Articles

Unaccountable Justice? The Decision Making of Magistrate Judges in the Federal District Courts
CHRISTINA L. BOYD AND JACQUELINE M. SIEVERT

Explaining Congressional Grants of Jurisdiction to the Federal District Courts
SETH W. GREENFEST

Blind Justice or Blind Ambition? The Influence of Promotion on Decision Making in the U.S. Courts of Appeals
JEFFREY BUJZIAK

Justices Denied: A County-Level Analysis of the 2010 Iowa Supreme Court Retention Election
ANDREW J. CLOPTON AND C. SCOTT PETERS

Comparing Between-Judge Disparities in Imprisonment Decisions Across Sentencing Regimes in Ohio
JOHN WOODREDGE, TIMOTHY GRIFFIN, AND AMY THISTLETHWAITE

Legal Note

SHENITA BRAZELTON

“Providing Leadership and Service to the State Courts”
Online at www.ncsc.org
Letter from the Editor-in-Chief

ROBERT M. HOWARD

Articles

Unaccountable Justice? The Decision Making of Magistrate Judges in the Federal District Courts

CHRISTINA L. BOYD AND JACQUELINE M. SIEVERT

Explaining Congressional Grants of Jurisdiction to the Federal District Courts

SETH W. GREENFEST

Blind Justice or Blind Ambition? The Influence of Promotion on Decision Making in the U.S. Courts of Appeals

JEFFREY BUDZIAK

Justices Denied: A County-Level Analysis of the 2010 Iowa Supreme Court Retention Election

ANDREW J. CLOPTON AND C. SCOTT PETERS

Comparing Between-Judge Disparities in Imprisonment Decisions Across Sentencing Regimes in Ohio

JOHN WOOLDRIDGE, TIMOTHY GRIFFIN, AND AMY THISTLETHWAITE

Legal Note


SHENITA BRAZELTON
Submissions to Justice System Journal

The Justice System Journal invites submission of articles and notes about courts and court administration that are likely to be of interest to practitioners and researchers in the field of judicial administration, broadly defined. Articles may draw on a variety of research approaches in the social sciences. The Journal is especially interested in manuscripts that have implications for justice system policy and that address problems faced by those with responsibilities for court administration. Short articles written by practitioners or researchers that describe management innovations in court systems are encouraged. Short manuscripts analyzing cases or legal issues of particular relevance to judicial administration are welcome and will be considered for the Legal Notes section. The Journal does not normally publish articles devoted to extended analysis of legal doctrine.

Inquiries to the editor about topic ideas that authors are considering are welcome.

Submissions to the Journal may not be under consideration by any other publication.

Manuscripts are evaluated by two or more reviewers; both practitioners and scholars participate in the review process. To facilitate the review process, references that reveal the identity of authors should be deleted from the manuscript and placed on a separate page; the author’s name, institutional affiliation, e-mail address, and any acknowledgments should appear only on a cover sheet.

A manuscript submitted for consideration should be accompanied by an abstract of approximately 100 words. Manuscripts should be double-spaced and prepared with ample margins, on 8 1/2” x 11” paper, one side per page. Use of footnotes extensive in number or length is discouraged; footnotes should be placed on a separate page, double-spaced. Submission of manuscripts in Journal style is preferred and accepted manuscripts must ultimately be revised by the author to conform to it. Accepted manuscripts must then be submitted either in Word 2000 or WordPerfect 8.0 (or higher) word-processing format, as an e-mail attachment or on a CD-ROM.

Two (2) electronic copies of all manuscripts should be submitted to Prof. Robert M. Howard, Department of Political Science, Georgia State University, 38 Peachtree Centre Avenue, Atlanta, GA 30303, e-mail: poljsj@gsu.edu. Inquiries concerning book reviews, including proposals for reviews, should be made to Prof. Chris W. Bonneau, University of Pittsburgh, Department of Political Science, 4600 Posvar Hall, Pittsburgh, PA 15260, e-mail cwbonneau@gmail.com. Inquiries concerning the Legal Notes section should be addressed to Pamela C. Corley, Assistant Professor, Department of Political Science, Southern Methodist University, 3300 University Blvd, Carr Collins Hall, Dallas, TX 75205; e-mail: pccorley@mail.smu.edu.

Photocopy requests: Up to fifty photocopies may be made without permission. These must be used exclusively for educational purposes, with no direct or indirect commercial advantage, and must include a notice of copyright. Requests for permission beyond these limits may be addressed to Charles F. Campbell, Managing Editor, Justice System Journal, National Center for State Courts, 300 Newport Ave., Williamsburg, VA 23185-4147; phone (757) 259-1838; fax (757) 564-2031; e-mail ccampbell@ncsc.org.

The contents of this journal are indexed in The Index to Legal Periodicals, The Social Sciences Citation Index, The Criminal Justice Periodical Index, and Westlaw.

Copyright 2013 by the National Center for State Courts. All rights reserved.

The Justice System Journal is printed by Thomson/West as a public service to scholars, practitioners, and others interested in judicial administration.
LETTER FROM THE EDITOR-IN-CHIEF

Please find our final issue of 2013, 34:3. This will also be the final issue published solely under the auspices of the National Center for State Courts. Beginning with volume 35, the journal moves to Routledge (Taylor and Francis) Publishing, and we will publish four, not three, issues per year. The link to the new journal website is: http://www.tandfonline.com/loi/ujjs20#.UnKaHvmsi-1. This will also be the last issue where I am the sole Editor-in-Chief. Mark Hurwitz will share EIC duties with me and over the course of next year will take over as the Editor-in-Chief. I will discuss this a bit more at the end of this letter.

The issue continues our pattern of presenting new and innovative research on both state and federal courts. We present three articles on federal courts and two articles on state courts. In our lead article, Christina L. Boyd of the University of Georgia and Jacqueline M. Sievert of the University of Buffalo examine a much understudied group in the federal judiciary—the U.S. magistrates. Using principal-agent theory they find that magistrates are constrained by the preferences of the district judges in and the institutional characteristics of their districts.

In our next article on the federal courts, Seth W. Greenfest of the College of Saint Benedict and St. John’s University examines the jurisdiction-granting powers of Congress. Professor Greenfest shows that separation of powers and ideological concerns often lead Congress to manipulate the jurisdiction of the federal district courts.

Our final article on federal courts studies another intriguing aspect of judicial behavior—namely, how the desire for promotion influences the behavior of federal appellate court justices. Jeffrey Budziak of Western Kentucky University finds that ambition influences courts of appeals judges. They alter their behavior in a manner consistent with attempting to secure a seat on the United States Supreme Court.

The next two manuscripts shift the focus to state courts. Our first examines the retention election of the three Iowa judges who voted to overturn Iowa’s ban on same-sex marriage. All three were ousted from office. Andrew J. Clopton of the University of Michigan Law School and C. Scott Peters of the University of Northern Iowa find that the results can be explained by many of the same factors that affect outcomes in partisan and nonpartisan judicial elections.

The last article, by John Wooldredge of the University of Cincinnati, Timothy Griffin of the University of Nevada at Reno, and Amy Thistlethwaite of Northern Kentucky University, looks at the impact of the implementation of the Ohio sentencing guidelines. They find that the guidelines were successful for reducing judicial effects on sentencing based on judges’ tenure on the bench, prosecutorial experience, and caseload composition.

Finally we offer a short legal note by Shenita Brazelton, who examines the future of affirmative action in higher education following the recent Supreme Court decision in Fisher v. University of Texas at Austin (2013).
I trust you will find these articles as interesting and informative as I have. Before I sign off, I would like to acknowledge the important contributions of some individuals who have assisted me greatly during my tenure as Editor-in-Chief. First, let me thank our Managing Editor, Chuck Campbell. Chuck’s editing and organizational skills have allowed me to take credit for a much better journal than would have been produced solely by my efforts. I cannot thank him enough and I will truly miss our frequent interactions. Good luck on the coasters.

I would also like to thank the Michigan State “Fraternity.” Mark Hurwitz has offered invaluable assistance, advice and counsel, first as Legal Notes Editor and then as the Associate Editor. Mark and I have been working with Routledge to ensure a smooth transition to a new publisher. Kirk Randazzo and Chris Bonneau, the other two-thirds of the MSU triumvirate, have provided invaluable editorial assistance and advice over the years. I thank them both and look forward to our continuing friendship even when I no longer have the authority to get their work published. I would also like to express my gratitude to our current Legal Notes Editor, Pam Corley, as well as the all of those who have submitted and reviewed manuscripts over my tenure.

I would like to thank Richard Schaufler and Tom Clarke of the National Center for State Courts. They provided crucial support through some difficult times as we moved the journal toward a much more secure future.

Finally, I would also like to thank Thomson Reuters (West) Publishing. They have offered free printing and support to the Justice System Journal over the years, and I thank them on my behalf and on behalf of the former editors and the working staff at the National Center for State Courts.

Robert M. Howard
Editor-in-Chief
Georgia State University, Atlanta
Unaccountable Justice? The Decision Making of Magistrate Judges in the Federal District Courts*

Christina L. Boyd and Jacqueline M. Sievert

Modern federal district courts delegate vast decision-making powers throughout criminal and civil cases to magistrate judges—judicial actors that, unlike other federal judges, serve without an Article III political appointment. Through the lens of principal-agency theory, this study seeks to rectify the relative paucity of systematic work on these actors by using original filing and motion-level district court data to examine magistrates’ decision making empirically. Our results support our expectations that magistrates are often constrained by the preferences of the district judges in and the institutional characteristics of their district while issuing reports and recommendations and serving as assigned judges by the consent of the parties.

In early 2011, the Washington Post argued that increasing case filings, an uptick in judicial retirements, and a high number of judicial vacancies have combined to form the perfect storm of judicial emergency in the lower courts of the federal judiciary (Markon and Murray, 2011). Nowhere are these concerns more salient than in the U.S. District Courts, the 94 trial courts in the federal judicial system, where there are now over 350,000 civil and criminal filings per year (Administrative Office of the U.S. Courts, 2011) and numerous unfilled judgeships. These increasing pressures on federal trial courts are well documented (Galanter, 2004), as are some of the techniques and mechanisms that government officials, lawyers, and scholars have presented to address them (Galanter, 1985; Resnik, 1982). We focus here on one of the most important, understudied, and, perhaps, controversial of these: the advent of the role and use of magistrate judges in federal district courts.

Over forty years ago, Congress authorized the creation of magistrate judges in federal district courts to help these courts more efficiently process cases. Over time, the responsibilities of these judges have increased, so much so that these judges are now disposing of over 1 million civil and criminal matters in district courts (including, e.g., motions and hearings) per year (Administrative Office of the U.S. Courts, 2011). As the quantity of delegation to magistrates increases, however, so too expands the potential for controversy. Unlike traditional district judges, magistrates are not Article

* We are grateful to the National Science Foundation, the Center for Empirical Research in the Law at Washington University, and the Baldy Center for Law and Social Policy at the University at Buffalo, SUNY for supporting our research. We also thank Jim Battista, Ryan Black, Lee Epstein, Pauline Kim, Andrew Martin, Lynn Mather, Margo Schlanger, Jim Spriggs, and the editor and anonymous reviewers at Justice System Journal for providing feedback on this and related work, and Sophie Fortin, Ben Guthorn, Shadi Peterman, Gregg Re, Annie Rushman, and Erica Woodruff for supplying excellent research assistance. Christina L. Boyd (cLboyd@uga.edu) is an assistant professor in the Department of Political Science at the University of Georgia, and Jacqueline M. Sievert (jm siever@buffalo.edu) is a Ph.D. candidate in the Department of Political Science at the University at Buffalo, SUNY.
III federal judges nominated by the president and confirmed by the Senate. This raises a number of important concerns about the decision making of these actors and the political and institutional accountability that they face—concerns that have, until now, gone virtually unaddressed in systematic political science scholarship.

Relying primarily on principal-agency theory, we argue that magistrates are agents to Article III district court judges and, as such, ought to be constrained in their decision making by these principals. To examine this, we use original filing-level civil law data from twenty federal district courts and analyze the decisions of magistrates in two important and very distinct contexts: 1) as a case’s primary assigned judge by the consent of the parties in the case and 2) as an assistant to a case’s assigned district court judge, providing a report and recommendation to the case’s district court judge on the outcome of important motions in the case. Our results support many of our expectations.

By way of preview, in our examination of the direction of case outcomes, we find that magistrates serving as a case’s assigned judge temper their decisions based on the ideological preferences of their district principals. Second, our results reveal that district judges are less likely to adopt the recommendations of magistrates when district judges have relatively low caseloads.

**Theorizing on the Role of Decision Making of Magistrates**

The relationship between a magistrate judge and the district court that he or she serves in can be described as one between an agent and a principal. Principal-agency theory is now widely used to explain hierarchical relationships within political science and the judiciary, including, for example, higher courts to lower courts (Songer, Segal, and Cameron, 1994; Benesh and Martinek, 2002; Randazzo, 2008; George and Yoon, 2003; Corley, 2009), justices to their clerks (Ditslear and Baum, 2001; Black and Boyd, 2011; Peppers, 2006), and courts to other political branches (Lindquist and Haire, 2006). While previous interdisciplinary work on magistrates has likened magistrates to bureaucrats or players in an organization (Seron, 1988) or has relied on proximate selectorate theory to describe their behavior (Carroll, 2004), we believe that principal-agency provides an excellent theoretical underpinning for the unique employment relationship between magistrates and their district court bosses.

**The Necessity of Delegation.** As Moe (1984) famously expounded, “[t]he principal agent model is an analytic expression of the agency relationship, in which one party, the principal, considers entering into a contractual agreement with another, the agent, in the expectation that the agent will subsequently choose actions that produce outcomes desired by the principal.” The resulting formal relationship proceeds with the delegation of tasks and responsibilities and the authorization of the agent to act on behalf of the principal. For principals, such delegation is desirable and necessary because they are “troubled by insufficient time and information” (Porter, 1975:40), and achieving their goals alone would be expensive if not impossible. By authorizing agents to act on their behalf, principals and their organizations can tackle a larger workload and achieve a more efficient processing of their tasks.
Applied here, the district court, via the active judges within the district, is the principal for magistrate judges. The very nature of the evolution of the role of magistrates recognizes the (ever-increasing) need of traditional district judges to delegate tasks to them. In 1968, in light of mounting caseload pressure and an extensive backlog of pending cases, Congress authorized the creation of a new federal judicial officer, the U.S. magistrate, who would assume all the former duties of commissioners and would conduct a wide range of judicial proceedings to expedite the disposition of the civil and criminal caseloads in the U.S. District Courts (Smith, 1990). In the years that followed, the potential powers and role of magistrates continued to grow (Smith, 1990; Carroll, 2004).

While the creation of magistrates came about via legislation, it is up to individual district courts to choose whether, how, and to what degree to use these actors. District judges, frequently and increasingly over time, have delegated large decision-making responsibilities to the magistrates in their courts. The current functions of magistrate judges in federal district courts include assisting and recommending actions to the district judge, conducting civil trials at the consent of the litigants, serving as a special master, and presiding over settlement conferences, ADR, and pretrial hearings. Magistrates can also perform a considerable role in federal criminal cases, including accepting criminal complaints, issuing arrest warrants and summonses, issuing search warrants, and conducting initial-appearance proceedings and detention hearings for criminal defendants (Baker, 2005). Delegation of vast responsibilities like these to magistrates arguably helps our judicial system and, specifically, district judges to process cases effectively and efficiently and ensure that litigants receive substantive and procedural justice. At the same time, however, this raises concerns about the outputs of magistrates within this system. Will these magistrates make decisions that will accord with their district principal? Or, instead, will these magistrates’ decisions be driven by independent and self-interested considerations?

The Tools of Magistrate Control. The concern with any extensive delegation and reliance on an agent is whether the agent will perform consistently with the expectations and preferences of the principal rather than disobeying, shirking, and acting in a self-interested fashion. Moe warns of the dangers of this implicit moral hazard: “there is no guarantee that the agent, once hired, will in fact choose to pursue the principal’s best interests or to do so efficiently. The agent has his own interests at heart” (Moe, 1984:756). Simply put, due to the information and interest asymmetries between the actors, delegation to agents puts principals’ preferred outcomes at risk.

To respond to this concern, the principal-agency model presents a principal with several methods to help prevent agents from engaging in undesirable behavior. Such tools are designed to “mitigate the asymmetry [between principal and agent] and thus minimize the problems of adverse selection and moral hazard” (Moe, 1984:766). Examples include ex ante control mechanisms, such as the screening and selection of agents, sanctions, incentives, and monitoring.

The screening and selection of agents is often the most direct and effective mechanism of agent control for principals (Van Houten, 2009). Not all principals
have the ability to select their agents, something that is particularly true in many judicial applications of this theory. When they do, however, careful agent selection, bolstered by signals from potential agents, presents an opportunity for principals to control the preferences of their agents, control their outputs, and limit or altogether remove the need for regular oversight (Calvert, 1989; Spence, 1974). While hiring that is careful and avoids adverse selection requires a substantial principal investment (Kiewiet and McCubbins, 1991; Miller and Moe, 1986), “both sides are better off if principals are able to identify those individuals who possess the appropriate talents, skills, and other personal characteristics prior to the establishment of the principal/agent relationship” (Kiewiet and McCubbins, 1991).

Sanctions and the threat of sanctions can also serve as an agent-control mechanism. Potential sanctions for shirking agents include, for example, termination, demotion, and lack of reappointment. While sanctions can be costly for some principals in some circumstances (Kiewiet and McCubbins, 1991), if the threat of them being carried out is credible, agents should be more likely to behave according to the interests of their principals. In addition, principals can use positive incentives, such as promotions, reputation building, and the promise of career advancement to induce agent compliance (Kiewiet and McCubbins, 1991; Black and Boyd, 2011; Peppers, 2006).

Where agent screening is limited and the threat of sanctions or promise of incentives are weak constraints on agent behavior, principals can turn to active monitoring to reduce information asymmetries and encourage agents to not act in their own self-interest. Monitoring through investigations, regular oversight, and reporting requirements can be difficult and costly to carry out and, particularly in the face of principals’ need to delegate to begin with, is often ineffective (Kiewiet and McCubbins, 1991). To overcome this inefficiency, principals can turn to fire-alarm oversight (McCubbins and Schwartz, 1984), where they depend on reliable third-party signals to indicate the presence of inappropriate agent activity.

How does this apply to magistrate judges and their district court principal? Unlike most principal-agency applications in a judicial setting, the principal here controls agent selection and can use this process to recruit trustworthy and like-minded magistrate agents. While district judges are Article III judges nominated by the president, confirmed by the Senate, and protected by lifetime tenure, magistrate judges are appointed by the district court judges in the district in which they will serve to an eight-year term. Potential magistrate judges are nominated through a selection panel consisting of lawyers and at least two nonlawyers in the community (Judicial Information Series No. 2, 2002). The selection panel (whose composition is voted on by the district court judges) evaluates the qualifications of the applicants and submits a list of five potential candidates to the court for consideration. The active district judges in the district then select their magistrate by a majority vote. Through this hiring process, district court judges can vet their agents for the qualities, demeanor, and
experience that will make them a good fit within the district. Because this selection process is inevitably a relative rare one in a district court, district judges are likely to be all the more able to devote the time necessary to be satisfied that their vote is carefully made.

Similarly strong are the sanctioning controls of district courts over magistrates. Like civil servants described by McCubbins, Noll, and Weingast (1987), magistrates can be fired (by majority vote) if they “stray too far” (p. 248) from the preferences of their district principals via, for example, incompetency, neglect of duty, or physical or mental disability (Judicial Information Series No. 2, 2002). Short of firing them, district judges can also relegate (temporarily or permanently) noncomplying magistrates to the district’s most menial and administrative tasks including, for example, prisoner litigation, bail review, and petty offenses. This is possible because of the vast discretion in the degree and content of delegation that district judges have when operating as a supervising body. With the expiration of a magistrate’s eight-year term, district judges can (and do) choose to not reappoint a magistrate they are not pleased with for another term.

Various monitoring mechanisms enable district courts to gain the information necessary to determine when removing delegated authority from a magistrate, or even sanctioning them, is appropriate. The potential for active district court monitoring of magistrate judge decisions and activity exist in nearly every context of magistrate activity. As noted above, many magistrate responsibilities place them in a role of formally or informally assisting a district judge, including through conducting pretrial hearings, ruling on preliminary case motions, and serving as a third-party neutral in settlement conferences. In other scenarios, like initial criminal proceedings involving arrest and search warrants, appearances, detention, and criminal complaints, magistrates hold temporary authority that can be swiftly removed, serve under the watchful eye of the district’s chief judge, and can have their decisions reviewed and reversed by the district judge that is assigned the case if it moves forward toward trial.

District judges also have an array of positive incentives available to induce desired magistrate compliance. These include, for example, the potential for reappointment, assignment to prime district responsibilities, and the development of a generally favorable reputation in the district that could lead to greater future political and judicial opportunities, including appointment as an Article III judge (and the higher status and salary benefits that such an appointment brings).

While specific details on the selection, monitoring, and sanctioning of magistrate judges rarely emerge from district court insiders, some anecdotal evidence preliminarily confirms our theory about this principal-agent relationship. For example, in 2008 South Carolina’s federal district court selected to not reappoint Magistrate Judge George Kosko following allegations that he made disparaging comments, including toward women and minorities, while serving on the bench (Smith, 2008). Interviews
conducted by Christopher Smith (1990) revealed that some district courts, rather than reject the reappointment applications of magistrates whose performance was not satisfactory, would send signals to those individuals encouraging them to retire at the conclusion of their term. In 2006, a chief district judge in the 11th Circuit was required to take action against a magistrate judge in his district after the magistrate had been found to have “engaged in conduct prejudicial to the effective and expeditious administration of the business of the courts” (Administrative Office of the U.S. Courts, 2006:Table S-22). Unfortunately, specific details on the nature of the magistrate’s offending conduct or the resulting sanction were not released to the public. Smith (1990) also found evidence that magistrate judges recognized the presence of district court monitoring, and, more important, felt constrained by it. One former magistrate judge interviewed said, “I just did not feel like a judicial officer, because the judges control everything, tell you how to do things, and tell you what to decide. There is often no room to make a judicial decision, even when the judge’s orders will lead to an unjust result” (Smith, 1990:73).

On the more positive, incentive-related side of our theory, magistrates across districts are regularly reappointed to their positions after their initial eight-year terms expire. We also know of one magistrate judge, Wallace Capel, who had developed such a good reputation for his work as a magistrate judge in the Eastern District of Michigan from 1999-2006 that, in 2006, he was appointed as a magistrate in the Middle District of Alabama (Almanac of the Federal Judiciary, 2011). Developing strong judicial reputations have also been valuable for many former magistrates who have gone on to receive appointments as Article III district judges. Speaking of Judge Leslie Kobayashi, a two-term magistrate judge and President Obama’s nominee in 2010 for a district court judgeship in Hawaii, Senator Akaka said that, as a magistrate judge, she was “an outstanding and well-respected jurist known for her fairness and timely rulings” (Inouye, 2010). Kobayashi was successfully confirmed in December 2010.

Overall, these examples confirm the nature, strength, and variety of the selection, monitoring, and sanctioning tools available to district courts over magistrate judges. They lead us to believe in the appropriateness of the application of principal-agency theory to the district court-magistrate relationship, particularly when the theory is viewed in light of its other, well-regarded judicial branch applications, many of which lack principals holding selection or sanctioning powers. We believe that these types of controls available to district courts over their magistrates, when viewed together, should enable them to select good agents and, when necessary, sanction and give them incentives to behave according to the district’s best interests. The strength of this resulting principal-agent relationship should greatly diminish the need for frequent monitoring of magistrate activities and decisions. As we describe in further detail below, if and when the need for direct oversight arises, district judges can turn to a third party—litigants—for fire-alarm-style signals.
DATA

For purposes of this study, we empirically analyze magistrate decision making within the civil case context in two specific areas: 1) service as a case’s sole assigned judge by the consent of the parties in the case and 2) the making of reports and recommendations on motion dispositions to assigned district court judges. While these activities represent only a portion of the workload of magistrate judges, they are some of the most salient and outcome-specific activities that occur in district courts and, as such, provide an excellent starting point for systematically examining the role of magistrates. In addition, as we describe in more detail below, the two areas of focus here are very different from each other in terms of the underlying work done by magistrates, yet each provides an opportunity to test principal-agency theory in action.

To study this area, we rely on an original database of civil cases terminated from 2000 to 2006 in 20 of the 94 federal district courts. These 20 districts were selected because their chief judges provided PACER (“Public Access to Court Electronic Resources”) fee exemptions for collecting and studying the case materials from their courts and, together, they provide circuit, partisanship, size, caseload, and geographic diversity (see Table 1). These 20 districts provide representation for 11 of the 12 Circuit Courts (excluding only the D.C. Circuit); range in size from very small (2 active judges), to medium (7 active judges), to very large (28 active judges); are diverse in terms of the partisanship of their sitting, active judges (as measured by the party of their appointing president); and vary in the number of criminal and civil filings per judge (from a low of 301 per judge in the Northern District of Oklahoma to a high of 724 per judge in the Eastern District of Louisiana). The result of this multifaceted district diversity in the data is that these districts, together, are likely to serve as good representatives of district courts more generally.

1 The expense of relying on PACER to do docket-level research on federal courts makes it impractical to make a broad, empirical study of all districts. Like this study, others have drawn their sample from districts where PACER fee exemptions were available (Coleman, 2008, 1 district; Kourlis and Gagel, 2008, 8 districts). Others, while not selecting their sample solely because of PACER fee waivers, have focused on a small number of districts in their non-opinion-based empirical study of district courts, but have made a number of generalizations from their findings to the broader study of district courts (e.g., Siegelman and Donohue, 1990, 1 district, small analysis with 7 districts; Rowland and Carp, 1996, 2 districts; Olson, 1992, 1 district). One may question whether selecting the 20 districts for study here based on the willingness of a district’s chief judge to grant a fee exemption amounts to a systematic selection bias related to magistrate judges. We believe, for many reasons, that it does not. First, the request for PACER fee exemptions, made in 2008, centered on a large-scale study of district courts, termination methods of cases, and the interaction between district courts and their hierarchical superiors, and as such, was not solely focused on magistrate judges. Second, some districts declined to provide fee exemptions because of a standing policy to decline all such requests, no matter the topic of study or identity of the requester. Third, other districts beyond the 20 studied here also provided fee exemptions. Because of practical research considerations, along with time and use limitations that certain districts placed on fee-exemption retrieved data, this study settled on the 20 districts included in the data here. Finally, the chief judges responding affirmatively to fee exemption requests in 2008 were relatively representative of district court chief judges more generally (see Table 1). Chiefs appointed by Presidents Ronald Reagan, George H. W. Bush, Bill Clinton, and George W. Bush are all included, although the overall percentage of appointed chiefs skews somewhat more Republican than the overall districts at that time. Still, what is more important here is that there is no reason to believe that the varying identities of these chiefs and the way that their districts were selected for inclusion in this study on magistrate judges should systematically bias its conclusions or its ability to generalize from 20 districts to all 94.
These data are derived from the Federal Judicial Center’s (FJC) Integrated Databases and a docket-level analysis from materials gathered from PACER. They consist of three major civil issue areas: personal injury disputes, civil rights cases, and business-related cases. For the starting database in this study, we have drawn a stratified random sample of 250 cases per district. Each of our two models described below draw from this database for their underlying analyses.

**The Votes of Magistrates**

We begin first with an analysis of magistrate voting behavior in cases where they serve as the assigned judge. Since much of the judicial decision-making literature focuses on

---

2 We use nature-of-suit (NOS) codes, which are selected by the plaintiff’s attorney at filing, to identify our issue areas. Our NOS codes include civil rights cases (440, 442, 443, 444), personal injury cases (310, 320, 340, 350, 360), and business-related cases (190, 820, 830, 840).
the merits voting of judges and the ideological direction thereof, it makes sense that we begin by examining the behavior of magistrates in this light. The language in 28 U.S.C. § 636 (c) indicates that magistrate judges may at times serve as the only official, assigned judges in a case:

Upon the consent of the parties, a full-time United States magistrate judge or a part-time United States magistrate judge who serves as a full-time judicial officer may conduct any or all proceedings in a jury or nonjury civil matter and order the entry of judgment in the case, when specially designated to exercise such jurisdiction by the district court or courts he serves.

While district court judges presiding over cases are likely to take into consideration a number of legal, hierarchical, and political concerns, they are not likely to be heavily constrained by the preferences of their colleagues, especially when it comes to the ideological direction of their decision making. This reflects the hierarchical nature of our judicial system, the internal independence of the judiciary, and the noncollegial decision-making environment of district court judges. ³

When serving as a consent-assigned judge, we expect that, because of selection and sanctioning constraints and the repeat, long-term game that they are playing in their district, magistrates will continue to serve as agents of their district judge principals. This means, among other things, altering their decision making and case guidance to accord with what they believe to be the preferences of their district principals, rather than making decisions based on their sincere, independence-minded preferences.

Expectations and Variables. This area provides a great opportunity to evaluate our principal-agency expectations since we can test whether magistrate judges, serving as assigned judges, decide cases differently than their district court principals. If principal-agency is truly in operation for magistrates, then we will see magistrates deciding cases in a way that clearly is constrained by the overall preferences of the Article III district judges, while similarly situated district court judges will make decisions relatively independent of that district force.

To carry out this analysis, we examine whether district court cases conclude with a conservative or liberal outcome, something that is coded, following conventions from other federal court empirical studies, based on the type of case, its primary issue area, and the winner (Spaeth, 2005). Cases in our database where no clear coding could be made on these dimensions, such as those that settled, were dropped from the analysis. Our dependent variable is dichotomous in nature, with 0 = conservative and 1 = liberal.

³ Despite this institutional environment, district judges may, through socialization, collegiality, and contact with their colleagues, be informally constrained by their colleagues (Carp and Wheeler, 1972). Even when this exists, however, we expect its pull to be far less strong and predictable than it would be for magistrates serving as agents to district judges.
Our independent variable of interest focuses on measuring the pull of the principal-agency relationship. To do this, we interact the type of judge assigned to the case (magistrate or district court judge) with the mean ideology score of the active judges in the district. Our expectation is that magistrates will be constrained by this district ideology and will conform their decision making in cases toward those managerial preferences. As a district’s central tendency becomes more conservative, magistrate judges should be less likely to have their cases terminate in a liberal fashion, and vice versa. However, given that traditional district judges are not principals to each other, this constraining relationship should not be present for them while making decisions.

