Empirical Research on Judicial Reasoning: Statutory Interpretation in Federal Tax Cases

Daniel M. Schneider

EMPIRICAL RESEARCH ON JUDICIAL REASONING: STATUTORY INTERPRETATION IN FEDERAL TAX CASES

DANIEL M. SCHNEIDER[*]

INTRODUCTION

Common wisdom and legal literature reveal how we think judges think. If you asked lawyers how judges decide cases, they would most likely respond with something about precedents, rules, reasoning, and arguments—notwithstanding what supporters of legal realism, critical legal studies, race theories, and feminism might say. But if you asked social scientists the same question, they would more likely inquire about the judge's socioeconomic background, education, or sex, assembling the data and sifting it for clues that predict behavior.

Literature is full of arguments about how we think judges interpret statutes. Approaches range from strict construction to nonliteralism, sauntering through pragmatism. All of these approaches resonate in tax. The methods are founded in theory grounded by analysis of discrete collections of cases. The importance of theory to statutory interpretation cannot be denied, but theory alone cannot explain how judges have actually interpreted the Internal Revenue Code (the Code).

There has been little empirical research in tax law about judges,[1] and none that explains how they interpret the Code.[2] In this article, I review a large number of

1. See, e.g., BERNARD WOLFMAN ET AL., DISSENT WITHOUT OPINION: THE BEHAVIOR OF JUSTICE WILLIAM O. DOUGLAS IN FEDERAL TAX CASES (1975) (reviewing Justice Douglas's federal tax decisions); James Edward Maule, *Instant Replay, Weak Teams, and Disputed Calls: An Empirical Study of Alleged Tax Court Judge Bias*, 66 TENN. L. REV. 351 (1999) (presenting original empirical research on alleged bias in the Tax Court); Beverly I. Moran & Daniel M. Schneider, *The Elephant and the Four Blind Men: The Burger Court and Its Federal Tax Decisions*, 39 HOW. L.J. 841 (1996) (reviewing the Burger Court's federal tax decisions).

2. While no one has yet empirically investigated how judges decide federal tax cases, there are some pieces of tax law literature that use empirical data. See, e.g., Wayne M. Gazur, *Do They Practice What We Teach?: A Survey of Practitioners and Estate Planning Professors*, 19 VA. TAX REV. 1 (1999) (surveying estate planning); Marjorie E. Kornhauser, *Love, Money, and the IRS: Family, Income-Sharing, and the Joint Tax Return*, 45 HASTINGS L.J. 63 (1993) (surveying couples' views regarding income sharing); Leandra Lederman, *Which Cases Go to Trial?: An Empirical Study of Predictors of Failure to Settle*, 49 CASE W. L. REV. 315 (1999) (presenting original empirical research regarding settlement of Tax Court cases); Martin J. McMahon, Jr. & Alice Abreu, *Winner-Take-All Markets: Easing the Case for Progressive Taxation*, 4 FLA. TAX REV. 1 (1998) (studying progressive taxation based on empirical research done by Congressional Budget Office); Beverly I. Moran & William Whitford, *A Black Critique of the Internal Revenue Code*, 1996 WIS. L. REV. 751 (questioning racial bias of Code based on empirical research done by Bureau of Census and other government agencies); Philip F. Postlewaite, *Life After Tenure: Where Have All the Articles Gone?*, 48 J. LEGAL EDUC. 558 (1998) (presenting original empirical research regarding publication by tenured tax law professors); Daniel M. Schneider, *Interpreting the Interpreters: Assessing Forty-Five Years of Tax Literature*, 4 FLA. TAX REV. 483 (1999) (giving original empirical research regarding tax literature); William J. Turnier et al., *Redistributive Justice and Cultural Feminism*, 45 AM. U. L. REV. 1275 (1996) (analyzing redistributive justice in tax by gender); see also works cited supra note 1.

The need for more empirical research in tax law has been noted. See, e.g., Michael Livingston, *Reinventing Tax Scholarship: Lawyers, Economists, and the Role of the Legal Academy*, 83 CORNELL L. REV. 365 (1998).

For examples of some more traditional articles about judges' decisions, see Stephen B. Cohen, *Thurgood Marshall: Tax Lawyer*, 80 GEO. L.J. 2011 (1992) (reviewing tax decisions of Justice Marshall) and Marvin A.

federal tax cases using social science methods to answer two questions not asked by
more traditional legal scholars:
- Have courts favored one method of interpreting the Code over others?
- What personal characteristics prompt a judge to use a particular method?

Tools of the social sciences lend themselves to analyzing large numbers of
decisions.[3] While social science research sometimes concedes the importance a
judge may assign to the legal precedent upon which lawyers rely in making a
decision, it also grapples with how a judge's socioeconomic characteristics—*e.g.*,
gender, education, politics—can influence her decisions.[4] This article draws that
research into federal tax law. Using a database comprised of a large number of Tax
Court and federal district court cases decided in three cities over a twenty-year
period, the article describes the methods of statutory interpretation judges employed
to justify their decisions and the effect judges' social backgrounds had on the
methods they used.

This empirical research differs from theoretical observations about interpreting
the Code. Instead of suggesting that particular methods were applied in isolation—a
common posture of scholars—the article establishes that judges were broad minded,
applying all methods and frequently combining them in a way that defies scholarly
purity. It also establishes that some social background characteristics may help to
explain why judges use certain methods.

This article is divided into four parts. First, it examines law review articles
discussing interpretation of the Internal Revenue Code and social science research.
Next, the article sets forth the methodology used in the research, discussing the
selection of cases and variables. Third, it describes and analyzes the empirical data.
Finally, it sets forth conclusions that may be drawn from the data.

I. LITERATURE

A. Interpreting the Internal Revenue Code

The broader debate about statutory construction, as well as the debate within the
field of tax law, can be reduced to those who favor a more literal method of
interpretation and those who do not.[5] Five methods of interpretation that are

---

Chirelstein, *Learned Hand's Contribution to the Law of Tax Avoidance*, 77 YALE L.J. 440 (1968) (reviewing Judge
Hand's tax decisions).

3. *See, e.g.*, James J. Brudney et al., *Judicial Hostility Toward Labor Unions? Applying the Social
Background Model to a Celebrated Concern*, 60 OHIO ST. L.J. 1675 (1999) (analyzing labor law decisions using
social sciences tools); Gregory C. Sisk et al., *Charting the Influences on the Judicial Mind: An Empirical Study of
Judicial Reasoning*, 73 N.Y.U. L. REV. 1377 (1998) (analyzing decisions about federal sentencing standards using
social sciences tools).

4. *See, e.g.*, HAROLD J. SPAETH & JEFFREY A. SEGAL, MAJORITY RULE OR MINORITY WILL: ADHERENCE
TO PRECEDENT ON THE U.S. SUPREME COURT (1999); HAROLD J. SPAETH & JEFFREY A. SEGAL, THE SUPREME
COURT AND THE ATTITUDINAL MODEL (1992); Howard Gillman, *What's Law Got to Do with It? Judicial
Behaviorists Test the "Legal Model" of Judicial Decision Making*, 26 LAW & SOC. INQUIRY 465 (2001) (examining
the behaviorist approach). Brudney et al., *supra* note 3, and Sisk et al., *supra* note 3, were among the first legal
scholars to conduct empirical research on judicial reasoning using logistic regression. *But cf.* Orley Ashenfelter et
al., *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257
(1995) (analyzing everyday, published, district court decisions, but using means, not logistic regression).

5. No particular theory for interpreting the Code is promoted here. Grounding the interpretation in what
judges have done is important, so I have tried to refer readers as much as possible to other scholars' legal research

predominantly reflected in scholarly literature also appear in the sampled cases. An overview of these methods follows.

1. Strict Construction

One approach to interpreting the Internal Revenue Code is to take the words of the Code literally. While his article, *Text as Limit: A Plea for a Decent Respect for the Tax Code*,[6] does not argue for a strictly literal approach to interpretation, Professor John Coverdale does suggest that the Code should not be read in an "antitextual" manner. He would not have "provisions of the Code [read] in other ways that fall outside the range of meanings that their text, taken in context, has in ordinary speech or in other provisions of the Code."[7] In other words, he perceives antitextual interpretations as those "that provoke us to think, 'If that is what Congress meant, it should have said so rather than saying something very different.'"[8] While the article does not take up whether "the statutory language so clearly compels a particular result that it is appropriate to ignore other sources of meaning,"[9] it does give such great weight to the text that there is little room for nonliteral interpretations.[10]

According to Coverdale, reliance upon the text of a statute requires respect for the integrity of our constitutional values, which give Congress, not the courts, the primary policy-making role.[11] Coverdale also suggests that tax is different than other substantive areas of law, and the highly formalistic rules Congress has laid down in tax, as opposed to vaguer standards used elsewhere, dictate greater deference to the Code's language.[12] As he writes,

> [t]he prohibition against antitextual interpretations of the Code...[for which I argue] could be defended on grounds that legislative intent and purpose are either meaningless concepts or illegitimate considerations not to be considered

---

about interpreting tax statutes for different theories of construction. I believe that the difficulty of understanding empirical research often impedes the use of such research by tax lawyers, so I have also attempted to lead readers to other scholars' legal research about the impact of judges' social backgrounds on the outcomes of their decisions in order to minimize technical discussion in this article.

    For example, this article uses logistic regression in order to explore the relationship between methods of statutory construction and social background factors because that is the most appropriate method for analyzing the relationship between dichotomous dependent variables (*e.g.*, method of construction) and several categorical or ordinal variables (*e.g.*, gender, politics, education). Readers interested in learning more about logistic regression, as well as less law-oriented research, are referred to discussion in this article, *infra* Part III.C. *See also* Brudney et al., *supra* note 3, at 1681-92; Sisk et al., *supra* note 3, at 1385-95; Lederman, *supra* note 2.

    6. John F. Coverdale, *Text as Limit: A Plea for a Decent Respect for the Tax Code*, 71 TUL. L. REV. 1501 (1997); *see also* Joseph Isenbergh, *Musings on Form and Substance in Taxation*, 49 U. CHI. L. REV. 859 (1982) (arguing against risks of nonliteral interpretation); *cf.* Mary L. Heen, *Plain Meaning, the Tax Code, and Doctrinal Incoherence*, 48 HASTINGS L.J. 771 (1997) (discussing the Supreme Court's increasing reliance on the "plain meaning" of a statute when deciding tax cases).

    7. Coverdale, *supra* note 6, at 1503 (footnote omitted).

    8. *Id.* at 1504.

    9. *Id.* at 1505.

    10. *Id.*

    11. *Id.* at 1515-16.

