

Gender-Based Favoritism Among Criminal Prosecutors

Stephanie Holmes Didwania¹

Abstract

Prosecutors enjoy wide discretion in the decisions they make but are largely unstudied by quantitative empirical scholars. This paper explores gender bias in prosecutorial decision-making. I find that defendants are charged more leniently when they are the same gender as their prosecutor as opposed to when the defendant and prosecutor are different genders. Such favoritism at charging translates into significantly lower sentences for defendants who are paired with an own-gender prosecutor. Further, this gender-based leniency is more pronounced in states with less prevalent sexism, cases in which gender is likely to be more salient, and on same-gender prosecutorial teams. However, gender match in defendant-prosecutor pairs is not strongly associated with differences in cooperation and bargaining. I conclude that prosecutors' social preferences are more likely to explain gender-based leniency than differences in how male and female prosecutors work.

¹ Assistant Professor of Law, Temple University Beasley School of Law. I am grateful to Emily Buss, Adam Chilton, Travis Crum, Dhammika Dharmapala, Vikas Didwania, Nate Ela, J.B. Heaton, William Hubbard, Benjamin Jones, Cree Jones, Emma Kaufman, Amanda Kleintop, Genevieve Lakier, Dorothy Lund, Anup Malani, Jonathan Masur, Tom Miles, Margot Moinester, Rachel Montgomery, Meghan Morris, Jeffrey Omari, Manisha Padi, Asad Rahim, Weijia Rao, Kyle Rozema, CJ Ryan, Max Schanzenbach, Roseanna Sommers, Lior Strahilevitz, and participants at the Conference on Empirical Legal Studies, the Annual Meeting of the American Law and Economics Association, the Marquette University Law School Faculty Workshop; the Midwestern Law and Economics Association, the UC Berkeley Law, Economics & Business Workshop; the University of Chicago Law School Faculty Works-in-Progress Workshop; and the University of Michigan Law and Economics Workshop for valuable comments and suggestions. Comments are welcome at didwania@temple.edu.

1. Introduction

Prosecutors are among the most powerful actors in our criminal legal systems. They have the sole power to initiate criminal cases, and they exercise important decision-making authority throughout the criminal process, for example by selecting which charges to bring against defendants and by engaging in plea bargaining. These decisions are largely unreviewable by courts and are subject to little public oversight. Although prosecutors exercise vast discretion in criminal cases, their behavior is largely unstudied by quantitative empirical researchers.²

This paper examines in-group favoritism in prosecutorial decision-making in federal criminal cases. In-group favoritism occurs when a decision-maker gives preferential treatment to those who share a salient trait with the decision-maker, such as being a member of their gender, racial, ethnic, or religious group (Everett, Faber, and Crockett 2015).³ This paper explores in-group favoritism on the basis of gender, a form of bias that has been documented in both laboratory and real-world settings (Hoyt, Simon, and Reid 2009; Lindeman & Sundvik 1995; Rudman & Goodwin 2004; Vial et al. 2017; Kennedy, McDonnell & Stephens 2017; Jannati et al. 2016).⁴

This paper presents evidence of gender-based in-group favoritism in the charging behavior of federal prosecutors. Using data that covers roughly 130,000 federal criminal defendants sentenced in 2002 through 2016, I find that while male and female prosecutors exhibit small and statistically insignificant differences in their treatment of defendants overall, they show relative favoritism towards defendants of their own gender. The leniency associated with gender matching in defendant-prosecutor pairs ultimately translates into significantly lower sentences for defendants who match their

² On the other hand, there is considerable theory on the promise and pitfalls of the way prosecutor offices self-police, incentives that prosecutors face, and prosecutorial ethics (Bibas 2009, Davis 2019, Leonetti 2012, Meares 1995, Ouziel 2017, Pfaff 2017, Sklansky 2017).

³ In this paper, I use the term *in-group favoritism* to represent a phenomenon: that decision-makers treat people relatively more favorably when the person is in their in-group. It is important to emphasize that I do not take a view about the underlying causes of this differential favoritism. For example, I do not take a position on whether the effect that I document is conscious or subconscious, nor do I have a view about the extent to which it is driven by out-group disfavoritism versus in-group favoritism.

⁴ Because of the way the data is coded, this paper classifies people as having a binary gender that is either female or male. Unfortunately, neither of the data sources used in the paper include information about transgendered defendants or defendants who do not identify as having binary gender. Rosenblum (2000) provides a thoughtful discussion of the legal issues facing transgendered prisoners.

prosecutor’s gender compared to those who do not. I also show that this gender-based leniency is not uniform across all geographic areas, case types, and defendants. As one might expect, the results are also stronger in cases in which gender is likely to be more salient, and among same-gender teams of prosecutors. The results are also most pronounced in states with below-median levels of sexism.

The paper proceeds as follows. Part 2 situates this paper in the prior literature. Part 3 examines the empirical setting, describing the key responsibilities of federal prosecutors and explaining how their offices are organized. Part 4 presents the empirical strategy. Part 5 describes the data and includes the main results. Part 6 attempts to shed light on the mechanisms responsible for the main findings by examining heterogeneity in the results along several dimensions. Part 7 discusses the findings, considers policy interventions, and suggests directions for further research.

2. Relevant Prior Literature

Limited quantitative work examines the behavior of criminal prosecutors, and the existing research—while insightful—leaves many unanswered questions. Very few papers leverage individualized data on line prosecutors—the career attorneys who carry out criminal law around the United States. One important exception is a recent working paper by CarlyWill Sloan, which documents an increased likelihood of conviction for property-crime misdemeanors, but not other types of misdemeanors, when the defendant and prosecutor are of different races (Sloan 2019). Such individualized analysis can shed light on the types of bias that might influence prosecutorial decision-making. Earlier empirical research about prosecutors, though largely lacking individualized data on line prosecutors, illustrates that prosecutors could be influenced by bias.⁵ While this prior work suggests that prosecutorial *offices* treat defendants in a biased way, it does not attempt to explain how bias enters the decision-making process of individual prosecutors.⁶ This paper begins to fill this gap.