The Magistrate (Consent) Assigned variable is simply a dichotomous variable, with values of 1 for magistrates assigned to the case by party consent, and 0 for cases with a traditional Article III judge presiding. This variable is coded from the case’s docket. Within our outcome data, approximately 5 percent of the cases have a magistrate serving as assigned judge by party consent.

District Ideology is the average ideology score for the active district judges at the time that a district case is terminated. The underlying ideology measurement for each judge ranges from -1 (liberal) to 1 (conservative) and is based on Judicial Common Space (JCS) scores using the methodology described in Giles, Hettinger, and Peppers (2001) and Epstein et al. (2007). The ideology scores, derived from the Poole and Rosenthal NOMINATE Common Space scores, account for the nomination and confirmation process, including the norm of senatorial courtesy that operates within lower-court judicial selection. The average district ideology in this analysis is -0.005.

Based on the extensive research findings that judge ideology is an excellent predictor for a decision’s ideological direction (Segal and Spaeth, 2002; Hettinger, Lindquist, and Martinek, 2004; Rowland and Carp, 1996), we control for the ideology of the case’s assigned judge in our data (District Judge Ideology). For traditional district judges, this variable is simply measured as their JCS score (see above). For magistrate judges, however, since they did not receive an appointment through the same political process, no JCS or comparable ideological measure exists. Instead, to code ideology for these judges, we took the average (active) district judge ideology score for the year that the magistrate judge was first appointed to the district (Boyd and Hoffman, 2010).4

Since district court decisions are subject to potential review (and reversal) from the courts of appeals, we also control for the median ideology of a district’s hierarchically superior circuit court in the year of a case’s termination. Based on the findings of Randazzo (2008), we expect that assigned judges in district courts will consider and

---

4 While this measure of magistrate ideology is not as precise as we would like, it does reflect our (as well as external reviewers’) desire to incorporate some measure of ideology into an outcome model. It is worth noting that estimating our model without this ideology measure does not change the results for our key variables. Future work will, we hope, be able to use ideal point estimation, bridging techniques (from magistrates serving across districts or being appointed to serve as district judges), or both to provide more-precise ideology measures for magistrates, something that is not likely to happen until much more district court data, across years and districts, are collected.
moderate their voting behavior based on the general ideological preferences of their circuit. These are measured using the JCS scores, obtained from Epstein et al. (2007).

Finally, we control for the general issue area of the case. The issue area of a case is likely to affect its baseline probability of being decided in a liberal direction (Rowland and Carp, 1996). As noted above, our data include cases in three general issue areas: civil rights, business, and personal injury torts.

**Results.** The results of our estimated logistic regression of whether the district court case has a liberal outcome are reported in Table 2. The data examined here include 1,808 cases that amount to all nonsettled, nonprocedural dismissals from the above sample.

Recall that our primary expectation from this model is that when magistrates are serving as the assigned judge in a case, they will be constrained by the preferences of the district judges, and that will affect the direction of the case outcome. To capture this conditional relationship, our covariate of interest is the interaction of the mean ideology of the district judges and the status of the assigned judge in the case (magistrate or district judge). Because of the nature of interactive terms in logistic regressions, we must turn to a plot of the relationship to understand its statistical effect on whether a case will be decided in a liberal direction. We do this in Figure 1, where we depict the predicted probability of a case receiving a liberal outcome in each of our

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Logistic Regression Results for the Outcome Direction in District Court Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>District Ideology</td>
<td>0.541</td>
</tr>
<tr>
<td>(0.81)</td>
<td></td>
</tr>
<tr>
<td>Magistrate (Consent) Assigned</td>
<td>1.032*</td>
</tr>
<tr>
<td>(0.41)</td>
<td></td>
</tr>
<tr>
<td>District Ideology x Magistrate Assigned</td>
<td>-1.881</td>
</tr>
<tr>
<td>(2.25)</td>
<td></td>
</tr>
<tr>
<td>District Judge Ideology</td>
<td>0.441</td>
</tr>
<tr>
<td>(0.28)</td>
<td></td>
</tr>
<tr>
<td>Circuit Court Ideology</td>
<td>-0.724**</td>
</tr>
<tr>
<td>(0.39)</td>
<td></td>
</tr>
<tr>
<td>Business Case</td>
<td>1.674*</td>
</tr>
<tr>
<td>(0.26)</td>
<td></td>
</tr>
<tr>
<td>Civil Rights Case</td>
<td>-0.916*</td>
</tr>
<tr>
<td>(0.30)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-2.047*</td>
</tr>
<tr>
<td>(0.23)</td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-651.816</td>
</tr>
<tr>
<td>Observations</td>
<td>1,808</td>
</tr>
</tbody>
</table>

Note: Statistical significance is represented with * (p<0.05, two-tailed) and **(p<0.10, two-tailed), and robust standard errors are reported in parentheses. Personal Injury Case is the baseline issue area.
three broad issue areas. The left-hand column figures illustrate the mean predicted probabilities for each of the three issue areas for magistrates and district judges as the assigned judge in a case, while the center and right-hand columns plot the 90 percent confidence intervals around those predictions.

As we can see from the plots, there is much about this statistical relationship that is meaningful. As expected, when a magistrate is, by the consent of the parties, the
(only) assigned judge in the case, the mean ideology of the magistrate’s district can have a constraining effect. Magistrates serving as the assigned judge by the consent of the parties in a liberal district have a predicted probability of deciding a civil rights case (plotted in the figure’s top panel) in a liberal direction of over 0.18. Conversely, magistrates serving in a conservative district decide civil rights cases liberally with a likelihood of approximately 0.07. Similar patterns of difference exist for business and personal injury tort cases, although the baseline likelihood of a liberal outcome differs in each (see Figure 1). For business cases, the probability of a liberal outcome moves from 0.75 (liberal district) all the way down to below 0.50 (conservative district); for personal injury cases, the change is from 0.35 in a liberal district to just over 0.15 in a conservative one.

We anticipated above that if principal-agency is the operating theoretical constraint, only magistrate judges will be affected in their decision making by the mean ideological pull of the assigned district judges. The dashed lines in Figure 1, which plot the predicted probability that a district judge will decide a case in a liberal direction, generally support this. District court judges are, as predicted, relatively unaffected by the political preferences of their Article III colleagues. The predicted probability of these district court judges deciding their cases liberally is subject to much less change based on the mean district ideology.

It is important to note that, as is clear from the size of the confidence intervals in Figure 1, many of these estimates have a relatively high level of uncertainty. This is particularly true when examining ideologically extreme districts, with either very liberal or very conservative average ideologies, where our sample sizes are quite small. Within our estimates, the 90 percent confidence intervals on the district judge and magistrate judge predictions overlap at these margins, indicating that there is not a statistically significant difference between the estimates of liberal decisions for these two types of judges. This is something that we would expect in a conservative-leaning district in our illustrations, but not in a liberal one (due to the unconstrained district judges). There is no confidence interval overlap, and thus there is a statistically significant difference between district judges’ and magistrate judges’ decision making, in districts ranging from about -0.2 to 0.05 in overall ideology—i.e., moderately liberal and moderate districts. It is in these districts that we can see that magistrate judges are significantly more likely than district judges to decide cases in a liberal direction in a statistically meaningful way. For example, in a district with a mean ideology of -0.20 (moderately liberal), magistrate judges have about a 0.35 higher probability of deciding their assigned business-related cases in a liberal direction than do similarly situated district judges. These results, when taken together, provide us modest support that magistrate judges, no matter the issue area in the case before them, alter their decision making to account for their perceptions of their bosses’ preferences.

Our control variables also behave largely as expected. The ideological preferences of the reviewing circuit court serve to influence the decision making of the assigned judges in the district courts. As a hierarchically superior circuit’s mean
ideology grows more conservative, the assigned decision makers in the district court, whether Article III or magistrate, become less likely to decide their case in a liberal fashion. However, judge ideology just misses statistical significance, with a p-value of 0.11.

REPORTS AND RECOMMENDATIONS

We turn next to an analysis of the delegation of decision-making authority to magistrates by district judges in one of its most formal and traditional forms. Although magistrate judges will sometimes be the primary judges in a case like described in the analysis above, a more common role for these actors is to serve as a secondary judicial actor in a case by formally assisting the assigned Article III district judge. Built into the laws governing magistrate judges is a provision that permits district court judges to delegate important pretermination decision making to magistrates while retaining supervision and final decision-making authority. Pursuant to 28 U.S.C. § 636 (b), a district judge may:

- designate a magistrate judge to conduct hearings, including evidentiary hearings, and to submit to a judge of the court proposed findings of fact and recommendations for the disposition, by a judge of the court.

Acting under this clause, magistrate judges issue preliminary rulings on important, and at times legally complex, pretrial dispositive motions, such as motions to dismiss, motions regarding discovery, and motions for summary judgment. Trial court litigation is complex and dynamic (Kritzer, 1986), and a judge has many open cases on his docket at once, meaning that there can be tremendous, and at times conflicting, demands on an assigned judge’s time. Because of this, this pretrial decision-making delegation to magistrate judges is likely essential for many district court judges.

The magistrate’s recommendation is, at least in theory, just that—a recommendation, accompanied by a written report and legal reasoning, about the outcome of the motion before him or her. As a check on the decisions that magistrate judges are making in this pretrial phase of a case, the assigned Article III district judge reviews the recommendation, after which he may decide to adopt it, reject it, or modify it in part or in its entirety. In practice, though, nearly all magistrate recommendations are adopted by the assigned district judge. But nearly all is not all. As we argue below, legal signals motivate district judges to abandon unfettered delegation and engage in a hard, independent review of magistrate recommendations in certain situations. When this scenario is combined with ideological incompatibility between a case’s assigned district judge and the rest of his district colleagues, we expect to see a substantial increase in the likelihood of non-adoptions of the magistrate’s recommendation.
**Expectations and Variables.** Reports and recommendations, along with their subsequent review, provide an excellent arena for us to formally test the effectiveness of magistrates as agents (or district judges as principals). Studying this report and recommendation process in a systematic, quantitative fashion, however, is no easy task. For each of the cases in our original database, we relied on the case docket as the source for the presence and details of reports and recommendations. Such a focus on dockets, rather than on, for example, published opinions, was necessary since rarely will something like a magistrate’s report and recommendation lead to a published opinion. By definition, a report and recommendation is not the final district court judicial disposition on a particular motion. With the case dockets in hand, we used text-analysis technology to search for reports and recommendations in the case dockets, and then we further hand coded relevant characteristics about the underlying motion, magistrate, district judge, and other case attributes.

Our independent variables of interest concern the relationship between magistrates and their district principals, both ideologically and institutionally, as well as legal standards that modify the operation of this relationship. While a district court serves as the principal for magistrates, the active judges on this court are not all like-minded. Magistrates are likely to recognize this and tailor their behavior toward the centralized preferences of the district rather than any outlier judges. To extend the earlier analogy about the principal-agency relationship between the Supreme Court and the courts of appeals, on most issues, we would expect a compliance-focused appellate court to be much more likely to target the center of the Court—today, Justice Kennedy—than an ideological outlier like Justice Ginsburg. In practice, then, we expect that magistrates that perceive the assigned judge to hold inconsistent preferences from the rest of the district judges will be more likely to make recommendations that will be contrary to the preferences of the assigned judge. To capture this, we take the absolute difference between the mean ideology of the district and the ideology of the case’s assigned Article III district court judge. Both of these measures were derived from the JCS scores described above.

While magistrate judges are likely to keep in mind their district judge principals, and their preferences, throughout the delegated decision-making process, the reverse is not likely to always be true. We anticipate that district judges will be uniformly predisposed to adopt magistrate’s recommendations (Seron, 1985), even when the district judge is an ideological outlier in his district (and, thus, the primary principals to the magistrate assisting him or her on a case). This deference to magistrates is likely a product of district judges’ limited time to conduct a thorough secondary review of every case motion, their trust of the delegation system and the restraint of magistrates within that system, and their understanding that some nontrivial number of motion outcomes are legally based, noncontroversial, and nonideological (Rowland and Carp, 1996).
This predisposition toward recommendation adoption is more likely to be set aside, however, when litigants signal to the court that the magistrate’s proposed outcome is controversial in some way. We liken this, in many ways, to the fire-alarm oversight that Congress engages in over executive-branch administrative agencies (McCubbins and Schwartz, 1984). To initiate the fire alarm here, a party may, before the judge’s decision whether to adopt the recommendation, file a written objection to the magistrate’s suggested outcome. When this happens, the district judge must provide de novo review to the magistrate’s proposed ruling, meaning that the magistrate’s recommended outcome may no longer receive a legal presumption of correctness on review.

It is when this objection takes place that we expect the danger of non-adoption to become real, since district judges will now be engaging in a thorough review. However, because magistrates making their original recommendations did so while accounting for the ideological preferences of their primary district principals, not all recommendations will be equally likely to not be adopted. Rather, we expect that once an objection is made, ideological congruity will matter. As the assigned judge’s preferences move further away from the mean ideology of the district (the principals that a magistrate was likely to be most concerned with in making his recommendation), the likelihood of adoption should drop.

There is also an institutional perspective to this application of principal-agency theory to the district judge-magistrate judge relationship. For magistrate judges, we expect that varying institutional pressures will affect the calculus of their decision making. After all, as indicated above, the need for and congressional authorization of magistrates came about because of a necessity to get help in processing cases in district courts. As these considerations become more salient and apparent in an individual district, the behavior of the magistrates and their supervising principals should change to account for this as well. This includes a rise in the caseload of district judges. These dynamics are not likely to be stable across time and districts. In particular, in districts where judges are overwhelmed by their cases and are not efficiently processing cases, the role of magistrates will be the most critical. During these times, magistrates will likely have more independence in their decision making, and we will see fewer non-adoptions, since the time and ability of district judges to engage in a careful review of recommendations will decline as caseload goes up. Thus, as the number of cases terminated each year by each judge rises within the district, we expect disagreement and the presence of non-adoptions to fall.

Our dependent variable in this analysis is whether the district judge disagrees with and fails to wholly adopt the magistrate judge’s recommendation. This variable is dichotomous, with complete agreement (adoption) being coded as 0 and any disagreement with the magistrate being coded as 1. This disagreement can come in the form of a wholesale rejection as well as through a modification of the recommendation in whole or in part. Within our data, just under 6 percent of the reports and recommendations led to disagreement.
Within our data, we measure the presence of an objection to a magistrate’s report and recommendation dichotomously, with objections coded as 1. Since motions and objections are recorded on a case’s docket, we use that document to code the existence of this variable. Over 41 percent of the recommendations in our data set received objections. We measure Ideological Distance: Assigned Judge and District as the absolute difference between a district’s (average) ideology and the ideology of the case’s assigned judge, both of which are coded in the year a case is terminated and using JCS scores. The average value of this variable is 0.26. Because this predicted relationship depends on both the presence of an objection and the ideological distance between the assigned judge and the district mean, we model this relationship as an interaction between these two key variables. We code the number of cases terminating in a district per year (Average Terminations per Judge) from the Administrative Office of the U.S. Court’s Annual Reporting (2007), with the average value in our data resting at approximately 457 per judge per year.

We also control for case, political, and judge-specific variables that might impact this relationship. Legally complex motions leading to reports and recommendations are more likely to lead to disagreement than those that are not complex. These motions, coded here with the presence of motions to dismiss and motions for summary judgment, require extensive briefing by the parties and, in most cases, independent research by the ruling judge. This process is likely to lead to hard second looks on review, some of which will lead to disagreement. Composing over 70 percent, these complex motions dominate our report and recommendation data.

We expect less agreement between magistrates and senior district judges than between magistrates and active district judges, since senior judges are no longer key decision makers in the district nor full-time members of their district community and, as such, should not be expected to have the same power over magistrates as the district’s active judges. In addition, senior judges are not as overburdened by their cases as their active colleagues, meaning that they are better positioned to engage in a thorough secondary review of magistrates’ recommendations. Senior judges serve as 17 percent of our report-and-recommendation-reviewing judges.

We also expect less agreement between the magistrate and assigned judge the longer the magistrate serves in the district. With additional years on the bench, magistrates are likely to gain more and more independence in their job. In addition, as that time increases, there is likely to be less likelihood that the judge will be elevated to a district court position, thus making concerns about a higher reversal rate relatively moot. The average service time of magistrates in these data is over nine years.

Finally, we control for the general issue area of the case. Based on the nature of suit codes (described above), these three categories are business cases (19 percent of reports and recommendations), civil rights cases (71 percent of cases), and personal injury tort cases (10 percent of reports and recommendations).
Results. The results of our estimated logistic regression of whether district judges do not adopt a magistrate’s report and recommendation are reported in Table 3. To conduct this analysis, we focus on the 343 cases from the above detailed sample that have a report and recommendation and a district judge response to that recommendation.\footnote{We do not include observations where the case settled or otherwise terminated before the district judge responded to the magistrate’s report and recommendation.}

Let us turn first to our principal-agency related covariates from this report-and-recommendation model. Our first key variable from this relationship is the interaction of objections to the magistrate’s recommendation and the distance between the ideology of the case’s assigned judge and the average ideology of the district judges. Because we are dealing with an interactive effect, we again turn to a plot of the predicted probability of the interaction impacting the dependent variable.

While the mean predicted probability of disagreement by a district judge with a magistrate’s recommendation varies greatly between cases with and without filed objections as the ideological distance between the case’s assigned district judge and the
district mean grows, these differences are generally not statistically significant due to large and overlapping confidence intervals on the predictions (see Figure 2). This relationship does near statistical significance in the differences when the ideological distance between the case’s assigned district judge and the district mean ranges from 0.17 to 0.43. This near-effect may still be notable due to the model’s relatively small sample size, the difficulty in systematically studying the report-and-recommendation-agreement process, and the strength of the trends that we find in our results, a subject that we return to in greater detail in the discussion section below.

Figure 2
Predicted Probability that a District Judge Will Disagree with a Magistrate’s Report and Recommendation

Based on whether the assigned judge in the case is a district judge or a magistrate judge serving by consent, the mean ideological score in the district, and the case’s issue area. Other variables are held at the median and modal values. Ninety percent confidence intervals are depicted around the mean probabilities in the center and bottom panels of the figures.
There is evidence for the institutional side of our principal-agency story in the context of reports and recommendations (see Table 3). Our variable capturing the average number of terminations per judge in a district is negative and significant, indicating, as expected, that as the caseload per judge goes up, the likelihood of disagreement between assigned judge and magistrate goes down. Since the coefficients in logistic regression models cannot be interpreted for the size of effect that they have on our dependent variable, we turn to the simulation of predicted probabilities for this task (see Figure 3).

This variable has a sizable substantive effect on our dependent variable (see Figure 3). Districts with very low levels of terminations per district judge have a disagreement rate of approximately 15 percent. This is likely a sign that the magistrates serving in these districts feel less independence in their decision making due to the lower caseloads in the district, the decreased sense of judicial crisis, and the increased ability and time for district judges to actively monitor magistrates. On the other hand, as the number of terminations per judge in the district rise, we see the magistrates and their corresponding district judges altering their behavior and disagreeing with each other far less often. As the figure reveals, as this termination reaches its maximum—over 900 per judge—the disagreement probability nears 0. In these cases, while some of the decreased probability of non-adoption may be due to magistrates respectfully altering their decision making out of deference to the high caseload and general sense
of crisis in the district, it is more likely that what we are seeing is caused by a decrease in the ability of district judges to actively minimize shirking.

Finally, we only find statistical support for one of our control variables. In particular, complex motions increase the likelihood of disagreement after a report and recommendation. The coefficients for senior district judges, senatorial courtesy, magistrate length of service, and issue area do not reach statistically significant levels.

**Discussion**

This study joins a small number of quantitative studies examining magistrate judge decision making and roles in federal district courts. To do so, it systematically examines, via district court dockets and case materials, the types of decisions that magistrate judges make and the constraining effect that district preferences have on those decisions in 20 of 94 district courts over a seven-year period in three broad issue areas. Our results from an analysis of both the decisions of magistrates while serving as the assigned judge in a case by consent and their reports and recommendations while assisting district judges indicate modest support for the existence of an effective principal-agent relationship between district judges and magistrate judges in federal district courts.

In the context of magistrates acting as assigned judges by the consent of the parties, our results indicate that in many situations where their overall supervising district court’s ideological identity is not conservative, magistrate judges are significantly more likely to decide cases liberally than district judges. This is exactly the type of behavior that we anticipate district courts want from their agents. Regarding the reporting and recommendation function of magistrate judges, while the story of this relationship is complex, we also find some interesting effects. For example, we find institutional effects that are related to principal-agency, monitoring, and shirking. In particular, as a district’s number of case terminations per judge (a proxy for caseload) rises, a district judge’s time and ability to effectively monitor a magistrate’s recommendations drops, meaning that the likelihood of non-adoptions goes down even further than the already low baseline probability. We did not find, however, strong statistical evidence to support our expectation that the presence of a party objection to a magistrate’s recommendation combines with an ideological outlier assigned district judge to increase the likelihood of non-adoption. Given our modest sample size and the inappropriateness of examining magistrate decisions and activity using traditional mechanisms (e.g., published opinions), we believe that these results are all the more impressive.

Overall, this research indicates that district courts have numerous effective mechanisms at their disposal that allow them to delegate vast swaths of decision making to magistrates while avoiding many of the pitfalls of the moral hazard of principal-agent relationships. As a centralized body, district courts can maintain control over the actions and outputs of magistrates through careful selection and screening of agents, sanctions such as firing and assignment demotions, incentives such as reap-
pointments and reputation building, and, where necessary, monitoring—particularly when permitted by institutional circumstances, such as a low-to-moderate caseload. Compared to many other applications of principal-agency theory to the judicial context, the strength and variety of these tools is particularly notable.

As a result, our findings are likely to give many district judges confidence in their increased use of and delegation to magistrate judges. More broadly, politicians and other outside court watchers are likely to take heart in this evidence. Despite magistrates not having Article III’s careful political vetting process, they are not, generally speaking, unaccountable, unresponsive, or abusive carriers of justice in our federal courts. As such, many are likely to continue to view magistrates as both necessary and effective actors at helping stave off the ongoing judicial crisis that has resulted from increasing caseloads, rising judicial retirements, and vacant Article III judgeships.

Ultimately, however, we believe that future systematic work is needed to more carefully examine magistrate decision making in a variety of contexts and across all district courts. As we note in the text above, magistrate judges serve in many capacities in district courts, and we only examine two, albeit very salient areas, of these here. It could be, for example, that while processing early criminal proceedings like arrest and search warrants, initial appearances, and detention hearings, magistrate independence is at its highest since districts have relatively small amounts of monitoring in place, district judges are less likely to have previously formed preferences over outcomes, and the decision-making tasks are more likely to be considered to be menial or sub-judicial.

At the same time, more work is needed to examine whether it is reasonable to assume that magistrate judges can accurately sense or assess a district’s true centralized ideology and preference regarding much of their decision making and activities. District court judges have many fewer occasions to work together in a formalized setting than, for example, a collegial court. Simultaneously, the empirical findings on effect of ideology in judicial decision making and outcomes is far less compelling and unified for district courts than it is regarding appellate courts, particularly when settlement and pre-termination decision making are added into the mix. The result of this may well mean that magistrate judges are unable to recognize what is expected of them from their overall district in all scenarios, an opening that could allow more room for personal preferences and judicial independence to operate for some magistrate judicial roles. Only with further detailed and systematic work on magistrate judges, in-depth examination of individual districts and their practices, and the dynamics of district court litigation more generally can we really begin to assess these things going forward.

jsj

REFERENCES


This article examines the conditions under which Congress passes jurisdiction-granting legislation, which expands the discretion of the federal district courts by designating them as policy venues. This project extends existing research that has demonstrated that Congress manipulates the parameters of jurisdiction by examining how Congress routinely engages in this activity. I construct and evaluate a comprehensive data set of laws in which Congress grants jurisdiction to the district courts for the period between 1949-2006 (81st–109th Congress). I consider the effect of separation-of-powers and ideological distance on the decision to grant jurisdiction. The results demonstrate that both separation-of-powers and ideological concerns influence such decisions.

Through grants of jurisdiction, Congress promotes judicial participation in the policy process by establishing federal courts as venues in which individuals and groups may bring their claims. Armed with the constitutional power to establish the jurisdiction of the lower federal courts (and the appellate jurisdiction of the Supreme Court), Congress confers on courts the authority to hear certain issues. The power of courts to hear cases is the result of a complex and ongoing process through which actors outside the courts, particularly members of Congress, shape and reshape the power of different courts to hear particular types of disputes. That process involves, among other things, efforts by Congress to limit or protect the discretion of regulatory officials, to alter which individuals or groups have standing to bring cases to court, and to assign particular types of disputes to particular courts by granting courts jurisdiction to hear certain claims.

In contrast, conventional explanations of federal court participation in the policy process often focus on the U.S. Supreme Court and center on judicial behavior and decisions made by individual justices as they determine which issues to focus on. That courts become involved in American politics is often viewed as a pure manifestation of judicial power and of the ability of justices to set their own agendas free from outside influence. The fact that federal courts are able to decide cases is usually taken for granted as a fixed feature of the institutional environment.

Courts, however, do not merely spring into existence. Constitutionally, Congress has a host of powers with respect to the federal courts, including the power to create lower courts, to staff them, to fund their operations, and to define their jurisdiction. Indeed, the Constitution is relatively silent regarding the federal judiciary as it does.*

* Prof. Seth W. Greenfest (sgreenfest@csbsju.edu) teaches courses in constitutional law, American politics, and law and society at the College of Saint Benedict and St. John’s University in St. Joseph, Minnesota.
not establish any federal courts besides the Supreme Court, and includes very little
detail about how even that court is structured. The Constitution instead specifically
delegates the power to establish components of the federal judiciary to Congress.
Courts do not appear fully formed and do not come armed with jurisdiction to hear all
types of disputes but, rather, depend on Congress to establish the conditions under
which they will operate.

Recognizing the manner in which Congress exercises its constitutional power with
respect to the federal courts is important for understanding how courts operate in
American politics. In one view, how courts operate can be understood by examining
the judges who staff the courts and exploring the factors that explain their behavior.
These accounts of judicial behavior often focus on the “backgrounds, attitudes, and ide-
ological preferences of individual justices” to explain judicial decision making (Gillman
and Clayton, 1999:1). In contrast, this article adopts an institutionally focused
approach in which “[s]cholars seeking to explore the broader cultural and political con-
texts of judicial decision-making . . . examin[e] how judicial attitudes are themselves
constituted and structured by the Court as an institution and by its relationship to other
institutions in the political system at particular points in history” (Gillman and Clayton,
1999:2 emphasis in original). In contrast to court-centered or behavioralist approach-
es, the theory presented here argues that Congress has played an integral role in influ-
encing the creation of opportunities for judicial decision making.

In focusing on institutional factors that influence whether federal courts become
involved in the policy process, this article builds on a body of scholarship that investi-
gates the conditions that lead to judicial empowerment. Judicial participation in the
policy process is routinely structured through the language of statutes as members of
Congress work to create opportunities for judges to exercise discretion (Graber, 1993;
Gillman, 2002; Frymer, 2003; Lovell, 2003; Crowe, 2007; Farhang, 2008, 2010). This
scholarship focuses on the “relationships between justices and elected officials,”
searching out ways in which elected officials “encourage or tacitly support judicial pol-
cymaking” (Graber, 1993:36-37). The result is a way of studying and explaining judi-
cial policymaking in which scholars investigate the numerous mechanisms through
which legislators share power, responsibility, and blame with judges (Graber, 1993:37;
Lovell, 2003:20). This scholarship has also been attentive to reasons why Congress
may create opportunities for judicial policymaking, which includes advancing political
and policy goals (Graber, 1993; Gillman, 2002; Lovell, 2003; Smith, 2005; Farhang,
2008, 2010). Legislators may wish to shift decision-making responsibility from the leg-
islature to the courts (Graber, 1993) or to lock in and advance policy goals (Gillman,
2002), some of which may be difficult for legislators to achieve on their own (Graber,
1993; Lovell, 2003). Legislators may additionally look to courts as institutions that
can provide an oversight function with respect to the executive branch (Shidan, 2000;
Smith, 2005).

The present study adds to our understanding of judicial empowerment by focusing
on an additional mechanism through which Congress empowers the judiciary.
Congress makes choices about the jurisdiction of the federal courts by passing *jurisdiction-granting legislation*. As defined here, jurisdiction-granting legislation explicitly expands judicial discretion by designating courts as venues through which certain specified categories of people or organizations may work to address their claims. As a policy venue, or an “institutional location where authoritative decisions are made concerning a given issue,” courts are established as institutions where certain policy questions are to be answered (Baumgartner and Jones, 1993:32).

Congress can take the opposite action as well by removing jurisdiction from the federal courts by passing jurisdiction-stripping legislation, a frequent subject of scholarship (see, e.g., Gunther, 1984; Weiman, 2005; Peabody, 2006; Chutkow, 2008; Olsen, 2009; Fallon, 2010). It is recognized that Congress has the constitutional power to remove jurisdiction, though there is disagreement over whether Congress’s power is plenary or limited by either the language of Article III or other provisions of the Constitution (Gunther, 1984:900, 908). With the exception of Chutkow (2008), studies assume that removal of jurisdiction from the Supreme Court rarely occurs (Epstein and Knight, 1998:143; Weiman, 2005:1678-79; Fallon, 2010:1045).

The expansion and contraction of federal court jurisdiction has not gone unnoticed in scholarship on the federal courts. Crowe (2007) links the passage of the Judges’ Bill of 1925 and its expansion of the Supreme Court’s discretionary appellate jurisdiction (through the *writ of certiorari*) to the institutional development of the judiciary. Curry (2007) examines the “politics of federal diversity jurisdiction,” finding that “jurisdictional issues present important opportunities for Congress and its members to oversee the federal courts” (p. 463). That Congress has the capacity and willingness to define the jurisdiction of the federal courts, either by granting or stripping jurisdiction, speaks to important institutional links between Congress and the federal courts. However, scholarship on the judiciary varies to the degree with which it recognizes these links between Congress and the courts. The attitudinal model, for example, assumes that Supreme Court justices are independent from Congress (and the president and public as well) and have fixed powers. Members of the Supreme Court, the focus of the attitudinal model, are “life-tenured . . . lack electoral or political accountability” and enjoy discretionary control over their docket (Segal and Spaeth, 2002:92). This allows Supreme Court justices to act free from outside influence at both the certiorari and merits stages of the judicial process, with their decisions being the product of their individual, personal deliberations and policy preferences.1 Scholarship additionally varies to the degree with which it recognizes the role that Congress performs in determining the jurisdiction of the courts. While scholars might acknowledge that Congress establishes jurisdiction and that Congress has control over, for example, the Supreme Court’s appellate jurisdiction (Segal and Spaeth, 2002:230),

1 As a point of contrast, consider Epstein and Knight’s strategic account, in which the authors argue that the justices must take into consideration the “preferences and expected actions” of other branches as well as the “greater social and political context of society as a whole” during the certiorari and decision stages of the judicial process (1998:138).
they do not investigate the implications of jurisdiction-granting for how courts participate in the policy process.