    12. *Id.* at 1521-25. *But see* Michael Livingston, *Practical Reason, "Purposivism," and the Interpretation of Tax Statutes*, 51 TAX L. REV. 677, 683-87 (1996) (noting that tax is not different than other areas of law); *cf.* John A. Miller, *Indeterminacy, Complexity, and Fairness: Justifying Rule Simplification in the Law of Taxation*, 68 WASH. L. REV. 1 (1993) (arguing that tax statutes should rely more on standards than on rules).

by a court in interpreting statutes. It does not, however, require either of these positions. The argument for avoiding antitextual results rests comfortably on the much narrower proposition that, in the search for congressional intent and statutory purpose, the enacted language of the statute should be the primary guide.[13]

On the opposite end of the spectrum from literal interpretations are nonliteral ones, casually used by tax lawyers—when they rely on a section's legislative history or on a regulation, for example—without thought, perhaps, other than how to best serve the client. After all, if the meaning of a statute were clear, why interpret it? (And, if the statute did not require interpretation, why hire a lawyer?) Many commentators argue for nonliteral interpretations of the Code, for various reasons, using one of the four approaches noted below.

2. Regulations

One method for interpreting a statute is to defer to regulations promulgated by the Treasury Department.[14] *Chevron, USA, Inc. v. Natural Resources Defense Council, Inc.*,[15] provides a recent anchor in that sea. *Chevron* marks the Supreme Court's deference to agency regulations in light of ambiguous statutes, holding that the regulation controls if Congress has not "directly spoken to the precise question at issue...[and if] the agency's answer...[is] based on a permissible construction of the statute."[16] Although decided in the environmental law area, *Chevron*'s application to other highly regulated areas, like tax, seems obvious.[17] Implicit in *Chevron* is a theoretical deference to Congress and then to the agency.[18]

3. Structure

Deference also may be given to the purpose of a statute as part of the structure of the larger Code in which it is situated. In *Interpreting Tax Legislation: The Role of Purpose*,[19] Professor Deborah Geier suggests that there is a "theoretical construct that overarches the sum total of the entire Internal Revenue Code and is intended to be captured by it...and that statutory structure could come within the umbrella of

---

13. Coverdale, *supra* note 6, at 1518-19.
14. *See generally* Michael Asimow, *Public Participation in the Adoption of Temporary Tax Regulations*, 44 TAX LAW. 343 (1991) (reviewing tax regulations); Moran & Schneider, *supra* note 1, at 907-27 (discussing the Burger Court's record on reliance on regulations in its tax decisions); Bernard Wolfman, *Note: Supreme Court Decisions in Taxation: 1980 Term*, 35 TAX LAW. 443 (1982) (reviewing use of regulations by Supreme Court).
15. 467 U.S. 837 (1984).
16. *Id.* at 842-43.
17. *Chevron* has never garnered much attention in tax cases. *See* Heen, *supra* note 6, at 784-86. Heen cites, among others, Ellen P. Aprill, *Muffled Chevron: Judicial Review of Tax Regulations*, 3 FLA. TAX. REV. 51 (1996) (discussing *Chevron*); Linda Galler, *Emerging Standards for Judicial Review of IRS Revenue Rulings*, 72 B.U. L. REV. 841 (1992) (same); Samuel B. Sterrett, *Suggested Approach for Judicial Interpretation of Regulations that Grant Discretion to Taxpayers*, 12 VA. TAX REV. 477 (1993) (discussing post-*Chevron* Supreme Court decisions). *Id.*
18. *See* Heen, *supra* note 6, at 781.
19. Deborah A. Geier, *Interpreting Tax Legislation: The Role of Purpose*, 2 FLA. TAX REV. 492 (1995).

statutory purpose."[20] For Geier, tax statutes should be interpreted consistently with that structure.[21] The Supreme Court has done precisely that in cases like *Crane v. Commissioner*[22] and *Tufts v. Commissioner*,[23] which analyze the income tax consequences of nonrecourse debt, *Corn Products Refining Co. v. Commissioner*[24] and *Arkansas Best Corp. v. Commissioner*,[25] which assess the meaning of a capital asset, and *Bob Jones University v. United States*,[26] which addresses the propriety of allowing charitable deductions for contributions to racially discriminatory colleges.[27] In each area, the Court fit these cases into a broader statutory purpose. To Geier, "purposivism" is "a joint effort between Congress and the courts...[in which] Congress's law...includes that larger statutory structure."[28]

4. Legislative History

In *Congress, the Courts, and the Code: Legislative History and the Interpretation of Tax Statutes*,[29] Professor Michael Livingston argues for the use of legislative history. To Livingston, "[t]he argument in favor of using legislative history, simply put, is that context is important to the ascertainment of meaning, and legislative history is an important part of context."[30] He details the important role congressional committee staffs play in writing tax legislation, running from the initial conceptual presentation of proposals to committee members through drafting committee reports.[31] Livingston also notes that while many cases use legislative history in conjunction with some other approach to statutory construction, others use it as an independent ground for their decisions.[32]

> The best way to view tax legislative history may be in institutional terms, as part of the evolving relationship between Congress, the courts, and administrative agencies in the making of tax law and policy. This approach treats legislative history not as an alternative to "literal" interpretation, but as one of several sources of contextual authority for courts deciding tax cases. It seeks to evaluate legislative history not (or not only) by assessing whether it reflects congressional intent, but by analyzing the function performed by the legislative history and the persuasiveness of legislative history in performing that function.[33]

---

20. *Id.* at 497, 502; *see also* Moran & Schneider, *supra* note 1, at 928-942 (speaking about the "deep structure" of the Code).

21. *Id.* at 493.

22. 331 U.S. 1 (1947).

23. 461 U.S. 300 (1983).

24. 350 U.S. 46 (1955).

25. 485 U.S. 212 (1988).

26. 461 U.S. 574 (1983).

27. These cases are drawn from Livingston, *supra* note 12, at 689-701.

28. Geier, *supra* note 19, at 508 (citation omitted). Geier gives simple examples of respect for the purpose of a statute, including deducing Congress's desire to avoid assignment of income from its imposition of progressive tax rates and its disallowance of an interest deduction to a taxpayer who has borrowed without a profit motive because the purpose of business and investment deductions requires such a motive. *Id.* at 500-01.

29. Michael Livingston, *Congress, the Courts, and the Code: Legislative History and the Interpretation of Tax Statutes*, 69 TEX. L. REV. 819 (1991).

30. *Id.* at 845 (footnote omitted).

31. *Id.* at 832-44.

32. *Id.* at 849-71.

33. *Id.* at 873.

Thus, Livingston is less worried about the influence of committee staff members in writing tax laws than he is about Congress exceeding its institutional competence where committee staffs draft extensive legislative histories.[34] The process notwithstanding, Livingston does not question the propriety of using legislative history.

5. Practical Reasoning

In *Practical Reason, "Purposivism," and the Interpretation of Tax Statutes*, Professor Livingston argues for "practical reasoning" when interpreting the Code.[35] He believes that difficulties abound with nonliteral interpretations that depend on statutory purpose, including the inability to construct a solitary purpose out of the minds of multiple legislators and the possibility that a statute might contain multiple, and conflicting, purposes.[36] To him, "purposivism" is a "questionable basis" for deciding the types of cases that reflect Geier's sense of the Code's structure.[37] Instead, Livingston's

> approach would consider statutory text, legislative history, and evolutive considerations—including judicial and administrative precedents and applicable current values—together with the consequences of alternate interpretations and the court's own policy sense. While these sources typically would be considered in descending order of priority, the precise mix of sources would depend on the nature of the provision in question and the facts of the case at hand. A practical reason approach provides both the right method of deciding tax cases and an accurate description of what courts actually do in such cases.[38]

B. Judges' Social Backgrounds

The legal profession's belief in the importance of doctrinal legal reasoning in judges' decision-making processes conflicts with social science's view that a judge's social background can influence her behavior. Establishing links between social background factors and judicial outcomes can diminish the relevance of more traditional doctrinal legal theories. After all, if a judge's elite law school education leads her to decide in favor of the Internal Revenue Service (the Service), what difference do the facts of a case make?[39]

Two recent law review articles survey the literature discussed above and illustrate the current state of research on the influence of judges' social backgrounds in areas outside of tax.[40] Both also review methodology, with an eye toward an audience of lawyers.

The more recent of the articles—Professors James J. Brudney and Deborah J. Merritt, and Ms. Sara Schiavoni's *Judicial Hostility Toward Labor Unions?*

---

34. *Id.* at 874-78.
35. *See* Livingston, *supra* note 12.
36. *See id.* at 681 (relying on William N. Eskridge, Jr. & Philip P. Frickey, *Statutory Interpretation as Practical Reasoning*, 42 STAN. L. REV. 321 (1990)).
37. *See id.* at 690-708.
38. *Id.* at 720.
39. *See* Brudney et al., *supra* note 3, at 1681.
40. *See id.* at 1682-92; Sisk et al., *supra* note 3, at 1385-96.

Applying the Social Background Model to a Celebrated Concern[41]—measures federal appellate decisions about adjudications of the National Labor Relations Board (NLRB) against judges' social background criteria. The article explores whether judges on the federal courts of appeals supported unions or management in their NLRB decisions and what factors might have influenced those decisions. As the article notes almost immediately,

> [e]mpirical research into judicial behavior acknowledges the relevance of case-specific facts and legal precedent, but posits that judges' personal traits, their education, training, and their pre-judicial activities can help explain court decisions. In particular, a number of social scientists contend that pre-court life experiences play a prominent role in shaping the personal values and policy preferences of judges, and that such biographical factors can be useful in predicting judicial decisions.[42]

An obvious criticism of attempting to link judges' social backgrounds and their decisions is that it undervalues "legal doctrine, in particular [by] failing to appreciate how judges develop that doctrine primarily through reasoned elaboration of language and precedent in written decisions, not through subconscious infiltration of life experiences."[43] The authors' response is to demonstrate how, empirically, the connections can be made between background and outcome. Stating the authors' conclusions does not give full breath to their article, but it nevertheless helps to understand the associations they drew between the social backgrounds of the judges who decided appellate NLRB cases and how the judges acted.