⁵ There is mixed evidence that minority defendants also fare worse in plea bargaining. Kutateladze, Andiloro, and Johnson (2014) and Kutateladze et al. (2014) find that Black and Latino defendants are more likely to receive a custodial plea offer than White defendants, but Kutateladze et al. (2014) also finds that Black and Latino defendants are more likely to have their cases dismissed. Both of these papers use data from New York City that is comprised primarily of misdemeanor defendants.

⁶ One important exception is a recent working paper in which the authors carried out a randomized, controlled experiment in which they surveyed prosecutors by presenting them with vignettes and asking them how they would charge in each situation. The vignettes were manipulated to change the race and class

Earlier work suggests prosecutors are influenced by career concerns. Prosecutors are more likely to take cases to trial in federal districts in which the local labor market for attorneys has above-average salaries (Boylan and Long, 2005).⁷ Boylan (2005) similarly examines the career paths of 570 United States Attorneys (the attorneys who lead federal prosecutorial offices) over a thirty year period and finds that the length of prison sentences imposed on defendants who were prosecuted by the office is the most important predictor of favorable career outcomes for U.S. Attorneys after leaving their office.

Other prior work has studied prosecutorial decision-making by analyzing charging decisions. Bjerck (2005) finds that prosecutors strategically charge mandatory minimums. Others have found that charging decisions are a critical source of gender- and race-based sentencing disparity (Rehavi and Starr 2014, Starr 2015, Yang 2014). Most race-based sentencing disparity begins at the charging stage, in large part because prosecutors are more likely to charge crimes carrying mandatory minimum sentences against Black defendants (Rehavi and Starr 2014). This prosecutorial tactic appears to respond to increased judicial discretion after the Supreme Court's 2005 decision, *United States v. Booker* (Starr and Rehavi 2013, Yang 2014, Yang 2015).

Empirical scholars also have long examined race- and gender-based bias by actors other than prosecutors in the criminal system, such as judges, law enforcement officers, and jurors. Most of this scholarship has examined bias with respect to race and ethnicity rather than gender, and the literature that investigates gender favoritism has produced mixed results. One recent article finds no strong evidence that judges either favor or disfavor defendants of their own gender (Lim, Silveira, and Snyder 2016). Earlier research reports that the presence of more female judges on a district court reduces the court's gender gap in sentence length, but because this analysis is aggregated to the

of the fictional perpetrator. The authors generally did not find evidence of bias in prosecutorial decision-making (Robertson, Baughman, and Wright 2019). One prior paper examines in-group favoritism among prosecutors on the basis of race, but only has racial information aggregated to the federal district level (Ward, Farrell, and Rousseau 2009). The authors find, among other things, that districts with more Black representation among prosecutors have smaller Black-White disparities in sentencing outcomes.

⁷ The findings presented in Boylan and Long (2005), however, are also consistent with an effect that is driven by defense counsel facing the same kinds of incentives that the authors attribute to prosecutors. The decision to go to trial is not solely a prosecutorial decision. It is the defendant's Sixth Amendment right to exercise, although prosecutors undoubtedly influence the decision during plea bargaining.

district court level, it does not constitute evidence of in-group favoritism (Schanzenbach 2005). Other earlier work finds that female judges show more gender disparity than male judges, but this study does not report whether this difference are statistically significant (Steffensmeier and Hebert 1999).⁸

Considerably more scholarship studies race-based favoritism in the criminal system, although evidence related to in-group favoritism is mixed. One recent article finds that judges exhibit in-group disfavoritism⁹ towards juvenile defendants of the judge's own race (Depew, Eren, and Mocan 2017),¹⁰ but another recent paper suggests that judges might slightly favor adult defendants of their own race (Lim, Silveira, and Snyder 2016). Others similarly find that African American judges exhibit smaller racial disparities in sentencing than their White counterparts, which is consistent with race-based in-group favoritism (Abrams, Bertrand, and Mullainathan 2012). In-group favoritism also presents to some extent along ethnic lines in pretrial detention decisions: one study finds that Arab and Jewish judges in Israel are less likely to detain defendants who share their ethnicity, but in-group favoritism does not affect the length of detention ordered (Gazal-Ayal and Sulitzeanu-Kenan 2010).

In the case of juror behavior, it appears that the diversity environment—and not just the race of the individual actor—affects outcomes. Researchers have found that juries formed from all-White jury pools convict Black defendants more often than White defendants, and that this gap is entirely eliminated when the jury pool includes at least one Black member, even when the Black jury pool members are not seated (Anwar, Bayer, and Hjalmarsson 2012). Recent work has also documented evidence of favoritism on the basis of race and gender in jurors' decisions to convict (Flanagan 2018). The effects of racial diversity among judges are more nuanced, but do not constitute strong evidence of race-based in-group favoritism (Schanzenbach 2005, Schanzenbach 2015).

⁸ The study also does not include fixed effects for geography or time, nor are standard errors clustered so it is hard to draw firm inferences from the results.

⁹ Depew, Eren, and Mocan (2017) refer to this phenomenon as “negative in-group favoritism” in their article. For clarity, this paper uses the term “in-group disfavoritism” to mean the same thing: that a decision-maker treats in-group members worse than out-group members.

¹⁰ Another recent contribution finds that judges treat more harshly defendants who share the same first initial of their name, which the author describes as an example of implicit egotism (Chen 2017).