The present study addresses a gap in our understanding of the relationship between congressional action and judicial empowerment by investigating the conditions under which Congress passes jurisdiction-granting legislation. While previous scholarship has tended to focus on periodic (though important) “Judiciary Acts,” in this article I evaluate the passage of a comprehensive set of laws that grant jurisdiction to the federal district courts between the 81st and 109th Congresses (1949-2006). Jurisdiction-granting laws are consistently and routinely passed (see Figure 1). The number of laws passed ranges as high as 50 in the 93rd Congress, and at least 12 laws are passed per Congress.

These data support the argument that Congress exercises its constitutional power to determine the jurisdiction of the federal district courts. Given the frequency with which it occurs, it illustrates the importance of focusing on the conditions leading to the passage of jurisdiction-granting legislation. As Congress frequently grants jurisdiction to the district courts, it is an active participant in deciding whether and how the courts will participate in the policy process. In the pages that follow, I build upon these initial empirical findings in an attempt to further demonstrate the importance of studying congressional efforts to empower the courts. I first discuss existing literature on why courts become involved in the policy process. Predominately, scholarship on the courts minimizes the relationship between Congress and the courts, obscuring the numerous ways in which the institutions are linked. This section works to establish a theoretical justification for examining congressional grants of jurisdiction. Then, two explanations that might explain the passage of jurisdiction-granting legislation are considered. I explore the implication of separation-of-powers and ideological concerns for the passage of jurisdiction-granting laws. Along with a new data set of jurisdiction-granting laws introduced above, I construct a measure of the median judge ideology per Congress (Poole, 1998, 2009; Giles, Hellinger, and Peppers, 2001; Epstein et al., 2007; Boyd, 2010). The results support the conclusion that separation-of-powers and ideological concerns are important to understanding judicial empowerment.

**Courts in the Policy Process**

Motivating scholarship on the federal judiciary is the question of how courts come to decide certain public policy questions. As noted in more detail below, the behavioralist and interbranch perspectives provide different accounts of how courts become involved in the policy process. For example, the behavioralist approach minimizes the role that Congress plays in the judicial agenda-setting process. With ongoing debates regarding the role of courts in American politics, debates that often concern whether courts are “counter-majoritarian,” it is important to examine assumptions underlying explanations of judicial participation in American politics.
In the behavioralist tradition, most often associated with the attitudinal model and studies of the U.S. Supreme Court, it is thought that Supreme Court justices come to decide certain public policy questions primarily because justices are in an institutional position that allows them to act on their preferences (e.g., Pacelle, 1991; Segal and Spaeth, 2002; Lanier, 2003; Baird, 2007). For the lower federal courts, the picture is somewhat more complicated as lower courts do not enjoy discretion in the way that Supreme Court justices do. However, even district court judges have discretion in their determinations of whether jurisdiction exists and if standing can be granted (Rowland and Todd, 1991:177; Kim et al., 2009:85). Furthermore, district court judges may become involved in the policy process in part because legislation may create a “fiduciary relationship” that courts may have to enforce or because judges themselves create new obligations, such as expanding access to courts through standing doctrines (Rowland and Carp, 1996:7).\(^2\) Whether scholarship investigates Supreme

\(^2\) In this manner, Rowland and Carp acknowledge the role of legislation in creating opportunities for judges to exercise discretion. According to the authors, “fiduciary jurisprudence has been characterized by legislation . . . that assigns standards of care rather than creating specific obligations” (p. 7). While the authors recognize a link between legislation and the exercise of judicial discretion, they do not explore the link between jurisdiction-granting legislation and judicial empowerment.
Court or lower-court policymaking, emphasis is placed on decisions made by judges as they determine which issues to focus on.

In a separate tradition of scholarship, scholars investigate how opportunities for courts to review the constitutionality of statutes or to interpret the meaning of legislative language arise, and conclude that members of Congress work actively to create such opportunities (Graber, 1993; Shipan, 2000; Lovell, 2003). In doing so, Congress empowers the federal judiciary to participate in the policy process. Constitutionally, Congress has a host of powers with respect to the federal courts covering whether courts will exist and in what form (creation), who will occupy the judges’ chair (staffing), what resources they will have to complete their tasks (funding), and what issues they will attend to (jurisdiction). In this view, Congress makes antecedent decisions that structure the ability of judges to participate in the policy process. Judges do make decisions regarding which issues their courts will attend to, but these decisions stand on a legal field constructed by Congress.

Following the tradition of examining judicial empowerment by Congress, the present study focuses on jurisdiction-granting legislation as one tool that Congress may use to empower the federal courts to participate in the policy process. Much like the drafting of ambiguous statutes (Graber, 1993; Lovell, 2003) or political party efforts to institutionalize policy gains through the creation and staffing of courts (Gillman, 2002), jurisdiction-granting legislation represents choices made by members of Congress regarding whether to empower the judiciary. These decisions are part of the legislative process (Smith, 2006). In drafting legislation, members of Congress decide whether the courts are to be granted jurisdiction and in what form such grants will take. Courts may come to play peripheral or central roles in the public policy process depending, in part, on how Congress envisions their role. While previous scholarship has focused on important moments in which Congress empowers the courts, the research presented here highlights the manner in which these decisions are in fact frequent and routine (see Figure 1). This is important as it speaks to the ongoing relationship between Congress and the courts in that Congress often takes affirmative steps to create opportunities for courts to participate in the policy process.

What emerges is a more complex picture of the reasons why courts are involved in deciding public policy questions. In this conception, judicial forays into the policy thicket are not the result of desire on the part of judges alone, but a function as well of legislatively created opportunities. Courts are able to review the language of statutes because Congress explicitly includes provisions granting them the authority to do so. Complexity arises from the process of tracing judicial decisions back to decisions made by other institutional actors to involve the courts in the policy process. Additionally, when courts do act it does not automatically follow that the effect of their decisions is to thwart the will of a legislative majority (Dahl, 1957; Gillman and Clayton, 1999; Lovell, 2003).

Building on this theoretical background, the next section explores in detail the reasons why Congress might grant jurisdiction to the federal district courts. It
examines the relationship between separation-of-powers concerns, as well as judicial ideology, to explain the passage of jurisdiction-granting legislation.

**THE DECISION TO GRANT JURISDICTION**

Under what conditions will Congress empower the federal courts? Expansions of judicial power have been studied elsewhere as “the sort of familiar partisan or programmatic entrenchment that we frequently associate with legislative delegation to executive or quasi-executive agencies” (Gillman, 2002:512). In treating the decision to empower the federal courts as analogous to the decision to delegate to administrative agencies, this article draws on parallels between the two situations. In both situations, Congress determines the costs and benefits of producing policy “in-house” versus the costs and benefits of allowing another entity to produce policy. Congress must overcome time, resource, and expertise challenges that would otherwise prevent the legislative body from writing complex, detailed legislation itself. Producing public policy is a time-consuming enterprise, and Congress’s attention is split between numerous activities. Institutional features, such as the committee system, provide members with tools to overcome cognitive limitations; nonetheless, Congress cannot attend to all issues equally, and producing legislation is a costly enterprise. Thus, for Congress, resolving policy questions might involve granting authority to a court to implement the policy given the constraints described above. Additionally, in making the decision to empower the courts, Congress must also take into account the relative ideological positions of entities that might become targets of judicial oversight, as well as the ideological positions of courts themselves. In the pages that follow, I build on the theoretical insight that delegating to federal courts mirrors the decision to delegate to administrative agencies.

**RESEARCH DESIGN**

**Jurisdiction-Granting Legislation.** Using HeinOnline’s U.S. Statutes at Large database, I identified public laws with provisions that grant jurisdiction to the federal

---

3 Initially, delegation could be justified as acceptable, as Congress would be delegating to experts with the knowledge and drive to produce sound public policy (Landis, 1938). However, scholars, politicians, and the public, after World War II, started to question fundamental assumptions that motivated earlier visions of delegation, especially that barriers against participation were low and that those wishing to participate in the agency process were able to participate. Administrative agencies were increasingly seen as overly beholden to the interests of industry and as no longer serving the interests of the public (Shapiro, 1988), and decision making was increasingly regarded as taking place behind closed doors, or behind the aegis of agency discretion (Stewart, 1975). In the 1960s and 1970s, members of Congress were motivated by a desire to provide more direction to agencies to prevent agencies from straying too far from the stated purpose of legislation, and to control generally the executive branch (Melnick, 1994:28-29). Delegation takes place in present times against this historical backdrop of looking to benefit from agency expertise while at the same time mitigating against agency capture.

4 Additionally, Congress might be motivated by a desire to transfer responsibility or blame from itself to the courts as alternate policymaking entities (e.g., Graber, 1993). This is a compelling argument, and it presents a possible explanation for judicial empowerment that I do not reject. However, the data collected for this study show that Congress routinely and frequently grants jurisdiction to the federal district courts, suggesting that Congress may take into account additional factors when granting jurisdiction.
district courts for the years 1949-2006 (81st–109th Congress). I identified such laws using the advanced search feature of the database to search for the terms “district court” and “district courts.” Through this method, I identified a comprehensive set of 747 laws that grant jurisdiction to the federal district courts. In constructing this set of laws and in the analysis that follows, I focus on the federal district courts. This reflects the importance of the district courts in American politics in that these courts decide more cases across more policy areas compared to other federal courts. As these decisions often go unreviewed by higher courts (Rowland and Carp, 1996:8), the decisions of the district courts affect all aspects of Americans’ lives. Furthermore, the federal district courts are important players in the public policy process, ones that often go overlooked in scholarship on the federal judiciary.

The dependent variable All Laws (N = 747 laws) is a count of the number of jurisdiction-granting laws passed per Congress. I searched each law for specific references to the federal district courts, and coded provisions of each law according to one of five categories through which the district courts’ jurisdiction was expanded by Congress. Congress has many options at its disposal for figuring out how courts will participate in the policy process. However, as discussed in more detail below, not all laws or provisions equally create a role for the courts, and disaggregating them allows for more nuanced analysis of their effects.

The specific subcategories are as follows. First, laws granted jurisdiction to the federal district courts by authorizing civil actions by the United States attorney general or by a federal agency. For example, the Housing Act of 1954 states that “[w]henever he finds a violation of any provision of this section has occurred or is about to occur, the Attorney General shall petition the district court of the United States or the district court of any territory,” which are granted jurisdiction to “hear, try, and determine such matter[s]” (Housing Act of 1954:612). The dependent variable Authorize (N = 738 provisions) is a count of the number of provisions similar to the one cited above.

Second, federal district courts were at times designated as venues in which persons affected by agency regulations could obtain judicial review of agency rules, regulations, or other agency decisions (DV: Review; N = 123 provisions). According to the Marine Mammal Protection Act of 1972, for example, “[a]ny applicant for a permit, or any party opposed to such permit, may obtain judicial review of the terms and conditions of any permit issued by the Secretary under this section or of his refusal to issue a permit. Such review, which shall be pursuant to chapter 7 of title 5, United States Code, may be initiated by filing a petition for review in the United States district court for the district wherein the applicant for a permit resides” (Marine Mammal Protection Act of 1972:1035). Third, in instances in which parties needed to figure out the value of goods and services, the district courts were granted jurisdiction to judicially determine those values (DV: Determine; N = 61 provisions). Thus, for example, the Special Health Revenue Sharing Act of 1975 states that the “United States shall be entitled
to recover . . . an amount bearing the same ratio to the . . . value (as determined by the agreement of the parties or by action brought in the United States district court for the district in which the center is situated)” as an amount of the project in question that constituted the federal share (Special Health Revenue Sharing Act of 1975:327). Fourth, a law could authorize a district judge to issue warrants, to seize goods for libel, or to enforce subpoenas (DV: WLS; N = 198 provisions). For example, “[a] fishing vessel … used in the commission of act prohibited by section 307 shall be liable in rem for any civil penalty assessed for such violation under section 308 and may be proceeded against in any district court of the United States having jurisdiction thereof” (Fisheries and Conservation Management, 1986:3714). Fifth, I coded provisions as falling into a general jurisdiction category when provisions granted jurisdiction without falling into one of the previous categories (DV: Jurisdiction; N = 359 provisions). For example, the Civil Rights Act of 1964 includes a provision that states that “the district courts of the United States shall have jurisdiction of proceedings instituted pursuant to this title” (Civil Rights Act of 1964:245).

While the data described above comprise a comprehensive set of laws and provisions that grant jurisdiction to the federal district courts, the measures do not capture the full range of ways in which Congress empowers the district courts. Other measures include expansions of the judiciary through increasing the number of judgeships or creating new courts (De Figueiredo and Tiller, 1996; Gillman, 2002); encouraging litigation through “private enforcement regimes” (Farhang, 2008, 2010); authorizing class-action suits (Barnes, 2007); or otherwise empowering the courts through deferring to the courts on policy questions (Graber, 1993; Lovell, 2003). Furthermore, these measures do not account for the relative importance of legislation. For example, the Civil Rights Act of 1964, and specific provisions of the act (mentioned above), is one of the most important pieces of legislation passed in American history, but the lack of a system by which to measure the importance of legislation means that the analysis that follows it is treated the same as all other legislation passed during the 88th Congress.5

A related issue is that some legislation may by design lead to certain types or even more litigation compared to other legislation. Some laws create mechanisms for expedited judicial review of their constitutionality.6 For example, the Bipartisan Campaign Finance Reform Act of 2002 creates a procedure for judicial review of the constitutionality of any provision of the act in the United States District Court for the District of Columbia, with appeal directly to the Supreme Court (Bipartisan Campaign Reform Act of 2002:113-14). Combined, the high-profile nature of the issues involved in acts such as this one and the specific procedures laid out in the act might encour-

---

5 One possible measure of the importance of legislation is Mayhew’s (1991) data set of important enactments by Congress. Mayhew uses contemporary and retrospective judgments of the importance of laws to construct a list of important legislation (see, e.g., Chapter 3). However, this data set covers solely the years between 1946-1990.

6 I thank one of the anonymous reviewers for this suggestion.
age litigation in the federal district courts. Other provision types, such as the power of the district courts to enforce subpoenas, may not similarly encourage more litigation. It is additionally possible that Congress may take certain factors into account for more important laws and may empower the district courts in situations in which its membership or even the public are more aware of policy considerations. Finally, these measures capture only expansions of judicial power and do not account for the numerous ways in which Congress may limit the power of the federal courts through stripping jurisdiction (Curry, 2007; Chutkow, 2008).

**Jurisdiction-Granting Legislation: Separation of Powers and Judicial Ideology.** The congressional decision to grant jurisdiction to the federal district courts comes against the background of a key characteristic of the U.S. political system: separation of powers (Burke, 2002; Kagan, 2003; Farhang, 2008, 2010). By separation of powers, I mean the existence of three coequal branches operating in a system of checks and balances in which power and responsibility are dispersed but in which each participates in the public policy process. This article builds on earlier findings that support the conclusion that separation of powers may encourage, when certain conditions are present, grants of power to the federal courts. These explanations generally focus on partisan congruence or ideological distance between the branches of government to explain judicial empowerment. Judicial empowerment may encompass a number of different, and sometimes contradictory, goals. The discussion that follows explores a number of these goals in the context of partisan congruence between Congress, the president, and the federal district courts, as well as ideological distance between these players.

**Alignment.** While divided government is a characteristic of our political system, alignment between the various players—the U.S. House of Representatives, the U.S. Senate, and the president—is also an empirical fact, and alignment has been associated with expansions of judicial power. De Figueiredo and Tiller (1996) explain the timing and size of expansions of the federal judiciary in terms of alignment between five players: the enacting House, Senate, and president, and the nominating president and confirming Senate. Gillman (2002) describes successful Republican Party efforts to expand the size and jurisdiction of the federal judiciary as stemming from their control of the House, Senate, and presidency. As with other expansions of judicial power, decisions to grant jurisdiction to the federal district courts might similarly rely on alignment between the House, Senate, and president.

At the same time, the presence of divided government may also be associated with expansions of judicial power as Congress may look to the courts to provide oversight with respect to executive-branch administrative agencies. Epstein and O’Halloran (1999) first discuss the implications for congressional delegation to the executive branch in times of divided government, finding that “Congress delegates less . . . under divided government” (pp. 77-81, 129-38). Noting, however, that delegation may still take place in times of divided government, the authors further examine the type of delegation that occurs and find that Congress does delegate to “non-executive”
entities, including courts, in times of divided government (Epstein and O’Halloran, 1999:154). In his examination of the potential oversight role of federal courts with respect to administrative agencies, Shiman (2000) argues that members of Congress may work to increase or decrease the participation of the courts in the policymaking process depending on their evaluation of the relative policy positions of Congress, courts, and agencies. Therefore, Congress may look to federal courts, including the district courts, to check the decisions made by executive-branch administrative agencies when there is divergence between the policy positions of the legislative and executive branches.

The literature suggests the following hypotheses:

H1: When branches of government are aligned, it will be more likely for Congress to pass jurisdiction-granting legislation.

H2: As the ideological distance between Congress and the president increases, Congress will be more likely to pass jurisdiction-granting legislation.

These hypotheses suggest contradictory outcomes—that both alignment and the lack of alignment can lead to judicial empowerment. However, recall Epstein and O’Halloran’s (1999) work in which they find that delegation takes place in periods of unified and divided government, but the entities to which Congress delegates are different under divided government. In the context of the work presented here, the delegation literature thus suggests that granting authority to the federal district courts is a routine and frequent occurrence but that motivations for doing so may differ when Congress and the president are aligned than when they are not. Additionally, alignment captures merely whether the branches are of the same party, which masks any independent effect of actual ideological convergence or divergence between the branches.

Judicial Ideology. As noted by Farhang (2008), the ideological position of the federal judiciary may influence legislators’ calculations regarding empowering the judiciary. This may manifest in two ways. First, when the judiciary is of the same “party” as Congress and the president, empowerment may be more likely as Congress and the president are more certain that the judiciary’s policy outputs will mirror their preferences. Second, the direction in which the judiciary is trending will affect legislator evaluations of the costs and benefits of granting policymaking authority to the courts (Chutkow, 2008; Farhang, 2008:826).

This suggests two hypotheses:

H3: When Congress, the president, and the federal district courts are of the same “party,” it will be more likely for Congress to pass jurisdiction-granting legislation.

H4: As the judiciary becomes more ideologically distant from both Congress and the president, it will be less likely for Congress to pass jurisdiction-granting legislation.

Data and Measures. I construct measures that use both ideological and partisan measures of the positions of Congress, the president, and the federal district courts
To measure the ideological position of the U.S. House (House Median) and U.S. Senate (Senate Median), I use Poole and Rosenthal’s DW Chamber median scores. I use Common Space Scores for the president (President). These scores were used to construct the following variables.

Judicial Ideology. Reflecting interest in the manner in which federal judges make their decisions, scholarship has identified a relationship between judges’ political ideology and their substantive decisions. This literature extends to the federal district courts. Opportunities for district court judges to exercise discretion do exist, over both procedural and substantive questions (Rowland and Todd, 1991:177; Kim et al., 2009:85). Political party affiliation has been correlated with standing decisions of liberal and conservative judges (Rowland and Todd, 1991) and with the “civil liberalism” of the federal district courts in times when ambiguity from the Supreme Court creates legal uncertainty (Rowland and Carp, 1980:294, see also Rowland and Carp, 1996).

When judges recognize the potential to exercise discretion, “they will be more likely to give vent to their own set of personal attitudes” (Rowland and Carp, 1996:41), attitudes that have been found to align with those of their appointing president (Rowland and Carp, 1980, 1983; Stidham, Carp, and Rowland, 1984).

However, the president does not act alone in picking nominees for the federal district courts. Senators, especially those of the same political party as the president and those who represent the state in which a district court appointment arises, may participate in choosing nominees (Rowland and Carp, 1996:Chapter 4). Senatorial courtesy is thought to be active when one or both of the home-state Senators are of the same political party of the president, and the assumption is that under this circumstance, the nominee reflects in part the senators’ input.

I constructed a measure of the ideological makeup of the federal district courts for the 81st–109th Congresses (1949-2006) using data made available by Boyd (2010), in which she calculated ideology scores for federal district judges. As I am interested in the ideological position of the federal district courts as a whole, I used the Federal Judicial Center’s History of the Federal Judiciary to identify the service dates of each federal district judge included in Boyd’s database. Averaging across the judges serving on the district courts per Congress produces the variable Judge Median. Negative values correspond with judicial liberalism and positive values correspond with judicial conservatism (see Figure 2).

I focus on the federal district courts and do not account for the ideology of circuit court judges or Supreme Court justices, both of which oversee decisions made by the district courts. It could be argued that the courts of appeals and the Supreme Court are participants on appeal in cases that stem from jurisdiction-granting legislation, meaning that Congress might consider the ideology of the courts of appeals and the Supreme Court when designing this legislation. Congress might

---

7 These data were downloaded from www.voteview.com, last visited February 19, 2012.
take into account where they stand on the ideological spectrum along with the ideology of the district courts, knowing that these courts might be involved in deciding certain policy questions. I argue, however, that a category of decisions are being made primarily by the federal district courts with little oversight from higher courts, justifying focusing on the ideology of federal district judges only. That is not to argue that the federal district courts always operate free from appellate review, but rather (as noted above) that it is more likely that their decisions will not be reviewed by a higher court (Rowland and Carp, 1996:8). I argue, therefore, that Congress considers the ideology of the district courts when passing jurisdiction-granting legislation because it is aware that these courts that will be making the bulk of policy decisions in these cases.

Alignment. The variable Alignment takes on a value of 1 when the House, Senate, and president, nominating president, and confirming Senate are controlled by the same party and 0 otherwise. Moving from the lack of alignment to alignment should result in the passage of more jurisdiction-granting legislation, a positive coefficient. The variable Judicial Direction takes on a value of 0 when the Congress, president, and federal district courts are of the same party and 1 otherwise (Farhang, 2008:831). As discussed above, when Congress and the president are of different parties, Congress may still grant jurisdiction to the federal district courts so that the district courts might serve an oversight function. Here, the coefficient should also be positive. In addition to these partisan variables, I calculated the ideological distance between Congress and the president (Congress-President Distance). As distance increases, the likelihood of passage of jurisdiction-granting legislation should increase (a positive coefficient). Congress-President-District Court Distance captures the relative ideological position of Congress and the president versus the federal district courts, with the measure of Congress's ideological position taken as the average of the Senate median and House median. For this variable, as distance increases, the number of grants of jurisdiction should decrease (Chutkow, 2008).

Controls. Congress may be less likely to grant new policymaking authority to the district courts if they perceive the courts as already dealing with a high number of cases. The model controls for the number of federal-district-court civil cases, denoted as Cases, filed annually, in thousands of cases (Administrative Office of the United States Courts, 1940-2006). Additionally, the model controls for the total number of laws passed per congressional session (Total Laws) to account for the possible effects of otherwise changing congressional activity on the number of jurisdiction-granting laws passed. In a similar vein, I control for the passage of Time to account for the possible independent effect of time on the incidence of granting jurisdiction to the district courts. Additionally, I separately control for the distance between the president (President-District Court Distance), Senate (Senate-District Court Distance), and House (House-District Court Distance) and the district courts. The House, Senate, and president may evaluate the ideological distance between the respective branches and the federal district courts when determining whether to enact jurisdiction-granting legislation.
Results and Discussion. Grants of jurisdiction are evaluated in light of separation-of-powers concerns and judicial ideology. Count data are used here. I employ a poisson regression to model the hypothesized relationships and, having found evidence of serial correlation, corrected standard errors when appropriate using the technique advanced by Newey and West (1987).

The results support the conclusion that separation-of-powers and ideological concerns inform the legislative choice to empower the federal district courts (see Table 1). For the passage of jurisdiction-granting laws in general (All Laws), a one-unit increase in the ideological distance between Congress and the president is associated with the passage of approximately eight additional jurisdiction-granting laws, all else equal. As hypothesized (H2), the results suggest that Congress may empower the federal district courts to perform an oversight function with respect to the executive branch in situations in which the ideologies of Congress and the president are diverging. For the subcategories, this relationship is significant across the board with the exception of laws granting courts the power to exercise judicial review (Review). This last finding is surprising because a function of judicial review is often to allow for the judiciary to examine executive-branch decisions, something that Congress may be
Table 1
Number of Jurisdiction-Granting Laws Passed 1949-2006 (81st–109th Congress)

<table>
<thead>
<tr>
<th>Alignment</th>
<th>All Laws+</th>
<th>Authorize+</th>
<th>Determine+</th>
<th>Review+</th>
<th>Jurisdiction+</th>
<th>WLS++</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congress-President</td>
<td>8.061***</td>
<td>19.177***</td>
<td>14.397*</td>
<td>5.702</td>
<td>12.041***</td>
<td>12.741***</td>
</tr>
<tr>
<td>Distance</td>
<td>(1.821)</td>
<td>(2.442)</td>
<td>(6.655)</td>
<td>(3.323)</td>
<td>(2.660)</td>
<td>(3.902)</td>
</tr>
<tr>
<td>Judicial Direction</td>
<td>-0.291</td>
<td>-0.661</td>
<td>1.037</td>
<td>0.326</td>
<td>0.417</td>
<td>0.415</td>
</tr>
<tr>
<td></td>
<td>(0.275)</td>
<td>(0.381)</td>
<td>(0.930)</td>
<td>(0.754)</td>
<td>(0.398)</td>
<td>(0.479)</td>
</tr>
<tr>
<td>Congress-President-District Court Distance</td>
<td>-7.987***</td>
<td>-18.768***</td>
<td>-13.371*</td>
<td>-5.516</td>
<td>-11.444***</td>
<td>-12.067**</td>
</tr>
<tr>
<td></td>
<td>(1.883)</td>
<td>(2.433)</td>
<td>(6.727)</td>
<td>(3.370)</td>
<td>(2.663)</td>
<td>(4.007)</td>
</tr>
<tr>
<td>President-District Court Distance</td>
<td>0.225</td>
<td>0.937*</td>
<td>0.128</td>
<td>0.956</td>
<td>-0.432</td>
<td>-0.959</td>
</tr>
<tr>
<td></td>
<td>(0.398)</td>
<td>(0.379)</td>
<td>(1.247)</td>
<td>(0.627)</td>
<td>(0.492)</td>
<td>(0.884)</td>
</tr>
<tr>
<td>House-District</td>
<td>-6.548***</td>
<td>-14.489***</td>
<td>-4.6</td>
<td>-7.961*</td>
<td>-5.476***</td>
<td>-8.818**</td>
</tr>
<tr>
<td>Court Distance</td>
<td>(1.27)</td>
<td>(1.697)</td>
<td>(3.92)</td>
<td>(3.213)</td>
<td>(1.568)</td>
<td>(3.083)</td>
</tr>
<tr>
<td>Court Distance</td>
<td>(1.018)</td>
<td>(1.757)</td>
<td>(3.967)</td>
<td>(1.381)</td>
<td>(1.594)</td>
<td>(1.887)</td>
</tr>
<tr>
<td>Cases</td>
<td>-1.88x10-6*</td>
<td>-5.22x10-6*</td>
<td>2.87x10-6</td>
<td>-1.16x10-6</td>
<td>-4.21x10-7</td>
<td>-2.73x10-6</td>
</tr>
<tr>
<td></td>
<td>(9.43x10-7)</td>
<td>(1.52x10-6)</td>
<td>(2.37x10-6)</td>
<td>(1.27x10-6)</td>
<td>(1.03x10-6)</td>
<td>(1.61x10-6)</td>
</tr>
<tr>
<td>Total Laws</td>
<td>0.0004</td>
<td>0.002</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0009)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.0007)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Time</td>
<td>0.049*</td>
<td>0.180***</td>
<td>-0.167**</td>
<td>0.133*</td>
<td>0.02</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.050)</td>
<td>(0.062)</td>
<td>(0.059)</td>
<td>(0.0256)</td>
<td>(0.44)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.906***</td>
<td>1.457</td>
<td>1.819</td>
<td>-2.111*</td>
<td>1.744***</td>
<td>3.033***</td>
</tr>
<tr>
<td></td>
<td>(0.357)</td>
<td>(0.841)</td>
<td>(1.446)</td>
<td>(0.860)</td>
<td>(0.549)</td>
<td>(0.591)</td>
</tr>
</tbody>
</table>

Notes: p<0.05*, p<0.01**, p<0.001***. +Poisson regression with uncorrected standard errors in parenthesis. ++Poisson regression with Newey-West corrected standard errors in parenthesis. N of Congresses = 29.

interested in under conditions in which Congress and the president are ideologically divergent.

Alignment between the House, Senate, president, nominating president, and confirming Senate failed to reach statistical significance for All Laws (H1) but was found to be associated with the passage of approximately 1 Authorize provision fewer and 2 additional Determine provisions, all else equal. While other studies have found alignment to be associated with increasing the size of the judiciary, the results reported here suggest that alignment is not associated with increasing the authority of the federal district courts.

The measure of Judicial Direction, which accounts for situations in which the federal district courts are of the same “party” as the rest of government (H3), failed to
reach conventional levels of statistical significance across the six categories. Compared to the ideological measures of distance, this was a rough measure of the partisan relationship between the district courts and the other branches of government. In contrast, for All Laws, a one-unit increase in the relative ideological distance between Congress, the president, and the federal district courts (with the courts moving farther away from Congress and the president) is associated with the passage of approximately eight fewer jurisdiction-granting laws, all else equal (H4). Additionally, a one-unit increase in this relative distance is associated with the passage of approximately 19 fewer Authorize provisions, 14 fewer Determine provisions, 11 fewer Jurisdiction provisions, and 12 fewer WLS provisions (with Review failing to reach statistical significance).

Together, these results suggest that Congress empowers the federal district courts through the mechanism of jurisdiction-granting legislation. Congress is more likely to pass such legislation in periods of divided government in which Congress may look to the district courts to examine executive-branch activity. Congress is less likely to pass such legislation when the judiciary is ideologically distant or becomes increasingly ideologically distant from Congress and the president.