> [O]ur analysis has identified numerous personal, political and professional background factors that are significantly associated with a judge's propensity to support or reject the union's legal position. Republican women, Democratic men, and Democratic women are about ten percent more likely than Republican men to vote for the union....Among all appeals from Board decisions favoring the employer, judges with...[management-side labor] experience are more than twice as likely as other judges to adopt the union's position and reverse the Board. In the category of divisive issues, some of the marginal effects are even larger, with Republican women four times more likely than other judges to adopt the union's position and judges with elected office experience almost twice as likely to do so. These findings strongly suggest that social background factors play a meaningful role in influencing judicial approaches to labor law issues.[44]

The second article, Professors Gregory C. Sisk, Michael Heise, and Andrew P. Morriss's *Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning,*[45] examines district court sentencing decisions. That article took the opportunity presented by enactment of the Sentencing Reform Act, which took effect in 1988, to determine how different courts reacted to the measures imposed

---

41. Brudney et al., *supra* note 3.
42. *Id.* at 1682 (footnotes omitted).
43. *Id.* at 1682 (footnotes omitted).
44. *Id.* at 1761.
45. Sisk et al., *supra* note 3.

on them.[46] The Sentencing Reform Act "established mandatory sentencing guidelines based on the offense and the characteristics of the crime and the offender."[47] The authors' concern that research frequently fails to deal with "incomparability"—the extent to which empirical research into judicial decisions might be criticized due to "differences in parties, time period, issues, and facts"—is answered by review of a single issue: the constitutionality of the Act's guidelines.[48]

The article observes,

> Our findings provide greater support to the behavioral model of judicial decisionmaking than we anticipated. While most of the social background variables we explored proved insignificant, some striking findings emerged that were consistent with a sociological or social construction model of decisionmaking, particularly with respect to the prior employment variable. For example, prior experience as a criminal defense lawyer was significant under several formulations of our dependent variables as an explanatory variable for opposition to the [guidelines]. On the other hand, prior experience as a state or local judge was related to upholding the [g]uidelines as constitutionally valid.[49]

Both of these articles use logistic regression when determining whether background factors could be associated with judges' decisions.[50]

II. METHODOLOGY

A. *Data and Variables*

1. Case Selection

I compiled a database consisting of all official Tax Court decisions rendered between 1979 and 1998 and all published federal tax decisions of federal district courts located in three major cities—Los Angeles, Chicago, and a portion of New York City—for the same time period.[51] I then sampled about fifteen percent of these cases. I thought that these four courts would give me a broad base of decisions, and by examining district court decisions as well as those of the Tax Court, I could

---

46. Sentencing Reform Act of 1984, Pub. L. No. 98-473, 98 Stat. 1837, 1987-2034 (codified as amended in scattered sections of 18, 28 U.S.C.).

47. Sisk et al., *supra* note 3, at 1381 (footnote omitted).

48. *Id.* at 1380-82. Sisk and his coauthors further discuss the benefit of looking at one case from the point of view of exploring judicial reasoning. *Id.* at 1410-11.

49. *Id.* at 1383.

50. Logistic regression is the appropriate multivariate technique for our analyses because our dependent variable (a vote for or against the union) is dichotomous....[T]he logistic regression method models the probability of an event—how likely the event is to occur. Thus, in our analysis we investigate judicial attributes that may impact the probability that a judge will vote for or against the union position.

Brudney et al., *supra* note 3, at 1708-09 n.119; *accord* Sisk et al., *supra* note 3, at 1431 n.230. *But see* Ashenfelter et al., *supra* note 4, at 281 (deciding not to use logistic regression and determining that social background factors did not influence judges' decisions).

51. The district court decisions were actually those rendered by judges in the Central District of California, the Northern District of Illinois, and the Southern District of New York. Other district courts also are located in these three metropolitan areas. For an examination of district court decisions, see Sisk et al., *supra* note 3, at 1415-16, and Ashenfelter et al., *supra* note 4, at 263-65.

ascertain differences, if any, that existed between judges in the two different tribunals. This database was comprised of 488 cases.[52]

A Tax Court judge sometimes adopted the decision of a special trial judge or a district court judge adopted the decision of a magistrate. I excluded the six decisions rendered by a special trial judge or magistrate that had not been adopted by another, more senior judge. On the other hand, I treated the adopting judge as having rendered the decision whenever adoption occurred.[53] After excluding the decisions that were not adopted, a total of 482 cases remained in the database—346 Tax Court decisions and 136 district court decisions.[54]

In addition to the exclusions noted above, I sometimes pared the fifty-one summary decisions from these 482 cases in order to examine only substantive cases. This facilitated analysis of the approach used in each case. There were 431 substantive cases.

Some authors draw distinctions between all decisions rendered by a court and the smaller portion of these decisions that are published. To some extent, the bias implicit in examining only cases that are selected for publication is avoided by this article's review of official Tax Court decisions, all of which are published. In contrast, the district court decisions in this database, all of which have been published, are part of the larger pool of all district court decisions—including unpublished ones. Finite resources dictated this more limited review. Nonetheless, I was interested in describing decision making in more important federal tax cases, and for that purpose this database was the most appropriate.[55]

2. Variables

I culled data about judges and their social backgrounds that reflected contemporary social issues, such as whether judges view legal questions differently

---

52. For each court, I randomly picked one of the first seven cases published by it in a particular year and then read every seventh case thereafter. For example, I might have picked the third case in the Tax Court Reports for 1986 and read every seventh case thereafter, or the fifth tax decision by the Central District of California for 1993 and every seventh case thereafter. Reading every seventh case enabled me to sample just less than fifteen percent of the database (1/7=14.28%).

Establishing the base for district court decisions was somewhat more difficult, but a LEXIS search ultimately proved most successful. That search, done during the latter half of 1999, used the following terms: "court (central district or central dist! or c.d. pre/5 california or cal.) and date is 1997." The court was varied to include the two other district courts and the other nineteen years. I did not include a case resulting from the search—even if it was the next, seventh case—if it was the decision of a bankruptcy judge, but I did include decisions about taxes that involved bankruptcy. Whenever I skipped the seventh case, I sampled the next available case.

53. *But cf.* Brudney et al., *supra* note 3, at 1696 n.8 (excluding cases that were summary affirmances).

54. The 346 Tax Court cases were decided by forty-two judges; the fewest number of opinions written by any Tax Court judge was one, and the greatest number was twenty-six. The 136 district court cases were decided by seventy-six judges; the fewest number of opinions written by any district court judge was one and the greatest number was six.

55. *Compare* Brudney et al., *supra* note 3 (examining all opinions), *and* Sisk et al., *supra* note 3 (same), *with* Ashenfelter et al., *supra* note 4 (examining only published opinions). *See generally* Elizabeth M. Horton, *Selective Publication and the Authority of Precedent in the United States Courts of Appeals*, 42 UCLA L. REV. 1691 (1995) (summarizing arguments regarding selective publication of cases by federal appellate courts). It might be noted that the only Tax Court cases excluded from this database are the memorandum decisions, which carry less weight than official decisions. *See* 1 GERALD A. KAFKA & RITA A. CAVANAGH, LITIGATION OF FEDERAL CIVIL TAX CONTROVERSIES ¶ 2.06 (2nd ed. 1997) (describing the hierarchy of Tax Court decisions and the differences between official and other Tax Court decisions).

if they are women, African-American, or the beneficiaries of elite educations. I also collected data about the cases, especially about the methods of statutory construction used by judges that could be influenced by social background.[56] These variables include the following:

I.   Data about the *judge.*[57]
     A.   The judge's *name.*
     B.   The judge's *sex.*[58]
     C.   The judges' *race or ethnicity* (district courts only).[59]
     D.   The judge's *educational background.*
          1.   How elite was the *college* the judge attended?[60]
          2.   Did the judge go to one of eighteen elite *law schools?*[61]

---

56.   For example, I thought that the eliteness of a law school education might make a judge less likely to use practical reasoning, so I collected data about law schools and about methods of interpretation judges used to explain their decisions.

57.   I used biographies of Tax Court judges set forth in the *Congressional Directory* and biographies of district court judges made available by the Federal Judicial Center. Since February 2000, these biographies have been published at the Center's website at http://air.fjc.gov/history.

58.   Gender of a judge is set forth in the Federal Judicial Center's information, but not in the *Congressional Directory.* Nevertheless, determining the gender of the Tax Court judges was not problematic.
Coding of the data is important for logistic regressions. Therefore, the coding of variables is given in the following notes where appropriate. For example, the coding for gender was male=0 and female=1.

59.   Racial and ethnic information is not available for Tax Court judges, either in the *Congressional Directory* or from the Court. Of the judges in the 136 district court cases, 120 were white, ten were African-American, four were Asian, and two were Latino. I categorized each ethnicity separately. *But cf.* Brudney et al., *supra* note 3, at 1702-03 (discussing their groupings, which combined Latino and Asian judges and treated African-American judges separately, and cautioning readers about reaching conclusions, given their small numbers of nonwhite judges).
     The coding for race was white=0, nonwhite=1.

60.   I used values created by ALEXANDER W. ASTIN, WHO GOES WHERE TO COLLEGE? 58-83 (1965). Astin estimated the relative selectivity of an undergraduate institution by dividing the "highly able students who" wanted "to enroll at the college...by the number of freshmen admitted." *Id.* at 55. A judge for whom no undergraduate institution was reported, either because she had not attended college or did not include it in her biography, was assigned a score of zero. *Compare* Brudney et al., *supra* note 3, at 1703 (using Astin's measurements), *with* Sisk et al., *supra* note 3, at 1417-29 (omitting undergraduate degree as a variable). Another method for determining the eliteness of an undergraduate institution distinguishes between public and private institutions and between Ivy League and other schools. SHELDON GOLDMAN, PICKING FEDERAL JUDGES Ch. 9 (1997).
     While it might appear that Astin's method is dated, it can also be argued that (1) these criteria accurately reflect those perceptions present when many of the judges went to college, and (2) perceptions about the eliteness of schools change slowly. Brudney et al., *supra* note 3, at 1703 n.104 (suggesting that the reputations of undergraduate institutions, like law schools, change slowly and that an older list may be more appropriate because many of the judges went to college when Astin's study was published). The latter point is even more striking in my database than in Brudney and company's because mine includes decisions as early as 1979, while theirs included decisions no earlier than 1986.

61.   The 1977 *Cartter Report* suggests that fifteen law schools are elite. These schools are Harvard University, Yale University, University of California at Berkeley (Hastings), University of Chicago, Columbia University, Cornell University, Duke University, University of Michigan, New York University, Northwestern University, University of Pennsylvania, Stanford University, University of Texas, University of California at Los Angeles, and University of Virginia. *The Cartter Report on the Leading Schools of Education, Law, and Business,* CHANGE, Feb. 1977, at 44, 46. Another study would substitute three other schools for three of the fifteen *Cartter Report* schools. *See* Diana Fossum, *Law Professor: A Profile of the Teaching Branch of the Profession,* 1980 AM. B. FOUND. RES. J. 501, 507 (1980) (substituting Georgetown, Iowa, and Wisconsin for Cornell, Duke, and UCLA); *see also* Sisk et al., *supra* note 3, at 1418-19 (providing authors' own list of seven schools); Brudney et al., *supra* note 3, at 1703-04 (using the *Cartter Report* without Fossum's changes and noting the slow pace at which law school reputations have changed). I treat all eighteen of the schools discussed above as elite and the rest as non-elite. The coding for law schools was non-elite law school=0, elite law school=1.

