

She'll Settle It?*

Christina L. Boyd
Assistant Professor
Department of Political Science
University at Buffalo, SUNY
cLboyd@buffalo.edu

August 28, 2011

*I am grateful to the Baldy Center for Law & Social Policy at the University at Buffalo, the Center for Empirical Research in the Law at Washington University in St. Louis, and the National Science Foundation for supporting my research. I also appreciate the helpful comments of Ryan Black, Lee Epstein, David Hoffman, Pauline Kim, Michael Lynch, Andrew Martin, Margo Schlanger, Joseph Smith, and James Spriggs and excellent research assistance from Sophie Fortin, Shadi Peterman, Gregg Re, Annie Rushman, and Erica Woodruff.

Abstract

This study seeks to gain new insight into the long-standing debate regarding the effect of judge sex. Relying on research that posits that female leaders and managers will be more likely than men to adopt a management style that favors participation, collaboration, and consensus-building, I argue that female district court judges, utilizing this style in their case management environments, should be more likely than their male colleagues to successfully foster intra-court case settlements. To test this, I compile data from over 23,000 cases terminated in four federal district courts across nine years. Rather than relying solely on standard regression modeling, this study turns first to matching methods. The results provide confirmation that the sex of a district court case's assigned judge matters, with female judges more likely than males to successfully foster settlement in their cases. In addition to having significance for litigants, these findings have broad implications for female decision makers across different institutions and organizations as well as the future of the judging profession and diversity appointments to the judiciary.

For political science and legal scholars, politicians, interest groups, and perhaps most importantly, for litigants themselves, the question over and over again has been whether the identity of a case's assigned judge or judges matters. By now, the evidence points clearly to judge ideology playing an important role in decision making in U.S. courts (e.g., Segal and Spaeth 2002; Epstein and Knight 1998; Hettinger, Lindquist and Martinek 2004; Kaheny, Haire and Benesh 2008), but for many other judicial characteristics, the story is far more complicated. One of these that has fascinated judicial scholars for years has been judge sex. And with good reason. In many of this country's business, profession, legal, and political arenas, the number of women involved in high-profile positions continues to rise, and the judiciary has been no exception. This has led many to ponder whether the presence of females leads to different judicial decisions, outcomes, and process, the results of which have been far less than homogeneous.

Surprisingly, within this large body of gender-centric judging work, little examination of judge sex has been done beyond merits voting that leads to published opinions. This is perhaps appropriate and understandable for appellate court judges, whose decision making role largely revolves around reviewing lower court decisions and interacting very little with a case's participants. However, it likely isn't suitable for trial courts, where regular and often years' long interaction between judges and the case's participants takes place while a case dynamically develops.

Indeed, trial judges are case managers across time rather than being decision makers at a single point, meaning that the trial court setting's characteristics call for a more process-oriented research design. This is particularly true with federal trial courts – called U.S. district courts – where civil caseloads have skyrocketed across time and these institutions, along with politicians, now encourage district judges to actively manage their cases and foster more intra-court settlements – rather than allow cases to advance toward trial – whenever possible (Resnik 1982; Galanter 2004; Kritzer 1986).

Ultimately, however, the discretion lies with individual district judges to determine how hands-on they wish to be in supervising their cases and informally mediating disputes assigned to them. Because of this discretion, judges are likely to develop or adopt case management styles, and it is here that their sex is perhaps most likely to play a role in determining whether they take the initiative to promote a settlement environment. A large body of research on females serving in leadership and management positions indicates that they are more likely to adopt a democratic, participative, and collaborative managerial style than similarly situated males, who tend to be more authoritative managers (e.g., Eagly and Johnson 1990; Eagly, Makhijani and Klonsky 1992; Burke and Collins 2001). If these styles of management map onto district judges serving in a management position over the parties in the trial court case, female judges should be more likely to provide an “affirmative push toward convergence” (Galanter 1985) than their male district court colleagues. And, as a result, cases assigned to female district court judges should be more likely to settle than those assigned to males.

To examine whether this expectation holds, I compile data from over 23,000 cases terminated in four federal district courts across nine years. Rather than relying solely on standard regression modeling, this study turns first to matching methods before estimating the effect of the sex of a district court case’s assigned judge on whether a case will settle. After matching, the results provide confirmation that the sex of a case’s assigned judge matters. Settlement is very common across the board, something that should come as no surprise to court watchers, but female judges are as much as 7 percentage points more likely to successfully foster settlement for their cases than male district judges.

Beyond this significant result regarding judge sex and the likelihood of case settlement, this article makes a number of important contributions to the gender-judging literature. It provides what is likely the first application of gender-specific management and leadership theory to the judiciary, something that has previously proven very valuable in explaining decision making differences between men and women in other political bodies such as legis-

latures and bureaucracies. This management focus helps highlight the unique institutional setting of trial courts and job description of trial judges, both of which combine to offer district judges a vast amount of discretion in supervising cases and interacting with parties. The result of this provides insight not just into differences in style between male and female judges but also more general aspects of trial court decision making, the dynamics of litigation, the multi-faceted role of the trial judge in this process, and how very different the responsibilities and decision making environment of trial judges are from their appellate court brethren.

In addition, this study supplies analytical and methodological clarity and helps reduce potential bias for two important aspects of the judge sex debate. Given the historical patterns regarding the selection of diverse judges in the U.S., there is a great potential for data imbalances in judicial gender studies along a number of key dimensions like judge ideology and age. Left untreated in the data, these could bias the results. As such, this project joins a small but growing literature in political science that implements propensity score matching in the data prior to estimating traditional regression models (Boyd, Epstein and Martin 2010; Imai 2005; Epstein et al. 2005; Greiner 2008). In addition, this project uses new district court data and focuses the quantitative analysis here not just on understudied trial courts, but on all types of terminations in these trial courts, including those that do not yield published opinions (Siegelman and Donohue III 1990). This project thus makes strides forward in reducing concerns about selection bias present in many other gender-judging studies, particularly those focusing on appellate courts and merits-level judge voting. This approach confirms that looking beyond votes yielded in authored opinions and instead turning to the earliest moments of judicial involvement in litigation may be the most fruitful, as well as the most meaningful, moment to uncover where judge sex – and other judicial and litigation characteristics – matters. Finally, this study’s results – that the sex of judges matters and does so beyond simply whether the judge in a case will vote more liberally than her colleagues in “women’s issue” cases – provide further evidence that judicial diversity

should remain at the forefront in our conversations and debates about judicial selection and even presidential elections.

The Female Judge Voice?

The scholarly quest to uncover whether modern women behave and make decisions differently from men — from communication to problem solving to leadership and beyond — has been ongoing for years. Those expecting these differences have posited a variety of rationales, from the biological “different voice” of men and women (Gilligan 1982), to female decision makers’ service as high-profile representatives of other women working to advance their group’s interests (Pitkin 1967), to female development of expertise in certain areas (like traditional “women’s issues”) because of the unique environment, challenges, and experiences that they face (Gryski, Main and Dixon 1986; Sullivan 2002). Laying the groundwork for much debate and scholarship, this work has led researchers to examine female behavior and decision making, and compare it to men’s in a variety of contexts including those fundamental to politics. One of these regards the potential differences between males and females serving in the judiciary, an area that has substantive implications for the outcome of litigation. In the sizable literature on this topic, scholars have connected the above general gender theory to the judicial profession, arguing, for example, that the social and professional backgrounds of women judges may lead them to make different decisions and otherwise behave differently on the court than their male colleagues (Songer, Davis and Haire 1994; Slotnick 1984), with “differently” meaning, for example, more empathetically (Brudney, Schiavoni and Merrit 1999) or more sensitively (Martin, Reynolds and Keith 2002).