CONCLUSION

These results suggest a number of conclusions regarding the relationship between Congress and the federal district courts and the federal judiciary. First, Congress frequently grants jurisdiction to the federal district courts. This suggests an active role for Congress in determining the boundaries of what courts are able to accomplish. The district courts are not self-organizing but, rather, depend on Congress to make choices regarding their jurisdiction. Congress frequently makes such choices, granting jurisdiction to the federal district courts in at least twelve laws per congressional session. Coupled with the frequency with which Congress strips jurisdiction from the federal courts (Chutkow, 2008), overrides the Supreme Court (Eskridge, 1991), or otherwise impacts the ability of individuals and groups to use the courts (Barnes, 2007; Farhang, 2008, 2010), the data collected for this study speak to the consistent attention given to the courts by Congress.

Second, by granting jurisdiction, Congress is taking action that empowers the federal district courts to participate in the policy process. Congress makes choices to establish courts as venues through which certain questions are to be asked and answered, either by establishing courts as primary venues to which individuals and groups originally turn or by creating an oversight role for courts as part of the regulatory process. Recognizing these causes of judicial participation in the policymaking process is important to understanding trends in courts’ involvement in the policy process. It has long been recognized that federal courts at all levels are increasingly involved in deciding important policy questions. Compared to earlier periods in United States history, courts today are involved in multiple aspects of Americans’ lives.
and even have taken up issues heretofore considered the sole provenance of the legislative and executive branches.

Known as the “judicialization of politics” (Vallinder, 1995; Clayton, 2002; Ferejohn, 2002), this process has resulted in the involvement of the federal courts in the “administration and operation of schools, prisons and jails, mental health centers, public housing authorities, and juvenile detention facilities” (Taggart, 1989:242) and in numerous policy areas, such as environmental law and abortion policy (Rowland and Carp, 1996:6-10). The present study argues that trends associated with the judicialization of politics might be attributed to decisions made by Congress regarding how courts participate in the policy process. If Congress plays an active role in granting courts jurisdiction, understanding what issues courts consider becomes more complicated than under a model in which courts are solely responsible for setting their agendas. The results here suggest an antecedent stage to the judicial agenda-setting process in which Congress’s choices influence who may sue, when they may sue, and which issues they may raise.

Third, decisions of whether to grant jurisdiction to the courts are part of the legislative process (Lovell, 2003; Smith, 2005). Congress, therefore, serves as an important link between the public, their elected representatives, and the courts. When courts participate in the policy process, they are doing so in part because of decisions made by elected officials. Congress has opportunities throughout the legislative process to consider the implications of its decisions regarding the participation of the courts in the policy process. While many scholars bemoan the lack of meaningful checks on the judiciary, the frequency with which Congress grants (and removes) jurisdiction demonstrates that Congress is not powerless when it comes to the federal courts.

The analysis presented here, while providing an explanation of an additional mechanism through which Congress empowers the federal district courts, leaves some questions unanswered. Most important, it is unclear exactly what motivates Congress to empower the district courts in periods of alignment versus divided government. A closer examination of the congressional record that compares the granting of authority to district courts across unified and divided government would provide more insight into this question.

This study contributes to a more comprehensive picture of the judicial agenda-setting process and argues for consideration of Congress’s role in establishing the jurisdiction of the federal courts. Judges set their agendas in response to a number of influences, which includes the jurisdiction granted to them by Congress. Without these congressional grants, courts would have no power or authority. With these grants, litigants, acting as individuals or as part of groups, take advantage of legislatively created opportunities and bring their cases to court with an understanding of what courts are able to do under the law. jsj
REFERENCES


**Statutes Cited**


Ambition is theorized to motivate many types of political behavior. However, as a motivation for federal judges, ambition has been largely ignored. Scholars have dismissed the possible influence of ambition by citing the low probability of receiving a promotion to a higher bench. Although a promotion is unlikely, judges may place substantial value on being promoted. Using a measure designed to capture a judge’s promotion potential, I investigate the possibility that ambition exerts an independent effect on judicial behavior. The results demonstrate that court of appeals judges alter their behavior in a manner consistent with attempting to secure a seat on the United States Supreme Court.

Find a judge who admits to political calculation or, worse, ambition, and you’ve found one who may shortly be blackballed for violating the lodge rules.

Donald Jackson, Judges

The first half of 2005 was an eventful period in the life of John G. Roberts, then judge of the U.S. Circuit Court of Appeals for the District of Columbia. On Friday July 2, 2005, Associate Supreme Court Justice Sandra Day O’Connor officially announced her resignation effective upon the confirmation of her successor. Immediately speculation began as to whom President Bush would nominate as a possible replacement. News organizations began generating unofficial “short lists” of candidates likely to be selected to replace O’Connor. John Roberts’s name was featured prominently on many.1

Speculation surrounding Judge Roberts was not unfounded. The Bush administration had viewed Roberts as a possible Supreme Court nominee long before Justice O’Connor’s resignation. By the end of March 2005, the Bush White House and the Department of Justice, fearing that an ill and aged Chief Justice William Rehnquist

might soon resign or pass away, began to interview possible Supreme Court replacements (Greenberg, 2007). John Roberts was first interviewed on April 1, 2005 by then Attorney General Alberto Gonzales. According to his questionnaire submitted to the Senate Committee on the Judiciary, Roberts recalls being reinterviewed on two separate occasions before Justice O’Connor’s resignation on July 1 and then “several” more times (Greenberg, 2007:55), including a July 15 interview with President George W. Bush, before ultimately being announced as the official nominee on July 19, 2005. His nomination would later be withdrawn and reintroduced to fill the vacancy left by Chief Justice William Rehnquist, who died on September 3, 2005.

Amid growing Supreme Court speculation, Roberts continued to fulfill his role as a judge on the D.C. Circuit. In December of 2004 Roberts had been assigned to a panel that would hear Hamdan v. Rumsfeld (415 F.3d 33 D.C. Cir.), the most prominent challenge to the constitutionality of the military commissions set up by the Bush administration to try detainees held at the Guantanamo Bay military base. The use of military commissions to try enemy combatants was perceived as crucial in furthering the goals of the Global War on Terror, one of the defining foreign-policy initiatives of the Bush administration. Roberts and two other court of appeals judges heard Hamdan’s appeal on April 7, 2005. The decision of the panel was not announced until July 15, 2005—the same day that Roberts met with President Bush concerning the vacant Supreme Court position. Roberts joined his two colleagues in unanimously upholding the constitutionality of the military commissions, providing a much-needed affirmation of the Bush administration’s policies concerning the War on Terror.

Although his vote was cast well before there was an actual vacancy on the Court, questions of impropriety were levied against Judge Roberts for his decision to support the preferences of the Bush administration amid the rampant speculation surrounding his possible future promotion to the Court (Gillers, Luban, and Lubet, 2005). Given his demonstrated conservative bona fides, there is little reason to doubt that Roberts would have likely voted for a similar outcome had he not been viewed as a possible candidate for the Supreme Court vacancy. However, this does not undermine the possible influence that Roberts’s ambition for the high court may have had on his decision. Commentators have speculated that “if Roberts had voted the other way, the president almost surely would have chosen a different nominee” (Baum, 2007:33). Some went as far as claiming that with the Hamdan decision Roberts had “proved himself worthy” of nomination to the Supreme Court (Toobin, 2007:277). Had Roberts preferred an alternative outcome, he likely would have been forced to choose between his preferred outcome and his possible promotion to the Supreme Court.

Although this example is uncommonly vivid, it may nonetheless be emblematic of a general pattern of behavior among court of appeals judges. Ambition is theorized to be a motivation for political behavior in a variety of contexts (e.g., Rhode, 1979; Squire, 1988; Shenkman, 1999; Maestas, Maisel, and Stone, 2005; Fulton et al., 2006; Jensen and Martinek, 2009; Williams, 2008). As a possible motivation for judges serving in the federal judiciary, however, ambition is largely ignored. How does career
ambition affect the behavior of political actors? Are judges serving in the U.S. Courts of Appeals influenced by a desire to be promoted to the United States Supreme Court? Does this ambition manifest itself in their decision-making patterns? I answer these puzzles in the analysis presented here. I argue that theories of political ambition support the hypothesis that ambition serves as both an inherent and operative goal for judges of the lower federal judiciary. Using a measure designed to capture a judge’s promotion potential, I find that court of appeals judges do alter their behavior in a manner consistent with the desire of being promoted to the Supreme Court. The results suggest that scholars should not be so quick to dismiss the possibility that ambition for a higher office has the potential to significantly affect the behavior of judges serving in the U.S. Courts of Appeals.

**Ambition as an Operative Motivation on the U.S. Courts of Appeals**

Ambition has played a central role in the study of political behavior. Schlesinger (1966) was among the first to undertake a systematic examination of political ambition in his seminal work on political recruitment. Of primary concern for Schlesinger was the notion of “opportunity structure” (1966:6). Schlesinger notes that political ambition is molded by the availability of opportunities for advancement, and such opportunities are determined by the political structure within which an ambitious actor operates. The political realities force officials to be responsive to changing “cost[s], benefits and opportunities” when considering career goals (p. 7). Writing several years later, Black (1972) relies heavily on the empirical data catalogued by Schlesinger to develop his own complete theory of political ambition (p. 144). Key to his theory are two premises:

The theory rests on the idea that office-seekers attempt to behave in a rational manner in selecting among alternative offices; that rather than being driven by excessive ambition, they tend to develop ambition slowly as the result of their changing circumstances (p. 145).

Ambition is, therefore, not a static trait, which exhibits a uniform effect, but rather is a dynamic motivation that grows in response to exogenous changes in the political environment. This framework has been used to examine the effect of ambition in many different political environments. Studies of recruitment (Moncrief, 1999), advancement (Deering, 1996), and retirement (Hall and Van Houweling, 1995) in Congress demonstrate the utility of this dynamic conceptualization. Such examinations have demonstrated that ambition is altered by many changing variables, including the likelihood of victory (Stone et al., 2010), general satisfaction with one’s position (Moore and Hibbing, 1998), and resource constraints (Goodliffe, 2001). Similar patterns are observed when examining ambition in elected state judiciaries. The decision of a challenger to run in a judicial election is affected by the quality of the incumbent (Bonneau and Hall, 2003; Hall and Bonneau, 2006), the amount of
fundraising needed to run an effective campaign (Bonneau, 2007), and gender (Williams, 2008; Jensen and Martinek, 2009).

Despite prominence in these contexts, ambition has played only a minimal role in the study of judicial behavior in the lower federal judiciary.² Scholars have devoted substantial theoretical and empirical attention to understanding decision making in the U.S. Courts of Appeals. Unlike studies of Supreme Court decision making, which tend to be dominated by attitudinal and strategic considerations (Segal and Spaeth, 2002; Epstein and Knight, 1998), this rich literature has produced a very complex picture of court of appeals decision making. Vote choices have been demonstrated to be a function of several competing considerations (Songer, Sheehan, and Haire, 2000; Hettinger, Lindquist, and Martinek, 2007; Cross 2007), most prominently including policy preferences (Giles, Hettinger, and Peppers, 2001; Hettinger, Lindquist, and Martinek, 2004), strategy (Cross and Tiller, 1998), legal precedent (Johnson, 1987; Klein, 2002), institutional hierarchies (Kastellec, 2007), gender (Songer, Segal, and Cameron, 1994; Boyd, Epstein, and Martin, 2010), and race (Farhang and Wawro, 2004). And while scholars have examined the determinants of promotion within the federal judiciary (Bratton and Spiller, 2004; Savchak et al., 2005), ambition as an explanation for judicial behavior has been largely ignored.

Hettinger, Lindquist, and Martinek (2007) borrow a conceptual framework developed by Baum (1997) to outline the primary objection to ambition as an influence on decision making. According to the authors:

... both circuit and district court judges are likely to have the inherent goal of elevation to a higher bench. But not all inherent goals become operative goals. “Of the various inherent goals that individual judges hold, some may be irrelevant to their work as judges” (Baum 1997, 14). Hence, while both circuit and district court judges might have the desire for elevation, the opportunity for elevation is practically nonexistent for the former, so the inherent goal of elevation is not readily translated into an operative goal (p. 24).

Several important considerations limit the persuasiveness of this objection. Despite the contention, the probability of promotion is not “practically nonexistent” for a substantial set of judges. Court of appeals judges have dominated the nomination process (particularly recently), leading many to view an apprenticeship on the courts of appeals as a near prerequisite for service on the Supreme Court (Yalof, 2005:696). Since 1900, nominees have been of a limited set of ages, ranging from 41 (Douglas) to 65 (Lurton) when they received their nominations. Democratic and Republican presidents rarely consider candidates outside of their own political party. Interest in racial and gender representation on the Court has greatly increased the likelihood of

² I focus my attention on the lower federal judiciary because promotion to a higher judicial position is unavailable to those already serving on the U.S. Supreme Court (see Segal and Spaeth, 2002:92-96). With that being said, ambition is also not regularly theorized to systematically affect decision making at the U.S. Supreme Court.
minority judges being chosen. So while the objective probability of any random court of appeals judge being chosen for the Supreme Court is in fact low, for many the probability of selection is substantially greater.

Even if the probability of promotion is low, decision makers have “limited sensitivity to low probability events” (Schoemaker and Kunreuther, 1979:616). Beyond some threshold, individuals are no longer sensitive to changes in either the likelihood or the unlikelihood of an event occurring. As a consequence, decision makers consistently overestimate the likelihood of very unlikely events. Even if the probability is objectively low, court of appeals judges likely overestimate the probability of being promoted. This may help explain why judges in other contexts have been shown to alter their behavior in response to objectively unlikely events (see, for example, Huber and Gordon, 2004).

Finally, the probability of an event occurring tells only part of the decision-making story. Expected utility theory predicts that a decision maker will choose between alternatives by considering both the probability of an outcome occurring and the value the decision maker places on the outcome. As Posner (2008) notes, “if a judge attaches enormous value to being a Supreme Court Justice, the expected utility of such an appointment may influence his behavior” (pp. 145-46). Even if a court of appeals judge believes that the probability of being elevated to the Supreme Court is low, if he or she places enough value on being a Supreme Court justice, ambition may influence his or her decision-making patterns.

Anecdotal evidence abounds suggesting court of appeals judges are likely to attach substantial utility to a promotion to the Supreme Court. Baum (2006) chronicles the feud that developed between J. Harvie Wilkinson and J. Michael Luttig, then members of the Fourth Circuit Court of Appeals, over their apparent competing statuses as candidates for a pending vacancy on the Supreme Court. For a short period, Wilkinson and Luttig engaged in calculated attacks on the other in their judicial opinions, which many commentators speculated was motivated by the desire for a possible promotion to the Court (Baum, 2006:1). This speculation was only encouraged by the unexpected resignation of Luttig shortly after being passed over for the position. Such a desire does not appear to be unique to these judges. Alex Kozinski, a member of the 9th Circuit Court of Appeals, is known to joke about an informal organization known as “OOPSSCA” or “the Organization of People Patently Seeking Supreme Court Appointments” (Bazelon, 2004:32). While few court of appeals judges would be likely to openly admit membership in this fictitious club, evidence suggests that many would place a high value on being promoted to the Supreme Court.

Evidence of Ambitious Behavior in the Courts of Appeals
The limited analyses investigating the influence of ambition in the lower federal judiciary offer evidence suggestive of ambitious behavior. Ambition has been shown

---

to affect court of appeals judges when engaging in activities other than actual
decision making. Gaille (1997) demonstrates that publication rates of articles and
books among court of appeals judges significantly declined shortly after Robert Bork's
failed nomination campaign. Bork's pre-nomination writings were thoroughly scruti-
nized by the Judiciary Committee and were a central source of controversy throughout
the course of nomination, creating “an incentive for the President to nominate people
without records, simply because they have not said anything controversial” (Strauss
and Sunstein, 1992:519). Examining publication rates by court of appeals judges
before and after the Bork hearing, the results demonstrate that judges may consider
the possibility of promotion when deciding to author articles. Reported article publi-
cation decreased throughout the 1980s, but decreased most strongly after the Bork
confirmation hearings. Of particular note was the drop-off in publications concerning
the highly controversial topics of race and abortion. These changes were most pro-
nounced among younger cohorts of judges—those most likely to be considered for a
promotion in the future.

While this analysis provides intriguing evidence that court of appeals judges may
act in response to their desire for promotion to the Supreme Court, it fails to answer a
central question: does ambition influence court of appeals judges in their decision
making? Most analyses addressing the direct influence of ambition on decision mak-
ing have focused on the district courts. The decision to focus on district court judges
makes a good deal of sense: the additional opportunities for promotion and increased
likelihood of receiving a promotion could seemingly induce ambitious behavior. Cohen
(1991, 1992) and Morriss, Heise, and Sisk (2005) show that an increase in a district
court judge’s “promotion potential” is accompanied by an increased likelihood of a
judge reaching a decision consistent with the preferences of both the Democratic and
Republican parties, and Taha (2004) demonstrates that district court judges with high-
er levels of promotion potential were more likely to publish such decisions.

Only Black and Owens (2012) investigate the influence of promotion on the
behavior of court of appeals judges. Using presidential “short lists,” the authors demon-
strate that “contender” court of appeals judges for a Supreme Court vacancy (i.e.,
those on the presidential “short list”) alter their behavior during vacancy periods in a
manner consistent with securing promotion to the Supreme Court. This provides an
important first step in testing the influence of ambition on judicial behavior in the
courts of appeals. However, the structure of the analysis limits what can be inferred
concerning the influence of ambition on judicial behavior. By only analyzing “con-
tender” judges, it is difficult to determine whether their results supporting ambitious
behavior are generalizable to all court of appeals judges. In addition, vacancy periods
are rare and represent only a small portion of time in which judges are making deci-
sions in the courts of appeals. The analysis is successful in demonstrating that certain
judges, under certain conditions, use their positions in an attempt to secure promo-
tion. However, it is not clear if this represents the universe of ambitious behavior
in the court of appeals or only a unique subset of a much larger, more general phenomenon.

Despite the varying contexts, each of these analyses supports the notion that ambition influences the behavior of judges serving in the lower federal judiciary. The analyses also demonstrate that ambition for federal judges manifests itself in a manner consistent with existing theories of political ambition (Schlesinger, 1966; Black, 1972). Rather than ambition having a uniform effect, judges respond to their changing political environment when deciding whether to engage in “ambitious” behavior. Those who perceive themselves as better candidates for promotion behave differently than other judges. While there appears to be substantial evidence of this pattern for district court judges, aside from the cases arising during a pending Supreme Court vacancy and before “contender” judges, the general influence of ambition on judicial decision making in the court of appeals remains unexplored.

MEASURING AMBITION AND PROMOTION-CENTRIC BEHAVIOR

If ambition is an operative goal for court of appeals judges, then judges ambitious for a position on the Supreme Court should behave in a manner consistent with securing a promotion. Testing this expectation requires two important measures. First, it requires a measure of a judge’s level of ambition. Ambition is a difficult concept to measure, particularly for the study of judges. As the introductory quotation indicates, any kind of self-interest would likely be considered among the least normatively acceptable judicial motivations. The result is strong professional pressure to disavow any possible self-interest when acting as a judge.

Second, testing the impact of ambition requires evidence of behavior consistent with attempting to secure a promotion (“promotion-centric”) to the Supreme Court. Much like measuring ambition, operationalizing this concept for court of appeals judges is difficult. Unlike institutions (Congress, state judiciaries, etc.) where theories of political ambition are commonly tested, vacancies on the Supreme Court are rare. They are not subject to open job competition. Candidates do not campaign for the opportunity to secure a position. In fact, purposefully attempting to secure a promotion may invite stronger criticism than merely admitting an interest in the possibility. Court of appeals judges cannot know when a position is likely to become available. For all these reasons, determining what does and does not constitute “promotion-centric behavior” is difficult.

Promotion Potential. I follow the insights provided by theories of political ambition (Schlesinger, 1966; Black, 1972) when developing a measure of judicial ambition to be used here. This perspective suggests political ambition does not remain constant over the life of an office seeker’s career but, rather, develops slowly and in response to changing circumstances and political fortunes. To build a dynamic measure of ambition for court of appeals judges, I rely on the two variables that best capture both the slow development of ambition over time and the responsiveness to changing
circumstances: age and years of court of appeals experience.\textsuperscript{4} Research examining promotion patterns suggests that age and experience play an important role in the likelihood of being promoted in the federal judiciary (Bratton and Spill, 2004:208; Savchak et al., 2005:487). Although many other considerations (such as gender in the appointment of Justice O’Connor or race in the appointment of Justice Thomas) have the potential to influence the selection process in idiosyncratic, nonsystematic ways, age and experience consistently emerge as important criteria in the modern selection process. President Eisenhower, for example, enforced a rigid age limit of sixty-two years for his Supreme Court nominees (Yalof, 1999:43), and the Clinton White House found Ruth Bader Ginsburg appealing as a candidate because of her experience on the D.C. circuit (Toobin, 2007:72). In regards to the motivating example detailed above, the Bush White House made both age and judicial experience important criteria when discussing candidates to fill the vacancy left by Justice O’Connor’s retirement. Nearly all of the candidates on the “short list” had experience on the courts of appeals. Of those candidates, age played a major role in the evaluation of their fitness for office. J. Harvie Wilkinson was considered as a possible candidate, but many were concerned that he was “already sixty” (Toobin, 2007:317). The two reported finalists for the position, John Roberts and J. Michael Luttig, were both fifty years old at the time of the vacancy, which the administration considered “the perfect age” (Toobin, 2007:317) for a nominee.

I therefore use a judge’s age and years of judicial experience to create a measure of “promotion potential.” Labeling the measure “promotion potential” is consistent with Cohen (1991), the first systematic examination of the influence of ambition on judicial behavior in the lower federal judiciary. Cohen measures “promotion potential” through factor analytic techniques and discovers two factors related to the likelihood of receiving a promotion from the district court to the courts of appeals. The first factor includes variables like the number of current vacancies on the court of appeals and the number of district court judges within that particular circuit. The second is what he terms an “age/seniority” factor (p. 192) and comprises the judge’s age and years of experience on the district court.\textsuperscript{5}

It is this second factor that I approximate here. While Cohen’s (1991) first factor is relevant to the promotion of judges from the district court to the court of appeals, many of the variables included are substantially less likely to be relevant for studying promotion to the Supreme Court. As discussed above, vacancies on the Supreme

\textsuperscript{4} The measure is based exclusively on experience in the court of appeals and not on total judicial experience (including district court experience or state court experience) because of the modern selection process. Nine of the last ten appointees to the Supreme Court had previously served in the court of appeals, while only one (Sonia Sotomayor) had ever served on the district court, and none had any experience as a state court judge. From this pattern it appears that modern presidents place substantial value on experience in the court of appeals, but little value on other forms. Judges interested in a promotion to the Court are, therefore, unlikely to view their experience in other judicial offices as substantially increasing their likelihood of promotion.

\textsuperscript{5} While I do not believe the fact detracts from the measure’s utility, it is worth noting that this factor is not significant in Cohen’s (1991) subsequent analysis.
Court are much fewer in number and occur with such infrequency that it would be nearly impossible to anticipate them with any degree of confidence. I therefore focus on the second factor. Before performing the analysis, I transform a judge’s absolute age and years of experience on the courts of appeals to capture each variable’s nonlinear effect on the likelihood of receiving a promotion. To do so, I take the absolute value of the distance of a judge’s age or years of experience on the courts of appeals from the average age of each justice (fifty-two) and average years of experience on the courts of appeals (seven) at the time of their appointment, beginning with Justice Scalia. Each resulting variable takes the value of 0 if a judge is currently exactly fifty-two years old or has exactly seven years experience on the court of appeals. This represents the “typical” candidate who has been promoted from the courts of appeals to the Supreme Court since the early 1980s. Each measure increases by one for each year a judge’s age or experience deviates from the typical nominee. I then performed an exploratory factor analysis. The transformed age and experience variables loaded onto a single factor with an eigenvalue greater than 1. For ease of interpretation I then rescaled the results so higher values are associated with greater promotion potential and lower values are associated with lower promotion potential. The final measure has higher scores for judges of moderately young ages with some levels of judicial experience. Judges who are very old or very young, or who have excessive or minimal judicial experience, receive comparably lower scores.

I refer to the measure as “promotion potential” because it best reflects what the variable is actually measuring: a candidate’s best assessment of viability for promotion. Promotion potential is, therefore, not a direct measure of judicial ambition. After all, not all judges who are likely to get promoted necessarily desire a promotion. There may be court of appeals judges who would reject a promotion to the Supreme Court. There may also be judges who so strongly desire a seat on the nation’s most prestigious court that they continue attempting to secure a promotion despite the fact that their advanced age precludes any possibility of ever reaching their goal. However, if court of appeals judges behave in a manner predicted by previous studies of political ambition, such behavior should be the exception. Judges serving on the courts of appeals are likely aware of the importance of age and experience. No matter what a sixty-five year old does as a court of appeals judge, history suggests that he or she will not be promoted to the Supreme Court. In addition, very old or very young judges may be less interested in a promotion to the Court. Young and inexperienced jurists may not feel prepared to serve on the Court. These judges may also not feel a sense of urgency to be promoted, knowing that future opportunities are expected to arise. Very old judges may not

---

6 I begin with Justice Scalia for two reasons. First, I use voting data from 1986-2002, making these justices the appropriate reference points for court of appeals judges attempting to gauge their likelihood of a promotion. Second, beginning with Justice Scalia through 2002 (the end of the data set), every new Supreme Court justice (five in total) was promoted directly from the court of appeals. This makes the current era a particularly appropriate one to study the influence of ambition—court of appeals judge may view their service on the circuit court as a necessary condition to receive the call to the Supreme Court.
desire to change their professional levels so late in their careers and may be more interested in reaching retirement. If these assumptions are valid, judges will behave “ambitiously” as their viability for promotion increases and less ambitiously as their perceived level of fitness begins to decline. The promotion potential variable created here captures this dynamic.

This conceptualization of “promotion potential” is undoubtedly an underinclusive measure of the likelihood of promotion from the courts of appeals to the Supreme Court. As noted briefly above, many other factors, including gender, race, religion, and confirmation politics, are likely to affect a president when selecting among potential nominees. However, the measure is designed to capture how court of appeals judges gauge their changing likelihood of promotion. It is not designed to be a predictor of presidential selection. For example, it would be difficult for a court of appeals judge to use likelihood of confirmation by the Senate to assess his or her chances of receiving a promotion. A court of appeals judge might know the likelihood of confirmation in the current climate, but to gauge the likelihood of a future confirmation, a judge would need to anticipate the future preferences of the Senate, the composition of the Supreme Court itself, and the broader political climate. Although a court of appeals judge may not be able to anticipate these variables, the judge can be confident that reaching a certain age or years of experience will change the probability of a promotion substantially. Designed to capture how court of appeals judges might reasonably appraise their political fortunes, the measure is theoretically appropriate.

A secondary concern of this operationalization of promotion potential is the possibility of time-bound results. By relying on a specific subset of Supreme Court justices to generate the promotion potential measure, there is the possibility that the measure lacks validity outside of the time period in question. Two considerations help to alleviate this concern. First, presidential motivation, in terms of these two factors, does not appear to vary substantially over time. For example, the “ideal age” for a candidate in the sample (1986-2002) is fifty-two years. Since 1900, the mean age of a Supreme Court justice at the time of appointment has been fifty-four years. The mean age of all Supreme Court justices at the time of appointment throughout history is fifty-three years. This suggests that while there might be some variation, most presidents have tended to look for similar candidates (again, in terms of age and experience) when selecting possible nominees. Second, even if presidents’ preferences were to change, there is no reason the measure could not be updated to reflect these changes. In this case, because the voting data is from 1986-2002, the specified set of justices represents the appropriate reference group. In the future, it will be possible to update the measure to reflect changes in the selection process.

**Ideologically Consistent Voting.** Testing the influence of ambition on court of appeals judges also requires evidence of behavior that would not be expected if ambition was absent as a motivation. How do ambitious judges behave differently than judges lacking in ambition for our highest bench? Gaille (1997) examines self-reported publication of law-review articles among court of appeals judges. Cohen uses
district court judges’ decisions on the constitutionality of the sentencing commission (1991) and antitrust-sentencing behavior (1992) to test for the influence of ambition because both political parties preferred similar outcomes in these cases. These possibilities, while intuitively pleasing, are limited. The publication of law-review articles (or lack thereof) may provide evidence of ambition as an inherent goal, but fail to address the primary question of whether ambition actually influences the decision-making process. Cohen’s (1991, 1992) analyses speak to the question of operative goals, but do so in only a small subset of cases. While his results are consistent with ambitious judicial behavior, it is difficult to determine whether the findings presented are generalizable to a broader set of cases.

To avoid many of these limitations, I analyze a different form of judicial behavior: ideologically consistent voting. Recently, scholars of judicial politics have begun to examine the conditions under which judges vote consistently. Epstein et al. (2007) demonstrate that, since 1937, most Supreme Court justices have become more liberal or conservative over time. Kaheny, Haire, and Benesh (2008) offer evidence that the ideological consistency of court of appeals judges also tends to vary over the course of judicial careers. Using a heteroskedastic voting model, the authors conclude that judges serving on the court of appeals from 1968-96 demonstrated greater degrees of ideological consistency at the very beginning and very end of their careers. In addition to examining voting consistency as an end, scholars have also examined consistency as a means to understanding judicial motivations. Examinations of ideological consistency have informed our understanding of how many factors, including ideological extremism (Collins, 2008), case salience (Collins, 2008), strategic behavior (Kim, 2009), and policy expertise (Miller and Curry, 2009), influence the decision making of federal judges.

Ideologically consistent voting is well-suited to the study of ambitious behavior. Examining ideologically consistent voting allows a much larger sample of judges to be analyzed across many diverse issue areas. Most important, there is clear and consistent evidence that presidents desire nominees who consistently adhere to a well-defined set of policy positions (Nemacheck, 2007). For example, many justice department officials in the Reagan administration championed Antonin Scalia to replace the recently promoted Justice Rehnquist because he was believed to adhere consistently to the president’s philosophy of judicial restraint. Consistency was particularly valued by the Reagan White House after the confirmation of Sandra Day O’Connor, who, after five years on the Court, caused administration lawyers “to have serious concerns about the depth and consistency of her commitment to the values of judicial restraint” (Greenberg, 2007:44).