E. The judge's *primary professional experience*. Had a judge primarily had a career in private practice, in government (such as a prosecutor or on a legislative committee's staff), as a judge, or as a teacher?[62]

F. The *political party of the president* who appointed the judge.[63]

G. Data related to the judge's *judicial* career (length of service on the bench at the time the case was decided).[64]

II. Data about the decision.

A. The *court* in which the case was decided.

B. The *year* of the decision.

C. The *area* covered by the decision (*e.g.*, if the case involved accounting generally, or the accrual method of accounting, and the primary Internal Revenue Code section under scrutiny, *e.g.*, §§ 446, 451 or 461).[65]

---

62. I treated whatever field the judge served the greatest number of years in as the judge's primary professional experience, regardless of when this time was served. While it might be argued that the three years in a state's attorney office served immediately before assuming the bench is more important than the ten years before that which a judge spent in private practice, I was unwilling and unable to draw that distinction. Most judges' careers were fairly clear, for example, where one spent most of his or her career at one or a few firms, so that he or she could be characterized as having primary professional experience in private practice. Thus, characterizing the primary professional experience was rarely a problem. Furthermore, I grouped careers such as work with a corporation, as a public defender, and with a nonprofit institution as private practice. Again, most judges' careers were fairly obvious, and they tended not to have worked for a corporation or in a public defenders office, so this variable was not problematic.

Other authors have drawn finer distinctions, especially in judges' other government work. *See, e.g.*, Brudney et al., *supra* note 3, at 1704; Maule, *supra* note 1, at 407-13; Sisk et al., *supra* note 3, at 1420. I did not, in part because Tax Court judges had substantially different experiences than district court judges (*e.g.*, work for the Internal Revenue Service), and simply categorizing for prior government work best accounted for these differences. Of the total number of judges, 279 judges came from private practice, 181 from government service, and twenty-two from both areas equally, or teaching, or being a judge. I coded background in a dichotomous manner, with government, equal service in private practice and government, teaching or being a judge=0 and private practice=1.

63. Presidents who had appointed district judges were listed in the Federal Judicial Center's biographies. *See supra* note 57. Presidents who had appointed Tax Court judges were not uniformly listed in the *Congressional Directory, supra* note 57, but I could inevitably interpolate an answer (*e.g.*, a judge appointed in 1985 must have been appointed by President Reagan).

The political views of the judge herself would be an even better measure for determining the effect of politics on decisions, but this information was not readily, and certainly not uniformly, available. Instead, the president's politics have been used as the measure of the judge's politics. Professor Goldman suggests that presidents from Eisenhower to Reagan appointed federal appellate judges from their own parties over ninety percent of the time. GOLDMAN, *supra* note 60, at 355-57. Only Carter, a Democrat, appointed fewer appellate judges who were Democrats, and even his percentage of Democratic judges was over eighty-two percent. *Id.* at 355; *see also* ROBERT A. CARP & RONALD STIDHAM, JUDICIAL PROCESS IN AMERICA 237-41 (4th ed. 1998) (supporting the assertion that presidents have followed their political ideologies when appointing federal judges). The coding for political party of the appointing president was Republican president=0, Democratic president=1.

64. I also gathered data about a judge's age when she rendered the decision in the database. Age has been used in other studies. *See* Brudney et al., *supra* note 3, at 1702 n.98. But seniority seemed to me to be a more intuitively meaningful, influential factor than age. *See* Sisk et al., *supra* note 3, at 1459-60, 1486-87 (using seniority).

65. The cases were later catalogued by section, not by area, because the former was easier to assemble. For discussion of the cataloging by section, see *infra* Part III.A.3.

A question raised by this and other variables is the accuracy of the cataloging: Has the case been accurately described as an accounting case or as a strict construction case? *See* Sisk et al., *supra* note 3, at 1410-12, 1434 (noting the difficulty of assessing opinions and safeguards employed to minimize mischaracterization). I

D. The case's *rationale*.[66] Possible rationales included:
1. Was it a *summary* decision? If so, no rationale was noted. A summary decision was one in which the judge discussed and probably even cited no cases, except in the briefest of manners.[67]
2. Was it some other, fuller decision? If so, on which rationale was the decision justified?
   a. *Strict construction* of the statute.[68]
   b. *Deference to the Internal Revenue Service*, either to its
      i. *regulations*,[69] or
      ii. *another Service proclamation*, such as a revenue ruling or revenue procedure.[70]
   c. Reliance on the *structure* or purpose of the Code.[71]
   d. Reliance on the *legislative history* of the section under scrutiny.[72]
   e. *Practical reasoning*. Inevitably, the courts relied on decisions rendered by other courts, so a fair amount of stare decisis was also present. But practical reasoning also makes sense in that the courts subscribing to this approach rendered opinions that were relatively unstructured and practical in their approach, as opposed, for example, to a court that relied solely on regulations or on legislative history.[73]
3. The *primary* and, if appropriate, the *secondary* and *tertiary* rationales used to explain the decision.[74] Every substantive, nonsummary, decision had at least one rationale. Some cases had more than one rationale.
4. *Who won*.[75]

---

reduced the risk of coding bias by entering my research without preconceived ideas about the outcome of the study. The database is in my possession and is available for inspection.

66. The method used sometimes had to be rendered into a dichotomous variable, *e.g.*, use of practical reasoning versus use of all other methods, use of strict construction versus use of all other methods, etc. The method used was always coded one, and all other methods were coded zero.

67. *See supra* text following note 54.

68. *See supra* Part I.A.1.

69. *See supra* Part I.A.2.

70. *See supra* Part I.A.2. Galler, *supra* note 17, at 857-69, questions whether *Chevron* should be applied to Service pronouncements. The cases frequently distinguished between reliance on regulations and pronouncements, and so I separated the two into different variables.

71. *See supra* Part I.A.3.

72. *See supra* Part I.A.4.

73. *See supra* Part I.A.5.

74. *See supra* text accompanying notes 47-49.

75. Sometimes the winner was clear; sometimes it was not. In cases where only one issue was before the court, a decision stated by the court to be in favor of the government or against the taxpayer was treated as a decision in favor of the government; conversely, a decision stated by the court to be in favor of the taxpayer or against the government was treated as a decision in favor of the taxpayer. Even if the "winner," as just defined, did not obtain from the court everything the taxpayer had sought, *e.g.*, the government "won" only sixty percent of the taxes it claimed the taxpayer owed, that party was still treated as the winner.

B. *Statistical Analysis*

I sampled cases from the Tax Court and three district courts in order to be able to contrast differences, if any, between two of the major forums of federal tax litigation and district courts in three different cities.[76] It is axiomatic that one should litigate in the court that presents the most favorable precedent.[77] Our federal tax system also permits one to choose whether to pay the tax and seek a refund in the district court, or not to pay and litigate in the Tax Court.[78] Therefore, choosing these two different types of courts enabled me to test for variables, other than precedent, that could enter into a decision—such as the judge's presumed politics, educational background, and gender—to see if they made a difference in outcome.[79]

Finally, I selected the district courts encompassing Los Angeles, Chicago, and New York City because I thought that, as courts in major metropolitan areas, they would offer a breadth of cases that would be at least as sophisticated as those drawn from other district courts. I also chose to review cases over a twenty-year period in order to minimize the risk of transitory aberrations and to offer a greater breadth of diversity—about gender, race/ethnicity, and approaches to interpreting statutes—than a shorter time span would allow.

Independent, or explanatory, variables included the judge's sex, years on the bench when the sampled case was decided, the appointing president's political party, eliteness of college and law school, primary professional experience before ascending to the bench, and, on the district courts, race/ethnicity. The dependent, or outcome, variable was the method of interpreting the Code section at issue with which the judge justified her decision in nonsummary cases.

Statistical significance for relationships created through comparisons between judges on the Tax Court and on the district courts (*e.g.*, Table 1) was

---

I resolved multiple-issue cases cautiously, much the same as I had single-issue cases where the winner might have obtained less than the full recovery the taxpayer had sought. For example, a party that had won on the more important issue but lost on a lesser issue (or issues) was still coded as the winner.

76. In addition to the reasons noted in the text, I did not choose the third forum, the Court of Federal Claims, because it is the only tax forum for which appeal is to a particular circuit, the Federal Circuit. In contrast, a Tax Court or district court decision could be appealed to the appropriate court of appeals, so that a taxpayer in Chicago would appeal to the Seventh Circuit or a taxpayer in New York City to the Second Circuit. *See* IRC § 7482(a)(1) (1994); 28 U.S.C. §§ 1291, 1295(a)(3) (1994). Therefore, the greater scope of appellate review was less likely to lead to more homogenous trial court decisions. The greater breadth of courts comprising the district courts also made them a more compelling case for study than the Court of Federal Claims.

77. *See, e.g.*, KAFKA & CAVANAUGH, *supra* note 55, at Ch. 1 (discussing where to litigate tax cases); Michael C. Durney, *Tax Litigation 1972-1992: A Retrospective*, 57 TAX NOTES 885, 886 (1992) (reviewing tax litigation culled from Internal Revenue Service figures); Marvin J. Garbis & Robert L. Frome, *Selecting the Court for the Optimum Disposition of a Tax Controversy*, 27 J. TAX'N 216 (1967) (examining factors to consider in tax litigation); Theodore Tannenwald, Jr., *The Tax Litigation Process: Where It Is and Where It Is Going*, 44 REC. ASS'N OF BAR OF CITY OF N.Y. 825 (1989) (examining tax litigation in light of proposals to centralize tax litigation).

78. *Compare* IRC § 6213(a) (1994) (describing Tax Court jurisdiction), *with* 28 U.S.C. § 1346(a)(1) (1994) (describing district court jurisdiction).

79. Choosing multiple district courts enabled me to contrast one with another, if that became appropriate. Differences among the district courts are noted. *See* Table 2A *infra* Appendix. For example, judges in the Central District of California led in elite college educations and percentage of women on the bench, judges in the Southern District of New York led in elite law school educations, and judges in the Northern District of Illinois led in appointments by Democratic presidents. Limited resources precluded me from engaging in further assessments.

established through statistically appropriate tests.[80] I used logistic regression to establish relationships between independent and dependent variables. Logistic regression permits the effect of each independent variable to be isolated while controlling for other influences. This method also facilitates the measurement of the magnitude of the influence.[81] Logistic regression, therefore, is well suited for the multivariate analysis of dichotomous, dependent variables undertaken in this article.

III. RESULTS

A. Who Were the Judges and How Did They Justify Their Decisions?

Analysis of the sampled cases revealed a number of facts about the judges, the approaches they used in the cases before them, and about the influence of social background.

1. Judges' Characteristics

Table 1[82] parses the judges who decided the cases by various characteristics, both as a group and by the Tax Court alone and the district courts alone.

The judges who decided the 482 sampled cases were mostly male[83] and, at least on the district court, mostly white. A majority of these judges had engaged primarily in private practice and had been on the bench for more than twelve years before deciding the sampled case. Slightly less than half had been appointed by Democratic presidents, and slightly less than half also had gone to elite law schools.