Within this field, the list of quantitative gender-judging studies is long, and the findings have been decidedly mixed. The majority of these studies focus on appellate court judicial decision making, where support for women judging differently than men has been

commonly found with traditional women’s issue-type cases like sexual discrimination and sexual harassment (e.g. Boyd, Epstein and Martin 2010; Peresie 2005; Crowe 1999; Farhang and Wawro 2004), but rarely elsewhere (e.g. Segal, Spaeth and Benesh 2005; Cox and Miles 2007). Notably, in those studies focusing on trial courts, outside of criminal cases, where some studies have found that female judges are punitive than males (Fox and Sichel 2000; Steffensmeier and Herbert 1999), very little evidence of differences in decision making between male and female judges has been found (Ashenfelter, Eisenberg and Schwab 1995; Sisk, Heise and Morriss 1998; Kulik, Perry and Pepper 2003; Walker and Barrow 1985; Schanzenbach and Tiller 2006; Segal 2000; Kritzer and Uhlman 1977). Nearly all of these quantitative studies, including those that center solely on trial courts, focus on recorded (published) outcomes – or judge “votes” – to search for sex differences (but see Ashenfelter, Eisenberg and Schwab 1995).

Given the notable selection of cases for litigation as well as the even more hefty censoring that takes place for cases to advance to trial, let alone appellate courts (Priest and Klein 1984; Atkins 1993), it is hardly a surprise that differences in the behavior of judges, based on their sex, have rarely been found by scholars. Cases advancing to trial or, to an even greater extent, going to a higher court on appeal, by no means represent all litigation. Perhaps more importantly, the judicial decisions that these activities yield constitute but a very small percentage of the business that judges engage in. Nowhere is this more true than for judges serving on federal trial courts, where the dynamic nature of litigation requires judicial participation and supervision throughout the life of a case. Turning away from a traditional utilization of judge votes and instead relying on a closer inspection of the trial judge’s role presents an ideal opportunity to examine how judge sex may affect litigants, litigation, and the law.

A Different Context: Why Judge Sex and Settlement?

Why might female federal trial court judges be more likely to successfully encourage case settlement than their male counterparts? The answer to this lies at the intersection of three areas: the management and leadership styles most likely to be held by females in positions of authority, the function of federal district courts, and the discretion afforded to judges that manage cases in these courts. To understand this, we can look first to insight from research analyzing the actions of females serving in leadership and management roles. Those working in this research area have long debated whether and to what degree women should be expected to behave differently than men while serving as leaders and managers in positions of authority. The generalized expectation of difference here is that women and men, armed with gender-specific qualities, upbringing, and/or experiences, bring a unique style to their management roles, and this style “offers a way to understand why managers, faced with seemingly identical situations, use such different decision processes” (Nutt 1990, 174). Because “women do not use or perceive their positions of power like their male counterparts” (Kathlene 1994, 561), a number of differences should emerge from management and leadership settings regarding communication, problem solving, and, ultimately, outcomes.

What differences in management and leadership style should we expect between men and women? When given discretion as to how to lead and manage, female decision makers should be more likely to adopt management styles that can be generally described as democratic whereas similarly situated males ought to tend toward autocratic styles (Eagly and Johnson 1990, 236). These styles of management will differ from one another on a number of dimensions. A female style of management is characterized by a decision making environment that is more likely to encourage participation among subordinates, democratic communication, collaboration, consensus-building, and the drawing of multiple-voices into deliberative processes (Rosener 1990; Eagly and Johnson 1990; Rosenthal 1998). In other words, in their position of authority, women managers will look for ways to build bridges,

empower participants, and reach conclusions that will meet the needs of those individuals that are involved (Burke and Collins 2001; Rosenthal 1998). Conversely, a traditionally male style of management is one that disfavors this type of participative environment and instead sees the male decision maker seizing opportunities to exercise his authority over outcomes (Eagly and Johnson 1990).

Quantitative studies have confirmed the presence of these gender-specific management styles in a variety of workplace settings, including in U.S. politics. Scholars of legislative behavior have found that women in leadership positions in legislatures, have more collaborative and team-oriented styles than similarly situated male leaders, evidence of gender differences that is often not visible at the legislative voting stage (Rosenthal 1998; Kathlene 1994). In state legislatures, women committee chairs are more likely to moderate and facilitate discussion rather than use their position to control outcomes (Kathlene 1994) while, on the other hand, males use “their position of power to control hearings in ways that we commonly associate with the notion of positional power and leadership” (Kathlene 1994, 572).

For many judges, particularly those serving on appellate courts, opportunities to serve in management and leadership positions are few and far between, resting perhaps only with chief judges and justices. Even there, judges generally serve as leaders over peers rather than subordinates, not to mention that the number of women serving in these leadership positions over time has been limited. However, due to unique institutional design and job descriptions, trial courts and the judges that sit on them serve as an important exception to this, and as such, provide an excellent opportunity to examine gendered management theory within the judiciary.

Federal trial courts are the courts of first instance, where litigation commences, develops a factual record through discovery, and provides the initial application of law to those facts. Within these courts, there is regular and often years’ long interaction between trial judges and case participants – i.e., case management – while the case dynamically develops.

Like U.S. Supreme Court justices and courts of appeals judges, federal district court judges are nominated by the president and confirmed by the Senate and, under Article III of the Constitution, receive life tenure. However, unlike their brethren in the federal appellate courts, district court judges decide cases almost exclusively alone rather than in a collegial setting and have a unique job description.

As a result of this design, federal district court judges engage in many diverse activities — from presiding over trials, to hearing and ruling on pre- and post-trial motions, to conferring and advising the parties in the stages before a trial is scheduled to take place. Through this role, the trial judge has an enormous amount of discretion to affect how a case develops and is resolved. In recent years, federal trial judges have been actively encouraged by their district courts and by politicians (through legislation like the Civil Justice Reform Act of 1990) to adapt their role to changing institutional pressures such as a rapidly growing caseload (growing from approximately 50,000 civil cases filed in 1962 to 283,000 in 2010 (Administrative Office of the U.S. Courts 2010; Galanter 2004)) by facilitating more intra-court settlements.

Viewed in a more global way, this push has been toward trial judges engaging in a new style of case management and judging called “managerial judging.” Managerial judging contrasts sharply with the traditional or classical style of judging, which advocates judges taking a very neutral stance toward their cases, so much so that they avoid direct involvement in the disputes and adopt a “disinterested pose” (Resnik 1982). On the other hand, with the new style of managerial judging, judges are actively involved in cases, where they are now “not only adjudicating the merits of issues presented to them by litigants, but also are meeting with parties in chambers to encourage settlement of disputes and to supervise case preparation. Before and after the trial, judges are playing a critical role in shaping litigation and influencing results” (Resnik 1982, 376-377).