Work that examines in-group favoritism among law enforcement officers has largely focused on racial rather than gender bias,¹¹ and—like the work on in-group favoritism among judges—has produced mixed results. Earlier work, aggregated to the police-force level, finds that police departments with more minority officers are more likely to arrest White suspects, with little impact on the arrests of non-White suspects (Donohue and Levitt 2001). Others similarly find that officers are more likely to conduct a search of a driver of a different race, which they attribute to preference-based discrimination after ruling out the possibility that officers are better at searching members of their own racial group or that the results are driven by the non-random assignment of officers to neighborhoods (Antonovics and Knight 2009). Others, on the other hand, do not find race-based differences in officers’ propensities to stop or arrest non-White suspects (Brown and Frank 2006, Sanga 2014).

Outside the criminal system, scholars have documented in-group favoritism in civil and extralegal settings. A study of in-group favoritism in professional basketball finds that NBA referees demonstrate race-based in-group favoritism towards players (Price and Wolfers 2010). In the civil context, Israeli judges favor defendants that share their ethnicity in small claims cases, and in-group favoritism increases at times when ethnicity is more salient (operationalized as years in which there was a recent terrorist attack in the vicinity of the court) (Shayo and Zussman 2011). Earlier work has similarly found that judges’ race and gender identities affect decision-making in cases in which race and gender are salient, such as affirmative action, discrimination, sexual harassment, and voting rights cases (Boyd, Epstein, and Martin 2010; Cox and Miles 2008; Farhang and Wawro 2004; Peresie 2005). Much prior literature also considers in-group favoritism in the employment and educational contexts, but it is not clear that these findings will translate to the criminal setting. Part 3 describes the research setting studied here—the federal criminal system.

¹¹ One exception is an unpublished paper that examines ticketing behavior on the basis of gender. The author finds that, compared to male police officers, female officers are less likely overall to ticket, but relatively more likely to ticket female drivers (Rowe 2009).

3. Research Setting

This section briefly describes the responsibilities of federal prosecutors and how they carry out their work. The federal U.S. Attorney’s Offices (USAOs) vary in size, organization, and case composition. This section describes the rules that are common to all USAOs, highlights instances in which heterogeneity is particularly prominent, and describes formal and informal checks on prosecutorial decision-making in federal criminal cases.

Each of the 94 geographically distinct federal district courts is associated with precisely one USAO, with one exception.¹² Figure 1 labels and depicts the boundaries of the federal district courts. The shaded districts indicate those districts that are represented in the data used in this paper. Each USAO is led by a United States Attorney. The lawyers who work in USAOs are called Assistant United States Attorneys (AUSAs). USAOs represent the United States as a party in both civil and criminal federal cases in their districts, and typically include separate criminal and civil divisions.¹³

USAOs function mostly autonomously and vary in size and organizational structure. For example, the U.S. Attorney’s Office for the District of Hawaii employs 27 AUSAs,¹⁴ and the criminal division of the office is divided into three sections: Drug and Organized Crime, Fraud and Financial Crimes, and Special Crime. In contrast, the U.S. Attorney’s Office for the Central District of California—based in Los Angeles—employs 264 AUSAs, and has a criminal division split into 10 sections.¹⁵ On average, a USAO is divided into roughly four sections.¹⁶

In handling criminal cases, USAOs are broadly tasked with prosecuting violations of federal criminal law in their jurisdictions. USAOs enjoy wide discretion in how they carry out this work, and most prosecutorial decisions are unreviewable by courts except

¹² The District of Guam and the District of the Northern Mariana Islands share a USAO.

¹³ I collected this information by hand by visiting the website of each USAO. The majority of USAO websites indicated that they have separate civil and criminal sections. I did not find any instances in which a USAO website stated that the USAO’s civil and criminal sections were combined. There were a handful USAOs for which it was unclear whether the civil and criminal sections were combined or separate.

¹⁴ Of these 27 AUSAs, 20 are assigned to the criminal division.

¹⁵ The ten sections are: Asset Forfeiture, Criminal Appeals, Cyber and Intellectual Property Crimes, Public Corruption and Civil Rights, Major Frauds, Organized Crime Drug Enforcement Task Force (OCDETF), Public Integrity and Environmental Crimes, and Violent and Organized Crime. The office also includes a national security division and a tax division, both of which handle civil and criminal cases.

¹⁶ This information was collected by hand by visiting the website of each U.S. Attorney’s Office.

in limited circumstances.¹⁷ In addition to charging defendants with crimes, prosecutors also engage in plea bargaining, which is the way that most criminal cases are resolved in the United States.

Federal prosecutors also advocate at various points during a criminal prosecution. Most notably, AUSAs advocate whether they believe the defendant should be detained pending trial, and recommend what they think is an appropriate sentence. Prosecutors can also support reductions or enhancements to a defendant's sentence. For example, under Federal Rule of Criminal Procedure 35(b) and U.S. Sentencing Guideline § 5K1.1, a prosecutor can ask the court to reduce a defendant's sentence if the defendant provides substantial assistance to the government.¹⁸

Despite limited oversight from the courts, a prosecutor's decision-making authority is not absolute. Individual AUSAs are subject to both formal and informal supervision from several sources, which might mitigate the effects of bias. First, like most employees, individual prosecutors are supervised within their workplace. For the time period covered by this paper, all AUSAs were governed by the United States Attorneys' Manual (the "Manual") which laid out detailed guidelines for how USAOs should be organized and how individual AUSAs should exercise discretion.¹⁹ For example, the Manual dictated that "[t]o ensure consistency and accountability, charging and plea agreement decisions must be reviewed by a supervisory attorney."²⁰

Second, the Manual included instructions for deciding which charges to file. During the time period covered by this paper, the Manual expressed a policy that federal prosecutors should usually charge "the most serious offense that is consistent with the nature of the defendant's conduct, and that will probably be sufficient to sustain a

¹⁷ One of the few ways a defendant can challenge their prosecution is on the ground that it was brought selectively—that is, based on a prohibited consideration such as the defendant's race or religion. *Oyler v. Boyles*, 368 U.S. 448, 456 (1962). In practice, these selective prosecution challenges virtually never succeed (McAdams 1998, Bibas 2009). Some argue that prosecutorial power is so unchecked that it raises separation-of-powers concerns (Barkow 2009).