Ideologically consistent voting is particularly important for ambitious judges because specific decisions that are deviant from expectations, and not general deviations in overall voting, tend to be what sink otherwise viable nominees. The case of Alberto Gonzales is instructive. Alberto Gonzales was recognized early in President
George W. Bush’s first term as a likely nominee for the Supreme Court. There were many reasons to suspect that Gonzales would be a potential nominee. Gonzales was a close, personal friend to President Bush. As governor, Bush had nominated Gonzales for a seat on the Texas Supreme Court. He afforded the president the opportunity to place the first Hispanic on the Supreme Court, a strong desire of President Bush (see Toobin, 2007:303), and was sufficiently young (forty-six when Bush took office). As a judge on the Texas Supreme Court, Gonzales had established a clear and consistent record as a judicial conservative.

Well, an almost entirely consistent conservative record. Despite many of the possible advantages to Gonzales as a nominee, prominent conservative groups quickly mobilized against him once a Supreme Court vacancy opened during the Bush administration. Why such conservative distrust of a seemingly ideal candidate? Most of the objections centered on Gonzales’s vote in In re Jane Doe 5 (43 Tex. Sup. J.910), a case before the Texas Supreme Court concerning a Texas parental-notification law preventing physicians from performing abortions on unaccompanied minors without parental consent. While requiring physicians to notify parents, the statute did allow the notifications to be “bypassed” for a variety of reasons, most concerning the interests of the minor. Of the seven decisions made by the Texas Supreme Court concerning the statute, Gonzales voted to uphold the “bypass” in one of them. This decision alone, and the subsequent furor from conservative interest groups that it caused, is commonly thought to have cost Gonzales a possible seat on the Supreme Court.

This example demonstrates the utility of ideologically consistent voting as a measure of promotion-centric behavior. One deviant decision can nullify the likelihood of promotion. Rather than focusing globally on a nominee’s general record, presidential administrations (and the interest groups who monitor them) tend to examine individual decisions to determine who can be considered reliable. Assuming that court of appeals judges recognize how the selection process operates in practice, those with strong ambitions for the court have incentive to vote consistently to at least keep themselves in the running as a possible nominee. Voting consistently will hardly guarantee that a judge will be considered as a possible nominee, but voting inconsistently will almost guarantee that a judge will not be seriously considered.

My primary hypothesis is, therefore, that promotion potential will be positively and significantly related to ideologically consistent voting. As court of appeals judges

---


8 The example is not perfect, as Gonzales was serving as a judge on the Texas Supreme Court and not in the courts of appeals when casting his vote. Judge Gonzales was likely not considering the possibility of promotion because state supreme court judges are rarely considered for Court vacancies. Even so, it does not undermine the usefulness of the example. The utility is not Gonzales’s lack of ambitious behavior, but the lesson it teaches: one vote can undermine the possibility of promotion. Court of appeals judges, who might view themselves as viable candidates, are undoubtedly aware of this lesson.

more closely mirror the ideal Supreme Court nominee (in terms of age and experience) the more likely they are to vote consistently. To test examine this relationship, I use voting data in non en banc cases from 1986-2002 collected in the judge-centered United States Courts of Appeals database. \(^{10}\) I categorize a judge’s vote as being ideologically consistent when he or she votes for the ideological outcome that a naïve attitudinal model would predict the party of that judge’s appointing president would prefer. The variable takes the value of one when judges appointed to the courts of appeals by Republican (Democratic) presidents vote for a conservative (liberal) outcome and zero when they vote for a liberal (conservative) outcome. Decisions on the ideological direction of the case outcome follow the rules set forth in the database’s codebook.\(^{11}\) Because the dependent variable is dichotomous, I employ logistic regression. I include sampling weights to account for the sampling procedure used to construct the database. I also employ robust standard errors clustered by judge. Circuit-fixed effects are included in the analysis but not displayed.

**Additional Variables.** A variety of factors other than ambition are likely to influence the extent to which judges vote consistently. Collins (2008) demonstrates that several considerations influence the extent to which a justice votes consistently at the Supreme Court. First, ideological extremity is inversely related to variance in decision making: justices with stronger ideological commitments show less variance in their voting patterns than their more moderate colleagues. Second, justices demonstrate less variance in salient cases. Third, justices exhibit less variance in civil-liberties cases. To account for these possibilities several control variables are included. To control for ideological extremism, I include Ideological Extremity, which takes the absolute value of each judge’s GHP score (Giles, Hettinger, and Peppers, 2001). Judges with higher scores are more extreme in their ideology than judges with lower scores and, therefore, should be expected to vote with greater ideological consistency. To account for case salience, I include several dummy variables. First, Amicus Participation takes the value of one if at least one amicus brief was filed in the case and zero if no briefs were filed. Amici briefs are relatively rare in the court of appeals (compared to the Supreme Court), so the presence of a brief should signal to judges the possible importance of a case. I also include dummies for whether a case included a dissenting opinion (Dissent) and whether the federal government was a party to the case (U.S. Government). Each of these variables is thought to raise the salience of a case in the courts of appeals.\(^{12}\) Finally, I control for case type by including a dummy variable, Civil Liberties, that takes the value of one if the case involves a question of civil liberties and zero if it does not.

---

\(^{10}\) The U.S. Courts of Appeals Database, Donald R. Songer (principal investigator), NSF# SES-89–12678. The database and documentation are available at the Ulmer Project at the University of Kentucky (www.as.uky.edu/polisci/ulmernetproject/appctdata.htm).

\(^{11}\) See www.as.uky.edu/polisci/ulmernetproject/appctdata.htm.

\(^{12}\) Alternatively to my expectations, it could be the case that presidential administrations only review whether or not a judge has voted consistently in a unique subset of salient cases. If this is the case (and judges are aware of
Collins (2008) examines voting consistency by Supreme Court justices. Institutional differences between the courts of appeals and the Supreme Court may also affect the likelihood of ideologically consistent voting among judges. Decision making in three-judge panels may influence the extent to which judges vote consistently. Studies of panel effects on the courts of appeals (Kastellec, 2007:2011; Kim, 2009; Sunstein et al., 2006) have demonstrated that judges on ideologically unified panels are less likely to deviate from their ideological commitments than judges on ideologically diverse panels. To account for this possibility I include a control variable, Panel Composition, which takes the value of one if the panel is ideologically unified and zero otherwise.\footnote{This is determined by the judges’ common space scores. Judges with common space scores greater than zero are considered “conservative,” and judges with common space scores less than zero are considered “liberal.”}

**Results and Discussion**

The results are presented in the Baseline Model (column 1) of Table 1.\footnote{For all of the models discussed, none of the circuit-fixed effects statistically differ from the baseline category (the D.C. Circuit Court of Appeals).} Most control variables behave as expected. Ideological Extremity is positively signed and statistically significant. Court of appeals judges with stronger ideological commitments are more likely to vote consistently than judges with weaker commitments to their ideological beliefs. Contrary to expectation, two of the variables designed to capture case salience, Amicus Participation and U.S. Government, fail to reach conventional levels of statistical significance. Only Dissent, which indicates the presence of a dissenting opinion in the case, is significant. Civil Liberties also fails to reach conventional levels of statistical significance. Although this variable has been found to reduce variability in the voting behavior of Supreme Court justices, it does not appear to influence the extent to which court of appeals judges vote consistently. This discrepancy is likely a function of the institutional differences between the two levels of the judiciary.

The importance of the institutional differences between the Supreme Court and the courts of appeals is further highlighted by the inclusion of Panel Composition, which reaches traditional levels of statistical significance. Judges serving on ideologically unified panels are significantly more likely to vote consistently than judges serving on ideologically diverse panels. This is consistent with expectations and demonstrates the importance of accounting for institutional differences when comparing the behavior of court of appeals judges and Supreme Court justices.

The primary variable of interest, Promotion Potential, is correctly signed and statistically significant. Court of appeals judges with higher levels of promotion potential are significantly more likely to cast ideologically consistent votes than their colleagues such practices) then the effect of promotion potential should be conditional on case salience. To test this possibility, I estimated three additional models, each including an interaction of Promotion Potential and one of the variables included to capture case salience (Amicus Participation, Dissent, U.S. Government). None of the interaction terms reached conventional levels of statistical significance. If ideologically consistent voting is only important in salient cases, court of appeals judges do not appear to have recognized this fact.
with lower levels of promotion potential. Consistent with theories of political ambition, judges appear to vary their behavior in response to changing political circumstances. As a judge grows to more closely approximate the typical candidate for the Supreme Court, the judge becomes more likely to cast an ideologically consistent vote. The substantive impact of this effect is substantial (see Figure 1).\textsuperscript{15} Moving from the

\textsuperscript{15} Predicted probabilities were calculated holding Ideological Extremity at its mean value and all other variables at their modal value.
minimum level of promotion potential to the maximum, the change in predicted probability of casting an ideologically consistent vote is approximately 25 percentage points. However, this likely overstates the substantive effect, as most of the judges in the sample are clustered in the moderate-to-high levels of promotion potential (as evidenced by the width of the confidence intervals). To more accurately gauge the substantive effect, refer to the portion of Figure 1 located between the two dashed vertical lines. These lines represent the 25th and 75th percentiles of levels of promotion potential. Judges with a promotion potential score at the 75th percentile were six percentage points more likely to cast an ideologically consistent vote than judges with a promotion potential score at the 25th percentile. Increasing the promotion potential score to the highest observed value (the right of the second dashed vertical line), the likelihood of casting an ideologically consistent vote is increased another three percentage points.\(^\text{16}\)

There are several possible questions that the Baseline Model fails to address. The first is the possible relationship between opinion writing and promotion potential.

\(^{16}\) The low predicted probability (between 50 and 60 percent for most of the data) of casting an ideologically consistent vote is likely a function of the choices made when calculating the predicted probabilities and the coding of ideologically consistent voting. All of the control variables hypothesized to positively correlate with ideologically consistent voting are held at zero (their modal category). More important, a substantial minority of the
The three-judge panels used by the courts of appeals and the workload facing many judges lead to the opinion author being perceived as having much greater influence over the content of the opinion than is true on other, larger appellate courts. If judges are aware of this perception, promotion-centric behavior might be more common only when the judge is the opinion author. To account for this possibility, I reestimate the Baseline Model by including a dummy variable for whether the judge was the opinion author and interacting that variable with the Promotion Potential measure. Neither the constituent term nor the interaction term (not reported) reaches conventional levels of statistical significance. There is no evidence to suggest that the effect of promotion potential is stronger when the judge engages in opinion authorship.

The second question is the possible conditional influence of ambition. As currently constructed, the variable Promotion Potential treats Republican and Democratic judges equally, regardless of the party of the current president. This seemingly assumes that judges appointed by Democratic and Republican presidents are equally likely to be appointed to the Court. This is, of course, not the case. The last president to appoint someone to the Court from another political party was Richard Nixon, when he appointed Lewis Powell. It is difficult to imagine a president in this era appointing a judge of a different political party. This does not, however, preclude ambition from affecting the behavior of court of appeals judges even when their preferred party does not control the presidency. Court of appeals judges are unlikely to expect a copartisan president to only review decisions made during the current presidential administration. Rather, the decisions rendered by a judge under a different president are likely to be of similar importance. Because court of appeals judges still could behave ambitiously under a president of either party, no party conditions have yet been included. To examine whether court of appeals judges engage in these calculations, I reestimate the model and include a dummy variable, Same Party, which takes the value of one if the judge shares the political party of the current president and zero otherwise.\(^\text{17}\) The Same Party measure is then interacted with the Promotion Potential measure. The results of the Conditional Model are presented in column 2 of Table 1.

All of the control variables and the Promotion Potential variable retain their sign and their level of statistical significance. However, the variable Promotion Potential X Same Party is not significant.\(^\text{18}\) This suggests that the effect of promotion potential on ideologically inconsistent votes are cast by Democratic judges in criminal cases (where a vote for the government is coded as a “conservative” decision). Even though these are properly coded as being ideologically inconsistent, many of these appeals are frivolous, limiting the ideological impact on judicial votes in these cases. Their frequency in the data likely overstates how often judges vote inconsistently. Figure 1 should, therefore, be read as an assessment of the substantive effect of Promotion Potential and not as a precise estimation of the true likelihood of casting an ideologically consistent vote.

\(^\text{17}\) A judge is considered to be a Democrat (Republican) if the president who appointed him or her to the court of appeals is a Democrat (Republican).

\(^\text{18}\) Interaction terms are not directly interpretable but are conditional on the changing levels of the constituent terms (Brambor, Clark, and Golder, 2006). To properly interpret their statistical significance, the relationship must be examined across all levels of the continuous variable. Such calculations (not reported) confirm that effect of Same Party is statistically insignificant across all levels of Promotion Potential.
ideologically consistent voting is not conditional on the partisan status of the current president. Judges with high levels of promotion potential continue to be more likely to cast ideologically consistent votes. The party of the sitting president does not appear to have any impact on this relationship. This pattern demonstrates the importance of examining the behavior of all court of appeals judges throughout their tenure and not only during periods when the likelihood of promotion is heightened.

A third question stems from symmetrical construction of the Promotion Potential measure. As currently operationalized, the measure assumes that all judges equidistant from the “typical” Supreme Court nominee behave similarly, regardless of which “side” of the nominee they are on in terms of promotion potential. This is consistent with Schlesinger (1966) and Black (1972), who theorize that ambition tends to emerge slowly over the course of a career. However, this may not be an accurate representation. Because presidential administrations thoroughly review a nominee’s entire record, young and inexperienced judges, whose promotion potential is going to increase in the short term, may have stronger incentives to engage in promotion-centric behavior than older, more experienced, judges whose promotion potential is only going to decrease.

To examine this possibility, I respecify the Promotion Potential variable to account for potential nonsymmetrical effects. I recode all judges under the age and years of experience as being “ideal” in these respective characteristics. All judges with ages or years of experience greater than the “typical” nominee are coded identically to the original Promotion Potential measure used in the Baseline Model. These alternative specifications were then factor analyzed, creating an alternative measure of Promotion Potential. The resulting measure assigns the maximum value of promotion potential to all judges under the age of fifty-two with less than seven years of court of appeals experience. Once a judge moves beyond these predetermined measures, their promotion potential is discounted in a manner identical to the original Promotion Potential measure. The results are presented in Alternative Model (column 3 of Table 1) and are consistent with those presented in the Baseline Model. All of the control variables retain their sign and level of statistical significance. Most important, Promotion Potential is still correctly signed and statistically significant. The evidence of ambitious behavior on the courts of appeals does not appear to be a function of the specification of Promotion Potential.

Judicial ambition appears to be the most plausible explanation for the presence of the observed relationship. But are there other possible explanations? Collins (2008) finds that as a justice’s length of tenure on the Court increases, less variability is observed in his or her decision making and attributes this finding to the presence of an “acclimation effect” (p. 868). Kaheny, Haire, and Benesh (2008) offer evidence that voting patterns change over the careers of court of appeals judges: judges exhibit less variation in the early and later stages of their careers. Age may also affect ideologically consistent voting in ways not yet discussed here. Research has demonstrated (Krosnick and Alwin, 1989) that older persons tend to have more stable political
beliefs than younger persons. Even judges with little or no judicial experience might be expected to possess more stable political beliefs than their younger colleagues.

Each of these analyses suggests that age or experience (or nonlinear transformations of each) has the potential to affect the behavior of court of appeals judges independently. The Promotion Potential variable created here has only tested the relationship between the common variance of age and experience and ideologically consistent voting. Perhaps the results are simply a function of the independent effects of age or experience? To examine this possibility I reestimate the model by dropping the Promotion Potential measure and including Age and Experience as independent predictors of ideologically consistent voting.19 The results are presented in column 1 of Table 2. As a further check, I also reestimate the model by dropping the Promotion Potential measure and including Age, Age^2, Experience and Experience^2 as independent variables. The results of that analysis are presented in column 2 of Table 2.

Looking first at the Linear Model (column 1), experience fails to significantly affect ideologically consistent voting. There is no evidence of a linear relationship between the years a judge has served on the court of appeals and propensity to vote in an ideologically consistent manner. Of greater interest is the result for our Age measure. A judge’s age does have a significant linear relationship with the likelihood of voting consistently, but in the opposite direction as would be expected. Older judges show less consistency in their voting than young judges, controlling for years of experience and ideological strength. The result suggests that the relationship between age and stable belief systems demonstrated in the general population may not be applicable to court of appeals judges (who differ from the general population in many important and meaningful ways). It may also be a function of the dependent variable itself. Ideologically consistent voting was coded in a manner that fails to reflect substantial changes in the ideological commitments of the major political parties. For example, all Democrats, regardless of appointing president, are expected to be equally likely to cast an ideologically liberal vote. The Age variable included here may be capturing presidential cohort effects that the measure of ideological extremity fails to capture. This might explain why Age is negatively signed; older judges (those appointed by Johnson, Ford, and Nixon in this sample20) may be less likely to vote ideologically, given the changing commitment of the political parties to judicial ideology over time. If those judges are also older judges appointed by the more recent presidents (Carter, Reagan, H.W. Bush, Clinton, and W. Bush) it should not be surprising to see a negatively signed coefficient.

Although presidential cohorts offer one possible explanation, the evidence presented in Table 1 suggests that this inconsistent finding is best explained by failing

---

19 Although age and experience are obviously correlated (.76), the correlation is not substantial enough to create problems for regression modeling, particularly in cases of large sample sizes such as this.

20 Judges nominated to the courts of appeals by Presidents Johnson, Nixon, and Ford make up 15.75 percent of our sample of judges. Obviously, these judges are disproportionately likely to be the oldest judges in the sample.
to account for the nonlinear effect of age on ideologically consistent voting. My expectation has been that ambitious judges will grow increasingly likely to vote consistently until reaching an age and level of experience when promotion is no longer a viable possibility. Once the possibility of promotion has disappeared, these judges may be more willing to deviate from their ideological commitments. Because court of appeals judges tend to serve on the courts of appeals for much longer after they are
considered viable for promotion than before, a disproportionate number of the votes in the data are likely taken from judges who are passed the age-viability threshold. The negative coefficient on the untransformed Age variable could be explained by this distribution.

Turning to the Nonlinear Model (column 2 of Table 2), when accounting for the possible nonlinear effects of age or experience, none of the variables reach conventional levels of statistical significance. The lack of a nonlinear effect of experience is particularly noteworthy as it is inconsistent with the results of Kaheny, Haire, and Benesh (2008). There are several possible explanations for this inconsistency. First, the time period under examination here (1986-2002) is different than that of Kaheny, Haire, and Benesh (1968-96). While there is some overlap, it is conceivable that the results of both studies could be time bound. Second, rather than examining vote choice directly, Kaheny, Haire, and Benesh (2008) employ a heteroskedastic probit model to model the variance in the decision-making process. The difference in estimation strategies could be driving these inconsistent results. Further research is necessary to determine the root cause of the discrepancy between the two analyses. What is clear is neither age nor experience independently predicts ideologically consistent voting as operationalized here.

The observed effect of the Promotion Potential measure does not appear to be a function of the effects of age or experience. The inclusion of age and experience (and their squared terms) fails to account for the unique requirements for each imposed by presidents considering possible nominees. The White House does not simply look for someone who is “middle-aged”; the requirements are much more specific. Only the Promotion Potential measure accounts for these requirements. Nor do the age and experience measures and their squared terms account for how age and experience might work together. Judges serving on the courts of appeals are likely aware that a president is not simply searching for a person of the correct age or with the correct years of experience; the president likely wants both. Again, only the Promotion Potential measure captures this dynamic. Taken together, the results of these several analyses are more consistent with ambitious behavior than other plausible explanations.

CONCLUSION

The analysis presented here represents the first systematic examination of ambition as a motivation for the behavior of judges serving in the U.S. Courts of Appeals. Building from theories of political ambition, I have developed a measure of promotion potential designed to capture the dynamic influence of ambition on the behavior of court of appeals judges. Relying on age and years of judicial experience, I have demonstrated

21 The data confirm that a disproportionate number of votes are cast by “older” judges. Sixty-five percent of votes in the data set are cast by judges older than sixty, which has been seen by members of several presidential administrations as the arbitrary “too old” cut-off for a nominee (Toobin, 2007:273). By comparison, only 3 percent of votes were cast by judges under the age of 42, the lowest age of a Supreme Court nominee in the modern era.
that judges who more closely mirror typical Supreme Court nominees are more likely to cast ideologically consistent votes. Ambition appears to be the best explanation of this relationship; a judge with a greater objective likelihood of being selected to the Supreme Court will change behavior in an effort to secure a promotion. Interestingly, ambition does not appear to be conditional on the current political landscape. Ambition appears to continue to influence the behavior of court of appeals judges even when their preferred party does not control the White House. This result suggests that court of appeals judges recognize that when presidential administrations inquire about the fitness of a nominee, they are unlikely to limit their investigation to any one subset of time. Rather, they behave as if their entire career is subject to review. Given the high stakes placed on selecting a Supreme Court nominee, this behavior appears to be consistent with conventional wisdom surrounding the selection process. I also examined other plausible explanations for the relationship between promotion potential and ideologically consistent voting (acclimation effects, age, and attitude stability) but have found no evidence to support them.

These results generate several possible avenues for future research. First, future research might be undertaken to better understand how ambition affects different forms of behavior of judges in all levels of the federal judiciary. I have examined only one type of “promotion-focused” behavior: ideologically consistent voting. Surely there are other ways that court of appeals judges could change their behavior in response to ambition for a higher office. Gaille’s (1997) work on publication of law-review articles provides an excellent example. Black and Owens (2012), in their examination of the decision to write separately, offer another. Other extrajudicial activity, such as involvement in ideological interest groups, could be undertaken with an eye on a Supreme Court seat. Future investigations will be able to provide some insight into these behaviors.

Understanding how district court judges respond to the possibility of promotion is an additional avenue available for future examination. While some work has been done on this question (Cohen, 1991, 1992), these analyses have focused on a unique set of judges facing a unique set of cases. Examining the possible systematic influence of ambition on the behavior of district court judges is particularly noteworthy because of the additional opportunities for promotion and increased likelihood of receiving a promotion for district court judges in comparison to court of appeals judges. If court of appeals judges are indeed affected by ambition for a seat on the Supreme Court, it stands to reason that district court judges would be equally, if not more strongly, affected by ambition for a seat on the courts of appeals.

These results also have important implications for the study of judicial behavior in the lower federal judiciary. Theories of judicial behavior have been developed largely to explain the voting behavior of Supreme Court justices. These theories prioritize securing legal policy consistent with preferences as the primarily motivation for Supreme Court justices. The unique institutional position of the Supreme Court is offered as an explanation for this policy-driven behavior. After all, Supreme Court
justices can select which cases they will decide on the merits. The justices are protected with lifetime appointments. The Supreme Court cannot be overruled by another court, and Supreme Court justices cannot be promoted within the federal hierarchy (for a discussion, see Segal and Spaeth, 2002:92-97). Only after these theories have been empirically tested at the Supreme Court are they transported to explain the decision making of judges in other levels—in this case the courts of appeals. While court of appeals judges are similar to Supreme Court justices in many respects, they operate in a different institutional context.

As a result, court of appeals judges have been assumed to prioritize creating legal policy consistent with their preferences at the expense of other goals unrelated to policy considerations. This assumption has been made despite the lack of institutional barriers at the courts of appeals (unlike the Supreme Court), which would render such motivations moot. The analysis presented here suggests that at least one nonpolicy motivation, ambition, may be influencing the work of court of appeals judges. This analysis will hopefully lead scholars to rethink dismissing the possible nonpolicy motivations of judges serving the lower federal judiciary and to consider the implications of economic theories of appellate court behavior more seriously.

This should also serve as motivation for scholars to begin developing theories of judicial behavior for institutional contexts other than the Supreme Court. Without recognizing the important institutional differences between the nine-member Supreme Court justices and the courts on which all other judges in the federal judiciary serve, our theoretical understanding of judicial behavior at these lower levels will remain incomplete. This is not to say that theories of judicial behavior of the Supreme Court cannot inform our understanding of judicial behavior in the lower federal judiciary—they certainly do. It only suggests that the work done by court of appeals and district court judges is significantly unique to require its own theoretical development. This examination is a first step in that direction. jsj

REFERENCES


In 2010 the Iowa electorate denied retention to all three Supreme Court justices on the ballot. We hypothesize that when retention elections feature vigorous campaigns, they can be explained by many of the same factors that affect outcomes in partisan and nonpartisan elections. In particular, we expect partisanship, the amount of campaigning (advertising), and political context (same-sex marriage) to explain the results in Iowa. We examine county-level election results, finding that the partisanship and concentration of evangelicals within a county, as well as the amount of advertising against the justices, explain the justices’ vote share. We also find evidence that ballot roll-off decreased most in counties with high concentrations of evangelical adherents and where advertising was present.

In 2010, Iowa voters denied retention to three justices of the Iowa Supreme Court: Chief Justice Marsha Ternus, Justice David Baker, and Justice Michael Streit. The vote was a result of an unprecedented mobilization by opponents of same-sex marriage, which a court decision had legalized in 2009. It also took place during a highly partisan gubernatorial election in which Republican Terry Branstad defeated an unpopular Democratic incumbent, Chet Culver.

Justices Ternus, Baker, and Streit were rare losers in a retention election system that almost always protects incumbent judges. Iowa had only rejected four lower-court judges since adopting merit selection in 1962; no supreme court justice had ever lost a bid for retention. The 2010 Iowa election was atypical in another way, as well. Though retention elections are marked by significant ballot roll-off (Bonneau and Hall, 2009), this one saw very little. The average ballot roll-off in Iowa Supreme Court races from 1990-2004 was 41.3 percent (Bonneau and Hall, 2009:24), but roll-off in 2010 was only 13.2 percent statewide.

We hypothesize that these results are best explained as the result of political mobilization that transformed a normal retention election into something much more like a contested election. We analyze county-level election results and ballot roll-off in the 2010 Iowa Supreme Court retention election. In particular, we expect partisanship, amount of campaigning (advertising), and the concentration of evangelicals to explain the results in Iowa. We examine county-level election results, finding that ballot roll-off decreased most in counties with high concentrations of evangelical
adherents and where advertising was present. We further find that partisanship and concentration of evangelicals within a county, as well as the amount of advertising against the justices, explain the justices' vote share. The results suggest that voting patterns routinely seen in highly competitive contested judicial elections may be replicated in retention elections, given the presence of a highly salient issue around which the campaign can be built. We believe that the nature of the 2010 Iowa retention elections underscores the challenges that retention elections might face in the coming years, as the tactics brought to bear in contested elections become more common in retention elections.

SAME-SEX MARRIAGE AND THE 2010 IOWA RETENTION ELECTION

In 2009 the Iowa Supreme Court extended same-sex couples the right to marry under the Iowa Constitution (Varnum v. Brien). The decision was rooted in the equal-protection clause, which dates to 1857 and broadly protects against discrimination: “All laws of a general nature shall have a uniform operation; the general assembly shall not grant to any citizen, or class of citizens, privileges or immunities, which, upon the same terms shall not equally belong to all citizens” (Iowa Constitution, Art. I, sec. 6). In its opinion, the court cited its long history of broadly applying this protection against discrimination based on race and sex. In light of this broad nondiscrimination principle, the court ruled unanimously that the state had no rational basis for denying gay couples the right to marry on the same terms as heterosexual couples. The Varnum court was composed of five justices appointed by Democrats and two appointed by Republicans.

Republican leaders in the state House and Senate immediately demanded that the legislature begin the lengthy process of amending the state constitution (McKinley, 2009; Paulsen, 2009), but the Democrats who controlled both houses refused to schedule votes on the issue. Bob Vander Plaats made overturning Varnum the centerpiece of his campaign for the Republican gubernatorial nomination throughout early 2010, vowing that if elected he would issue an executive order to defy implementation of same-sex marriage in the state. After losing soundly to a more moderate candidate (former governor Terry Branstad), numerous press reports indicated that Vander Plaats was mulling an independent bid for governor.

Instead, he announced on August 11, 2010 that he would lead a campaign against the retention of the three Supreme Court justices on the ballot in November. Though promoted as a grassroots effort, subsequent press reports and disclosure documents established that the group he formed, Iowa for Freedom, was almost entirely funded by out-of-state money raised by the national group the American Family Association (Caldwell, 2011). The National Organization for Marriage funded a large advertising campaign, spending over $600,000, two-thirds of which paid for television ads (Hancock, 2010). Presidential contenders got into the mix as well, looking to increase their profile within the state in advance of the 2012 caucuses. Newt Gingrich helped raise hundreds of thousands of dollars early on, in part by providing at least
$125,000 from his own PAC (Hamburger and Gold, 2011), and every presidential candidate was asked his or her position on the issue. All told, groups supporting the justices’ ouster spent nearly one million dollars (Hancock, 2010; Aspin, 2011).

Meanwhile, though they did step up their public appearances around the state, the justices refused to raise money or mount a true campaign to keep their seats. This left only a loose coalition of good-government, gay-rights, and ideologically liberal groups to defend them. The Fair Courts for Us Committee, which emerged as the dominant pro-retention group, was an umbrella organization that did not coalesce until mid-October. Though it was ultimately able to spend over $350,000 to support the justices, this money went primarily to radio ads and mailers. As nearly half a million dollars was being spent on broadcast television ads to defeat the justices, pro-retention interests spent nothing (Aspin, 2011:230). In an election year that heavily favored Republicans, and propelled by such interest-group advertising and heavy grassroots efforts, the electorate denied an additional term to the justices. Chief Justice Ternus received only 45.1 percent of the statewide vote, and Justices Baker and Street each received 45.7 percent.

INFORMATION AND PARTICIPATION IN RETENTION ELECTIONS

The normative arguments about selection and retention of judges typically revolve around concerns over judicial impartiality. Many argue that elections threaten the impartiality of the judiciary; such means of control “could cause judges to disregard the law and implement the preferences of those who threaten or control them, to the detriment if not the demise of decisional independence” (Geyh, 2008:1260). Burnett (2007) takes matters further, arguing that interest-group involvement in judicial elections is “bluntly inconsistent” with judicial impartiality (p. 275). On the other hand, Hall (2011b) points out that appellate court judges exercise considerable discretion and make important policy decisions, but that they do so within a schema of institutional and normative constraints. Elections, she argues, are not necessarily inconsistent with the rule of law, especially if the public is enforcing constraints that the incumbent judge is supposed to follow. The Iowa retention elections, which revolved around the court’s same-sex marriage decision, reflected many of these normative concerns. Most of the groups that organized to try to defend the justices made arguments about the threat to impartiality posed by the prospect of the majority punishing judges for protecting minority rights. Rhetoric of those in favor of ousting the judges stressed their belief that the judges had violated key norms of judicial restraint and, in so doing, had usurped power from the legislature and the people.