Despite this move toward more managerial judging being institutionally-encouraged,

the decision of how involved to be as managers ultimately lies within the discretion of the judges, meaning that some trial judges can (and do) continue to practice traditional case management rather than choosing to be “settlement judges.” And when this happens, there is little-to-no institutionalized enforcement power in place should this be their choice. For trial judges that do choose to engage in the newer, managerial style of case management, judicial case supervision is much more hands-on and geared toward encouraging the parties to find common ground. As a result, settlement negotiations can be actively encouraged or even brokered by the case’s assigned judge (Galanter 1985). Not surprisingly, attorneys practicing litigation before district judges recognize these different case management styles among their judges and how these styles affect the resulting settlement environment. For example, a judge in an Illinois district court was lauded by an attorney for her people skills and patience that were “ideal to achieve settlement,” while an attorney in New York noted that his case’s judge was “no good at settlement. He’s a bully. That’s not the attitude that gets cases settled” (*Almanac of the Federal Judiciary* 2005).

Male and female district judges are generally free to adopt case management styles of their choosing – either the more traditional, hands-off approach or the newer managerial judging approach. The adoption of a judicial management style is likely to affect the resulting litigation environment, with, of course, a major consequence of this being the underlying likelihood of case settlement. As such, this wide grant of discretion and independence provides for the application of generalized gendered leadership and management theory to trial court judges and their case management duties such that we can begin to see more concrete expectations emerge based on the sex of a case’s assigned judge.

Female district court judges, with management styles that are inclined toward collaboration and participation and the discretion that their job offers to actively implement these things, should be more likely to fit the mode of a managerial judge. As such, female judges should be inclined to foster negotiation and consensus between the parties in their cases. The resulting environment created by this case management style should increase the

overall likelihood that cases will settle when assigned to a female judge.¹

Conversely, the above theory would predict that male district court judges will be more likely to behave autocratically and provide less promotion of a collaborative environment while managing cases. In other words, male judges should be more likely to adopt a more traditional model of case management rather than managerial judging. The result of this should be less encouragement of collaboration and informal negotiation under the watch of the judge while simultaneously leaving the judge in a stronger position to control the outcome of the case, whether that be through a trial or ruling on a dispositive motion like summary judgment or involuntary dismissal. Like Resnik's (1982) traditional, detached judge or the "bully" judge described by a lawyer above, this style of judicial case management should be less likely to lead to settlement, since it places the burden much more squarely on the parties to broker a pre-trial settlement if they choose to do so.

Data and Methods

To examine whether judge sex affects the likelihood of settlement in federal trial court cases, this project studies cases terminated in four federal district courts (District of Massachusetts, Eastern District of Missouri, Western District of Washington, and the Northern District of Texas) from 1996 to 2004. The data include all cases in two large issue civil law areas: civil rights disputes (including jobs, accommodations, and welfare) and personal injury torts (such as those involving motor vehicles, airplanes, and boats). The examined data include observations from these four districts from the Administrative Office of the United States Court's ("AO") terminations database, which compiles information on every case that is filed in the federal district courts. For each case, the assigned judge's

¹Unlike merits-level gender judging models, this management-style theory is not issue-area specific. In other words, while female appellate court judges might be predicted to vote more liberally than males only for "women's issue" cases, female district judges should be expected to adopt a pro-managerial judging stance regardless of the type of case in front of them.

name was then collected from the individual district court’s case filing materials available for purchase through PACER (“Public Access to Court Electronic Records”).²

The dependent variable for the study – settlement or not – was coded based on a recoding of the AO’s disposition variable. The AO directs court clerks from individual districts to classify the method of termination of each case in their district in one of over twenty possible ways. I have reclassified these termination methods to reflect whether settlement took place in the case (Hadfield 2004).³ Dispositions that involved non-terminal outcomes or were irrelevant to the project were dropped from the analysis,⁴ with the resulting full (pre-matched) database including 23,072 cases.

The data on this study’s key covariate – judge sex – as well as other judicial characteristic control variables were collected from the Federal Judicial Center’s Biographical Database (Federal Judicial Center 2011). Judicial ideology was tabulated using the Judicial Common Space methodology (Epstein et al. 2007; Giles, Hettinger and Peppers 2001; Poole 1998; Boyd 2010).

To ensure that the assignment of cases to judges within the four districts of this study approximates random assignment, I test for random assignment in 1999 filings for two different case characteristics in my dataset — namely, personal injury cases and civil rights cases — for each of the four districts. Using the χ^2 (chi-squared) test statistic, I compare the number of cases heard by individual judges in each category to the expected number of cases that those judges would hear if assignment was indeed random. The results,

²Cases assigned to U.S. magistrates by the consent of the parties under 28 U.S.C. §636(c) rather than a district judge are excluded from this study. Magistrate judges are not only not selected and protected through Article III like district judges, thus making the coding of a number of key variables about them in this study impossible, but they also face a number of different constraints in their decision making environment from district judges.

³Cases coded in the AO data as “dismissed: voluntarily” were coded as settled based on Hadfield’s (2004) auditing finding that the plurality of dispositions under this were settled. Cases coded as “dismissed: other” were dropped from the analysis based on Hadfield’s findings that the coding of dispositions in this category was excessively inconsistent.

⁴The following dispositions were dropped from the analysis: transferred to another district, remanded to state court, dismissed for want of prosecution, judgment on default, multi-district litigation transfer, remanded to U.S. agency, statistical closing, and stayed pending bankruptcy.

reported in Table 1, indicate that for none of the categories in any of the districts can we reject the null hypothesis that assignment of cases to judges is independent of the case being a personal injury case or a civil rights case. Thus, short of being in the clerk’s office when case assignment is taking place in these districts, this test provides some evidence, at least for the case types in this study, that random assignment has taken place in the four districts of interest in this study. The presence of the equivalent to random assignment of cases within district courts allows this study to assume that differences within cases that could be potentially important to their likelihood of settlement – like “hard” cases or cases with repeat player litigants and/or attorneys – are randomly distributed among judges (Ashenfelter, Eisenberg and Schwab 1995).

[Table 1 about here]

Matching Methods

When working with any observational data, such as is the case here, imbalances are common, frequently caused by the treatment being systematically related to other covariates (Rubin 2006; Epstein et al. 2005; Holland 1986; Ho et al. 2007). Historical realities about the diversity of judges in the United States provide ample evidence as to why this might be particularly problematic when the treatment is the sex of the judge (Boyd, Epstein and Martin 2010; Greiner 2008) and require careful consideration in the present study. Through the late 1970s, very few female judges had been appointed to the federal courts, from the district courts through the U.S. Supreme Court. President Carter made diversifying the courts a priority, such that with his presidency, the first sizable influx of women came onto the federal bench. Since Carter, the federal courts have continued to add more female judges, but the rate of increase has been anything but constant over time, with Democratic presidents (Carter, Clinton, and now Obama) having stronger records for appointing women

to the bench than the Republican presidents serving immediately before them (Scherer 2005).

This abbreviated historical account of judicial diversity points to at least two covariates of interest in judicial behavior studies on judge sex that are likely to occur unevenly among male and female judges: ideology and age. Male judges have been relatively equally likely to be nominated by Democratic and Republican presidents, and since an appointing president's party closely aligns with judge ideology, these male judges are likely to span the ideological spectrum. For female judges, however, this isn't the case. Far more Democratically-appointed females serve in the federal courts than those appointed by Republicans, meaning that the ideological distribution of women serving in the courts is likely to be decidedly right-skewed. Figure 1's kernel density plot depicts just this expected pattern for the ideologies of male and female judges serving in this study's district court dataset. While male judges (in black) are relatively equally present in the ideologies ranging from most liberal to most conservative (left to right), female judges (in grey) are heavily distributed on the liberal side of the figure. And, in fact, there is a complete lack of common support for the most conservative ideologies of the data.