The main differences between the two courts are noted below.

- The judges in the 136 district court cases were drawn from more elite law schools and less elite colleges than the judges in the 346 Tax Court cases.

- The Tax Court judges were more senior than the district court judges.

- District court judges were more likely to be drawn from private practice, academia, and the judiciary, while more Tax Court judges had previous experience working for the government.[84]

---

80. The X^2 test was used for cross-tabulations of discrete variables (*e.g.*, gender, method of construction). Differences between the means were used to establish significance for continuous data (*e.g.*, years on the bench when deciding the case), and the Wald test was used for logistic regression. Statistical significance is "designed to allow us to make statements about the probability that hypothetical relationships actually occur;" it permits the inference that two variables are related—in the population and in the sample—and not merely the result of random association. GEORGE W. BOHRNSTEDT & DAVID KNOKE, STATISTICS FOR SOCIAL DATA ANALYSIS 22-23, 158 (3d ed. 1994); Schneider, *supra* note 2, at 492 n.20. *See generally* Brudney et al., *supra* note 3, at 1709 n.121 (discussing 0.05 level of testing for significance and 0.10 level for approaching significance).

81. Brudney et al., *supra* note 3, at 1680.

82. In this and all following tables, the number of cases that comprises the percentage is set forth in parentheses. Thus, 217 of the judges in the 482 cases, or forty-five percent, attended elite law schools.

83. *See* Table 1A *infra* Appendix (setting forth additional data about the differences revealed by the sex of the judges). This table reveals that women were frequently junior to their male colleagues, attended less elite colleges and law schools, had lesser tenure when deciding cases, and came from private practice less often than men. They came from teaching, being a judge, or from private practice and government at least as often as men did and were more likely to have been appointed by Republican than Democratic presidents.

84. Other data may be of interest, even though not included in the analysis. The average (mean) ages of the judges were as follows: all judges, 59.78 years old; judges appointed by Republican and Democratic presidents,

TABLE 1. JUDGES' CHARACTERISTICS

Characteristic	Both Courts Combined (N=482)	Tax Court (N=346)	District Courts (N=136)
Mean score of "Astin" number for college[a] attended*	59.46	60.18	57.60
Percent that attended elite law schools*	45% (217)	40% (137)	59% (80)
Mean of tenure when rendering sampled decision**	11.69 years	12.73 years	9.05 years
Percent of judges who were women	14% (66)	13% (44)	16% (22)
Percent appointed by Democratic presidents	45% (215)	45% (156)	43% (59)
Experience before becoming judge in:* a) government	38% (181)	45% (154)	20% (27)
b) private practice	58% (279)	53% (184)	70% (95)
c) other: both activities equally, teacher, judge, etc.	5% (22)	2% (8)	10% (14)
Ethnicity—percent who were not white[†]	–	–	12% (16)

[a] *See supra* note 56 (discussing Astin's formula for measuring college eliteness). This comparison of the means assumes equal variances, as does the comparison of the means of tenure. KRISTIN E. VOELKL & SUSAN B. GERBER, USING SPSS FOR WINDOWS: DATA ANALYSIS AND GRAPHICS Ch. 12 (1999).

*$p \leq .10$
**$p \leq .05$
[†]p not applicable

56.66 and 63.62 years old, respectively; male and female judges, 60.59 and 54.65 years old, respectively; Tax Court and district court judges, 60.02 and 59.17 years old, respectively; and white and nonwhite district court judges, 59.58 and 56.13 years old, respectively.

     When each case was decided, the length of tenure of the judge deciding the case ranged from one year—the tenure of judges in forty cases—up to forty-six years, the tenure of a judge in only one case. When a case was decided, the youngest age of the presiding judge was forty-one and the oldest age was eighty-eight. More of the judges were younger, *e.g.*, no more than fifty years old (twenty-one percent), than older, *e.g.*, eighty years or older (three percent). For a summary of data in another study about Tax Court judges and their professional backgrounds, see Maule, *supra* note 1, at 408-13.

All of these observations were statistically significant—or at least approached significance. Only sex and the political party of the appointing president lacked any type of statistical significance.[85]

2. Judges' Justifications When Interpreting the Internal Revenue Code

What approaches did judges use when explaining decisions in the 431 sampled, substantive, cases that required analysis of the Internal Revenue Code? In examining the database to determine what approach judges used, I found that 431 cases had at least one approach to statutory construction, 206 had a second approach to it, and 43 cases had a third approach. Table 2[86] sets forth the primary approach used in these cases (e.g., how many cases were explained by the court's primary reliance on structure, on strict construction, etc.).

What do these statistics tell us about the judges' approaches to interpreting the Internal Revenue Code in the sampled cases? First, Table 2 reveals that judges rely on diverse approaches. While commentators, in arguing for a particular approach for theoretical reasons, implicitly argue for that approach's exclusive use as well,[87] even the diverse approaches encompassed by practical reasoning do not seem to permit *judges* to rely exclusively on just one of the approaches. Each commentator gains some support from the data—and because two thirds of all cases rely primarily on practical reasoning, Livingston gains the most support—but no one method is used exclusively.

Table 2 also shows that the judges strongly favor nonliteral approaches, especially practical reasoning. Judges in the district court cases relied on practical reasoning even more than judges in the Tax Court cases. Perhaps the greater breadth of the subject matter jurisdiction that district court judges face leads them to justify decisions more pragmatically than judges on the Tax Court, with its narrower jurisdiction. Alternatively, perhaps Tax Court judges feel more comfortable with sophisticated approaches. The approaches other than practical reasoning are not unique to tax and should not differentiate the two courts. Still, the expertise Tax Court judges possess regarding tax might lead to their greater reliance on approaches other than practical reasoning.[88]

---

85. The non-continuous variables in Table 2 were tested to see whether the different groups noted in Table 1, *e.g.*, men versus women, Republican versus Democratic appointees, approached interpreting the Internal Revenue Code differently. While some differences were to be expected—for example, judges with law degrees from elite schools used practical reasoning fifty-six percent of the time, while those from non-elite schools used this method forty-four percent of the time—the differences were never statistically significant. Therefore, these differences are not set forth.

86. The values in the columns are set forth in percentages.

87. *See supra* Part I.A; *see also* Coverdale, *supra* note 6 and accompanying text (arguing against any nonliteral interpretation of the Internal Revenue Code, leaving only strict construction); Livingston, *supra* note 12, at 683-87 (arguing against purposivism in promoting practical reasoning).

88. *Cf.* Sandra Jo Craig, *Federal Income Tax and the Supreme Court: The Case Against a National Court of Tax Appeals*, 1983 UTAH L. REV. 679, 680-85 (summarizing arguments for a national court of tax appeals, including expert knowledge of judges on such a court). *But cf.* Livingston, *supra* note 12, at 683-87 (arguing that tax is not a special area).

TABLE 2. PRIMARY APPROACH USED**

Approach	Both Courts Combined (N=431)	Tax Court (N=334)	District Courts (N=97)
Strict construction	6% (25)	7% (22)	3% (3)
Nonliteral interpretation: practical reasoning	66% (283)	60% (200)	86% (83)
Legislative history	9% (37)	10% (34)	3% (3)
Regulations	12% (50)	14% (47)	3% (3)
Other IRS pronouncements	2% (8)	2% (6)	2% (2)
Structure	6% (27)	8% (25)	2% (2)

**$p \leq .05$

The tables contrasting each case's primary approach against other, subsequent approaches have been placed in an appendix to this article due to their lengths.[89] Review of these tables underscores the practicality judges used to explain their interpretations of the Code. Judges often rendered substantive decisions using a second approach on top of the first.[90] Because judges should not be able to construe a statute both literally and nonliterally, one surprising finding is that strict construction is frequently associated with other approaches. Review of these tables shows that practical reasoning was a popular second approach, and, to a lesser extent, legislative history as well.[91]

---

89. *See* Tables 3A–8A *infra* Appendix.

90. The type of case in which a second approach was least likely to be used was the practical reasoning cases, where only thirty-one percent (89/283) of the cases had a second approach. The next closest primary approach was strict construction, where sixty-eight percent of the cases had a second approach.

91. I did not catalogue the cases for the occurrence of single versus multiple issues. Assuming that some of the sampled cases were multiple-issue cases, this might explain multiple approaches used by the same court. It seems reasonable to conclude that not all the sampled cases were multiple-issue decisions; in which case, use of multiple approaches by a court faced with a single issue suggests a relaxed approach to interpreting the Code. Even if a sampled case was a multiple-issue decision, the other possibility—use of a different approach for each issue—still reveals a more relaxed approach to interpreting the Code than the literature would suggest. *See* Sisk et al., *supra* note 3 (reviewing the approach to statutory construction in single-issue cases).

3. Areas of the Decisions

When I catalogued the sampling by specific reference to the primary Code section at issue,[92] the cases dispersed as indicated in Table 3.[93]

TABLE 3. AREAS OF SAMPLED CASES

Type of Case	Both Courts Combined (N=407)	Tax Court (N= 308)	District Courts (N=99)
Individual	32% (132)	38% (118)	14% (14)
Compensation	3% (14)	4% (12)	2% (2)
Accounting	6% (25)	8% (24)	1% (1)
Exempt organizations	5% (22)	6% (19)	3% (3)
Business tax	11% (46)	14% (42)	4% (4)
Estate planning	5% (20)	5% (16)	4% (4)
Foreign tax	4% (15)	5% (14)	1% (1)
Capital assets	4% (15)	5% (14)	1% (1)
Procedure or criminal	29% (118)	16% (49)	70% (69)

Two aspects of Table 3 worth noting are the dispersion of the cases' subject areas and the courts in which the cases were litigated. Most of the cases were individual or procedural/criminal cases (thirty-two and twenty-nine percent of all cases), and a greater percentage of the district courts' cases were procedural or criminal than the Tax Court's (seventy and sixteen percent, respectively).

---

92. I necessarily made choices about how to group Code sections. Some groupings were easier to make, such as "foreign tax," which, among the catalogued cases, included §§ 861-994, and § 1402. Other groupings were more difficult to categorize, such as §§ 41-280, and §§ 1001-1031, for example, which I characterized as "individual."

Sections of the Code were organized as follows, and those groupings appear as row captions in Table 3: (1) individual, §§ 41-280, 1001-1031; (2) compensation, §§ 401-422, 3121, 4973-75; (3) accounting, §§ 448-82, 1341-48; (4) exempt organizations, §§ 501-12, 4911-45; (5) business tax, §§ 301-68, 533-613, 702-818, 1366-72, 1502, 4081-4481, 4987-97; (6) estate planning, §§ 641-75, 2013-2518; (7) foreign tax, §§ 861-994, 1402; (8) capital assets, §§ 1211-81; (9) procedure or criminal, §§ 6012-7609. The exclusion of a section from this list, such as § 1, means that § 1 was not at issue, or primarily at issue, in any sampled case. The presence of a section at the beginning or end of a group means that it was at issue, and the inclusion of a section somewhere within a group does not necessarily mean that it was primarily at issue in any case either. For example, among the "individual" Code sections, both § 41 and § 280 were at issue at least once; § 61 and § 71 also were at issue, but §§ 62, 63, and 72 were not.