¹⁸ The Guidelines allows a sentencing judge to give the defendant a reduced sentence (including a sentence below the mandatory minimum) "upon motion of the government stating that the defendant has provided substantial assistance in the investigation or prosecution of another person who has committed an offense." U.S. SENTENCING GUIDELINES MANUAL § 5K1.1 (2016).

¹⁹ As noted in footnote 9, the United States Attorney's Manual was replaced by the Justice Manual in September 2018. Because AUSAs were governed by the United States Attorneys' Manual for the time period covered by this paper, I cite this version of the Manual.

²⁰ UNITED STATES ATTORNEYS' MANUAL § 9-27.300.

conviction.”²¹ In theory, this directive promotes uniformity in prosecutorial charging behavior, which might mitigate bias. On the other hand, the Manual left room for an AUSA to deviate from this policy by conducting an “individualized assessment” of the defendant’s offense conduct and history.²²

Third, although courts are “hesitant to examine the decision to prosecute,”²³ they exercise oversight during plea bargaining and sentencing. For example, judges have the power to reject certain kinds of plea agreements reached by the prosecutor and defendant.²⁴ In federal court—the setting studied here—judges enjoy considerable discretion in choosing a defendant’s sentence, especially after the Supreme Court’s decision in *United States v. Booker*.²⁵ It is possible that judges sentence in a way that neutralizes any prosecutorial bias.

Finally, prosecutors are subject to outside pressures from the legislative branch, media, advocacy organizations, and the general public. Prosecutors nonetheless tend to play an outsized role in criminal justice policy-making (Barkow 2013). In the federal setting studied in this work, prosecutors are not subject to electoral pressures, and are in some ways insulated from political pressures.²⁶ In contrast, nearly all states elect the District Attorneys that lead their prosecutorial offices (Lantigua-Williams 2016).

4. Data and Empirical Strategy

This paper estimates gender favoritism using regression analysis. There are two hurdles to inference, which are described in more detail below. The first is that prosecutors are not randomly assigned to cases. The second is that underlying criminal conduct is not directly observable.

²¹ UNITED STATES ATTORNEYS’ MANUAL § 9-27.300.

²² UNITED STATES ATTORNEYS’ MANUAL § 9-27.300.

²³ *Wayte v. United States*, 470 U.S. 598, 608 (1985).

²⁴ Fed. R. Crim. P. 11(c)(3)(A), 11(c)(5).

²⁵ 543 U.S. 220 (2005). In *Booker*, the Supreme Court held that the U.S. Sentencing Guidelines are advisory; that is, that Congress may not require district judges to sentence defendants within the Guidelines range. However, a district judge must always calculate a criminal defendant’s Guidelines range before sentencing and articulate any reasons for deviating from that advisory range. 543 U.S. at 259-60; *see also Peugh v. United States*, 569 U.S. 530, 530 (2013) (“District courts are required to being their sentencing analysis by looking at the ranges in the guidelines; a judge must have a good reason for deviating from those ranges”).

²⁶ Historically, U.S Attorneys have been insulated from political pressure, although some argue that the positions have become increasingly politicized over the last decade. In 2006, President George W. Bush fired seven U.S. Attorneys, and in 2017, President Donald Trump dismissed all U.S. Attorneys that had been appointed by President Barack Obama.

4.1. Data

This paper uses data from two distinct sources: (1) the United States Sentencing Commission’s annual sentencing data files (the “Commission data”); and (2) federal case data from the Legal Information Office Network System, which is published by the Executive Office of United States Attorneys in response to Freedom of Information Act requests (the “LIONS data”). This section describes the two data sources and explains how I combined them.

The United States Sentencing Commission (the “Commission”) annually publishes data files of defendants sentenced in federal district courts.²⁷ The Commission extensively reviews the data it receives for completeness and quality before making the data files available to the public on its website.²⁸ The Commission’s individual-level data files include detailed case information, including: demographic characteristics of the defendant; the defendant’s criminal history; the statutes under which the defendant was convicted; the defendant’s recommended Guidelines range; the sentence imposed; and any reasons for an out-of-range sentence.

This paper uses Commission data for defendants sentenced in 2002 through 2016. The primary outcome variable used in the analysis is the defendant’s base offense level, which reflects the severity of the offense with which the defendant was charged. The paper also uses defendant- and case-specific information, both as control variables and for matching defendants between the two data sets.

The Commission data is insufficient to estimate gender-based favoritism because it does not identify the prosecutor(s) that worked on each case. Nor does it identify the courthouse in which the defendant was sentenced—it simply identifies the federal district court. Roughly 80 percent of federal district courts comprise more than one courthouse, to which defendants, judges, and prosecutors are non-randomly assigned,

²⁷ The Commission reports data for defendants convicted of felonies and Class A misdemeanors. Notably, the Sentencing Commission data files do not include information about cases involving: juvenile offenders, defendants convicted of Class B and C misdemeanors, and death penalty cases (Reedt, Semisch, and Blackwell 2013).

²⁸ The Commission cross-checks its data for completeness and quality with data from another source—the Administrative Office of U.S. Courts. It pays special attention to cases with out of range values, logical inconsistencies, and sentences outside the Guidelines range.

making it impossible for an empirical researcher to fully account for intra-courthouse correlation using Commission data. The LIONS data fills these gaps.

The Executive Office for United States Attorneys (EOUSA) regularly publishes case data from the 93 USAOs located throughout the United States. This data originates from the Legal Information Office Network System (LIONS), which is the computer program that the EOUSA uses to track cases. The breadth of the LIONS data is substantial—the data covers all cases in which a USAO was involved, including those cases that are the primary responsibility of another agency and those the USAO declines to prosecute.