Within the context of this normative debate, retention elections have typically been seen as a compromise between judicial impartiality and accountability. (For early investigations into the merit plan and retention elections, see, e.g., Watson and Downing, 1969; Carbon, 1980; Dubois, 1980a.) As Aspin and Hall (1987) explain, the retention election was intended, first, to:
ensure that judges would be able to serve for long periods of time by insulating them from political influences and by making removal from office fairly difficult. Second, it would provide a small measure—but a measure nonetheless—of public accountability (p. 704).

Tarr (2009) refers to this compromise as an “uneasy” one, in which retention elections were never truly intended to provide much electoral accountability. In reviewing the history of merit selection, Tarr notes that supporters of retention elections did not expect that voters would have enough information at their disposal to form opinions about judges, thus ensuring that incumbents would be retained in office unless their record was extraordinarily bad (2009:608-10).

If this was the intention, then retention elections have fully lived up to it. Between 1980 and 1994, only six state supreme court justices were not retained in office (Hall, 2001:319; Bonneau and Hall, 2009:84). Incumbents in state supreme court elections typically earn between 65 to 75 percent of the vote, with the average “yes” vote since 1980 hovering around 71 percent (Hall, 2001:318; Bonneau and Hall, 2009:82). Larry Aspin is perhaps the most thorough chronicler of retention election outcomes; his work, both solo and with co-authors, uses a database that includes the result of every retention election at all levels since 1964 (see, e.g., Aspin, 1999, 2007, 2011; Aspin and Hall, 1987; Aspin et al., 2000). Largely descriptive, this work is invaluable in understanding trends in retention elections. His most recent review (Aspin, 2011) notes, for example, a slight decline in levels of ballot roll-off since their height in the early 1970s. In terms of vote share, though incumbents remain very safe, there is a slight downward trend since 1964 in retention margin (p. 219).

Aspin (2011) further observes that the share of “yes” votes received is markedly smaller in states in which campaigns are mounted against the justices. This suggests the importance of the availability of information in retention-election voting. Retention elections rarely offer clear cues to voters. Party, of course, is the most significant cue that can be present on a ballot, and in judicial elections partisan elections tend to see higher levels of participation (see, e.g., Dubois, 1980a; Lovrich and Sheldon, 1983; Bonneau and Hall, 2009). Studies of judicial elections (e.g., Klein and Baum, 2001) as well as low-information elections more generally (e.g., Squire and Smith, 1988; Schaffner, Streb, and White, 2001; Schaffner and Streb, 2002), support the importance of the partisan cue by demonstrating that the presence of a candidate’s party on the ballot allows voters to make choices that more accurately reflect their own policy preferences. Without such partisan cues, then, voters are left to grasp at whatever information is available. Incumbency tends to become an important cue to candidate experience in low-information elections (Byrne and Pueschel, 1974; Dubois, 1979; Jewell and Breaux, 1988; Schaffner, Streb, and Wright, 2001; Streb and Frederick, 2009), including low-information judicial races (Baum, 1983, Dubois, 1980b, 1984). Aspin and Hall (1987b) suggest a “friends and neighbors” effect at work in local retention elections, with voters having more information about candidates more closely tied
to where they live. Voters in low-information elections may also base their vote on gender, racial, or ethnicity cues (McDermott, 1997, 1998; Matson and Fine, 2006) or on ballot position (Matson and Fine, 2006).

But what if highly salient, even partisan, voting cues were present in a retention election? Once a campaign is launched against a sitting judge or justice, retention elections may be influenced by some of the same political pressures that shape contested court elections today. After all, today’s judicial elections are not the judicial elections of the middle 20th century. The “new-style” (Gibson, 2008) judicial campaigns of today present different types of political pressures than those that the merit plan were designed to resist. As a latter-day progressive reform, the merit plan was primarily aimed at shielding judicial selection from partisan politics, i.e., control by party bosses who wielded influence through their ability to control information and deliver votes through horse trading and corruption. Today, though, electoral politics tends to be more candidate centered. Judicial elections have followed this broader trend. In states that allow contested elections for state high courts, incumbents today are challenged over 70 percent of the time compared to about 50 percent of the time as recently as 1990 (Bonneau and Hall, 2009:80). Partisan elections are particularly competitive, with incumbents losing roughly one-third of elections (Bonneau and Hall, 2009:84). This increased competitiveness has been spurred by increased spending (Bonneau, 2005; Bonneau and Hall, 2009), which pays for more professionalized campaigns that include a significant amount of television advertising, direct mail, and other sophisticated forms of campaigning (Peters, 2008). If a new-style campaign were mounted against judges standing for retention, one would expect voters to behave in ways comparable to more salient, competitive state supreme court elections.

Indeed, history suggests that retention elections have little defense against such mobilization. Interest groups successfully targeted justices over the years in California in 1986, Nebraska and Tennessee in 1996, and Pennsylvania in 2005 (see, e.g., Kritzer, 2007; Reid, 1999; Culver and Wold, 1986; Wold and Culver, 1987). The campaigns in California, Nebraska, and Tennessee revolved almost entirely around opposition to justices’ rulings on matters of criminal law (the death penalty in California and Tennessee and the definition of second-degree murder in Nebraska). Justice Nigro in Pennsylvania may stand out as exceptional here: he was the victim of circumstance, standing for election soon after the legislature approved a pay raise for all elected officials. Without state legislative elections on the ballot, the judges got the brunt of voter fury. Justice Nigro was defeated, and another justice on the ballot narrowly won retention.

Given these findings and these recent historical examples, we expect that if voter information is more readily available, money is spent, partisan cues are relied upon, and advertising is present—in short, if a true campaign is mounted—voters can be mobilized in retention elections as effectively as they can in contested elections.
**Voter Mobilization in a Salient Retention Election.** As “down-ballot” races, retention elections are susceptible to voters “rolling off” by simply not casting votes. Aspin and Hall (1987) noted that, while the amount of roll-off seemed relatively stable overall between 1964 and 1984, close retention elections saw less roll-off than those with lopsided “yes” votes (Hall and Aspin, 1987:15). Aspin (2011) noted a sharp decrease in roll-off from 2008-2010 that was mostly attributable to opposition campaigns. This continues a general downward trend in roll-off since 1964, the cause of which is unclear.\(^1\) Roll-off declined four percentage points nationally in 2010 from the 2008 average, its lowest value in forty years (Aspin, 2011:221). In the seven states where roll-off decreased more than 3 percentage points, five saw opposition campaigns (Aspin, 2011:221). Aspin’s findings suggest a relationship between the availability of information via opposition campaigns and the likelihood of voters participating in the election, but Aspin does not directly test this relationship.

This relationship has been persuasively demonstrated by findings on contested elections. Bonneau and Hall (2009) examine roll-off in partisan and nonpartisan elections as well as in retention elections, which they conclude “overall are not effective agents of voter mobilization.” But vigorous campaigning can contribute to more participation in judicial elections, and there is no reason to think that this cannot apply to retention elections. (In fact, television advertising in the Iowa 2010 retention election included an instruction to “turn over your ballot.”) Bonneau and Hall (2009) find that partisan elections and campaign spending are strongly negatively related to roll-off. In other words, the more campaigning there is, and the easier it is for voters to determine the likely policy positions of candidates, the more likely they are to vote in these elections. More recently, though, Streb and Frederick (2011) have questioned whether this same effect might be seen in low-salience elections. Studying 172 contested races for intermediate appellate court seats over five election cycles, they find that spending had no effect on ballot roll-off. As they conclude, these elections “are substantially less expensive than supreme court elections, so much so that it does not create conditions for voters to participate easily in the lower court elections” (p. 680).

Retention elections are nonpartisan affairs. But where campaigns can appeal to partisan themes, exploit partisan cues, and gain support from partisan figures, they are more likely to mobilize support. The 2010 election in Iowa was highly partisan. Republicans saw an opportunity to take control of the governor’s mansion, the House, and the Senate, which were all controlled by Democrats. With an unpopular incumbent governor at the head of the Democratic ticket, Republicans were by far the most motivated voters in the state. Almost 69 percent of all registered Republicans in the state turned out to vote in the election, while only 56.5 percent of Democrats and 36.5 percent of those registered as “no party” turned out (Iowa Secretary of State, 2011).

\(^1\) Aspin offers conjecture that it could be due to more opportunities for voters to evaluate judges via judicial performance evaluations. A recent paper by Casey, Dube, and Endersby (2011) supports this possibility, finding a positive relationship between bar poll results and retention margin, though the paper does not directly investigate roll-off.
Together with the one-sided nature of the campaign, with pro-retention messages being scarce and poorly communicated, this leads us to believe that partisan cues were more likely to mobilize Republicans. Thus, we expect roll-off to be negatively related to the share of the county electorate that is registered Republican.

Opponents of the judges mobilized because of opposition to a particularly salient court decision on a divisive social issue with significant religious dimensions. Evangelicals are powerful within the Iowa Republican Party (as witnessed by the success of Pat Robertson, Mike Huckabee, and Rick Santorum in the presidential caucuses). Views on gays and lesbians (Wilcox and Wolpert, 2000) in general, and views on same-sex marriage in particular, have historically been linked very closely to religious beliefs (Brewer and Wilcox, 2005; Wilcox et al., 2007). Green, Conger, and Guth (2005) found that, despite some schisms within the movement, Christian conservative activists in 2004 strongly opposed same-sex marriage and demonstrated widespread antipathy to the courts. One study of Defense of Marriage (DOMA) referenda across the states (Camp, 2008) notes that “given a particular configuration of circumstances—namely where a salient electoral issue cross-cuts already existing party cleavages—ballot referenda can become tools by which political parties ‘poach’ members of an opposing coalition” (p. 713). Using county-level data, Camp finds that DOMA ballot items mobilized not only conservative, white Protestants, but also African-Americans and Hispanics.

Because the mobilization against the three Iowa justices was prompted by same-sex marriage, opponents of the justices had a ready issue around which to mobilize voters. One clear path to defeating the justices was to try to transform the election into a referendum on same-sex marriage, thus providing a highly salient issue that voters could understand and that would motivate them to cast a ballot. If opponents of retention could increase the salience of the election in this way, we expect that the percent of county residents who belong to evangelical congregations will be negatively related to ballot roll-off.

The 2010 retention election saw significantly more advertising by the opposition campaign than by those urging retention. As detailed above, a number of studies of judicial elections have found a link between spending and roll-off, with more campaigning leading to less roll-off. Bonneau and Hall (2009b) find that attack ads decrease ballot roll-off in contested elections. Put another way, as potential voters are exposed to more negative information about incumbent justices, they become more likely to cast votes in the state supreme court election. Thus, we expect that levels of advertising in the Iowa retention election are negatively related to ballot roll-off.

**Vote Choice in a Salient Retention Election.** Griffin and Horan’s (1982) examination of results from Wyoming retention elections found that information was unrelated to voters’ choice to retain incumbent judges. Absent information to the contrary, voters seem to default to “yes” in the voting booth. With media coverage of retention elections typically low to nonexistent, voters often relied upon personal knowledge of the candidates (Griffin and Horan, 1979; Dubois, 1980b).
But if partisan cues are present, voters are more likely to be able to vote for candidates based on their own beliefs and attitudes. Baum (1987) surveyed voters following the 1984 Ohio Supreme Court election, asking them to cite the reasons for their vote choice. Party affiliation was a particularly important factor, even though the Ohio general election ballot does not identify the candidate’s party affiliation. A number of studies on contested elections (e.g., Bonneau and Hall, 2009a; Peters, 2009) show that partisan elections result in lower vote shares for incumbents. Such partisan cues allow voters to distinguish between their options and make informed choices.

As described above, the partisan cues of retention opponents were quite strong in Iowa’s election. Hence, we expect that the justices’ vote share will be negatively related to the percentage of voters in the county who are registered Republicans.

Similarly, as we have reviewed, views on same-sex marriage are closely linked to religious beliefs. There is evidence that these views are also tied to views on the courts (Green, Conger, and Guth, 2005). Thus, we expect justices’ vote share to be negatively related to the concentration of evangelicals within a county.

Baum (1987) also notes the importance of media, including advertising. The Republican candidates in the race he analyzed enjoyed large advertising advantages, and evidence suggests that Republican voters responded to this advertising (Baum, 1987:369). More recently, Bonneau and Hall (2009a) note that “increased spending in elections to state supreme courts has the effect of substantially enhancing citizen participation in these races” (Bonneau and Hall, 2009a:46). Bonneau and Hall (2009b) demonstrated that advertising, especially negative advertising, can help mobilize voters. Hall (2011a) likewise provides evidence that attack ads can diminish the incumbency advantage, particularly in nonpartisan elections where voter information is typically lower. Given that the only television ads broadcast in Iowa were attack ads, we expect that the amount of advertising against the justices will reduce the justices’ vote share within the county.

**DATA AND MEASUREMENT**

Below we estimate the effects of party, religion, and advertising on county-level results from the 2010 Iowa Supreme Court retention election. We analyze three dependent variables: ballot roll-off, change in ballot roll-off from an earlier election, and vote share (see Table 1). Although we conducted separate analyses for each justice, for simplicity’s sake we present single models that average the results for the three justices.² We measure ballot roll-off as the percentage of voters within each county who cast a ballot in the gubernatorial race but did not vote in the retention election.³ We also construct a similar measure for roll-off in the 2002 election and measure the difference

² We ran separate models of roll-off, change in roll-off, and vote share for each individual justice. In every case, the results were substantively identical to the combined models reported here. In fact, the results were so similar that the values of the coefficient estimates were almost identical.

³ (Votes cast for governor) – (Avg. number votes cast among the three justices)/(Votes cast for governor). We ran separate models of roll-off for each individual justice, and the results were substantively identical.
between roll-off in 2002 and 2010. We selected the 2002 election because it offered the best comparison to the 2010 election: it was also a gubernatorial election in a non-presidential election year, and it included on the ballot two of the justices defeated in 2010. We measure the justices’ vote share as the mean “yes” vote among the three justices within each county.

**Independent Variables**

*Partisanship.* It is well known that party is the most important cue available to voters. When this cue is absent, voter mobilization tends to be lower and voters lack information about the candidates (see, e.g., Dubois, 1980a; Lovrich and Sheldon, 1983). Comparing across electoral systems, Bonneau and Hall (2009a) report less voter roll-off in partisan systems, along with higher rates of campaign spending, higher likelihood of contestation, and lower vote shares for incumbents.

Clearly, partisanship is an important factor in mobilizing voters and in voter choice. We measure partisanship by calculating the percentage of voters within a county who are registered Republicans. Data on registered voters came from the Iowa Secretary of State’s Web site.

*Religion.* The key issue in the Iowa election was clearly the court’s same-sex marriage decision, which raised considerable religious cues. We measure the concentration of evangelicals within a county by identifying the number of evangelical adherents within each county using data from the Association of Religion Data Archives and dividing it by the total population from that county, as determined by the U.S. Census Bureau.

*Advertising.* As discussed, we expect that vigorous campaigning informs voters, transforming retention elections into something more akin to contested elections with

---

**Table 1**

Operationalization of Variables
(All Variables Measured at County Level)

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Roll-off</th>
<th>Change in roll-off</th>
<th>Vote share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Votes cast for governor) – (Avg. number votes cast among the three justices) / (votes cast for governor)</td>
<td>2010 roll-off – 2002 roll-off</td>
<td>Mean share of the vote received by the three justices within a county</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Republican registration</th>
<th>Evangelical adherents</th>
<th>Advertising</th>
<th>Education</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent of all registered voters registered as Republicans</td>
<td>Percent of county population who attend an evangelical church</td>
<td>Number of television ads broadcast in county</td>
<td>Percent of population over 25 holding at least a bachelor’s degree</td>
<td>Number of inhabitants in the county</td>
</tr>
</tbody>
</table>
a strong challenger. To measure advertising, we obtained the data from the Campaign Media Analysis Group (CMAG), which was made available to us by Justice at Stake. The data are reported by media market; using a map of media markets provided on CMAG’s Web site, we determined which counties fall in each media market. The variable is simply a count of the number of anti-retention ads that aired in the county.

Demographic Factors. We also control for several other factors that could influence the retention results. First, we expect that population density could be related both to more tolerance toward same-sex marriage (see., e.g., Smith, DeSantis, and Kassel, 2006), as well as to more support for merit selection (Dubois, 1990), both of which should increase the vote share of justices. Thus, we include a measure of the total population of the county, which we collected from 2010 Census Bureau data. Age is also known to be related to attitudes toward same-sex marriage. For example, a May 20, 2011 nationwide poll showed that among 18-34 year olds, 70 percent supported same-sex marriage, while only 39 percent of those 55 and above did so (Newport, 2011). We therefore include a measure of median age of the county gathered from Census Bureau data. Similarly, education levels are associated with more tolerance toward gay marriage. A 2010 survey from the Pew Center for the People and the Press found 52 percent of college graduates support same-sex marriage; support among those who had no more education than a high-school degree was only 34 percent (Pew Center, 2010). We determine the level of education at the county level by taking the percentage of county population that holds a bachelor’s degree or higher (U.S. Census,
We expect this to be positively related to support for the justices, but given the link between education and likelihood of voting we expect it to be negatively related to roll-off.

**Findings**

*Describing the 2010 Iowa Supreme Court Vote.* Figure 1 shows the retention election roll-off figure from 2010, with darker shades indicating more roll-off. Statewide, ballot roll-off was just 13.19 percentage points. Nineteen counties had roll-off values below 10 percentage points. To appreciate how unprecedented such high rates of participation are, we can compare the 2010 election to the 2002 election, when Justices Ternus and Streit were up for retention in a gubernatorial election year. In 2002, ballot roll-off was 38.73 percentage points. (Justices Streit and Ternus were retained overwhelmingly with an average of 73.8 percent of the vote.)

One key part of the roll-off puzzle, then, is explaining the unusually low roll-off in 2010. Figure 2 shows the difference in roll-off between the two elections (2002 and 2010), with darker shades indicating greater reductions in ballot roll-off. Statewide, the 2010 election marked a decrease in roll-off of 25.54 percentage points. Comparing Figure 2 with the maps of partisanship and religion in Figures 3 and 4, one can see the greater changes in roll-off among those counties with relatively high percentages of Republicans and evangelicals. In fact, heavily evangelical counties did not necessarily have lower ballot roll-off than other counties in 2010, but they did see a much greater
Figure 3
Registered Republicans as a Percent of County Population
2010 Iowa Supreme Court Election

Figure 4
Evangelical Adherents as a Percent of County Population
2010 Iowa Supreme Court Election
reduction in ballot roll-off from 2002 than other counties. In Sac County, in the
northwestern corner of the state, roll-off declined by a remarkable 81.86 percentage
points (94.24 percent roll-off in 2002; 12.38 percent in 2010). Other counties had
similar results, such as Sioux County (-61.73), Ida County (-50.87), and Crawford
County (-48.90). No counties had increases in roll-off; the lowest change was regis-
tered in Louisa and Pottawattamie counties at -11.82 and -13.11 points, respectively.

There appears to be quite a bit of overlap between change in roll-off and adver-
tising. Figure 5 shows the counties where anti-retention ads aired. Ad airings occurred
in 69 of the 99 counties in Iowa, with the number of ads ranging from 192 to 227.
When compared to the map of roll-off change from 2002, a pattern begins to emerge.
The counties with the least change in roll-off were generally focused in extreme southern,
northern, and eastern parts of the state, where no ads aired. In the central parts
of the state where ad airings were concentrated, we see much higher levels of change
in roll-off.

Figure 6 illustrates the justices’ vote share by county, with darker shades indicat-
ing more support. The justices’ support is primarily found in the central and eastern
parts of the state, where nearly all of the population centers are located and which
tends to be more heavily Democratic. In only nine out of ninety-nine counties did
a majority vote to retain the justices: Black Hawk (Waterloo/Cedar Falls), Clinton
(Quad Cities), Jefferson, Johnson (Iowa City), Linn (Cedar Rapids), Muscatine (Quad
Cities), Polk (Des Moines), Story (Ames), and Winneshiek. These nine counties
include five of the ten largest counties in the state. Even here, however, the majorities were narrow ones. For example, the justices obtained only 51.8 percent of the “yes” vote in Black Hawk and just 51.59 percent in Polk County.

As we examine the relationship between partisanship (Figure 3) and support for the justices, a clear pattern emerges: heavily Republican counties galvanized against the justices. Republican strongholds are primarily in the rural areas, particularly the western half of the state. As one would expect, rural areas outside of traditionally liberal, Democratic strongholds showed much less support for retention. This opposition was particularly strong in northwest Iowa, where heavily Republican counties voted overwhelmingly against retention. In Sioux County, for example, 73.49 percent of voters identify as Republican (only 8.84 percent as Democrat, with the balance as “no party”) and only 15.21 percent of voters voted to retain the justices. Other noteworthy counties include Lyon (62.57 percent Republican; 23.77 percent voting to retain), O’Brien (54.05 percent Republican; 24.62 percent voting “yes”), and Osceola (57.80 percent Republican; 24.62 percent for retention). As noted above, even some heavily Democratic counties voted not to retain the justices (e.g., Dubuque, where only 47.68 percent voted “yes”).

We can also see a relationship between the justices’ vote share and evangelical adherents within a county (Figure 4). The highest evangelical concentrations are found in western Iowa. If one were to overlay Figure 4 with Figure 6, one would notice a significant overlap between the concentration of evangelical adherents within a
county and the county’s lack of support for the justices. Many of these are in heavily Republican western Iowa. But also noteworthy is the presence of evangelical pockets in other parts of the state, including Decatur County (24.58 percent) in the south, Hardin County (20.06 percent) in the center of the state, and Iowa County (20.90 percent) in the east.

Explaining Ballot Roll-Off. Table 2 reports results of the OLS regression using ballot roll-off as the dependent variable. The model improves over a null model and explains about 32 percent of the variance in county-level roll-off. Neither age nor education, which have been found to be significant in some statewide studies of roll-off in supreme court elections (e.g., Bonneau and Hall, 2009b) were significant here (but see Streb and Frederick, 2011, where education levels did not explain roll-off in intermediate appellate court races). This may be due to less variance in these variables within a single state than among the two dozen states that have contestable judicial elections. And it may be due to the fact that roll-off was low across-the-board in Iowa in 2010.

Although the coefficients for the partisanship of the county and the concentration of evangelicals are signed as expected, neither approaches statistical significance. Mobilization, it appears, was not as one-sided as we expected it to be. Advertising, on the other hand, is highly significant and in the expected direction. If two hundred ads were aired within a county, the county’s roll-off decreased by an average of 4 percentage points.
Next we turn to examining how the 2010 election differed from a more typical retention election for Iowa’s high court. Table 4 presents results of a regression on change in ballot roll-off from 2002 to 2010. We calculated this by subtracting 2002 roll-off from 2010 roll-off. Thus, a lower value indicates less roll-off in 2010 than in the previous election, and therefore higher levels of voter mobilization.

The model performs moderately well, explaining 31 percent of variation in the change in roll-off. Most importantly, two of the three items we hypothesize to be responsible for voter mobilization in 2010 are significant: concentration of evangelicals and advertising. While roll-off was lower in all counties in 2010, that effect was greatest in counties with more evangelicals—the higher the concentration of evangelicals within a county, the less roll-off the county saw in 2010 than in 2002. For each 10-percentage-point difference in the concentration of evangelicals within a county, roll-off declined by an average of 5.6 percentage points in 2010 compared to 2002.

Advertising was also effective in mobilizing voters to cast a ballot in the retention election. Those counties in which the opposition ran ads (roughly 200 airings) saw about a 7-point reduction in ballot roll-off in 2010. Thus, it appears that the successful mobilization of voters to defeat the justices began with grassroots appeals to evangelicals and over-the-air, broad-based commercial appeals that convinced more people than normal to turn over their ballot and vote in the retention election.

These results suggest that the political environment, centering on same-sex marriage, allowed opponents of the justices to mobilize evangelicals to cast votes in
Table 4
Determinants of Average County-Level Vote Share of Incumbents in 2010 Iowa Supreme Court Retention Election (OLS Regression with Robust Standard Errors)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Robust s.e.</th>
<th>t</th>
<th>p-value</th>
<th>Standardized Coefficient (β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.5189</td>
<td>0.0757</td>
<td>6.85</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Republican registration</td>
<td>-0.4223</td>
<td>0.0474</td>
<td>-8.91</td>
<td>0.0000</td>
<td>-0.493</td>
</tr>
<tr>
<td>Evangelical adherents</td>
<td>-0.2041</td>
<td>0.0620</td>
<td>-3.29</td>
<td>0.0014</td>
<td>-0.193</td>
</tr>
<tr>
<td>Median age</td>
<td>-0.0001</td>
<td>0.0015</td>
<td>-0.60</td>
<td>0.5522</td>
<td>-0.043</td>
</tr>
<tr>
<td>Ad airings</td>
<td>-0.0001</td>
<td>0.0000</td>
<td>-3.63</td>
<td>0.0005</td>
<td>-0.204</td>
</tr>
<tr>
<td>Education</td>
<td>0.5762</td>
<td>0.0689</td>
<td>8.36</td>
<td>0.0000</td>
<td>-0.476</td>
</tr>
</tbody>
</table>

R² = 0.743
Adj. R² = 0.729
Root MSE = .040
F(5, 93) = 7.17
Prob > F = 0.00
N = 99

the 2010 retention election. The concentration of evangelicals within a county contributed to lower levels of ballot roll-off. Moreover, even in an election that was highly salient and in which roll-off declined across the board, counties with more evangelical adherents witnessed a steeper decline in roll-off than did other counties. The mobilization was not wholly grassroots, however. Broadcast advertising also played a role in driving down roll-off compared to earlier elections, contributing to the mobilization of voters in much the same way that negative ads drive down roll-off in partisan and nonpartisan judicial elections (Hall, 2011).

Explaining the Retention Rate. Table 3 reports results of the OLS regression using the average percent of the “yes” vote as the dependent variable. Overall, the retention rate model explains about 75 percent of the variance in county-level vote share. The percentage of registered Republicans, percentage of evangelical adherents, and number of ad airings were all negatively and significantly related to the average “yes” vote. Level of education and population were significant and in the expected direction. The standardized coefficients reveal that partisan registration within the county has the largest impact on the justices’ vote share. The concentration of evangelicals and the broadcast of anti-retention advertising also have a sizable negative effect.

As noted, the results of the model suggest that partisan balance within the county is key to explaining the justices’ vote share. Even after controlling for other factors, the justices lost an average of 4 percent of the “yes” vote for every 10-percentage-point increase in registered Republicans within a county. The mean percentage of registered
Republicans across the 99 counties is .349, with a standard deviation of .09. Thus, a swing from one standard deviation above the mean to one standard deviation below the mean is tantamount to nearly an 8-percentage-point vote swing against retention.

The influence of evangelicals is also visible. For every 10-percentage-point increase in evangelical adherents, the justices lost an average of 2 percent of the “yes” vote, after controlling for other factors including partisanship. A county one standard deviation above the mean level of evangelical adherents would have a “yes” vote total, on average, about 2.8 percentage points lower than a county one standard deviation below the mean (mean = .13, s.d. = .07).

As expected, anti-retention advertising also helped to drive down the justices’ vote share. Though there was not much variation among the levels of advertising (totals among the three media markets that saw advertising range from only 192 to 227), roughly one-third of counties saw no over-the-air-advertising. If we compare two counties, one of which saw no ads and one of which featured roughly 200 negative ads, vote totals in the county with advertising would be, on average, about 3 percentage points lower.

Education and population density contributed positively to the justices’ vote share. For every 10-percentage-point increase in highly educated voters (bachelor’s degree or higher), the justices gained an average of 5.95 percent of the “yes” vote. Finally, for every 1,000 additional county residents, the justices gained an average of 2 percent of the “yes” vote.

Thus, our hypotheses regarding the 2010 Iowa retention election are well supported. The analysis of election results suggests a highly partisan affair in which social views on a highly controversial issue were made salient by activists and advertising, triggering partisan cues that influenced the vote.

CONCLUSION

Although the norm for retention elections involves high levels of ballot roll-off and high margins of victory for incumbents, these elections can be transformed by campaigns that inform and mobilize voters. When this happens, their outcomes may be characterized by some of the same phenomena that explain the outcomes of partisan and nonpartisan elections: the vigor of the campaign and the political context of the election. Our findings suggest that in Iowa’s 2010 retention election, opponents of the justices were effective at defining the election as a partisan issue, giving Republican voters a simple cue to determine their vote choice. The justices’ opponents were also able to mobilize evangelical Christians to turn out in greater numbers than in the past and to vote against the justices. Finally, there is evidence that the substantial resources devoted to broadcast ads paid off for opponents of retention. Even after controlling for other factors, advertising helped to mobilize voters and to spur them to vote against the justices.

Some might argue that Iowa’s 2010 retention election was a one-off event driven entirely by the uniquely salient issue of same-sex marriage. If this is the case, then
studying a single election of this sort may seem of limited utility, inapplicable beyond the place and time it happened. If a case study is not representative, after all, the conclusions drawn from studying it may not be validly applied to other cases. We believe, however, that our results have clear implications for practitioners and scholars.

The lesson of these results for groups that wish to target sitting merit-plan judges is clear: launching a campaign that plays on current political issues can increase the amount of information available to voters, activate dormant partisan attitudes, and potentially mobilize groups that would otherwise ignore the retention election. The lesson for targeted justices is equally clear, if one assumes that their primary motivation is keeping their job: raise money and campaign. In Iowa, the justices refused to campaign on their own behalf, and pro-retention advocates were not well enough organized or financed to mount a successful defense on their behalf. By contrast, in neighboring Illinois, Chief Justice Kilbride was challenged in his retention bid by pro-business groups, but he raised $2.8 million to the $688,000 spent by the groups opposing him (Skaggs et al., 2011). He received 65 percent of the vote. In addition to Illinois and Iowa, justices in Alaska, Kansas, Florida, and Colorado faced opposition in 2010. Even though their opposition was substantially less well organized and financed, the campaign in Alaska was nearly successful: Chief Justice Fabe was reelected with only 55 percent of the vote.