[Figure 1 about here]

Similarly, with women not entering the federal judiciary in significant numbers until Carter's presidency, we should expect less variance in the ages of women serving on the courts. Older men, born as early as the 1910s and first appointed in the 1950s, 1960s, and early 1970s continued to serve while the oldest women were first appointed in the late 1970s. Figure 2's kernel density plot of judicial birth year confirms this for a group of district court judges serving in 1996-2004. Among these federal trial court judges, men of all ages, including a large number in their 60s and 70s, were serving in relatively equal numbers whereas for women, the vast majority of those serving were in their late 40s.

[Figure 2 about here]

Without question, concerns about data imbalances exist in this gendered judging arena, the very presence of which may lead to misleading results through quantitative analysis (Boyd, Epstein and Martin 2010). To overcome this, matching methods allow for the comparison of cases assigned to judges of similar ages, ideologies, and other key covariates, so that the final results provide insight into whether a judge’s sex is what makes a difference in the outcome of case proceedings. After creating a summary of pretreatment variables (called a propensity score), I use nearest-neighbor matching to do just this (Ho et al. 2006, 2007).⁵ Figure 3 depicts the distribution of propensity scores in the data analyzed here (see above), before and after matching. As expected, the matching exercise appears to address the data imbalances that are blatantly apparent in the left panel of the figure. This is confirmed in Table 2, which provides the summary statistics on the data, pre- and post-matching. These summary statistics indicate that the matching leads to balance improvements on all but one pre-treatment variable. While the results of this exercise lead to some discarded data points, the more compact emerging dataset greatly reduces the inherent imbalance in the data and allows for a more careful consideration of the effect that judge sex – rather than sex that is inescapably entangled with variables like ideology or age – has on case management and the likelihood of settlement.⁶

[Figure 3 about here]

[Table 2 about here]

⁵The components for the “best” propensity score model included the following: appointing president’s party, birth year, birth year squared, race, prosecutor, state judge, magistrate, and district.

⁶It is worth noting that the key question posed here is *not* one of a causal nature (Rubin 2006). This is the case for two reasons:

First, because the treatment is the sex of the judge, most would say that it fails to meet the “no causation without manipulation” standard. Second, the other covariates relevant to this question—whether centering on the judge’s attributes (e.g., ideology and age) or the case’s details (e.g., direction of the lower court decision)—occur *after* the sex of the judge is determined. With only post-treatment covariates, we cannot estimate a causal effect (Boyd, Epstein and Martin 2010).

Despite not being causal, the inquiry into differences between the sexes in this context remains important and interesting.

Analysis

With these newly collected federal district court data in hand and matching methods completed, we can now estimate the effect of judge sex on the likelihood that a case settles. Table 3 reports the results of four estimations of this subject of interest. Estimations are performed on both the full and matched data, and for each, both a bivariate and multivariate model is fit.⁷

[Table 3 about here]

Within each model, the presence of a female judge has the expected effect – positive and significant – on whether a case will settle in a federal district court. The average treatment effect (ATE) of this from these four models is plotted in Figure 4. It indicates that the presence of a female judge increases the likelihood of case settlement by as much as 0.07. Notably, while this effect is positive and statistically significant even in the unmatched data, the size of the effect of a female judge is at its strongest when the estimation is performed using only the pre-processed matched data. In other words, only with matching were we able to see the size of this effect at its fullest.

The evidence, as expected, supports the theory that, when all else is equal, a female district judge is more likely to provide a litigation environment that, through urging consensus and participation among the opposing parties, leads to intra-court case settlement of litigation. While a sizable number of the male-assigned cases are also settling, they are doing so at a significantly lower rate than female-assigned cases. This certainly seems to support the idea that, while many male judges will not necessarily oppose the parties taking the initiative amongst themselves to settle their dispute while it is ongoing, they will also be

⁷As previous research has indicated, there is a lack of consensus among proponents of propensity score matching as to the methodology to utilize post matching. Some encourage full, multivariate models while others suggest that bivariate models will be sufficient given the newly balanced, pre-processed data (Smith 1997; Ho et al. 2007; Boyd, Epstein and Martin 2010). With an aim toward consistency, this project implements both methods.

more willing to avoid taking an active, managerial role in the case that is designed to foster settlement.

[Figure 4 about here]

The substantive importance of the presence of a female district court judge is particularly clear after looking at Figure 5, which plots the predicted probability that a case will settle based on the sex of case’s assigned judge (top and bottom panels) and the changing ideology score of that judge (most liberal to most conservative in each panel). From this figure, we can see that a case assigned to a liberal male judge has just a 65% chance of settling while one assigned to liberal female judge does so nearly 75% of the time. The effect is even more drastic when one compares the same liberal male judge to a conservative female judge – whose cases have around an 80% chance of settling. Therefore, when it comes to district court cases and their likelihood of settling, just as has been seen in studies of other courts and in other settings, there is no escaping that the identity of a case’s judge – including his or her sex – matters quite a lot.

[Figure 5 about here]

Discussion

This article has offered a new understanding of the important effect that female judges can have in the judiciary. While a great deal many pages have been written and models have been estimated on the importance of judge sex, the previous focus has been rather limited. For the first time, I have shown here that judge sex can affect the litigation environment, and particularly the likelihood of settlement, regardless of whether the cases deal with, for example, traditional “women’s issues.” As the literature on females in leadership and management positions predicts, female district judges, like other women serving in managerial

and leadership positions, should be more likely than their male colleagues to foster collaboration, bridge-building, and negotiation in their case management environments. With the inherent discretion given to district court judges regarding the style that they use in the management of their cases, this should lead to more settlements for female-assigned cases than male-assigned cases. And, indeed, after matching methods and a careful analysis of federal district court case dispositions over nine years, the results indicate that, in comparison to actions taken by her male colleagues in the cases that are assigned to them, for the female district court judge we can say that, consistently, across districts and after controlling for a slew of other factors, “she’ll settle it.”

The results here also point to some broader conclusions. For one, they indicate that female judges bring a different perspective to the role of judging – one that likely fosters increased party participation and voice in their case outcomes and actively encourages compromise, even within the walls of a courthouse. Viewed in combination with other recent findings regarding female judges including, for example, that women can affect their male colleagues’ decisions while serving on panels with them (Farhang and Wawro 2004; Boyd, Epstein and Martin 2010), these findings seem to hint at a trend that an increase in the numbers of women serving in the judiciary may not only affect outcomes in certain cases but actually may be responsible for helping to alter the very profession of judging. The potential consequences of this, including for parties litigating cases and politicians selecting judges, are immense.