As noted above, the data was not catalogued for single versus multiple issues. See supra Part III.A.2. Nevertheless, it is my impression that even multiple-issue cases were more likely to revolve around a single area, such that the presence of multiple issues had a less adverse impact on the discussion of the area than the topic of the secondary approach to interpreting the Code used by a court.

93. The study population is made up of 482 cases. The column values are set forth in percentages. Four hundred and seven cases could be characterized by Code section. Seventy-five had no predominate Code section because they were about discovery or some other non-tax matter. Of these seventy-five, the Tax Court decided thirty-eight, and thirty-seven were decided by the district courts.

Theoretically, the preponderance of district court cases in the procedural/criminal area might be attributable to its exclusive jurisdiction over criminal tax matters,[94] but only two district court cases dealt primarily with criminal tax statutes. Thus, the table reveals litigants' preference for resolving procedural matters in the district court. Arguably, litigants might also turn to the Tax Court, with its presumed expertise in tax matters, for areas requiring specialized knowledge, such as business planning (fourteen and four percent of all Tax Court and district court cases, respectively), and the statistics bear out that conclusion. But other areas requiring expertise, such as estate planning (five percent of all Tax Court and four percent of all district court cases), contradict the suggestion that litigants turn to the Tax Court for its skill in tax matters.

Less can be said about the meaning of the subject areas of the sampled cases. A benefit of the single-subject Sentencing Reform Act study is that it was a more controlled experiment than this one, which examined diverse cases over a broad spectrum of issues.[95] The large number of individual or procedural/criminal tax cases could not be easily characterized as inherently factual and, therefore, more likely to lead to litigation than other issues like capital assets or compensation. Even a highly structured area, such as business tax, has its pockets of factual inquiry—debt versus equity for example. Thus it is difficult to generalize about the subject areas of the cases.[96]

Table 4[97] sets forth the use of different approaches to interpretation of the Code in each of the enumerated subject matter topics. It illuminates whether certain primary approaches were used more by judges examining issues in one area than in others.

Judges justified their decisions with practical reasoning in all areas. This should not be surprising, given the overwhelming primary reliance on the practical reasoning approach.[98] The more interesting results reflected in this table are the areas in which other approaches appear. For example, deference to regulations was used more in accounting, exempt organizations, business tax, and foreign tax, a result consistent with the popular perception that these areas are highly regulated. Legislative history was used more in foreign tax, capital assets, exempt organizations, and compensation. Structure was used more in compensation and estate planning. Strict construction was used more in foreign tax and estate planning.

---

94.   18 U.S.C. §§ 3231, 3237(b) (1994) (giving district courts exclusive jurisdiction over federal crimes and providing venue for tax crimes, respectively).

95.   *See supra* note 48.

96.   *See, e.g.*, BORIS I. BITTKER & JAMES S. EUSTICE, FEDERAL INCOME TAXATION OF CORPORATIONS AND SHAREHOLDERS ¶ 4.03 (6th ed. 1998) (discussing differences between debt and equity); William T. Plumb, Jr., *The Federal Income Tax Significance of Corporate Debt: A Critical Analysis and a Proposal*, 26 TAX L. REV. 369 (1970-71) (reviewing how case law distinguishes between debt and stock).

The areas of the sampled cases contrast with the topics about which people have written. In a study of articles during a forty-five year period in some of the more significant tax journals, the most frequently analyzed area seems to have been business tax; litigation, which included procedural and criminal tax matters, was the subject of far fewer articles. *See* Schneider, *supra* note 2, at 519-523.

97.   Populations and percentages are given for each row. Information about the relative number of cases using different primary approaches is set forth in Table 2, *supra* Part III.A.2 (using primary approach).

98.   *See* Table 2 *supra* Part III.A.2.

TABLE 4. AREAS AND FIRST APPROACH TO INTERPRETING THE CODE

Area/Approach	Strict Construction	Regulations	Other IRS Pronouncements	Structure	Legislative History	Practical Reasoning
Individual (N=128)	3% (4)	11% (14)	5% (6)	5% (6)	9% (11)	68% (87)
Compensation (N=12)	8% (1)	8% (1)	8% (1)	25% (3)	17% (2)	33% (4)
Accounting (N=24)	4% (1)	33% (8)	—	4% (1)	—	58% (14)
Exempt organizations (N=22)	5% (1)	27% (6)	—	—	18% (4)	50% (11)
Business tax (N=46)	7% (3)	24% (11)	—	11% (5)	11% (5)	48% (22)
Estate planning (N=19)	16% (3)	—	—	26% (5)	5% (1)	53% (10)
Foreign tax (N=15)	20% (3)	27% (4)	—	7% (1)	20% (3)	27% (4)
Capital assets (N=15)	—	—	—	—	20% (3)	80% (12)
Procedural-criminal (N=99)	7% (7)	5% (5)	—	3% (3)	7% (7)	78% (77)
No area (N=50)	4% (2)	2% (1)	2% (1)	6% (3)	2% (1)	84% (42)

The possible relationship between subject matter and statutory construction has not yet generated discussion. A benefit of setting forth data about subject matter might simply be to pave the way for additional research about applying specific approaches to specific areas. Why, for example, is strict construction used more in foreign tax and estate planning? Why is structure used in compensation and estate planning?

B. *Logistic Regressions—Why Judges Used Different Interpretations*

The effect of independent variables on dependent, dichotomous variables is best ascertained using logistic regression.[99] In my analysis, the five different primary methods of statutory construction with which judges explained their decisions made up my dependent variables.[100] I performed regressions on each dependent variable for the Tax Court alone, for the district courts alone, and for all courts combined.

My independent variables were gender, years on the bench when the decision was rendered, the appointing president's political party, education (both undergraduate and law school), the judge's primary experience before ascending to the bench (in private practice, government, or elsewhere), and, for district court judges, race/ethnicity.[101] Because there are so many dependent variables, and they are used in regressions for three sets of courts, only the relationships that are statistically significant or approaching significance are noted below.[102]

Tables 5 and 6 examine practical reasoning as the primary method of interpretation. When both courts were combined, there were no associations between that method and any independent variable. When the courts were examined separately, these tables show that the only independent variable that could be associated with use of practical reasoning as the primary method of interpretation was the elite nature of the judge's college education. An elite college education had a negative impact on Tax Court judges' use of practical reasoning (Table 5[103]),

---

99. *See, e.g.*, Brudney et al., *supra* note 3; Sisk et al., *supra* note 3.

100. Each dependent variable consisted of one type of construction as opposed to all other methods. For example, one dependent variable (practical reasoning) was coded one, while all other methods (strict construction, legislative history, regulations and other Internal Revenue Service pronouncements, and structure) were coded zero. Each dependent variable needed a dichotomy in order to facilitate analysis.

101. For descriptions of these variables, see *supra* Part II.A.2. These variables, or ones like them, have been used in other studies. *See* Brudney et al., *supra* note 3; Sisk et al., *supra* note 3.

Multicollinearity was not problematic here. *See* Sisk et al., *supra* note 3, at 1432, n.233, (suggesting "no firm 'rule' exists" about what constitutes collinearity and choosing to adopt a model excluding independent variables, the bivariate correlation of which exceeded .500). None of the correlations between independent variables in regressions for both courts exceeded +/- .372; none exceeded +/- .426 in regressions for the Tax Court, and none exceeded +/- .339 in regressions for the district courts.

102. Thus, for example, none of the other independent variables, like gender or appointing president's politics, could be associated with the dependent variable of primary use of the practical reasoning method for the Tax Court. Nor could any of the independent variables be associated with the dependent variable of primary use of the practical reasoning method for both types of courts combined.

103. The population of this regression and all of the following regressions involving the Tax Court is comprised of 334 cases; the population for regressions involving the district courts is made up of ninety-seven cases, and the population for regressions involving both courts combined is comprised of 431 cases. In every regression, the approach indicated in the caption to the table—practical reasoning, for example, in Tables 5 and 6—is coded one and all other approaches are coded zero.

The columns labeled "β" and "Exp (β)" state the same fact differently. Beta (β) indicates the effect of an independent variable on the logarithmic probability of the dependent variable occurring. A number in the exponentiated beta (Exp (β)) of more than one indicates that something is more likely to occur than not, and a number of less than one indicates that it is less likely to occur than not. All of the tables except Table 8 are single, independent variable cases. Thus, in Table 8, the table involving two independent variables, both eliteness of college and prior professional experience (specifically in private practice) led to the lesser likelihood that the district court would use regulations or other Service pronouncements. *See* Table 8 *infra* Part III.B. As between these two variables, a judge's having primarily prior private practice experience was even less likely to lead to use of regulations or other Service pronouncements than having gone to an elite college, either because the prior

leading them to avoid practical reasoning, and the contrary effect on district court judges, causing them to use this approach (Table 6).

TABLE 5. LOGISTIC REGRESSION: PRIMARY USE OF PRACTICAL REASONING BY THE TAX COURT

Independent Variable	β	Standard Error	Significance	Exp (β)
Eliteness—college[a]	-.031	.014	.029	.969

[a] The more elite a college was, the higher its score. *See supra* note 56.

TABLE 6. LOGISTIC REGRESSION: PRIMARY USE OF PRACTICAL REASONING BY DISTRICT COURTS

Independent Variable	β	Standard Error	Significance	Exp (β)
Eliteness—college	.036	.014	.011	1.036

Tables 7 and 8 review the primary use of regulations/pronouncements. Like practical reasoning, no association existed between the primary use of regulations and other Internal Revenue Service pronouncements when the Tax Court and the district courts were combined. When the two courts are viewed separately, however, length of tenure had a negative impact on Tax Court judges' use of this approach (Table 7[104]). An elite college education had a negative impact on district court judges' use of this approach, and primary professional experience before becoming a judge—other than in private practice—had an even stronger negative impact on them (Table 8).[105]

These findings are restated below.
- The only methods of interpretation influenced by social background were practical reasoning and deference to regulations/ pronouncements.
- With reference to practical reasoning:
  - Controlling for all other independent variables, Tax Court judges who had more elite college educations were more likely to avoid practical reasoning.
  - District court judges who had elite college educations were more likely to explain their decisions using practical reasoning.

---

experience beta was lower than the elite college beta, or because the prior experience exponentiated beta was lower than the elite college exponentiated beta.

