguity instead binds the judge to accept a wide range of agency action, even if the judge perceives that action as unjust. Ambiguity ceases to be an engine of judicial authority and becomes an engine of uncabined executive power.

In the absence of the usual forces pulling and pushing judges toward different interpretive approaches, judges who typically are influenced by very different normative interpretive traditions adopt roughly equivalent understandings of how ambiguous is ambiguous enough at step one.²¹⁸ This does not mean that all judges will agree in every case, of course (although

217. Cf. FRANK, *supra* note 3, at 67-68 (discussing how legal indeterminacy expands judicial capacity to create proper remedies that can solve otherwise intractable problems).

218. Given how often panels of circuit judges hear *Chevron* questions, it seems natural that most judges would tend over time to come to a roughly similar understanding of how high a hurdle is posed by *Chevron*'s first step. The frequency with which judges apply *Chevron* is remarkable. In the D.C. Circuit, for example, circuit judges are likely to face several dozen *Chevron* issues a year. Even limiting ourselves to those *Chevron* applications that met the methodology requirements of this study and appeared in a published opinion, the D.C. Circuit decided 77 *Chevron* applications in 1995 and 1996. Ignoring en banc proceedings, which would also increase the number of judge-applications, that amounts to 115 judge-applications per year. There were 11 judges sitting during the time of the study, which means that D.C. Circuit judges each decide at least 10 or 11 *Chevron* issues per year in the subset of cases leading to published opinions.

most cases are unanimous), but it does mean that no one set of judges will be led to adopt a particularly different vision of *Chevron*.²¹⁹

Conclusion

Given the pervasive influence of legal realism, it should not be surprising that the *Chevron* doctrine has been a highly popular subject of critique among academic commentators.²²⁰ More than most recent doctrinal innovations, *Chevron* evokes a certain formalism. As construed by some, it is a more or less mechanical rule, designed to draw a sharp line between the sphere of “law” mandated by Congress and the sphere of “policy” delegated (by legal fiction) to agencies.²²¹ This is just the kind of legal doctrine that it was the mission of the realists to deconstruct.²²²

The surprising finding of this study is that *Chevron* withstands the challenges posed by the contextual, political, and interpretive critiques reasonably well in the courts of appeals. The political model mounts the most effective attack: The results of the study show that judges are more likely to defer to agency interpretations that support judges’ personal political preferences than they are to interpretations that oppose their personal political preferences. This inroad is modest, however, because there is no evidence that this likelihood is greater than that fostered under the doctrinal regime that *Chevron* replaced. Further, the remainder of the challenges posed by the three models are generally unsuccessful. Contrary to the claims of the contextual model, the presence of pre-*Chevron* traditional factors does not exert a strong influence on results; their impact is weak, and perhaps nonexistent. Contrary to the claims of the interpretive model, courts of appeals judges associated with textualist approaches to statutory interpretation do not apply *Chevron* or defer less frequently than others; in fact, judges from very different normative interpretive traditions appear to apply the doctrine in essentially the same way. Even the political model fails to describe the doctrine accurately at a structural level: Its claim that judges use the doctrine to uphold the executive decisions of “allies” and reject the decisions of “enemies” is unfounded.

219. Cf. Harry T. Edwards, *The Judicial Function and the Elusive Goal of Principled Decisionmaking*, 1991 WIS. L. REV. 837, 856 (“Perhaps the most fundamental reality of D.C. Circuit decisionmaking is that, contrary to popular belief, circuit judges rarely disagree with one another over the disposition of particular cases.”).

220. See, e.g., Joseph William Singer, Review Essay, *Legal Realism Now*, 76 CAL. L. REV. 465, 467 (1988) (“We are all legal realists now.”).

221. See, e.g., DAVIS & PIERCE, *supra* note 2, at 113 (“[P]olicy disputes within the scope of authority Congress has delegated an agency are to be resolved by agencies rather than by courts. Courts and agencies are instructed by *Chevron* to distinguish policy disputes from disputes with respect to issues of law by determining whether Congress resolved the dispute.”).

222. See MORTON J. HORWITZ, *THE TRANSFORMATION OF AMERICAN LAW 1870-1960*, at 170 (1992) (“[A]bove all, Realism is a continuation of the Progressive attack on the attempt of late-nineteenth-century Classical Legal Thought to create a sharp distinction between law and politics and to portray law as neutral, natural, and apolitical.”).

Although the three models have attempted to reveal the *Chevron* doctrine's true dynamics, this study suggests a reality somewhat out of step with their critical preconceptions. Oddly, the best guide for predicting judicial outcomes under *Chevron* is probably the test itself.