Looking at these findings regarding female judges in light of the larger gender-management literature also provides new insight into the broader study of women in leadership, management, and decision making roles. Within many organizations, research has posited that women leaders seeking effectiveness may be less likely to exercise their preferred collaborative, feminine leadership styles in favor of more accepted and traditional authoritative ones (Eagly and Johnson 1990, 248). The results here, however, show no apparent signs of this in federal district courts, something that could be attributed in large part to not only

the independence that federal trial judges are afforded under Article III of the Constitution but also the modern federal courts' active institutionalized *encouragement* of settlement, compromise, and collaboration. Participating in settlement is now, as Galanter (1986, 257) puts it, "a respectable, even esteemed, feature of judicial work." Under these circumstances, it is no surprise that female judges show no empirical signs of being constrained in their management style over cases. With evidence like this, albeit indirect in nature, that this institutional support works for female judges, other organizations may see value in adopting this model in their unique settings.

We can also begin to see the addition of a new layer of complexity to the ongoing debate over the continued move toward trial court-encouraged settlements. The above discussion reiterates something that we may have already known about the case management stage of judicial decision making: settlement dominates. As we can see from all of the models' estimates, judges, no matter their sex, are far more likely to dispose of cases by settlement than by any other means. This, combined with the growing overall numbers of women serving in federal courts and the evidence that female judges are more successful at fostering settlements than men, is likely to be encouraging to some. This includes politicians and court reformers who have, over the last thirty or more years, tried to bolster the attractiveness of managerial judging and court-induced settlement options. In a different way, it will likely provide support to those scholars who favor consensual legal outcomes and the procedural justice and fairness that they promote (Lind et al. 1990). On the other side, however, those who have long argued against the normative benefits of removing judges from their traditional, neutral position during litigation, will undoubtedly be concerned with the trend emerging from these results and what it hints at for the future of trial courts.

More generally, this work provides some of the first systematic evidence in the political science literature on trial judge actions beyond final, published votes on the merits of cases. This further points to what political scientists have long ignored – that trial court judges, via cases management, have very different job descriptions than appellate court judges –

and that these differences provide an excellent new arena for examining judicial behavior. While terminations-level data that include settlements are now systematically available electronically – something that wasn't the case less than 10 years ago – the gathering of these data across a sizable sample of cases and judge-specific variables remains expensive and technologically onerous. Overcoming this hurdle, like attempted in this study, helps to reduce concerns about selection bias while simultaneously opening up the examination of many new questions about, among other things, judges, litigation, and judge sex by political scientists.

References

- Administrative Office of the U.S. Courts. 2010. “Judicial Business of the United States Courts.” available at <http://www.uscourts.gov/Statistics/JudicialBusiness/JudicialBusiness2010.aspx>.
- Almanac of the Federal Judiciary*. 2005. New York: Aspen Publishers.
- Ashenfelter, Orley, Theodore Eisenberg and Stewart J. Schwab. 1995. “Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes.” *Journal of Legal Studies* 24(2):257–281.
- Atkins, Burton M. 1993. “Alternative Models of Appeal Mobilization in Judicial Hierarchies.” *American Journal of Political Science* 37:780–798.
- Boyd, Christina L. 2010. “Federal District Court Judge Ideology Data.” available at: <http://cLboyd.net/ideology.html>.
- Boyd, Christina L., Lee Epstein and Andrew D. Martin. 2010. “Untangling the Causal Effects of Sex on Judging.” *American Journal of Political Science* 54(2):389–411.
- Brudney, James J., Sara Schiavoni and Deborah J. Merrit. 1999. “Judicial Hostility Toward Labor Unions? Applying the Social Background Model to a Celebrated Concern.” *Ohio State Law Journal* 60:1675–1766.
- Burke, Sarah and Karen M. Collins. 2001. “Gender Differences in Leadership Styles and Management Skills.” *Women in Management Review* 16(5):244–256.
- Cox, Adam B. and Thomas J. Miles. 2007. “Judging the Voting Rights Act.” *Columbia Law Review*, forthcoming.
- Crowe, Nancy. 1999. “The Effects of Judges’ Sex and Race on Judicial Decision Making on the U.S. Courts of Appeals, 1981-1996.” University of Chicago Dissertation.

- Eagly, Alice H. and Blair T. Johnson. 1990. "Gender and Leadership Style: A Meta-Analysis." *Psychological Bulletin* 108(2):233–256.
- Eagly, Alice H., Mona G. Makhijani and Bruce G. Klonsky. 1992. "Gender and the Evaluation of Leaders: A Meta-Analysis." *Psychological Bulletin* 111(1):3–22.
- Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, & Organization* 23(2):303–325.
- Epstein, Lee, Daniel E. Ho, Gary King and Jeffrey A. Segal. 2005. "The Supreme Court During Crisis." *NYU Law Review* 80(1):1–116.
- Epstein, Lee and Jack Knight. 1998. *The Choices Justices Make*. Washington, D.C.: CQ Press.
- Farhang, Sean and Gregory Wawro. 2004. "Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making." *Journal of Law, Economics, & Organization* 20:299–330.
- Federal Judicial Center. 2011. "History of the Federal Judiciary: Federal Judges Biographical Database." available at: <http://www.fjc.gov/public/home.nsf/hisj>.
- Fox, Richard and Robert Van Sickel. 2000. "Gender Dynamics and Judicial Behavior in Criminal Trial Courts: An Exploratory Study." *Justice System Journal* 21:261–280.
- Galanter, Marc. 1985. "'... A Settlement Judge, not a Trial Judge.'" Judicial Mediation in the United States." *Journal of Law and Society* 12(1):1–18.
- Galanter, Marc. 1986. "The Emergence of the Judge as a Mediator in Civil Cases." *Judicature* 69(5):257–262.
- Galanter, Marc. 2004. "The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts." *Journal of Empirical Legal Studies* 1(3):459–570.

- Giles, Michael W., Virginia A. Hettinger and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54(3):623–641.
- Gilligan, Carol. 1982. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge: Harvard University Press.
- Greiner, D. James. 2008. "Causal Inference in Civil Rights Litigation." *Harvard Law Review* 122(2):533–598.
- Gryski, Gerard, Eleanor C. Main and William J. Dixon. 1986. "Models of State High Court Decision Making in Sex Discrimination Cases." *Journal of Politics* 48:143–155.
- Hadfield, Gillian K. 2004. "Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases." *Journal of Empirical Legal Studies* 1(3):705–734.
- Hettinger, Virginia A., Stefanie A. Lindquist and Wendy L. Martinek. 2004. "Comparing Attitudinal and Strategic Accounts of Dissenting Behavior on the U.S. Courts of Appeals." *American Journal of Political Science* 48(1):123–137.
- Ho, Daniel E., Kosuke Imai, Gary King and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3):199–236.
- Ho, Daniel, Kosuke Imai, Gary King and Elizabeth Stuart. 2006. "MatchIt: Nonparametric Preprocessing for Parametric Casual Inference." R package version 2.2-11, available at: <http://gking.harvard.edu/matchit>.
- Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396):945–970.