In 2012 an opposition campaign mounted against Iowa Supreme Court Justice David Wiggins, another member of the Varnum court, failed to unseat him. Opponents spent less than $500,000 this time around, and supporters spent nearly $350,000 promoting the merit plan and encouraging retention. Thus, pro- and anti-retention coalitions were on much more equal footing in 2012 than in 2010, where those opposed to retention outspent supporters by nearly 3 to 1. Ballot roll-off was not as low as it was in 2010, but it was still much lower than typical: 21.2 percent. Wiggins received roughly 54.5 percent of the vote, far below the 75 percent of the vote received by the three new justices running in their first retention race. He outperformed President Obama in the state by only 2 percentage points; further analysis will likely reveal that partisanship was a driving factor in his vote share.

In addition to the effort in Iowa, the Republican Party of Florida, along with a Koch brothers-sponsored independent group, targeted three Florida Supreme Court justices running for retention. But according to Justice at Stake (2012), they were wildly outspent by the justices themselves, who, combined, raised $1.5 million, and by an independent pro-retention campaign that spent at least $2.7 million on television ads. State high court justices in Indiana and Arizona also saw challenges from Tea Party groups, though both of them won retention by a large margin.

It is possible that Iowa’s 2010 retention election represented nothing more than an exception to the general rule of low-salience retention elections. After all, we have seen periodic campaigns against justices before. It is also possible, however, that 2010 represents the year that “new-style” campaigning in state supreme court elections began to bleed into retention elections. If this is true, merit-plan justices who play by
the old rules, like the Iowa justices did, will be at significant risk for defeat. The merit plan may have offered a defense against perceived threats to judicial independence via the selection process a century ago, but today’s selection-and-retention process operates under a new style of politics, where interest groups exercise influence via campaigning against retention and mobilizing voters by providing them with negative information and partisan cues. Over the past twenty years, interest groups, many with national affiliations, like the Chamber of Commerce, have actively injected themselves into contested state supreme court elections. The playbook here is well known and proven: identify a state where control of the court could be at stake, identify a decision an incumbent has made that could resonate with voters (often criminal cases where a “criminal goes free”), and run attack ads.\(^4\) Retention elections offer incumbents less protection against such tactics than do contested elections. To the extent that opponents of retention can seize on a salient issue and use partisan cues to their advantage, justices in retention elections may need to do what Justice Kilbride in Illinois did in 2010 or what the Florida justices did in 2012: raise money and run campaigns.

Defenders of retention elections argue that this would harm the legitimacy of the courts. Cann and Yates (2008) report that, while retention elections have not had any effect on diffuse support for the courts, partisan elections may negatively affect their legitimacy. Gibson et al. (2010), on the other hand, find that elections may have a net positive effect on the legitimacy of courts. There may be a parallel here between salient retention elections and U.S. Supreme Court confirmation battles. Gibson and Caldeira (2009) have explored the effects of such confirmation fights on citizen assessment of the Court’s legitimacy. Among their findings: advertising may decay assessment of institutional legitimacy by shifting the frame of reference from one based in legal symbols to one rooted in pure politics. The comparison to retention elections is not perfect, but the similarities cannot be ignored. In both cases, the decision is yes or no, and, typically, the nominee/incumbent does not do much speaking for himself or herself. Instead, groups both for and against attempt to mobilize supporters or opponents. The justices in Iowa who refused to campaign frequently cited their perception of the importance of the legitimacy of the courts—they felt that they would undermine that legitimacy (not to mention play into their opponents’ hands) if they were seen as the very politically motivated actors that they were attacked as being. Given Gibson and Caldeira’s findings, they may be right about that threat. jsj

\(^4\) Identifying an opportunity for changing the balance of the court may be key. In Iowa in 2010, the incumbent governor, Chet Culver, was very unpopular and had little chance of winning his reelection bid. Would same-sex marriage opponents have been as eager to mount a campaign against the justices if they did not know that a Republican would be appointing their successors?
REFERENCES


A COUNTY-LEVEL ANALYSIS OF THE 2010 IOWA SUPREME COURT RETENTION ELECTION


Comparing Between-Judge Disparities in Imprisonment Decisions Across Sentencing Regimes in Ohio*

John Wooldredge, Timothy Griffin, and Amy Thistlethwaite

Empirical studies of judicial effects on the use of imprisonment have yet to estimate changes in these effects under more-structured sentencing schemes. Findings are presented from a multilevel analysis of whether the implementation of Ohio’s presumptive guidelines in 1996 was effective for reducing inter-judge differences in the distribution of non-suspended prison sentences and in defendant-level effects on imprisonment. These data for seven urban courts provide a unique opportunity to estimate changes in judicial effects across sentencing regimes. Results suggest that Ohio’s guidelines were successful for reducing judicial effects on sentencing based on judges’ tenure on the bench, prosecutorial experience, and caseload composition.

Scholarly interest in judicial effects on sentencing (such as whether the use of imprisonment for convicted defendants varies by a judge’s sex, race, years on the bench, etc.) has generated a number of studies on the subject (e.g., Anderson and Spohn, 2010; Collins and Moyer, 2008; Frazier and Bock, 1982; Gruhl, Spohn, and Welch, 1981; Johnson, 2006; Kritzer and Uhlman, 1977; Meyer and Jeslow, 1997; Muhlhausen, 2004; Myers and Talarico, 1987; Scott, 2010; Spohn 1990a, b; Steffensmeier and Britt, 2001; Steffensmeier and Hebert, 1999; Tiede, Carp, and Manner, 2010; Ulman, 1978; Ulmer, 2005; Welch, Combs, and Gruhl, 1988). All of these studies have found significant differences across judges in the use of imprisonment, although findings are mixed for specific types of judicial effects such as a judge’s sex and race. Over the past decade, more attention has been paid to inter-judge disparities under sentencing guidelines (United States Sentencing Commission, 2005), with discussions of how these disparities might be curbed with more-structured sentencing schemes. However, the most recent research on this topic has focused primarily on the federal sentencing guidelines (with a couple of exceptions such as Johnson, 2006; Ulmer, 2005) and have consisted mostly of post-guideline analyses that cannot determine whether changes in sentencing schemes actually coincided with reductions in judicial effects.

* Supported, in part, under award No. 98-CE-VX-0015 from the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this report are those of the authors and do not necessarily represent the official position of the U.S. Department of Justice. Thanks to Travis Pratt and Fritz Rauschenberg for their assistance on this project. John Wooldredge (John.Wooldredge@uc.edu) is a professor in the Division of Criminal Justice, University of Cincinnati. Timothy Griffin (tgriffin@unr.edu) is a member of the faculty of Department of Criminal Justice at the University of Nevada at Reno. Amy Thistlethwaite (thistlethwai@nku.edu) is an associate professor of criminal justice on the Political Science, Criminal Justice, and Organizational Leadership faculty at Northern Kentucky University in Highland Heights.
To address judicial effects on sentencing under state guidelines and whether these effects were stronger before their implementation, we took advantage of examining an older data set tapping sentencing decisions across Ohio Common Pleas Courts immediately before and shortly after the implementation of presumptive sentencing guidelines in 1996 (Wooldredge et al., 2002). An analysis of judicial effects on sentencing has never been performed with these data because of the absence of information on judges in the original set (which were added to conduct the analyses presented here). While findings may not generalize to sentencing across the state of Ohio today, given that the presumptive guidelines became voluntary in 2006, these data provide a unique opportunity to more directly examine whether changes in the structure of sentencing decisions might correspond with greater consistency in the use of imprisonment across judges, regardless of their personal characteristics.

The question of whether more-structured sentencing schemes might help to curb inter-judge disparities in sentencing and reduce judicial effects based on factors such as a judge’s race, sex, or tenure on the bench is a different question from whether defendant-level effects based on extralegal factors tend to dissipate under more-structured sentencing. Defendant-level effects might appear to be null when pooled across judges, yet null effects could result from differences in the strength of defendant-level effects across judges. To more fully understand whether more-structured sentencing is successful for reducing disparate sentencing, it is necessary to examine changes in the uniformity of sentencing across judges and whether greater uniformity also reflects the absence of considerations of defendants’ personal characteristics. To shed light on these issues, judicial effects on decisions to send convicted felons to prison were compared before and after the implementation of presumptive sentencing guidelines in Ohio. Samples of convicted felons were examined from the seven largest counties in the state.

DIFFERENCES IN SENTENCING PRACTICES ACROSS JUDGES

The utility of examining judicial effects on sentencing disparities depends on the existence of significant differences in sentencing between judges. A greater use of multilevel modeling in related studies has permitted more reliable estimates of these differences, and recent studies of U.S. District Courts, as well as Pennsylvania and Ohio state courts, have consistently produced evidence of significant between-judge variance in the use of prison sentences, the length of imprisonment, or both (Anderson and Spohn, 2010; Collins and Moyer, 2008; Johnson, 2006; Scott, 2010; Tiede, Carp, and Manner, 2010; Ulmer, 2005; United States Sentencing Commission, 2005; Wooldredge, 2010). On the other hand, these significant between-judge differences vary in magnitude across studies, such as Johnson’s (2006) observation that 5 percent of the variance in imprisonment decisions fell between judges from all Pennsylvania trial courts of general jurisdiction versus Wooldredge’s (2010) estimate of 20 percent of the variance in imprisonment falling between judges in a single urban Ohio jurisdiction. Although these different estimates were statistically significant, they reflect a
relatively broad range in the magnitude of inter-judge differences in sentencing and underscore the need to estimate such differences to place judicial effects on sentencing disparity in their proper context (i.e., different judicial effects across jurisdictions could easily result from differences in variance estimates because such effects are more likely to be null when inter-judge variance in sentencing is limited to begin with). In short, while it is common to tacitly assume between-judge differences in sentencing exist, estimating the magnitude of these differences is important for understanding why particular judicial effects might be relevant in some jurisdictions but not others.

Related studies conducted to date have produced mixed findings regarding the relevance of judicial characteristics for understanding sentencing disparities. Some of these anomalous results could stem from methodological differences across studies or differences in the courts examined, which might contribute to or inhibit inter-judge disparities. This body of research has focused primarily on whether the background characteristics of judges influence their sentencing decisions, such as whether minority or female judges are more lenient than those who are white or male (for analyses of these and other characteristics, see Anderson and Spohn, 2010; Collins and Moyer, 2008; Frazier and Bock, 1982; Gruhl, Spohn, and Welch, 1981; Johnson, 2006; Kritzer and Uhlman, 1977; Meyer and Jesilow, 1997; Muhlhausen, 2004; Myers and Talarico, 1987; Scott, 2010; Spohn 1990a, b; Steffensmier and Britt, 2001; Steffensmier and Hebert, 1999; Tiede, Carp, and Manner, 2010; Ulman, 1978; Ulmer, 2005; Welch, Combs, and Gruhl, 1988).

Among some of the findings related to a judge’s race, Johnson (2006) found that minority judges in Pennsylvania were less likely to send someone to prison and, in cases where they did send offenders to prison, they also administered shorter sentences. A different race effect was uncovered by Muhlhausen (2004), who also conducted an analysis of Pennsylvania judges, where African-American judges administered longer prison terms than white judges. Regarding sentencing differences based on a judge’s sex, Spohn (1990b) discovered longer prison terms distributed by female judges to defendants convicted of sexual assaults in Detroit. A more general effect was uncovered by Steffensmier and Hebert (1999) in their statewide analysis of Pennsylvania judges, where female judges administered harsher sentences relative to male judges in general. A qualification to these findings was offered by Tiede, Carp, and Manner (2010), who found that female U.S. District Court judges distributed harsher sentences only when they were appointed under Republican presidents. Contrary to all of these findings for a judge’s sex, Meyer and Jesilow (1996) found that female judges were less likely than male judges to distribute jail sentences in a California municipal court.

Most studies of judicial effects on sentencing have focused on the role of a judges’ race, sex, or both. A few studies have focused on other potentially relevant factors, such as a judge’s age, caseload, and tenure (time on the bench). For example, Johnson (2006) found that older judges in Pennsylvania state courts were less likely to send someone to prison and, in cases where they did send offenders to prison, they also
administered shorter sentences. He also found that judges with heavier caseloads of violent offenses and drug offenses tended to administer longer prison terms. By contrast, Anderson and Spohn (2010) found that a U.S. District Court judge’s overall caseload of sentenced offenders was inversely related to prison sentence length. Whereas Johnson (2006) found marginal support for positive relationships between a judge’s tenure (time on the bench) and both the odds of imprisonment and prison sentence length, Anderson and Spohn (2010) found no relationship between tenure and sentence length in the federal system.

The general conclusion to date is that judicial backgrounds hold little predictive power when it comes to explaining between-judge variance in sentencing (for exceptions, see Anderson and Spohn, 2010; Johnson, 2006; Meyer and Jesilow, 1996). This must be tempered, however, with the observation that related studies are few in number.

**GREATER UNIFORMITY WITH MORE-STRUCTURED SENTENCING**

Proponents of more-structured sentencing have argued that determinate- or presumptive-sentencing schemes should help to reduce judicial discretion and, in turn, inequities in sentences for similarly situated offenders (i.e., persons convicted of the same charges and with similar criminal backgrounds). A study conducted by the United States Sentencing Commission (2005) produced some evidence in favor of this idea, where inter-judge variance in sentencing was reduced between one-third to one-half before versus after the implementation of the federal guidelines. The reduction varied by type of crime, however, and regional differences in outcomes for drug-trafficking cases actually increased under the guidelines (USSC, 2005). The federal sentencing guidelines shifted from mandatory to voluntary after the U.S. Supreme Court ruled in *Booker* and *Fanfan* (543 U.S. 220) that the guidelines are no longer binding on judges and may not be the only factor considered in sentencing decisions. Scott (2010) compared inter-judge differences in sentencing before versus after *Booker/Fanfan* and found an increase in between-judge differences post-*Booker/Fanfan*, also consistent with the idea that more-structured sentencing should coincide with greater uniformity in sentencing across judges. In theory, restricting the factors that judges are allowed to consider in their decisions in conjunction with narrowing the range of minimum and maximum prison terms for a given charge should generate more-similar outcomes in similar cases and across judges.

Despite the findings above regarding reduced disparities under the federal mandatory guidelines, significant disparities still existed under the guidelines even before they became “advisory” (Anderson and Spohn, 2010; Scott, 2010; Tiede, 2009; USSC, 2005). Ulmer, Light, and Kramer (2011) actually found no meaningful differences pre- versus post-*Booker/Fanfan* in defendants’ extralegal effects on imprisonment. All of these studies suggest that guidelines might be limited in their impact on reducing sentencing disparities, at least among certain judges. Some scholars have posed that more structured sentencing schemes might not actually translate into more
uniform sentencing because of plea bargaining and the ability of judges to approve particular guilty plea agreements to maneuver around guideline restrictions (Miethe, 1987). Prosecutors can also avoid mandatory prison sentences in particular cases through their decisions regarding the severity of formal charges and whether to subsequently drop some or all formal charges against a defendant (Lagoy, Hussey, and Kramer, 1979).

The efforts of court actors to preserve individualized justice have been discussed in the context of “substantive rationality” in decision making (Dixon, 1995). A substantively rational decision might be very different than a formally rational decision that is dictated solely by the rules of law. Whereas formal rationality under sentencing guidelines might dictate a mandatory prison sentence for anyone convicted a second time on a more serious felony, a substantively rational decision is influenced by both rules of law, as well as other factors falling outside the parameters set by those rules. For example, two defendants arrested for the same crimes might provoke different concerns among court actors, where a judge and a prosecutor might be more concerned with community protection in one case versus preserving the other defendant’s community ties (Steffensmeier, Ulmer, and Kramer, 1998). Whereas formally rational decision making might lead to the imprisonment of both defendants, a prosecutor might be more willing to offer a plea agreement involving reduced charges for the defendant who fits a stereotype of a “good risk” offender (e.g., a female with dependent children and a stable work history).

These arguments for why extralegal disparities in the application of non-suspended prison sentences may not be reduced with more structured sentencing cannot be tested directly with these data. They are presented here only as possible reasons for the persistence of such disparities under sentencing guidelines.

BEFORE/AFTER STUDIES OF MORE-STRUCTURED SENTENCING
The limited number of studies examining changes in disparate sentencing under different sentencing schemes have focused on short-term effects only (Moore and Miethe, 1986; Parent et al., 1996; Bales, 1997; Wooldredge, Griffin, and Rauschenberg, 2005), as well as on both short- and long-term effects (Stolzenberg and D’Allessio, 1994; Koons-Witt, 2002; USSC, 2005; Ulmer, Light, and Kramer, 2011; Wooldredge, 2009). Except for the USSC (2005) study, however, all of these studies have focused on changes in the magnitude of defendant-level effects on sentencing pooled across multiple judges. This approach is useful for determining whether defendant-level effects involving extralegal factors have generally waned over time, but it cannot reveal possible differences in how individual judges respond to more-structured sentencing and whether guidelines are more or less effective for weakening particular judicial effects based on a judge’s race, sex, experience, etc.

There are three possible questions related to a focus on changes in judicial effects under more-structured sentencing. First, are differences in sentencing practices
between judges reduced under guidelines? Even if not eliminated, significant reductions in these differences could reflect some success in generating greater uniformity in sentencing across judges. Second, are between-judge differences in the magnitude of specific defendant effects reduced under guidelines? For example, Johnson (2006) found significant between-judge differences in defendants’ race effects on prison sentences, suggesting that some judges are more prone than others to treat defendants differently based on their race. If so, then might these differences disappear or become smaller under more-structured sentencing? Finally, do the effects of judges’ personal characteristics on sentencing tend to dissipate under guidelines? For example, if male judges were more prone to administer harsher sentences than female judges before the guidelines, did this sex effect disappear under the guidelines?

The analysis described here focused on whether Ohio’s presumptive guidelines were effective for reducing or eliminating the disparities described above. This focus moves beyond the question of whether defendant-level effects become weaker under more structured sentencing by recognizing that such effects may be judge-specific and, if so, whether inter-judge differences in these effects are reduced under more-structured sentencing. Another important aspect of examining the sentencing scheme implemented in 1996 is that these presumptive guidelines were declared unconstitutional by the Ohio Supreme Court in 2006 and became strictly “voluntary” (State v. Foster, 109 Ohio St.3d 1, 2006-Ohio-856). Public officials have voiced concerns about greater sentencing disparities and more crowded prisons under the voluntary guidelines (Worner, 2006). Although sentencing disparities under the voluntary guideline scheme are not examined here, the comparison of sentencing practices before versus under presumptive guidelines provides insight into whether these concerns might be warranted. A fairly thorough description of Ohio’s guidelines can be found in Wooldredge et al. (2002), but the Appendix presents a brief summary of the scheme to place the analysis and findings in their proper context.

**METHODS**

The analysis described here focused on one outcome measure tapping whether a convicted felon received a non-suspended prison sentence.¹ Related to inter-judge differences in sentencing and judicial effects on imprisonment, the following research questions were examined for each study period (before versus after the implementation of Ohio’s sentencing guidelines):

1. What proportion of total variation in imprisonment was explained by the sentencing judge?
2. Did the magnitude of defendants’ extralegal effects (race, sex, age, children, employment status) on imprisonment differ significantly across judges?

¹ Sentence length for an imprisoned felon was also examined but generated the same substantive conclusions regarding inter-judge disparities in sentencing. The findings for sentence length are not described here due to space limitations and because of the argument that related analyses require statistical adjustments for sample selection bias associated with examining samples of imprisoned offenders (Stolzenberg and Relles, 1997). A selection bias adjustment void of collinearity with other predictors in the models examined could not be created. Findings for the “uncorrected” models are available upon request.
3. Did differences in judicial characteristics (race, sex, tenure on the bench, etc.) account for significant proportions of between-judge variation in imprisonment?

Examining these questions for the pre- and post-guideline periods separately permitted examination of whether the disparities reflected in these questions were reduced under presumptive guidelines. More specifically and in order of the above hypotheses:

A. Were inter-judge differences in the use of imprisonment reduced under guidelines (i.e., was the magnitude of inter-judge variation in imprisonment smaller for the post-guideline period relative to the pre-guideline period)?

B. Were inter-judge differences in the magnitude of specific defendant effects reduced under guidelines?

C. Did judicial effects on the use of imprisonment dissipate under guidelines?

**Samples.** The pre- and post-guideline samples examined here were cross-sections of defendants convicted on felony charges in the seven largest urban counties in Ohio (Cuyahoga, Franklin, Hamilton, Montgomery, Lucas, Summit, and Lorain). These cases were selected as part of a larger study of case processing and sentencing in twenty-four counties (Wooldredge et al., 2002). The analysis described here was restricted to the seven largest counties due to the focus on judicial effects and the need to include enough cases per judge to estimate the models of interest. In other words, none of the other (smaller) counties generated enough cases per individual judge to support the analysis of judges in those counties.

The original sampling procedures were designed to produce two separate cross-sections of indicted felony suspects reflecting the pre- and post-guideline populations of defendants, for the purpose of comparing multivariate models predicting case outcomes for each group. The Ohio sentencing guidelines were implemented on July 1, 1996, and persons indicted between July 1, 1995 and June 30, 1996 (pre-guidelines) as well as persons indicted between January 1, 1997 and December 31, 1997 (post-guidelines) were targeted for the samples.

Systematic sampling of prosecutors’ case files was used to ensure proportionate representation by case year within each county. Except for Lorain County, a 5 percent sample of indictments was selected from each of the urban counties during the two time periods. Although still a metropolitan statistical area, the smaller population of Lorain required a larger proportionate sample, and so a 15 percent sample of indictments was selected from this county. The first codefendant listed on any file with multiple defendants was selected since the unit of analysis was the individual.

The sampling procedure yielded 3,693 indicted felons from the seven counties (pre-guidelines = 1,262; post-guidelines = 1,238). Only those defendants who were convicted on felony charges were examined here. To provide enough degrees of freedom for the statistical models, only cases processed by judges who presided over at least twenty cases during the study period were included (for a similar strategy, see Johnson, 2006). A relatively small number of Mexican-American defendants were also
excluded because the breakdown of these individuals by judge resulted in too few cases per judge to support a reliable analysis of the effects of a defendant’s ethnicity on sentencing. The final samples for the analyses described here included 2,500 convicted felons processed under 107 judges (pre-guidelines = 1,262 felons under 104 judges; post-guidelines = 1,238 felons under 99 judges).

**Data and Measures.** Data on defendants were gathered from district attorneys’ offices, probation departments, and the Ohio Department of Rehabilitation and Correction (ODRC). Data on judges were gathered primarily from the Ohio State Bar Association, annual case-processing reports of each common pleas court provided to the Ohio Supreme Court, and county Web sites offering descriptions of the judiciary. Some information was gathered from newspaper articles highlighting the contributions of particular judges, Web sites for law schools from which these judges graduated, and obituaries appearing in various forums. The defendant and judge measures included in the subsequent models are described in Table 1. At the defendant level, some of the original ratio scales were logged due to coefficients of variation larger than 1.0 (Fox, 2008).

The measures of defendants’ extralegal characteristics included a defendant’s race (African-American or white), sex, age (18-24, 25-39, 40+), whether employed at arrest, and whether she or he had dependent children. All of these measures were dummy coded. The age groups roughly follow Steffensmeier, Kramer, and Ulmer’s (1995) discussion of differences in sentence severity between very young, relatively young, and older defendants. They found that the youngest and the oldest defendants were treated less severely than those falling in between the two groups, and so the reference category for the analysis described here was defendants aged 25-39. (Steffensmeier and his colleagues examined different age groups including 18-29, 30-49, and 50+, but half of the defendants examined here fell between 19 and 29 years of age, whereas only 5 percent were 50 or older.)

Absent from this list of extralegal measures are interaction terms reflecting a defendant’s race (x) sex (x) age, which are often examined. These interaction terms could not be examined here because several product terms would have had to be included in the models, and the number of measures would have reduced the degrees of freedom to a level that excluded many judges from the analysis (i.e., with defendants “nested” within judges, the limited numbers of sampled cases processed by some of these judges would have forced their exclusion from the analysis).²

Legally relevant measures at the defendant level were included in the models to reduce the odds of finding spurious extralegal effects, in situations where legal measures were correlated with the extralegal measures, and to lessen the chances of missing significant extralegal effects, in cases where legal correlates of the outcome only were not controlled (Zatz, 1987; Berk, Li, and Hickman, 2004). These predictors

---

included the most serious felony charge a defendant was convicted on (a series of dummy coded measures), number of counts convicted on, number of prior prison terms served by a defendant, whether a defendant was incarcerated as a juvenile, any history of drug or alcohol addiction, whether the defendant pled guilty (the relevance
of which has been argued by LaFree, 1985; Albonetti, 1990; Steffensmeier, Ulmer, and Kramer, 1998), and whether she or he had a public defender (Albonetti, 1990, 1991). Other legally relevant measures explored for the analysis included specific felony offense measures (homicide, manslaughter, rape, aggravated sexual assault, etc.), whether a defendant was convicted on any gun specifications (carrying mandatory prison terms), and whether she or he obtained pretrial release. These measures were ultimately excluded because their inclusion in the models did not result in significantly different estimates of the extralegal effects (i.e., estimates of extralegal effects in the model with these legal factors fell within 95 percent confidence intervals for the estimates of extralegal effects in the model without these legal factors), and because their inclusion reduced degrees of freedom and resulted in smaller samples of judges and defendants available for model estimation.

The binary measures of charge severity were felony 1 through felony 5, tapping the most serious charge convicted on. The felony 5 category did not exist before the guidelines, so reference groups for the pre-guideline and post-guideline periods were felony 4 and felony 5, respectively.

At the second (judge) level of analysis, the measures of judicial characteristics included a judge’s sex, race (African-American or white), number of years on the bench at the time of study, number of years served as a prosecutor, whether the judge graduated from a first-tier law school (based on U.S. News and World Report’s ranking of the top 50 law schools in the United States), the judge’s felony caseload during the year examined, and the proportion of each judge’s caseload consisting of drug crimes. Also explored were a judge’s age, ethnicity (Latino or white Anglo), number of years since receiving the juris doctorate (JD), and the proportions of a judge’s caseload consisting of property crimes and personal crimes. These last five measures were ultimately dropped because age and years since earning the JD were collinear with tenure on the bench; there was only one Latino judge in the sample; and the “proportion” measures did not improve prediction or alter the remaining estimates. With the exception of law school ranking, all of these measures (or variants of) were examined by Johnson (2006) in his multilevel study of judicial effects on sentencing in Pennsylvania trial courts. Anderson and Spohn (2010) also examined years on the bench and judicial caseloads, although they measured the latter as the number of convicted defendants sentenced. As previously described, studies aside from Anderson and Spohn (2010) and Johnson (2006) focused primarily on a judge’s sex and race.

**Statistical Analysis.** Multilevel modeling was used for the analysis due to the nested nature of the data, with 2,500 defendants grouped within 107 judges from 7 jurisdictions. Aside from the need to group defendants by judge to examine inter-judge variability in sentencing and the judicial effects of interest, judges, in turn, were grouped by jurisdiction to control for possible jurisdiction differences in sentencing practices via group mean-centering of the “judge” measures (see Raudenbush and Bryk, 2002, for other examples of and rationales for nesting data). The statistical software used for the analysis was HLM7.0 (Raudenbush, Bryk, and Congdon, 2011).
Binary logistic regression models were estimated due to the dichotomous outcome measure of whether a convicted felon went to prison. Unconditional models (with no predictors) provided the baseline estimates of the proportion of variation in imprisonment falling between judges during the pre- and post-guideline periods. Next, random coefficient models with only the defendant measures included revealed the significant defendant-level effects on sentencing and whether any of these effects varied significantly across judges during each time period. Finally, intercepts-as-outcomes models were estimated to identify significant judicial effects on sentencing. These models controlled for the defendant effects from the previous stage of the analysis to control for differences in caseload composition across the judges.

These models were estimated separately for the pre- and post-guideline periods so that the defendant and judge effects on sentencing could be compared for any changes over time. An equality-of-coefficients test was used to determine whether each estimate changed significantly in magnitude over time (Clogg, Petkova, and Haritou, 1995). Two-tailed significance tests were used for the analysis.

RESULTS

The information displayed in Table 1 permits comparisons between the pre- and post-guideline samples of defendants from the seven Ohio counties. These two samples are very similar in the legal and extralegal factors examined, with the exception of the proportions convicted on felony 3 and felony 4 charges. These particular differences stem from the felony 5 classification being unique to the post-guideline period and from the reclassification of certain felony 3 offenses to felony 4 offenses. None of the remaining factors differed significantly between the two samples \( (p > .05, \text{ based on two-sample } t\text{-tests}) \). The judge samples are also very similar in the characteristics examined, but this was expected given that over 70 percent of these judges were the same between the two study periods.

Relevant to the first research question, Table 2 displays the information on inter-judge variability in the distribution of non-suspended prison sentences across the seven counties. The first pair of models in Table 2 displays the unconditional models (no predictors) for the pre- and post-guideline periods.

There were significant differences in the use of imprisonment across judges during both time periods, as revealed by the significant between-judge variance estimates \( (\tau_{00}) \). The intraclass correlation coefficients \( (\rho) \) indicate that 38 percent of the total variance in imprisonment fell between judges before the sentencing reform versus 21 percent after the reform. Both figures are substantial, but the drop in sentencing variability between judges over time was also substantial and could reflect the impact of the guidelines on reducing inter-judge disparities in the use of incarceration.

Despite adding the explanatory variables to the models, the same theme emerged from the second set of models in Table 2. That is, the between-judge variance remained significant and substantive in the post-guideline sample. Also important is
that none of the extralegal effects changed significantly in magnitude over time (based on the Clogg, Petkova, and Haritou, 1998, \(z\)-test mentioned earlier), despite changes in the statistical significance of some of these coefficients. The absence of any substantive changes in extralegal effects over time suggests that Ohio’s scheme did not significantly reduce any preexisting extralegal disparities in imprisonment for the seven jurisdictions examined, even if it was successful for reducing some of the between-judge differences in their use of imprisonment. Significant extralegal effects during the post-guideline period included a defendant’s sex, age (18-24), and employment status. A defendant’s race, on the other hand, maintained no significant direct effect on imprisonment during either period across these particular counties.