The LIONS data includes abundant case-specific information but little defendant-specific information. Crucially, though, the LIONS data provides limited information about the USAO staff members who work on each case. LIONS anonymously identifies staff members and labels their role in a case as, for example, lead attorney, co-counsel, paralegal, or victim witness coordinator. Roughly 84 percent of staff assignments in the data are coded as “lead attorneys.” The LIONS data also includes a field that provides the staff member’s salutation—such as “Mr.” or “Ms.” I use the salutation field to deduce the staff member’s gender. A complete list of all salutations and how each was coded is produced in Appendix A.

It is important to note two shortcomings of the LIONS data. First, unlike the Commission data, there is no evidence that the LIONS data is cross-checked for accuracy with court documents. Second, the LIONS data contains many more missing values than the Commission data.

The data used in this paper comprises a sample that is merged between the Commission and LIONS data sources of federal defendants sentenced in 2002 through 2016. The sample is restricted to U.S.-citizen defendants in district courts that do not border Mexico.²⁹ In total, 73.6 percent of such defendants from the Commission data are

²⁹ I remove non-citizens and defendants in border districts for two reasons. First, non-citizen defendants who are charged with immigration offenses—as most non-citizen defendants are—are often eligible for reduced sentences as part of the Department of Justice’s Fast Track program (Cole 2012), so the extent to which prosecutors exercise charging discretion in these cases is very different from cases involving U.S. citizen-defendants. Second, and consequently, because the border districts include many non-citizen defendants charged with immigration offenses and receiving identical sentences, these defendants are intrinsically difficult to match between the Commission and LIONS data sets. The results, however, are largely robust to including citizen-defendants from the border districts (results on file with the author).

identified in the LIONS data. After merging, the sample is reduced in several ways. The most important restriction is removing defendants whose lead prosecutor is of unknown gender—a little more than half of defendants.³⁰ This significant reduction does not compromise the representativeness of the sample, as Table A.1 shows.³¹

I also remove defendants from four federal district courts whose courthouse identification variables appear to be unreliable;³² defendants who are coded as different genders in the LIONS and Commission data sets (1.5 percent of defendants); and defendants who are missing entries in any of the following variables in the Commission data: gender, race, Hispanic ethnicity, age, educational attainment, criminal history points, offense type, sentencing fiscal year, and base offense level (3.5 percent of defendants).

Because I control for the year in which a case was commenced and the Guidelines version used at sentencing, I exclude defendants whose cases commenced before 1999 because there are few defendants in these early years. Because I include courthouse-prosecutor fixed effects in the main regression analysis, I exclude courthouse-prosecutors that have fewer than 25 observations and those that prosecute only male or only female defendants. Finally, I exclude offenders charged with sex offenses or offenses involving child exploitation, a decision that is discussed in subsection 4.3.

The sample used in the main analysis includes 131,462 defendants sentenced in 64 of the 94 federal district courts in the United States, comprising 173 unique courthouses, 1,520 unique prosecutors, and 1,615 unique courthouse-prosecutor

³⁰ A lead prosecutor’s gender is unknown if the salutation field in the data is empty, or if the salutation entry does not convey gender information (for example, if it is “AUSA”). Appendix A contains a complete list of how salutations are coded. Missingness in the salutation field is largely driven by missingness between—rather than within—USAOs. Among those USAOs with any gender information, most have gender information for at least 75 percent of AUSAs.

³¹ Appendix Table A.1 assesses the representativeness of the sample by comparing the variable means of defendant and case characteristics presented in Table 1 to those: among all Commission defendants in federal districts represented in the data (column (2)); among all defendants in the merged sample (column (3)); and among the full universe of Commission defendants (column (4)). Table A.1 allows readers to not only assess how well the data used in this work represents the full universe (column (1) versus column (4)), but also to assess how well the matching process worked (column (3) versus column (4)). In all columns, I restrict to U.S. citizen-defendants and exclude districts that border Mexico so as to compare equivalent populations. Table A.1 demonstrates that data used in this paper is extremely representative of the districts represented in the data, the matched sample as a whole, and the universe of all Commission defendants.

³² Data from the Northern District of Alabama, the Northern District of Iowa, the District of Nebraska, and the District of South Carolina contain many more courthouses than actually exist in those districts.

combinations. Table 1 presents summary statistics of the data. Roughly 72 percent of cases have a male lead prosecutor. Eighty-three percent of defendants are male. Sixty-nine percent have graduated from high school but only seven percent from college. Around 42 percent of defendants are Black, 12 percent are Hispanic ethnicity, and 41 percent are non-Hispanic White. The average defendant age in the sample is 36 years old.

In terms of case outcomes, the vast majority of defendants—roughly 87 percent—receive sentences of some incarceration, and the mean sentence among all defendants is 62 months. Average sentence length is considerably lower than the mean Guidelines range, which is 86 months. This is because roughly half of defendants receive a sentence below their recommended Guidelines range, while just two percent receive a sentence above their recommended Guidelines range.

4.2. Regression Design

I obtain estimates of gender-based favoritism using the following empirical specification:

$$y_{ipfct} = \beta_0 + \beta_1 MaleD_i * MaleP_{ip} + \beta_2 MaleD_i + \beta_3 MaleP_{ip} + \theta X_i + \gamma_{fc} + \delta_t + \epsilon_{ipfct} \quad (1)$$

where y_{ipfct} is an outcome for defendant i , assigned prosecutor p , charged with offense of type f , in courthouse c , in year t ; $MaleD_i$ is an indicator variable that equals one if defendant i is male; $MaleP_{ip}$ is an indicator variable that equals one if defendant i 's prosecutor p is male; X_i is a vector of observable characteristics of defendant i and their case; and γ_{fc} and δ_t are courthouse-section type and time fixed effects, respectively.