- Imai, Kosuke. 2005. "Do Get-Out-The-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments." *American Political Science Review* 99(2):283–300.
- Kaheny, Erin B., Susan Brodie Haire and Sara C. Benesh. 2008. "Change over Tenure: Voting, Variance, and Decision Making on the U.S. Courts of Appeals." *American Journal of Political Science* 52(3):490–503.
- Kathlene, Lyn. 1994. "Power and Influence in State Legislative Policymaking: The interaction of Gender and Position in Committee Hearing Debates." *American Political Science Review* 88(3):560–576.
- Kritzer, Herbert M. 1986. "Adjudication to Settlement: Shading in the Gray." *Judicature* 70:161–165.
- Kritzer, Herbert M. and Thomas M. Uhlman. 1977. "Sisterhood in the Courtroom: Sex of Judge and Defendant in Criminal Case Disposition." *Social Sciences Journal* 14:77–88.
- Kulik, C.T., E.L. Perry and Molly Pepper. 2003. "Here Comes the Judge: The Influence of Judge Personal Characteristics on Federal Sexual Harassment Court Decisions." *Law and Human Behavior* 27:69–86.
- Lind, E. Allan, Robert J. Maccoun, Patricia A. Ebener, William L. F. Felstiner, Deborah R. Hensler, Judith Resnik and Tom R. Tyler. 1990. "In the Eye of the Beholder: Tort Litigants' Evaluations of Their Experiences in the Civil Justice System." *Law & Society Review* 24:953–996.
- Martin, Patricia Yancey, John R. Reynolds and Shelley Keith. 2002. "Gender Bias and Feminist Consciousness among Judges and Attorneys: A Standpoint Theory Analysis." *Signs* 27(3):665–701.

- Nutt, Paul C. 1990. "Strategic Decisions Made By Top Executives and Middle Managers with Data and Process Dominant Styles." *Journal of Management Studies* 27:173–194.
- Peresie, Jennifer L. 2005. "Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts." *Yale Law Journal* 114(7):1759–1790.
- Pitkin, Hanna. 1967. *The Concept of Representation*. Berkeley: University of California Press.
- Poole, Keith T. 1998. "Recovering a Basic Space From a Set of Issue Scales." *American Journal of Political Science* 42(3):954–993.
- Priest, George L. and Benjamin Klein. 1984. "The Selection of Disputes for Litigation." *Journal of Legal Studies* 13(1):1–55.
- Resnik, Judith. 1982. "Managerial Judges." *Harvard Law Review* 96(2):374–448.
- Rosener, Judy B. 1990. "Ways Women Lead." *Harvard Business Review* 68:119–125.
- Rosenthal, Cindy Simon. 1998. "Determinants of Collaborative Leadership: Civic Engagement, Gender or Organizational Norms?" *Political Research Quarterly* 51(4):847–868.
- Rubin, Donald B. 2006. *Matched Sampling for Causal Effects*. New York: Cambridge University Press.
- Schanzenbach, Max M. and Emerson H. Tiller. 2006. "Strategic Judging Under the U.S. Sentencing Guidelines." *Journal of Law, Economics & Organization* 23:24–56.
- Scherer, Nancy. 2005. *Scoring Points: Politicians, Activists, and the Lower Federal Court Appointment Process*. Stanford, CA: Stanford University Press.
- Segal, Jeffrey A. and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge University Press.

- Segal, Jeffrey A., Harold J. Spaeth and Sara C. Benesh. 2005. *The Supreme Court in the American Legal System*. New York: Cambridge University Press.
- Segal, Jennifer A. 2000. "Representative Decision Making on the Federal Bench: Clinton's District Court Appointees." *Political Research Quarterly* 53(1):137–150.
- Siegelman, Peter and John J. Donohue III. 1990. "Studying the Iceberg from Its Tip: A Comparison of Published and Unpublished Employment Discrimination Cases." *Law & Society Review* 24:1133–1170.
- Sisk, Gregory C., Michael Heise and Andrew P. Morriss. 1998. "Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning." *New York University Law Review* 73(5):1377–1500.
- Slotnick, Elliot. 1984. "The Paths to the Federal Bench: Gender, Race and Judicial Recruitment Variation." *Judicature* 67:370–388.
- Smith, Herbert L. 1997. "Matching With Multiple Controls to Estimate Treatment Effects in Observational Studies." *Sociological Methodology* 27:325–353.
- Songer, Donald R., Sue Davis and Susan Haire. 1994. "A Reappraisal of Diversification in the Federal Courts: Gender Effects in the Courts of Appeals." *Journal of Politics* 56(2):425–439.
- Steffensmeier, Darrell and Chris Herbert. 1999. "Women and Men Policymakers: Does the Judge's Gender Affect the Sentencing of Criminal Defendants?" *Social Forces* 77:1163–1196.
- Sullivan, Kathleen M. 2002. "Constitutionalizing Women's Equality." *California Law Review* 90(3):735–764.
- Walker, Thomas G. and Deborah J. Barrow. 1985. "The Diversification of the Federal Bench: Policy and Process Ramifications." *Journal of Politics* 47(2):596–617.

| District | Personal Injury Cases | | Civil Rights Cases | |
|--------------------------------|-----------------------|------------|--------------------|------------|
| | χ^2 | p -value | χ^2 | p -value |
| Massachusetts | 7.68 | (1.00) | 5.17 | (1.00) |
| Eastern District of Missouri | 2.64 | (1.00) | 1.03 | (1.00) |
| Northern District of Texas | 7.14 | (0.99) | 5.84 | (1.00) |
| Western District of Washington | 11.58 | (0.17) | 7.30 | (0.51) |

Table 1: χ^2 (chi-squared) statistic for random assignment of cases to judges in the four districts in the database. Results are for personal injury cases and civil rights cases. The χ^2 statistic represents how close the number of cases heard by individual judges in that district come to the number they would be expected to hear if assignment was random. The number in parentheses is the p -value (one-tailed).

| | Full Data ($N = 23072$) | | | Percent Improvement | Matched Data ($N = 10142$) | | |
|-------------------------------|---------------------------|----------|-------|---------------------|------------------------------|----------|-------|
| | Mean | Mean | eQQ | | Mean | Mean | eQQ |
| | Treated | Control | Mean | | Treated | Control | Mean |
| Propensity Score | 0.429 | 0.161 | 0.268 | 71.48 | 0.429 | 0.362 | 0.076 |
| Case Termination Year | 1999.613 | 1999.590 | 0.137 | 100.00 | 1999.613 | 1999.613 | 0.000 |
| Minority Judge | 0.121 | 0.157 | 0.036 | 64.29 | 0.121 | 0.134 | 0.013 |
| Party of Appointing President | 0.706 | 0.356 | 0.350 | 94.09 | 0.706 | 0.686 | 0.021 |
| Ideology Score | -0.178 | 0.107 | 0.284 | 71.64 | -0.178 | -0.153 | 0.081 |
| Birth Year | 1943.962 | 1940.017 | 4.518 | 31.86 | 1943.962 | 1942.103 | 3.079 |
| Confirmation Year | 1989.510 | 1988.136 | 2.500 | 0.00 | 1989.510 | 1991.954 | 3.197 |
| State Judge (experience) | 0.567 | 0.426 | 0.141 | 53.21 | 0.567 | 0.633 | 0.066 |
| Magistrate (experience) | 0.322 | 0.065 | 0.257 | 60.54 | 0.322 | 0.221 | 0.102 |
| Prosecutor (experience) | 0.390 | 0.377 | 0.013 | 42.65 | 0.390 | 0.383 | 0.008 |

Table 2: Propensity score matching summary statistics. The left portion of each table provides results for the full, unmatched data, while the right portion displays results after matching has taken place. eQQ Mean is the mean difference in the empirical quantile-quantile plot. Percent improvement records the balance improvement between the eQQ Means for the full and matched data.