104. There were too few cases—only eight—in "Other Internal Revenue Service Pronouncements" to give an accurate regression. *See* Table 2 *supra* Part III.A.2. Therefore, these cases were combined with the fifty "Regulation" cases in order to run the logistic regression.

105. Again, there were too few cases, so regulations and pronouncements were combined. *See supra* note 104. *See supra* note 103 (indicating that such prior experience would have been in government, as a judge or teacher, or equally in government and private practice).

TABLE 7. LOGISTIC REGRESSION: PRIMARY USE OF REGULATIONS OR OTHER INTERNAL REVENUE SERVICE PRONOUNCEMENTS BY THE TAX COURT

Independent Variable	β	Standard Error	Significance	Exp (β)
Length of tenure[a]	-.032	.020	.107[b]	.969

[a] The longer the judge had served on the bench, the higher was the number of years in the length of tenure variable. *See supra* note 64.
[b] Length of tenure fails to approach significance by 0.07%, but I have included this variable anyway to view the data more expansively.

TABLE 8. LOGISTIC REGRESSION: PRIMARY USE OF REGULATIONS OR OTHER INTERNAL REVENUE SERVICE PRONOUNCEMENTS BY THE DISTRICT COURTS

Independent Variable	β	Standard Error	Significance	Exp (β)
Eliteness—college	-.040	.024	.095	.960
Prior professional experience (not private practice=0, private practice=1)	-2.511	1.439	.081	.081

- With reference to reliance on regulations or Service pronouncements:
  - ○ Controlling for all other independent variables, Tax Court judges were less likely to rely on regulations/ pronouncements as they became more senior.
  - ○ District court judges who had more elite college educations were less likely to give deference to regulations and pronouncements. This tendency is even more pronounced in district court cases decided by judges whose primary work before becoming a judge was in the government, as a judge, as a teacher, or equally split between private practice and government work. Stated somewhat differently, district court judges were less likely to rely on regulations or pronouncements as the primary interpretive approach if they had less elite college educations or had come from a "non-private practice" prior work experience.

C. Interpreting the Logistic Regressions

Other authors who have used logistic regression to analyze the relationship between social background factors and judges' decisions have concluded that such

factors have some "effect" on judicial reasoning.[106] The preceding analysis suggests that the following social background factors may be at play in outcomes in this article's database.

1. Education

Education broadcasts socioeconomic signals. Undergraduate education has long been perceived as a socioeconomic marker, and even legal education has been viewed similarly, although in a more attenuated manner.[107] Students attending more elite colleges tend to come from more privileged backgrounds than those who study at other colleges.[108]

Here, district court judges with more elite college educations used practical reasoning and avoided giving deference to regulations and other Internal Revenue Service pronouncements. The effect of undergraduate education on the district court judge's choice of interpretive method is consistent with Sisk, Heise, and Morriss's observation that elite law school education does not lead to conceptual decision making. They found that judges with elite law school educations applied practical reasoning in making sentencing decisions. This finding surprised them because they had expected judges from elite law schools to favor conceptual reasoning.[109] They conclude that "this finding undermines the hypothesis that judges educated at elite law schools are more attracted towards conceptual reasoning than judges educated at non-elite law schools."[110] Arguably, because both types of degrees mark socioeconomic standing, willingness to use practical reasoning could rest with either marker.

If practical reasoning can be characterized as non-abstract, then it follows that the other area in which undergraduate education was important—deference to regulations or other Service pronouncements in district court cases—is more abstract. The regression coefficient in this latter area is negative, so district court judges with elite educations were less likely to rely primarily on regulations and other Service pronouncements.

On the other hand, the opposite result in the practical reasoning area among judges in the Tax Court cases negates these observations. A less elite college education is associated with Tax Court judges' primary reliance on the practical reasoning approach. Certainly, the opposing regressions in the two types of courts simply may not be reconcilable. If they are, however, a way in which to rationalize

---

106. *See, e.g.*, Brudney et al., *supra* note 3, at 1761 ("Notwithstanding...[several] caveats, our analysis has identified numerous personal, political, and professional background factors that are significantly associated with a judge's propensity to support or reject the union's legal position."); Sisk et al., *supra* note 3, at 1498-99 ("as is often the case with empirical research, our study provides both comfort and challenges to all camps...the behavioral model is simultaneously bolstered and buffeted").

107. For brevity, I have emphasized these two articles as points of reference in the text. There is, however, a wealth of additional literature, some of which I have also noted. The literature is exhaustively reviewed in Brudney et al., *supra* note 3, at 1739-59, and Sisk et al., *supra* note 3, at 1451-98. For an article in which legal education was found to be important, see Gerard S. Gryski & Eleanor C. Main, *Social Backgrounds as Predictors of Votes on State Courts of Last Resort: The Case of Sex Discrimination*, 39 W. POL. Q. 528, 532 (1986) (finding that judges were more liberal if they had attended public or out-of-state law schools).

108. *See, e.g.*, Brudney et al., *supra* note 3, at 1750-51 nn.228-29 (citing various sources).

109. Sisk et al., *supra* note 3, at 1464.

110. *Id.* at 1465.

them is to note the source of the socioeconomic marking. Eliteness in law schools—possessed by the judges on the district courts more than on the Tax Court—arguably promotes non-abstract reasoning, while that eliteness in college—possessed by the judges on the Tax Court more than on the district courts[111]—negates non-abstract reasoning. This logic, however, is tortured, and literature does not reinforce such a conclusion.

2. Seniority

Seniority led Tax Court judges to give less deference to regulations and pronouncements. Perhaps seniority might lead to a diminishing desire to follow the strictures of regulations or other Service pronouncements, at least relative to other methods of interpreting the Code. Such a conclusion would be consistent with the limited research, both in law and in social science, regarding the effect of age or seniority on how judges decide.[112]

3. Prior Work Experience

Judges were less likely to use regulations or other pronouncements when they had worked for the government (the greatest number in this group), taught at a law school, been a judge, or worked equally for the government and in private practice before becoming district court judges. Brudney, Schiavoni, and Merritt's findings that management-side experience tended to lead to a judge's later support of the union position[113] echoes this discovery. They believe that the NLRB's history—its enactment during the New Deal and its fostering of collective action—"may signify that familiarity with the Act breeds greater respect for its protective doctrinal scope."[114] In tax, however, familiarity appears to have bred contempt.

4. Gender

There was no association between the method of statutory construction employed and a judge's sex. Differences between the sexes might have been expected because, for example, of Professor Carol Gilligan's "different voice" theory that women view themselves as more community-oriented, while men tend to view themselves as more autonomous.[115] Professor Suzanna Sherry's assertion of

---

111. *See* Table 1 *supra* Part III.A.1.

112. *Compare* Sisk et al., *supra* note 3, at 1486-87 (finding that more senior judges did not increasingly uphold sentencing guidelines), *with* Brudney et al., *supra* note 3, at 1755 n.238 (finding that age was significant, reinforcing the "adage that with old age comes a reluctance to deviate from the status quo"); *see also* Sue Davis et al., *Voting Behavior and Gender on the U.S. Courts of Appeals*, 77 JUDICATURE 129, 133 (1990) (stating that the age of a judge affects the outcome of a case); Sheldon Goldman, *Voting Behavior on the United States Courts of Appeals Revisited*, 69 AM. POL. SCI. REV. 491, 499 (1975) (finding that age made appellate judges more conservative, but not in the "government fiscal" area, which included tax); Gryski & Main, *supra* note 107 (noting modest evidence that age affects judges' decisions).

113. *See* Brudney et al., *supra* note 3, at 1741-50.

114. *Id.* at 1745; *see also id.* at 1471-74. For an example of a social science article where prior career experience was important, see C. Neal Tate & Roger Handberg, *Time Binding and Theory Building in Personal Attribute Models of Supreme Court Voting Behavior, 1916-88*, 35 AM. J. POL. SCI. 460 (1991).

115. CAROL GILLIGAN, IN A DIFFERENT VOICE: PSYCHOLOGICAL THEORY AND WOMEN'S DEVELOPMENT (1982).

a feminine attitude in which women stress "relationships to others," while men emphasize "rights" might have led to the same expectation.[116] Studies about the impact of gender on judges' decisions are divided in their conclusions, some perceiving a difference and others, not.[117]

As has been observed,

> [w]hether because gender-based theories of differences are wrong or overstated, because the judicial recruitment process selects only women compatible with the views of the appointing president, or because "differences between men and women judges are neutralized by the very nature of law and legal process," "[t]he weight of the evidence demonstrates that most female judges do not decide cases in a distinctively feminist or feminine manner."[118]

The same might be said of sex and methods of interpretation used to justify judges' federal tax decisions. Tax cases may present few opportunities for a distinctly feminine or feminist point of view.[119] At best, justifying decisions in tax cases is a complicated area in which a feminine or feminist approach cannot easily be applied.

5. Race

My study did not reveal any relationship between race and reasoning in tax case decisions. Despite the importance of race in other empirical research—especially in the social sciences—it might, like gender, not be easily imported into tax cases.[120] In other words, the importance of race in how judges sentence criminal

---

116. Suzanna Sherry, *Civic Virtue and the Feminine Voice in Constitutional Adjudication*, 72 VA. L. REV. 543, 582 (1982); *see also* Michael E. Solimine & Susan E. Wheatley, *Rethinking Feminist Judging*, 70 IND. L.J. 891 (1995) (reviewing literature about effect of gender on judging and arguing for more women judges).

117. *Compare* Brudney et al., *supra* note 3, at 1756-59 (finding that gender affected outcome among Republican women), *with* Sisk et al., *supra* note 3, at 1451-54 (finding that gender had no influence).

Social science literature has also been divided about the effect of gender. *See, e.g.,* Davis, *supra* note 112, at 133 (finding that male and female judges act differently in employment discrimination cases, but not obscenity cases); John Gruhl et al., *Women as Policymakers: The Case of Trial Judges*, 25 AM. J. POL. SCI. 308 (1981) (noting that male and female judges act the same, *e.g.,* in leniency of sentence, but differently, *e.g.,* when sentencing male and female defendants). Most articles focus on how judges act in the criminal area. One article concluded that female judges supported federal economic regulation more than their male colleagues. *See* Thomas G. Walker & Deborah J. Barrow, *The Diversification of the Federal Bench: Policy and Process Ramifications*, 47 J. POL. 596, 614 (1985). That finding contrasts with the conclusion here that sex had no impact on interpreting the Code.

118. *See* Sisk et al., *supra* note 3, at 1453-54 (footnotes omitted).

119. *But see* Gwen Thayer Handelman, *Sisters in Law: Gender and the Interpretation of Tax Statutes*, 3 UCLA WOMEN'S L.J. 39 (1993) (discussing interplay between "feminism" and statutory construction). *See also* Symposium, *Critical Tax Theory: Criticism and Response*, 76 N.C. L. REV. 1519 (1998) [hereinafter Symposium] (recounting recent debate about tax and feminism).