In this formulation, β_1 captures in-group favoritism: average differences in the outcomes for female defendants relative to male defendants prosecuted by female versus male prosecutors.³³ In other words, Equation (1) is a difference-in-differences design³⁴ in which, holding all else equal, the predicted differences in base offense level for each defendant-prosecutor gender combination relative to female-female pairs are:

³³ The paper uses ordinary least squares regression throughout for interpretive clarity. All regressions in which the outcome variable is binary are robust to using logit regression.

³⁴ This empirical specification follows Shayo and Zussman (2011).

	Male Pros	Female Pros	Difference (M-F)
Male Def	$\beta_1 + \beta_2 + \beta_3$	β_2	$\beta_1 + \beta_3$
Female Def	β_3	–	β_3
Difference (M-F)	$\beta_1 + \beta_2$	β_2	β_1

4.3. Estimation and Identification

The central challenge in estimating equation (1) is that whether a prosecutor and defendant match on gender might be correlated with unobservable defendant or case characteristics that also affect case outcomes. Controlling for many observable defendant and case characteristics and relevant fixed effects will help, but might not fully capture unobservable differences between gender-matching and gender-mismatching cases.

There are several ways in which the assignment of cases to prosecutors might be non-random in a way that biases the coefficient estimate of β_1 . Most obviously, the estimate of β_1 will be biased upward (downward) if prosecutors are likely to be assigned cases with less serious (more serious) defendants with whom they match on gender. Without data on each defendant’s underlying criminal conduct, it is impossible to be sure that differences in case assignment are not masquerading as gender-based favoritism.

There are several reasons to think case assignments are likely to be orthogonal to gender match (especially conditional on observables). First, it is likely that the factors that influence an AUSA’s charging decision are not known by a supervisor at the time of case assignment. If supervisors do not have a nuanced picture of a defendant’s social history and specific offense conduct at the time case assignments are made, it would be difficult for such assignments to be made strategically. For this reason, the paper codes the prosecutor’s gender based on the first prosecutor assigned to a case. Prosecutors who join a case later might be more strategically assigned, but they are not part of the main analysis.³⁵

Second, one might worry that female prosecutors could disproportionately work in areas of federal criminal law in which female defendants are treated even more leniently

³⁵ In the data, 99.8 percent of cases have just one prosecutor assigned on the first day. For the few cases in which more than one prosecutor is assigned to a case on the first day, I code the prosecutor as female if all prosecutors assigned on the first day are female, male if all prosecutors assigned on the first day are male, and remove the case from the data if both male and female prosecutors are assigned on the first day. In subsection 5.3 below, I examine the behavior of same-gender and mixed-gender teams of prosecutors.

than usual. The main analysis also includes prosecutor-specific fixed effects in the regression equations, which should address this concern. I also exclude cases involving sexual offenses and child exploitation. These cases exhibit above-average gender disparities, involve very few female defendants, and are more likely than other types of cases to be staffed by female prosecutors. I exclude this offense category from the main analysis, although the results are robust to including these cases. I do not find evidence of gender disproportionality in the share of prosecutors in any other offense type. Among the five offense types remaining in the data after sex offenses are removed, the percentage of lead prosecutors that are female ranges from 27 to 29 percent. In contrast, among sex offense cases, more than 40 percent of lead prosecutors are female.

Third, I look for quantitative evidence to shed light on whether prosecutors are assigned observably different defendants by gender. Table 2 presents results of regressing the prosecutor's gender on a host of case characteristics that were fixed at the time the case began.³⁶ Columns (1) and (2) present results in the whole sample, while columns (3)-(4) separate the cases by the defendant(s)' gender.

In column (1), prosecutor gender is regressed on case characteristics without any time or geographic controls. In these regressions, male prosecutors are associated with an increased likelihood of prosecuting male defendants and a decreased likelihood of prosecuting defendants of color. Once courthouse-section fixed effects are included in column (2), male and female prosecutors no longer appear to differ in the racial compositions of their caseloads. This change demonstrates the crucial role of courthouse-section fixed effects—it suggests that female prosecutors are more prevalent in geographic regions and offense types in which non-White defendants are more prevalent.

Even after including courthouse-section fixed effects in column (2), the coefficient estimate on defendant gender remains positive and highly significant. The point estimate—0.021—suggests that a male defendant's probability of being assigned a male prosecutor is two percentage points larger than a female defendant's probability of being assigned a male prosecutor. Thus, it appears that supervisors might be influenced (consciously or not) by gender in making case assignments, which is not surprising.

³⁶ The offense type categories are: drug offenses, fraud and other White-collar offenses, regulatory offenses, and violent offenses.

Researchers studying the assignment of law enforcement officers to neighborhoods similarly report that officers are more likely to be assigned to areas in which the majority racial group is the officer's racial group (Antonovics and Knight 2009, Donohue and Levitt 2001). None of the other case characteristics are statistically significant predictors of whether the prosecutor is male.

Given that defendants are slightly but statistically significantly more likely to be assigned a prosecutor of their own gender, it is crucial to check whether prosecutor gender seems to be correlated with any defendant and case characteristics that are fixed at the time of the case, *conditional on defendant gender*. Table 2, column (3) reports results of regressing prosecutor gender on all the covariates listed above among cases with female defendants, while column (4) reports results among male defendants. The point estimates presented in columns (3) and (4) are all close to zero and all but one are statistically insignificant, suggesting that female and male prosecutors see similar mixes of defendants *within* each gender group.

It is also important to remember that USAOs likely vary in their assignment procedures. In many courthouse-sections, prosecutor gender is not significantly correlated with defendant gender. Subsection 5.4 restricts attention to courthouse-sections in which prosecutor gender is not significantly related to any defendant characteristics that are fixed at the time of the case—including gender—and finds that the results are quite similar (if not stronger) than those in the main analysis, offering more evidence that the results are not driven by differences in prosecutors' caseloads.