| Covariates | Full Data | | Matched Data | |
|--|-----------|--------------|--------------|--------------|
| | Bivariate | Multivariate | Bivariate | Multivariate |
| Female Judge | 0.203* | 0.288* | 0.251* | 0.377* |
| | (0.038) | (0.045) | (0.046) | (0.051) |
| Ideology | | -0.037 | | 0.302* |
| | | (0.049) | | (0.102) |
| Minority Judge | | 0.173* | | -0.015 |
| | | (0.050) | | (0.104) |
| Years on Bench | | -0.015* | | -0.024* |
| | | (0.003) | | (0.004) |
| Prosecutor (experience) | | 0.051 | | 0.030 |
| | | (0.036) | | (0.062) |
| State Judge (experience) | | 0.098* | | 0.026 |
| | | (0.037) | | (0.076) |
| Magistrate (experience) | | -0.462* | | -0.423* |
| | | (0.058) | | (0.080) |
| Eastern District of Missouri | | -0.131* | | -0.075 |
| | | (0.051) | | (0.074) |
| Northern District of Texas | | -0.118* | | -0.280* |
| | | (0.048) | | (0.095) |
| Western District of Washington | | 0.005 | | 0.167* |
| | | (0.047) | | (0.077) |
| (Intercept) | 1.055* | 1.217* | 1.008* | 1.348* |
| | (0.017) | (0.064) | (0.032) | (0.109) |
| <i>N</i> : | 23072 | 23072 | 10142 | 10142 |
| Standard errors in parentheses; * $p < 0.05$ | | | | |

Table 3: Logistic regression estimates for the study of the effect that a judge’s sex has on whether a federal district court case settles. The average treatment effects depicted in Figure 4 are derived from these estimates. The District of Massachusetts is the omitted baseline district court. To conserve space, estimates for year of case termination are omitted.

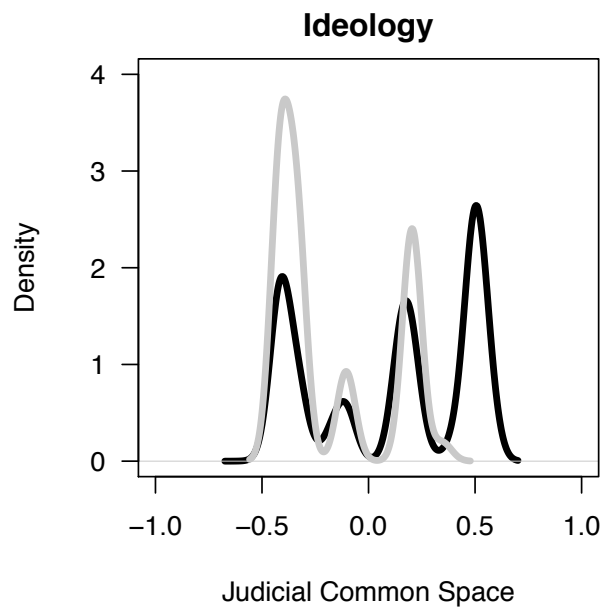


Figure 1: Kernel density plot of the distribution of ideology (measured with Judicial Common Space scores) for male and female district court judges in the data. Male judges are depicted in black; females in grey. Judicial common space scores run from most liberal on the left (-1) to most conservative on the right (1).

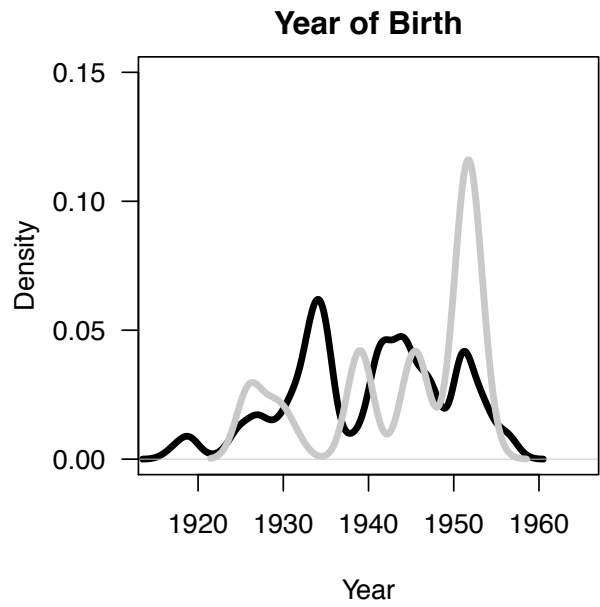


Figure 2: Kernel density plot of the distribution of birth year for male and female district court judges in the data. The black line indicates male judges, while the grey line shows the density distribution for female judges.

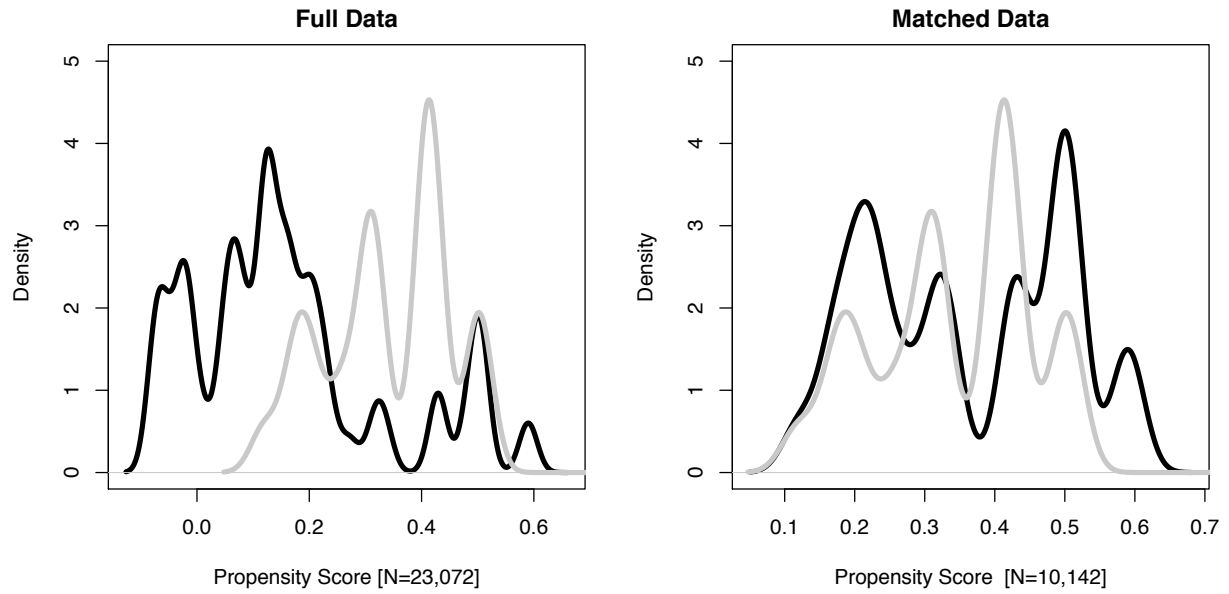


Figure 3: Kernel density plots of the estimated propensity score for the full and matched data. The black lines depict the density for the male judge-assigned cases; grey lines for female-assigned cases. The left-hand panel represents the propensity scores for the full dataset while the right-hand panels display the same for the matched data only.