120. There is limited research about tax and race. *See* Moran & Whitford, *supra* note 2. *Compare* Sisk et al., *supra* note 3, at 1453-54 (finding minority judges more likely to invalidate guidelines, but no association between race and outcome), *with* Brudney et al., *supra* note 3, at 1755 (finding African-American judges more likely to support a union than white judges, but no association between race and outcomes). *See generally* Symposium, *supra* note 119 (recounting recent debate about tax and critical race theory). There are, however, more numerous examples of social science literature dealing with the impact of race. *See, e.g.,* Cassia Spohn, *The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities*, 24 L. SOC. 1197, 1206-08 (1990) (finding that black and white judges generally do not sentence criminal defendants differently); Thomas M. Uhlman, *Black Elite Decision Making: The Case of Trial Judges*, 22 AM. J. POL. SCI. 884, 893 (1978) (finding that black and white judges come from the same elite legal background); Susan Welch et al., *Do Black Judges Make a Difference?*, 32 AM. J. POL. SCI. 126, 134 (1988) (finding that black and white judges do sentence criminal defendants differently).

defendants notwithstanding, the race of the judge does not matter when interpreting the Code. As with gender, perhaps a judge's race is neutralized by the nature of the legal process, especially in tax.

6. Politics

Finally, there was no relationship between politics and outcome. Like race, politics has been an important factor in other empirical research about how judges decide issues, especially in the social sciences. While tax is not free from politics, its failure to influence the methods of construction judges use to justify their decisions suggests that politics, like gender, may simply be neutralized by the legal process.[121]

IV. CONCLUSIONS

Articles about interpreting the Internal Revenue Code consistently theorize about using particular approaches. Theory is certainly important, but it cannot explain who the judges are who have interpreted the Code, how 'they have interpreted it, or why they have justified their decisions with the approaches they have used.

The judges who decided the cases in this database were mostly men, drawn more frequently from private practice than government work, somewhat more likely to have been appointed by a Republican than a Democratic president and, on the district court, overwhelmingly white. Nonliteral modes of interpretation, especially practical reasoning, were favored by these judges, and practical reasoning was favored more by judges on the district courts than on the Tax Court. Some social background factors—college education, seniority, and prior practice—played into judges' decisions, and the presence of these factors may explain, in a way that traditional legal analysis does not, judges' choices about interpreting the Internal Revenue Code.

---

121. Neither Brudney et al., *supra* note 3, at 1737-39, nor Sisk et al., *supra* note 3, at 1465, found any association between a judge's politics and outcome. Social science literature, however, frequently finds politics to be important. *See, e.g.*, Jilda M. Aliotta, *Combining Judges' Attributes and Case Characteristics: An Alternative Approach to Explaining Supreme Court Decisionmaking*, 71 JUDICATURE 277, 280 (1988) (finding a judge's politics important); Jon Gottschall, *Reagan's Appointments to the U.S. Courts of Appeals: The Continuation of a Judicial Revolution*, 70 JUDICATURE 48, 54 (1986) (same); Jon Gottschall, *Carter's Judicial Appointments: The Influence of Affirmative Action and Merit Selection on Voting on the U.S. Courts of Appeals*, 67 JUDICATURE 165 (1983) (same); Stuart S. Nagel, *Political Party Affiliation and Judges' Decisions*, 55 AM. POL. SCI. REV. 843 (1961) (same); Donald S. Songer & Sue Davis, *The Impact of Party and Region on Voting Decisions in the United States Courts of Appeals, 1955-1986*, 43 W. POL. Q. 317 (1990) (same); Tate & Handberg, *supra* note 114 (same).

APPENDIX

Table 1A sets forth judges' characteristics by sex. In each cell, women's statistics have been set forth first; then men's statistics are set forth in brackets. References "only for women" or "only for men" are intended to indicate statistical significance. For example, "% appointed by Democratic presidents**, only for women" means that statistical significance was found only for the women judges appointed by Democratic presidents.

TABLE 1A. JUDGES' CHARACTERISTICS

Characteristic	Both Courts Combined (N=482)	Tax Court (N=346)	District Courts (N=136)
Percent of judges who were women [men]	14% (66) [86% (416)]	13% (44) [87% (302)]	16% (22) [84% (114)]
Mean score of "Astin" number for college attended*, only for women	57.79 [59.72]	59.95 [60.22]	53.45 [58.40]
Percent attended elite law schools**	39% (26) [46% (191)]	30% (13) [41% (124)]	59% (13) [59% (67)]
Mean of tenure when rendering sampled decision**	9.05 years [12.12 years]	11.27 years [12.94 years]	4.59 years [9.92 years]
Percent appointed by Democratic presidents**, only for women	23% (15) [48% (200)]	11% (5) [50% (151)]	46% (10) [43% (49)]
Percent appointed by Republican presidents**, only for women	77% (51) [52% (216)]	89% (39) [50% (151)]	55% (12) [57% (65)]
Experience before becoming judge in: ** only for men a) government	38% (25) [38% (156)]	43% (19) [45% (135)]	27% (6) [18% (21)]

* $p \leq .10$

** $p \leq .05$

† p not applicable

TABLE 1A (CONTINUED)

b) private practice	46% (30) [60% (249)]	39% (17) [55% (167)]	59% (8) [72% (82)]
c) other: both activities equally, teacher, judge	17% (11) [3% (11)]	18% (8) [–]	14% (3) [10% (11)]
Ethnicity—percent who were not white†	–	–	18% (4) [10% (12)]
Ethnicity—percent who were white†	–	–	82% (18) [90% (102)]

* p≤.10
** p≤.05
† p not applicable

Table 2A sets forth judges' characteristics for the different district courts.

TABLE 2A. JUDGES' CHARACTERISTICS BY DISTRICT COURT

Characteristic/Court	C.D. Cal.	N.D. Ill.	S.D.N.Y.
Mean score of "Astin" number for college attended**	59.26	55.39	58.91
Percent attended elite law schools**	37% (13)	56% (30)	79% (37)
Percent women	20% (7)	13% (7)	17% (8)
Mean of tenure when rendering sampled decision	8.74 years	7.74 years	10.81 years
Percent appointed by Democratic presidents	46% (16)	50% (27)	34% (16)
Experience before becoming judge in:** a) government	23% (8)	24% (13)	13% (6)
b) private practice	66% (23)	61% (33)	83% (39)
c) other: both activities equally, teacher, judge	11% (4)	15% (8)	4% (2)
Ethnicity—percent who were not white**	26% (9)	6% (3)	9% (4)

** $p \leq .05$

Tables 3A–8A set forth information about subsequent approaches used by a court in association with the approach that it first used, e.g., of the 23 cases that relied primarily on structure, how many relied on each of the other remaining approaches, etc.

TABLE 3A. WHEN MOST IMPORTANT APPROACH IS STRICT CONSTRUCTION[a]

Next Approach Used	Both courts combined (N=17)	Tax Court (N=15)	District Courts (N=2)
Regulations	6% (1)	7% (1)	—
Structure	18% (3)	20% (3)	–
Legislative history	47% (8)	40% (6)	100% (2)
Practical reasoning	29% (5)	33% (5)	—

[a]The population of cases in which the first approach used was strict construction was twenty-five. Eight of these cases had no second approach, seven of which were decided by the Tax Court and one by a district court.

None of the secondary approaches possesses significance, except for those in Table 4A, in which the primary approach is regulations. The lack of significance may be due to the large number of empty cells in the cross-tabulations done in the X^2 text for those tables. *See* ALAN AGRESTI, AN INTRODUCTION TO CATEGORICAL DATA ANALYSIS 194 (1996) ("When cell counts are so small that chi-squared approximations may be inadequate, one could combine categories of variables to obtain larger counts.").

TABLE 4A. WHEN MOST IMPORTANT APPROACH IS REGULATIONS[a,**]

Next Approach Used:	Both Courts Combined (N=42)	Tax Court (N=40)	District Courts (N=2)
Strict Construction	2% (1)	3% (1)	—
Other IRS pronouncements	7% (3)	5% (2)	50% (1)
Structure	7% (3)	5% (2)	50% (1)
Legislative history	17% (7)	18% (7)	—
Practical reasoning	67% (28)	70% (28)	—

[a] The population of cases in which the first approach used was reliance on regulations was 50. Eight of these cases had no second approach. Of those that had no second approach, seven were decided by the Tax Court and one by district court.

** $p \leq .05$

TABLE 5A. WHEN MOST IMPORTANT APPROACH IS OTHER IRS PRONOUNCEMENTS[b]

Next Approach Used	Both Courts Combined (N=8)	Tax Court (N=6)	District Courts (N=2)
Regulations	25% (2)	17% (1)	50% (1)
Structure	25% (2)	17% (1)	50% (1)
Legislative history	13% (1)	17% (1)	—
Practical reasoning	38% (3)	50% (3)	—

[b] The population of cases in which the first approach used was other IRS pronouncements.

TABLE 6A. WHEN MOST IMPORTANT APPROACH IS STRUCTURE[a]

Next Approach Used	Both Courts Combined (N=23)	Tax Court (N=21)	District Courts (N=2)
Regulations	22% (5)	19% (4)	50% (1)
Other IRS pronouncements	4% (1)	5% (1)	–
Legislative history	22% (5)	19% (4)	50% (1)
Practical reasoning	52% (12)	57% (12)	—

[a] The population of cases in which the first approach used was structure was 27. Four of these cases had no second approach. Of those that did not have a second approach, all were decided by Tax Court.

TABLE 7A. WHEN MOST IMPORTANT APPROACH IS LEGISLATIVE HISTORY[b]

Next Approach Used	Both Courts Combined (N=27)	Tax Court (N=25)	District Courts (N=2)
Regulations	11% (3)	12% (3)	—
Other IRS pronouncements	11% (3)	8% (2)	50% (1)
Structure	26% (7)	28% (7)	—
Practical reasoning	52% (14)	52% (13)	50% (1)

[b] The population of cases in which the first approach used was legislative history was 37. Ten of these cases had no second approach; nine were decided by the Tax Court and one by a district court.

TABLE 8A. WHEN MOST IMPORTANT APPROACH IS PRACTICAL REASONING[*]

Next Approach Used	Both Courts Combined (N=89)	Tax Court (N=78)	District Courts (N=11)
Strict construction	5% (4)	5% (4)	—
Regulations	35% (31)	36% (28)	27% (3)
Other IRS pronouncements	17% (15)	14% (11)	36% (4)
Structure	21% (19)	23% (18)	9% (1)
Legislative history	23% (20)	22% (17)	27% (3)

[*] The population of cases in which the first approach used was practical reasoning was 283. In 194 of these cases, there was no second approach. Of these 194 cases, 122 were decided by the Tax Court and 72 by the district courts.