4.4. Selection into the Sample

Because this paper uses data on defendants sentenced in federal district courts, it necessarily excludes defendants who are not ultimately convicted. In this subsection, I present evidence that suggests it is unlikely that sample selection biases the results of the paper. All quantitative work on the criminal system must grapple with some form of the selected sample problem: researchers typically only observe defendants who formally enter into the criminal system (or, in the case of many data sources, are eventually sentenced). It is plausible that selection into the sample could bias the results if there is

gender-based favoritism on the margin of selection.³⁷ On the other hand, one might not expect selection into the sample to pose a serious threat in the federal setting because nearly all federal felony defendants—around 94 percent—are ultimately convicted. Only 0.4 percent are acquitted after trial. The remaining 5.6 percent of cases that do not result in conviction are those voluntarily dismissed by the government (Motivans 2017).

I am able to make progress on the sample selection problem because the LIONS data includes defendants whose charges the government declines to prosecute, voluntarily dismisses, or of which the defendant is acquitted. Therefore, I can investigate whether there appears to be gender-based favoritism on the margin of whether a defendant is convicted. Because the Commission data only includes sentenced defendants, this subsection uses the LIONS data on its own.

Using the LIONS data alone has several drawbacks, which is why it is not the default for the rest of this paper. First, the LIONS data does not include any of the outcome variables used in this paper except the defendant's sentence length. Second, the LIONS data lacks most of the demographic information used as control variables in the paper, such as the defendant's race and ethnicity, level of education, age, and criminal history. And although LIONS includes a variable that indicates the defendant's gender, it contains many missing values. Third, the LIONS data does not appear to be quality-checked for completeness and accuracy the way the Commission data has been.

To create the LIONS-only subsample I apply the same restrictions to the data that I used when constructing the sample used in the paper, as described in subsection 4.1. The LIONS-only sample includes 89,117 defendants, of whom 83,790 are sentenced. This translates into a 6.0 percent charged-but-not-convicted rate, which is very similar to Bureau of Justice Statistics reporting that in 2014, 6.4 percent of federal felony defendants were not ultimately convicted (Motivans 2017). Appendix Table A.2 looks for in-group favoritism on the margin of whether a person is ultimately sentenced (i.e., enters the data used in the paper). The outcome variable is a binary variable that equals one if a defendant's charges were acquitted, declined or dismissed and zero if the

³⁷ In particular, if prosecutors display gender-based favoritism in the decision to decline or dismiss cases, the paper's estimates of gender favoritism would understate the phenomenon. On the other hand, if there is in-group disfavoritism in the decision to decline or dismiss cases, the paper's estimates could overstate the magnitude.

defendant was convicted and sentenced. The coefficient estimates on the interaction term are statistically insignificant in all specifications. The point estimates on the interaction term, although statistically insignificant, are all greater than zero, which suggests that if anything, gender-based favoritism on the margin of selection into the sample would likely attenuate (rather than bolster) the findings of favoritism later in the case that are documented in this paper.

5. Results

5.1. Raw Data

The defendant's base offense level is the primary dependent variable used in the analysis. The base offense level is a natural number that captures the severity of the charges of which the defendant was convicted. It ranges from 0 (least serious) to 43 (most serious).³⁸ This paper focuses on the defendant's base offense level because it reflects the severity of the charged offense before any adjustments occur as the case proceeds.³⁹ I therefore consider it to be the variable in the data that most closely quantifies the prosecutorial charging decision.⁴⁰

Increasing by one base offense level corresponds with a roughly ten percent increase in the Guidelines-recommended sentencing range. Figure 2 presents a histogram of the base offense level variable in the data. Due to the way the Guidelines manual assigns base offense levels to offenses, base offense levels that are even numbers are more prevalent than odd numbers, which Figure 2 illustrates.

Before turning to regression estimates, Figure 3 illustrates the average base offense level for the four possible gender pairings of prosecutor and defendant relative to the average across all defendants within the same courthouse. Cases prosecuted by male prosecutors are solid columns and cases prosecuted by female prosecutors are striped

³⁸ In the data, eight out of the 131,462 defendants have base offense levels that are above 43 (ranging from 44-47). I assume these are erroneous entries.

³⁹ In federal criminal cases, a defendant's sentencing advisory Guidelines range is determined based on two factors: their final offense level and their criminal history score. The final offense level is computed by adjusting the base offense level to reflect the defendant's offense conduct and their behavior during the case. For example, under U.S. Sentencing Guideline 3E1.1, a defendant can earn a reduction to their base offense level by quickly pleading guilty.

⁴⁰ More precisely, the base offense level will capture charging severity at sentencing—which will reflect any bargaining between prosecutor and defendant. The results presented in the main analysis in subsection 4.3, however, are robust to alternative measures of charging severity. These results are presented in Appendix Table A.3.

columns. Figure 3 suggests gender-based favoritism: female defendants earn lower base offense levels (i.e. lesser charges) when assigned a female rather than male prosecutor, while male defendants earn a lower base offense level when assigned a male rather than female prosecutor. When all defendants are pooled, the overall difference between male and female prosecutors are smaller than when one compares the averages within each defendant gender.

5.2. Main Results

As described in the previous section, the base offense level is the primary outcome variable used in the analysis. Table 3 presents results of estimating equation (1) with ordinary least squares regression. Standard errors are clustered at the prosecutor level to account for likely intra-prosecutor correlation in the error term (Bertrand, Duflo, and Mullainathan 2004).