### Table 2
Random Intercepts Models of Imprisonment

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Unconditional Models</th>
<th></th>
<th></th>
<th>Conditional Models</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Guidelines</td>
<td>Post-Guidelines</td>
<td>Pre-Guidelines</td>
<td>Post-Guidelines</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
<td>s.e.</td>
</tr>
<tr>
<td>Male</td>
<td>.48**</td>
<td>.19</td>
<td>.70***</td>
<td>.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>.17</td>
<td>.15</td>
<td>.26</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 18-24</td>
<td>-.06</td>
<td>.16</td>
<td>-.39*</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 40+</td>
<td>-.09</td>
<td>.19</td>
<td>-.08</td>
<td>.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children</td>
<td>-.52*</td>
<td>.21</td>
<td>-.20</td>
<td>.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>-.33</td>
<td>.18</td>
<td>-.55**</td>
<td>.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felony 1</td>
<td>1.86***</td>
<td>.41</td>
<td>3.10***</td>
<td>.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felony 2</td>
<td>1.31***</td>
<td>.28</td>
<td>1.82***</td>
<td>.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felony 3</td>
<td>.72***</td>
<td>.15</td>
<td>.74***</td>
<td>.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felony 4</td>
<td>——</td>
<td>——</td>
<td>.03</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td># counts</td>
<td>.24**</td>
<td>.08</td>
<td>.24***</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior prison terms</td>
<td>.89***</td>
<td>.14</td>
<td>.75***</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juvenile incar.</td>
<td>.55*</td>
<td>.28</td>
<td>.26</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance abuse</td>
<td>.59**</td>
<td>.19</td>
<td>.25</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pled guilty</td>
<td>-1.39*</td>
<td>.60</td>
<td>-.25</td>
<td>.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public defender</td>
<td>.06</td>
<td>.14</td>
<td>-.01</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \tau_{00} \] = Inter-judge variance (or variance in the outcome existing between judges at level 2).

\[ \rho \] = Intraclass correlation coefficient (proportion of variance in the outcome at level 2).

Notes: Parameter estimates (b) and standard errors of estimates (s.e.) reported.

* \( p < .05; ** p < .01; *** p < .001 \) (two-tailed tests).
Regarding legally relevant factors, the effects of the most serious charges convicted on increased significantly in magnitude for felony 1s ($z = 2.10; p < .05$), but not for felony 2s or felony 3s. These findings are consistent with the mandatory prison terms set in place under SB2 for defendants convicted a second time for felony 1s in conjunction with the presumption of imprisonment for first-time offenders convicted on these charges. A significant increase in effect size also might have been expected for felony 2 convictions given the presumption of imprisonment for these as well, but the post-guideline effect was not significantly greater than the pre-guideline effect. On the other hand, there is no presumption of imprisonment for felony 3s under the newer scheme, and so judges were likely following previous decisions in these cases (note the very similar estimates for felony 3s during both periods). Estimates for the number of charges convicted on and prior prison terms also did not change significantly in magnitude, which is consistent with the lack of change in how these factors were to be considered under the guidelines.

A few legal predictors that were significant during the pre-guideline period became nonsignificant during the post-guideline period even though the estimates did not drop significantly in magnitude over time. These predictors included histories of juvenile incarceration and substance abuse, and whether a defendant pled guilty. The nonsignificant effect of pleading guilty versus going to trial could reflect a greater consistency in outcomes under guidelines, regardless of the mode of disposition, which might be expected under more structured sentencing. The nonsignificant impact of a defendant’s history of substance abuse might also reflect reduced discretion under Ohio’s guidelines if these histories tend to correspond with drug offenses, most of which do not carry the presumption of imprisonment. Despite these pre- versus post-guideline differences in the significance of certain legal effects on imprisonment, similar to findings for the extralegal factors, the general theme is that Ohio’s guidelines did not significantly alter the magnitude of either legal or extralegal effects on the odds of imprisonment (with the exception of the impact of felony 1 convictions on imprisonment).

Returning to the issue of between-judge differences in their use of imprisonment, the random coefficient models displayed in Table 3 provide some insight into why there was greater variability across the judges in their use of imprisonment during the pre-guideline period versus the post-guideline period (pertinent to the second research question). This information reveals which of the legal and extralegal effects on imprisonment differed significantly in magnitude across judges. The corresponding coefficients from Tables 2 and 3 are very similar despite some modest raw differences, and both sets of coefficients generate the same conclusions as those described above regarding the relevance of extralegal versus legal effects on imprisonment and the changes (or lack thereof) in those effects over time.

Two of the extralegal effects varied significantly across judges during the pre-guideline period ($p < .05$), whereas none varied significantly during the post-guideline period. It appears that the effects of a defendant’s sex and family status were
significantly stronger among some judges relative to others during the first study period but not during the second period. Even though a defendant’s sex effect was more uniform post-guidelines, this effect remained a significant predictor of imprisonment across all judges. By contrast, both the significant general pre-guideline effect and the significantly varying pre-guideline effect of a defendant’s family status were rendered nonsignificant under guidelines.

Also relevant are the two significantly varying legal effects on imprisonment during the pre-guideline period that did not vary significantly during the post-guideline period. The significantly varying effects of felony 3 and # counts became nonsignificant over time (although the effects of juvenile incarceration varied significantly

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Pre-Guidelines</th>
<th>Post-Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>s.e.</td>
</tr>
<tr>
<td>Intercept</td>
<td>.09</td>
<td>.12</td>
</tr>
<tr>
<td>Extralegal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>.51*</td>
<td>.22</td>
</tr>
<tr>
<td>African-American</td>
<td>.20</td>
<td>.16</td>
</tr>
<tr>
<td>Age 18-24</td>
<td>-.02</td>
<td>.18</td>
</tr>
<tr>
<td>Age 40+</td>
<td>-.08</td>
<td>.22</td>
</tr>
<tr>
<td>Children</td>
<td>-.60*</td>
<td>.24</td>
</tr>
<tr>
<td>Employed</td>
<td>-.39</td>
<td>.21</td>
</tr>
<tr>
<td>Legally Relevant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Felony 1</td>
<td>2.11***</td>
<td>.49</td>
</tr>
<tr>
<td>Felony 2</td>
<td>1.24***</td>
<td>.36</td>
</tr>
<tr>
<td>Felony 3</td>
<td>.72***</td>
<td>.16</td>
</tr>
<tr>
<td>Felony 4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td># counts</td>
<td>.35*</td>
<td>.11</td>
</tr>
<tr>
<td>Prior prison terms</td>
<td>.95***</td>
<td>.16</td>
</tr>
<tr>
<td>Juvenile incar.</td>
<td>.63</td>
<td>.35</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>.61**</td>
<td>.23</td>
</tr>
<tr>
<td>Pled guilty</td>
<td>-1.59*</td>
<td>.70</td>
</tr>
<tr>
<td>Public defender</td>
<td>.11</td>
<td>.15</td>
</tr>
<tr>
<td>** ρ **</td>
<td>.44</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Parameter estimates (b), standard errors of estimates (s.e.), and between-judge variance (s²_b) reported.

* p < .05; ** p < .01; *** p < .001 (two-tailed tests).

Significance of the between-judge variance in parameter estimates (s²_b) determined by chi-square tests based on 103 degrees of freedom for the pre-guideline sample, and 98 degrees of freedom for the post-guideline sample (Raudenbush and Bryk, 2002:63-64).

ρ = Intraclass correlation coefficient.
during both periods). The guidelines might have been effective, therefore, in generating greater uniformity in judicial considerations of felony 3 convictions and total charges convicted on. Considered in conjunction with the findings for extralegal effects, it appears that the guidelines might have been successful for creating greater uniformity in imprisonment decisions across judges.

The existence of significant between-judge variance in imprisonment justifies examination of possible effects of judicial characteristics on the use of imprisonment (the third research question). The dependent variable for the model of judicial effects is the proportion of convicted felons sent to prison by each judge (more specifically, the logit transformation of this proportion). Whether between-judge differences in the proportions of incarcerated defendants corresponded with differences in judicial characteristics can be assessed with the information displayed in Table 4. These estimates were derived from models including both defendant and judge predictors. The defendant effects are not displayed in Table 4 because they were very similar to those displayed in Tables 2 and 3.

A few of the judicial effects were statistically significant during the pre-guideline period only, including a judge’s tenure on the bench, number of years served in a prosecutor’s office, and the proportion of a judge’s caseload consisting of drug offenses. Judges who had served on the bench longer, served more years in a prosecutor’s office, and had proportionately fewer drug cases were more likely to use imprisonment as a sentence. In contrast to these significant pre-guideline effects, there were no significant effects of a judge’s sex, race, law school, or felony caseload during either period examined.

The pre-guideline model accounted for 34 percent of the between-judge variance in imprisonment, whereas the post-guideline model included no significant predictors and explained none of the between-judge variance. The coefficients for prosecutorial experience and drug caseload also decreased significantly in magnitude over time (equality of coefficient z values = 2.03 and -2.15, respectively; $p < .05$). Although the coefficient for a judge’s tenure on the bench did not decrease over time, the guidelines might have had an impact on reducing disparities in imprisonment based on these other judicial factors.

**Discussion and Implications**

Results from the analysis of inter-judge disparities in imprisonment rates before versus after the implementation of presumptive sentencing guidelines in Ohio revealed that these disparities decreased under guidelines but remained statistically significant in the seven urban courts examined. A judge’s identity still accounted for roughly 20 percent of the variance in imprisonment under guidelines (versus roughly 40 percent of the pre-guideline variance), even after controlling for compositional differences in case-loads across judges. This figure is considerably larger than Johnson’s (2005) post-guideline estimate of 5 percent for Pennsylvania, although his figure represented the entire state as opposed to only the most-urban jurisdictions as were examined here. Even so,
the drop in sentencing variability between judges over time was substantial and could reflect the impact of the guidelines on reducing inter-judge disparities in the use of incarceration. Greater consistency in the use of imprisonment across judges post-guidelines is consistent with the finding of the United States Sentencing Commission (2005) regarding the federal guidelines. Moreover, that study revealed a one-third to one-half reduction in inter-judge variance after the guidelines were implemented, versus a 44 percent reduction in our analysis. The finding of significant inter-judge variance during both periods is also consistent with Scott’s (2010) before-after analysis of the federal guidelines (see also Anderson and Spohn, 2010; Tiede, 2009). In short, despite the more discretionary scheme implemented in Ohio relative to the federal scheme, the impact of each on inter-judge disparities in the use of imprisonment was strikingly similar.

As was consistent with Ulmer, Light, and Kramer’s (2011) analysis of whether extralegal effects changed when federal district courts shifted from mandatory to voluntary sentencing guidelines, we also found no substantive changes in defendants’ extralegal effects on the odds of imprisonment before versus after the implementation of Ohio’s guidelines. Scholars have argued that guideline restrictions can be maneuvered around by prosecutors in their charging decisions (Lagoy, Hussey, and Kramer, 2010).
1979) and in plea agreements with the defense (Miethe, 1987), so greater prosecutorial discretion post-guidelines may have compensated for reduced judicial discretion in Ohio.

Despite increased uniformity in incarceration rates across judges in these courts, therefore, the guidelines were not successful in reducing disparities in imprisonment based on some of the extralegal factors examined. There was no change in the significant effect of a defendant’s sex over time, and disparities based on a defendant’s age and employment status emerged only after the guidelines were in place. Although the significant pre-guideline effect of family status became nonsignificant under the guidelines, the magnitude of the estimate did not change significantly over time. These observations imply that while judges might have become more uniform in the proportions of convicted defendants they sent to prison under the guidelines, this did not mean that judges were “uniformly” ignoring extralegal factors.

Consistent with the finding of reduced disparity between judges in the odds of sending convicted felons to prison, we also found that several extralegal and legal effects that varied significantly across judges during the pre-guideline period did not vary significantly under the guidelines. These changes involved the effects of a defendant’s sex and family status in addition to felony 3 convictions and total counts convicted on. None of the defendant effects varied significantly across judges during the post-guideline study period, further suggesting that the guidelines might have been successful for creating greater uniformity in imprisonment decisions across judges. However, this uniformity did not necessarily reflect greater equity in the treatment of defendants based on extralegal factors, as evidenced by the significant general effects of a defendant’s sex and employment status during the post-guideline period. Greater uniformity in sentencing behaviors across judges does not necessarily imply greater equity for similarly situated defendants if this uniformity actually reflects a more general advantage for a particular group of defendants (such as females being less likely to go to prison, controlling for felony levels convicted on). On the other hand, regarding whether a defendant had dependent children, both the significant general pre-guideline effect and the significantly varying pre-guideline effect were rendered nonsignificant under guidelines. These changes in the effects of a defendant’s family status are more in line with what might be anticipated under more structured sentencing.

The absence of significant differences between judges in the effects of a defendant’s race on imprisonment decisions stands in contrast to Johnson’s (2006) finding of a significantly varying race effect across Pennsylvania judges. His analysis included all judges in the state, however, and our more restricted focus on urban jurisdictions with a smaller sample of judges could underlie this difference in findings. The race effect in these Ohio counties did not vary across judges even during the pre-guideline period, whereas Johnson’s (2006) analysis focused on post-guideline sentencing when variation in such effects should be diminished (in theory) due to more structured sentencing.

Also consistent with the reduction in inter-judge variance in imprisonment under Ohio’s guidelines, the sentencing reform coincided with significant reductions
in disparities in imprisonment based on a judge’s prosecutorial experience and the proportion of his or her caseload consisting of drug offenses. Although the significant pre-guideline effect of a judge’s tenure also became nonsignificant under guidelines, the magnitude of this effect did not decrease significantly over time. Nonetheless, the guidelines may have been successful for reducing some differences in imprisonment likelihoods based on judges’ characteristics.

The significant pre-guideline effect of a judge’s tenure on the use of imprisonment is consistent with Johnson’s (2006) finding of a modest effect for a similar measure (odds ratio = 1.01, \( p < .10 \) in his study; odds ratio = 1.03, \( p < .05 \) in this study). Anderson and Spohn (2010), however, found no significant effect of a U.S. District Court judge’s tenure on the length of imprisonment, although this outcome differed from the in/out decision examined here and by Johnson (2006). Johnson did not find prosecutorial experience and drug caseload to be significant predictors of incarceration among Pennsylvania judges, in contrast to our findings. The remaining judicial effects (sex, race, law school, felony caseload) were null during both study periods, and the null race effect stands in contrast to both Johnson’s (2006) and Muhlhausen’s (2004) findings of significant effects for a Pennsylvania judge’s minority status. The null sex effect also counters significant sex effects uncovered by Meyer and Jesilow (1996) in California, Spohn (1990b) in Detroit, Steffensmeier and Hebert (1999) in Pennsylvania, and Tiede, Carp, and Manner (2010) in the federal system. And whereas Anderson and Spohn (2010) found a significant inverse effect of caseload size on sentence length, no significant effect of caseload was found here.

Whereas the more general conclusion from previous studies of judicial effects is that judicial background factors are weak predictors of sentencing, our pre-guideline findings are more consistent with the few studies that have uncovered substantive judicial effects (Anderson and Spohn, 2010; Johnson, 2006; Meyer and Jesilow, 1996). Even so, unlike this smaller group of studies, the substantive judicial effects found here were limited to Ohio’s pre-guideline period only. This area of empirical research is still in its infancy, however, and more studies are needed before drawing firm conclusions about common themes in findings across a majority of studies.

Overall, our findings suggest that the implementation of Ohio’s presumptive sentencing guidelines contributed to greater uniformity across urban trial court judges in their use of imprisonment even though significant between-judge differences remained. Similarly, various defendant-level effects on imprisonment (both extralegal and legal) were more consistent across judges post-guidelines even though a couple of extralegal effects remained significant predictors of imprisonment across all judges. Finally, all of the significant pre-guideline judicial effects on the use of imprisonment became nonsignificant post-guidelines, although one of these effects (judge’s tenure) did not change significantly in magnitude over time. While each of these observations includes a caveat regarding the “success” of Ohio’s scheme for reducing judicial sentencing disparities, similar exceptions have also been found in extant studies of the federal guidelines (as previously noted).
The caveats above are still important, however, because they underscore important areas that might only be addressed effectively with a different focus. For example, greater uniformity in sentencing across judges may not promote more equitable treatment of defendants if all judges treat particular groups more severely, similar to the findings described earlier regarding higher odds of imprisonment for men relative to women. Such uniform severity might indicate a problem with the new sentencing laws themselves if particular group biases are unwittingly structured into these laws, such as the implementation of harsher drug laws for particular drugs (e.g., crack) more commonly used by minorities (Mauer, 1999).

Some of the reduction in pre-guideline judicial disparities in the use of imprisonment might be attributed to the stronger effect of felony 1 convictions on prison sentences resulting from the presumption of imprisonment for defendants convicted for a first time on this charge, in addition to the mandatory prison terms for defendants convicted a second time. Important to emphasize, however, is that the long-term effects of Ohio’s guidelines may have differed from the short-term effects, in light of research findings suggesting that short- and long-term effects of sentencing guidelines are not necessarily consistent (Koons-Witt, 2002; Miethe and Moore, 1985; Stolzenberg and D’Allesio, 1994; Wooldredge, 2009).

The evidence of reductions in certain forms of sentencing disparity under Ohio’s presumptive guidelines, implemented in 1996, brings up the question of whether the shift to voluntary guidelines in 2006 may have increased disparity to pre-guideline levels. The state prison population of Ohio rose by 15 percent (from roughly 43,000 to 50,000 inmates) during the two years surrounding the switch from presumptive to voluntary guidelines, and state officials attribute this growth, in part, to “changes in sentencing structure” (Ohio Department of Rehabilitation and Correction, 2007).

**Directions for Theory and Research**

Theories of disparate sentencing such as “perceptual shorthand” (Hawkins, 1981), “uncertainty avoidance” (Albonetti, 1991), and “focal concerns” (Steffensmeier, Ulmer, and Kramer, 1998) have focused on judges’ perceptions of defendants and the impact of those attributions on sentencing. Some of these perspectives differ, however, in their predictions of more or less severe sanctions for defendants with particular characteristics. For example, sentences might be more severe for male or unemployed offenders if judges perceive them as higher risks for future criminality (e.g., Hawkins, 1981; Albonetti, 1991; Steffensmeier, Ulmer, and Kramer, 1998), or sentences may be less severe for these same offenders if judges consider criminal activities to be more “normal” for males (relative to females) and the unemployed (e.g., Swigert and Farrell, 1977; Farrell and Swigert, 1978; Myers, 1980; Steen, Engen, and Gainey, 2005). The findings from our study introduce the notion that each perspective might be an overgeneralization and that the “reality” lies in different perspectives applied to different types of judges. Granted, the overall findings for defendant-level effects on sentencing generally favored the perspective of harsher sanctions for “lower status” offenders, but
the significant between-judge variance in several of the extralegal effects also suggested that some of these effects varied in direction across judges, such as more or less leniency shown to males, offenders with dependent children, and the unemployed. Attribution theories need elaborations on why and how particular defendant characteristics are considered by some judges but not others, not to mention how differences in these considerations can exist across judges operating under the same sentencing scheme.

Judicial differences in how defendant-level factors are considered might exist under Ohio’s guidelines because judges are allowed to consider different reasons for incarceration such as “the need for incapacitating the offender, deterring the offender and others from future crime, rehabilitating the offender, and making restitution to the victim of the offense, the public, or both” (Ohio Revised Code §2929.11). Under these circumstances, attribution theories predicting harsher sanctions for lower-status offenders might apply to judges who place greater priority on protecting the community through incapacitation (possibly due to stereotypes of lower SES offenders as “high risk”), while theories predicting less-severe sanctions for lower-status offenders might apply to judges who emphasize rehabilitation (if lower SES offenders are more likely to be perceived as products of their environment).

Related to the above, existing theories do not address why a particular trait might be considered by two judges but to varying degrees. From a “focal concerns” perspective, for example, is it possible that different combinations of focal concerns might lead judges to consider the same extralegal characteristics but with different priorities? A defendant’s sex might be relevant to two judges if both are concerned with protection of the community and the implications of their sentencing decisions for defendants’ families. The priority of each concern might differ between them, however, resulting in a stronger effect for the judge placing heavier emphasis on incarcerating the “higher-risk” males to protect the community, versus a weaker (but still substantive) effect for the judge who wishes to avoid separating women from their dependent children.

More studies of the topic are needed beyond this limited focus on seven jurisdictions in a single state. Analyses of how different judges weigh specific defendant characteristics in their decisions will be relevant for understanding the applicability of particular attribution theories to different judges (see also Anderson and Spohn, 2010; Johnson, 2006; Tiede, Carp, and Manner, 2010; Ulmer, 2005). Related studies of larger samples of judges from more jurisdictions will provide a more comprehensive understanding of how judges differ in their sentencing decisions. Such a pursuit is necessary for targeting the primary contributors to extralegal disparities in sentencing in a jurisdiction and, ultimately, for assessing and informing related theories. jsj
REFERENCES


APPENDIX

The Ohio Criminal Sentencing Commission (OCSC) drafted the presumptive sentencing guidelines implemented in 1996 under Senate Bill 2 (SB2). Before SB2, the last major rewrite of Ohio's criminal code had been in 1974 based on the Model Penal Code and followed indeterminate sentencing. Under the newer scheme, judges were “guided” by presumptions, factors, and other required findings in their decisions to send convicted defendants to prison and to determine the length of imprisonment. Although purposely designed to be more flexible than other guideline schemes (see Wooldredge, 2009, for the OCSC’s rationales for greater flexibility), commission members hoped it would reduce judicial discretion relative to the previous scheme.

Under these guidelines, prison sentences were presumed for first- and second-degree felonies, with mandatory imprisonment for offenders convicted a second time on these charges. Judges were guided against prison for certain fourth- and fifth-degree felonies, and there was no guidance for third-degree felonies. Minimum prison terms were presumed for offenders incarcerated the first time, with exceptions in particularly serious cases. Imprisoned offenders served the full terms as declared by the judges (following “truth in sentencing”). The only indeterminate sentences were life sentences. Sentences were to be dictated by the offenses and felony levels convicted on, in addition to criminal history. There were multiple goals of sentencing that judges were allowed to consider, including retribution, incapacitation, deterrence, and rehabilitation. Judges were permitted to depart from guideline requirements with adequate justifications (Ohio Criminal Sentencing Commission, 1996).

Funding for intermediate sanctions also increased under SB2 and generated a relatively broad range of community alternatives to incarceration, including day reporting and day fines. Convicted felons for whom prison was not mandatory or “presumed” were eligible for these other sanctions. Ohio’s sentencing reform, therefore, involved the implementation of guidelines that still permitted a fair amount of judicial discretion (although intended by the OCSC to reduce pre-guideline levels of discretion).
In 1978 the Supreme Court decided the landmark case of *Regents of the University of California v. Bakke*, 438 U.S. 265, holding unconstitutional an admissions policy setting aside sixteen seats for certain minorities at the medical school. However, according to the Court, race is a factor that can be considered in college admissions. Twenty-five years later, the Court rendered a decision in the undergraduate admissions program at the University of Michigan in *Gratz v. Bollinger*, 539 U.S. 244 (2003) and Michigan’s law school admissions program in *Grutter v. Bollinger*, 539 U.S. 306 (2003). Specifically, the Court ruled that the use of affirmative action in college admissions is constitutional if it treats race as one factor among many, its purpose is to achieve a “diverse” class, and it does not substitute for individualized review of applicants, but is unconstitutional if it automatically increases an applicant's chances over others simply because of his or her race. Justice Sandra Day O’Connor, writing for the Court, opined that the Court expected that racial preferences would be unnecessary in another twenty-five years. However, the Court only waited ten years to review another higher-education affirmative-action case in *Fisher v. University of Texas at Austin*, 570 U.S. ___ (2013). The Court held (7-1) that the Fifth Circuit failed to examine the undergraduate admissions policy under the strict-scrutiny test\(^1\) and remanded the case for further proceedings. Justice Kennedy wrote the opinion of the Court, while Justice Ginsburg was the sole dissenter. Justices Scalia and Thomas each wrote concurring opinions. Justice Kagan did not take part in deciding the case since she was the solicitor general when the case was heard in the federal district court. On remand, the Fifth Circuit should uphold the admissions policy as constitutional.

The University of Texas at Austin’s undergraduate admissions program was challenged by Abigail Fisher and Rachel Michalewicz who sought admissions to the University in Fall 2008. UT Austin admits the top 10 percent of all Texas high-school graduates who apply for admissions, pursuant to a law mandated by the Texas legislature with the goal of increasing diversity in its universities. UT Austin reserves 90 percent of its seats for Texas residents; however, a large majority of those seats are awarded on the basis of automatic admittance in accordance with the top 10 percent rule.

Texas residents who do not graduate in the top 10 percent compete for seats in a separate pool of applicants. The university considers admissions by computing

\(^{1}\) Strict scrutiny is the most stringent standard of judicial review. To pass strict scrutiny, the policy must be justified by a compelling government interest, and the policy must be narrowly tailored to achieve that interest.
a Personal Achievement Index (PAI) and an Academic Index (AI). The AI is calculated by an applicant’s standardized test scores and class rank. Admissions can be granted or denied on the basis of this score alone. If the applicant’s score is not high enough to grant admissions but not low enough to deny it, then the university considers an applicant’s PAI. The PAI is derived from a number of factors that include one score for two essays and a personal achievement score (PAS). The PAS is derived from a number of factors that include evidence that an applicant overcame obstacles associated with being in a lower socioeconomic status, demonstrated leadership abilities, and participated in community service. One factor includes race. Each factor in the PAS is not assigned a particular score; rather, admissions officers consider all PAS factors as a whole and assign a score. The university states that the consideration of race not only helps minority applicants, but also can benefit a white applicant, such as one who demonstrates leadership abilities in an environment where that student is a racial minority in his or her school or community. UT Austin notes that a high PAI score cannot compensate for a low AI score. UT Austin created this policy under the confines of the Supreme Court’s decisions in Gratz and Grutter.

University admissions policies that consider the race of applicants must withstand strict scrutiny: the policy must be narrowly tailored to achieve a compelling governmental interest. Unlike the plaintiffs in Gratz and Grutter, the parties do not contest that diversity, including racial diversity, is a compelling governmental interest. However, the question remains whether UT Austin’s goal is narrowly tailored. That is, the means by which UT Austin achieves diversity must be the least restrictive.

The Supreme Court in Grutter held that an admissions policy is narrowly tailored when a university considers race as a “plus” factor and not a predominate factor in assessing an applicant for admissions. In Grutter, the Court upheld the law school admissions policy of the University of Michigan that considered race as a factor among several others in admissions decisions. The law school assigned no specific weight to the race of an underrepresented minority applicant. The Court noted that other factors, such as traveling abroad, speaking other languages, and overcoming adversity, were also considered to significantly contribute to the diversity of the law school. However, the Court held unconstitutional the admissions policy in Gratz because the policy was not narrowly tailored. Michigan’s undergraduate admissions policy automatically assigned 20 points to underrepresented minorities out of 100 points necessary for admittance. The Court held the points scheme unconstitutional because the award of 20 points rendered race as a primary factor, considering one-fifth of the score of an underrepresented minority was awarded on the basis of race. Therefore, an applicant’s race was often a decisive or primary consideration for an extension of admissions.

Unlike the policy in Gratz, UT Austin’s policy considers race as a factor among others in its goal of creating a diverse student body. Admissions officers consider factors that indicate an applicant’s propensity to contribute to the diversity of the
university other than race, such as whether the applicant speaks more than one language or is a first-generation college student. Race does not outweigh such factors. Even more, the policy does not assign a specific weight to race as did Michigan’s policy in *Gratz*. UT Austin’s consideration of an applicants’ unique contribution to the diversity of the university steers away from wholesale notions of diversity that primarily relies on race, as did the scheme upheld in *Grutter*.

The Supreme Court remanded the case because the Fifth Circuit failed to assess whether UT Austin’s admissions policy considering race was narrowly tailored. In short, the Fifth Circuit failed to apply strict scrutiny. The Fifth Circuit stated that it is “ill-equipped” to “second-guess the merits” of the university’s decision to use race in considering admissions. 570 U.S. at __ citing 631 F.3d at 231. Thus, the Fifth Circuit presumed that the school had acted in good faith and gave Fisher the burden of rebutting that presumption. The Supreme Court held that it is the university’s burden to provide that its admissions program is narrowly tailored to obtain the benefits of diversity. Specifically, the Court stated that it is the purview of the judiciary to assess whether the use of race was necessary to achieve the benefits of a diverse student body.

The Fifth Circuit’s reluctance to assess UT Austin’s admissions policy is grounded in a long-standing dilemma of federal courts assessing policies and schemes and mandating solutions that it is, in fact, ill-equipped to address (Horowitz, 1977). Courts are generalists, not specialists. Particularly, courts are not experts creating and assessing policies seeking to maximize the educational benefits of students in colleges and universities. The Supreme Court stated that it is indeed ill-equipped to second-guess a university’s assessment that a diverse student body is a necessary component of its educational mission. However, when courts determine whether the use of race is the least restrictive means of achieving diversity, courts walk a fine line in adjudicating the constitutionality of an admissions plan and rendering a policy judgment assessing which plan is best. However, courts routinely make policy judgments wrapped in legal arguments (Segal and Spaeth, 2002). Even more, the analysis of the least-restrictive alternative under a narrowly tailored review could place university admissions in a precarious predicament: it affirms the use of ethnic and racial diversity as a compelling governmental interest, but asks universities to narrowly tailor this goal, if at all practically feasible, without affording race significant weight.

These notions are reflected in the Fifth Circuit’s questions to the parties on remand. Among others, the Fifth Circuit asks whether “workable alternatives” other than considering race are available but not currently employed by UT Austin. Furthermore, it asks whether UT Austin has reached a “critical mass” of minority students and whether its decision that the university does not have a “critical mass” is due any deference. In *Grutter*, the Supreme Court afforded some deference to Michigan’s assessment of attaining a “critical mass” in determining whether it had achieved its goal of creating a diverse student body. Particularly, the Court noted that Michigan’s attention to statistics to ascertain whether it had created a racially diverse class did not
transform the scheme into an unconstitutional quota system. In *Fisher*, UT Austin provided evidence that its admissions officers do not afford weight to periodic reports tracking admissions of minority groups when deciding to admit other applicants.

After the Fifth Circuit renders its decision, the case could reach the Supreme Court a second time. The separate opinions provided in *Fisher* by Justices Scalia, Thomas, and Ginsburg provide a preview of their positions in *Fisher* after remand. Justices Scalia’s concurrence reiterates his position in *Grutter* that the Constitution proscribes any use of race in university admissions. Justice Thomas’s concurrence equates segregationalists’ arguments with arguments in support of affirmative action but fails to draw an essential distinction between segregationalist and proponents of affirmative action: segregationalists believe that African-Americans are inferior. In Justice Ginsberg’s dissent, she aptly notes that the requirement of a university to create a diverse student body without considering race is an oxymoron. The Court is likely to render a 5-3 decision, with the conservative justices holding the UT Austin policy unconstitutional and the liberals dissenting (with Justice Kagan recusing herself).

**REFERENCES**