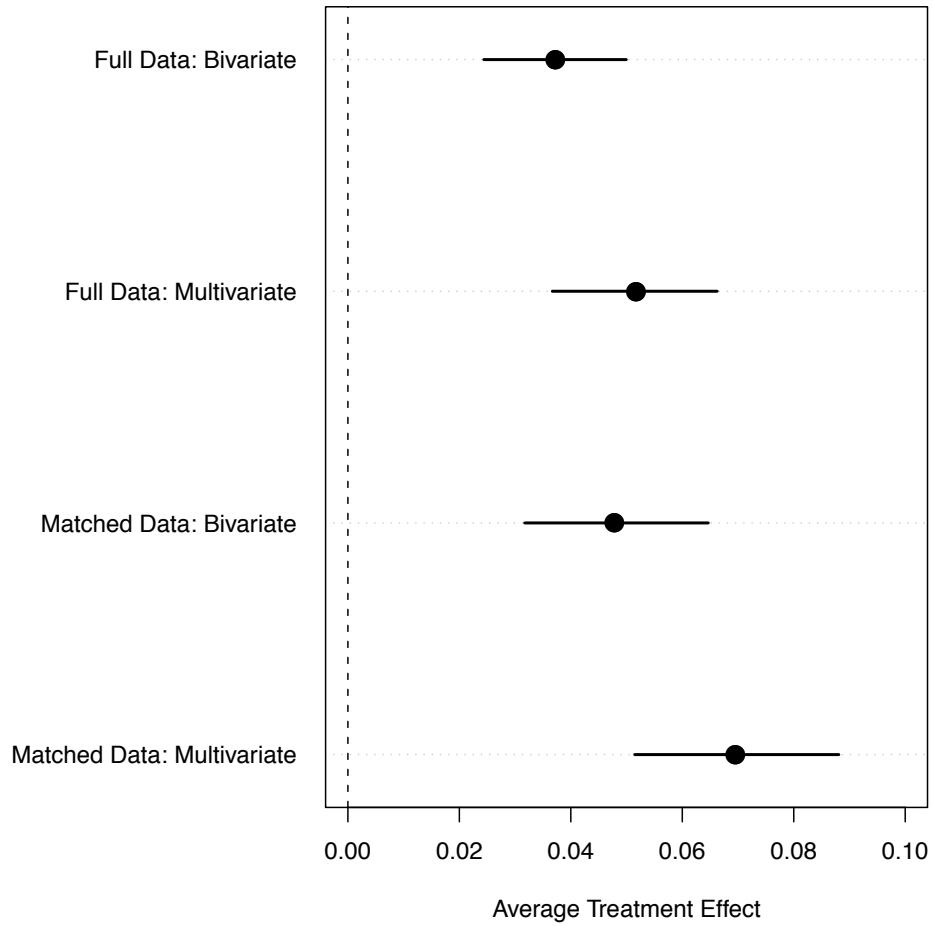


Figure 4: Dotplots of average treatment effects (ATEs) for four estimated regression models. The horizontal lines represent 95% confidence intervals for the ATE. The first two models are logistic regression models fit to the unbalanced, unmatched full dataset. The first of these is a bivariate model (with the treatment – judge sex – as the only covariate) and the second is a full, multivariate model. The third and fourth models are the same (bivariate and multivariate) but are estimated only with the smaller, post-matching dataset.

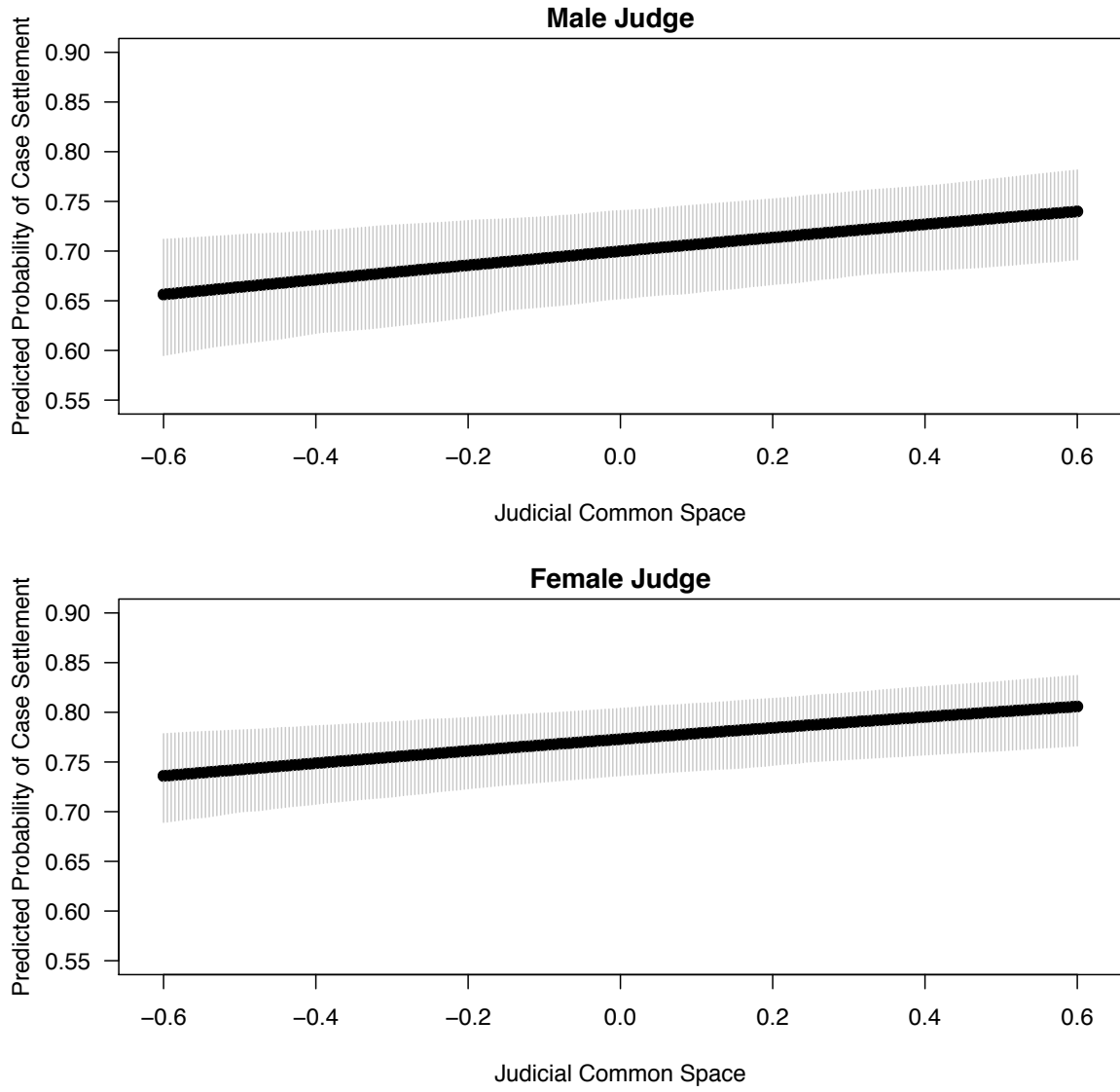


Figure 5: Predicted probability that a case will settle given the sex of a judge and his or her ideology (measured using Judicial Common Space scores). The top panel depicts the probability that a case will settle when assigned to a male judge, with his ideology ranging from (left to right) most liberal to most conservative, while the bottom panel does the same for a female judge.