The coefficient on the interaction term—in the top row of Table 3—represents gender-based favoritism. In the regression with demographic and case controls, year fixed effects, and courthouse-section fixed effects presented in column (4), the coefficient estimate of the interaction term is -0.492 and is statistically significant at the one-percent level. When prosecutor fixed effects are included in column (5), the prosecutor gender variable is not identified and the coefficient estimate of the interaction term is -0.456, which is statistically significant at the one-percent level. This result can be interpreted as the differential effect of prosecutor gender on male defendants relative to female defendants. In the remaining analysis in this paper, I use the specification with individual prosecutor fixed effects (column (5)) when estimating results in the full sample. When estimating results in subsample of the data, I use the specification with courthouse-section fixed effects (column (4)) to maximize statistical power.

Consistent with in-group favoritism, the third row of Table 3 demonstrates that prosecutor gender is a significant predictor of charging severity when the defendant is female (columns (1)-(4)). However, in column (6), which does not include the interaction term, the coefficient on prosecutor gender is small and is statistically insignificant—indicating that prosecutor gender does not influence prosecutorial behavior on average.

Appendix Table A.3 reports similar results when equation (1) is estimated with alternative measures of charging severity.⁴¹ The first such alternative is the defendant’s statutory minimum across all counts of conviction. While the statutory minimum reflects charging severity untethered from judicial interference, it is a coarse measure. Roughly two-thirds of defendants have no statutory minimum, and some of these defendants are still sentenced to life in prison (suggesting that they committed serious offenses despite not facing a mandatory minimum). The second alternative measure of charging severity is the leave-one-out average sentence for all defendants within the same courthouse who were sentenced under the same Guideline as the defendant, excluding the defendant themselves. This is likely a noisy measure of charging severity given that sentences are highly dependent on other factors (such as criminal history). Both of these alternative measures of charging severity, however, produce estimates of gender-based favoritism that are statistically significant.

5.3. Additional Outcome Variables

The prior analysis uses the defendant’s base offense level as the outcome of interest. While the base offense level plausibly captures charging severity, it is not a case *outcome*—it is simply an input into a defendant’s ultimate sentence. Table 4 presents results of estimating equation (1) with nine additional outcome variables. I organize additional outcome variables into two categories. The first are outcomes that are largely a function of prosecutorial discretion. As explained above, I believe the defendant’s base offense level almost purely captures a prosecutorial decision.⁴² Other variables of this nature include, for example, whether the government recommended a substantial

⁴¹ Some might object that the base offense level is an imperfect measure of charging severity because it can include adjustments for facts that are found by the court. This concern is largely mitigated by the fact that the paper uses the value of the base offense level *before* Chapter Two adjustments for specific offense characteristics. The Commission data reports both the base offense level (inclusive of Chapter Two adjustments) and the Chapter Two adjustments themselves, so it is straightforward to measure the base offense level prior to Chapter Two adjustments, which is the outcome variable used in Table 3.

⁴² To be more precise, the base offense level will also reflect any bargaining over the charges between the defendant and the prosecutor. Subsection 6.4 further discusses the bargaining dynamics between prosecutors and defendants. The Justice Manual (formerly the U.S. Attorneys Manual), however, attempts to limit charge-bargaining by instructing prosecutors to “charge and pursue the most serious, readily provable offenses,” which are defined as “those that carry the most substantial guidelines sentence, including mandatory minimum sentences.” Justice Manual 9-27.300. The Manual does not completely foreclose charge bargaining, making clear that “there will be circumstances in which good judgment would lead a prosecutor to conclude that a strict application of the above charging policy is not warranted.” *Id.* In such a situation, the individual AUSA must get approval from a supervisor to bring a lesser charge. *Id.*

assistance reduction. The second type of outcome variable is one that is a function of both judicial and prosecutorial discretion. The defendant's ultimate sentence is an example of such an outcome: a prosecutor can influence the sentence through the charging decision, the application of enhancements and reductions, and their own recommendation, but, ultimately, the district judge sentences the defendant.

Table 4 presents regression results estimating equation (1) with nine additional outcome variables. Panel A considers early, prosecutorial outcomes—variables that capture decisions that occur prior to sentencing and over which the prosecutor exercises large influence. These early outcomes include: (1) the final offense level; (2) whether the defendant was charged with a mandatory minimum;⁴³ (3) the defendant's final Guidelines Range (represented by its mean, transformed by the inverse hyperbolic sine); and (4) whether the defendant received a substantial assistance reduction. Panel B considers outcomes over which the prosecutor's influence is more limited: (5) whether the defendant was released pending trial; (6) the defendant's sentence in months (transformed by its inverse hyperbolic sine); (7) same as (6) but with the base offense level added as a control variable; (8) the defendant's sentence as a fraction of their mean Guidelines range; and (9) whether the defendant received a sentence below their recommended Guidelines range. If gender-based favoritism shapes prosecutorial decision-making, one would expect coefficient estimates of the interaction term to be significant for outcome variables that are under the prosecutor's control, but insignificant for outcome variables that are largely determined by the defendant's judge.

As predicted, the results in Panel A suggests that gender match—represented by the coefficient on the interaction term—is associated with reductions in: the defendant's final offense level, the likelihood of facing a mandatory minimum, and the defendant's Guidelines range. Defendant-prosecutor gender match does not appear to affect the probability that a defendant receives a substantial assistance reduction—the coefficient estimate is very close to zero and statistically insignificant. This finding suggests that favoritism is concentrated at the beginning of the case, perhaps because there is less

⁴³ This variable equals one if the defendant's statutory minimum was greater than 12 months.

involvement from other actors (like defense counsel) or because early decisions are made with less information about the defendant.⁴⁴

Panel B, which considers outcomes in which a judge exercises decision-making authority, finds that the prosecutor-defendant gender match plays a smaller role. Sentence length reported in column (6)—which is a function of both prosecutorial and judicial discretion—is the only outcome that appears to be influenced by gender match. The coefficient estimate on the interaction term, -0.077, suggests that female defendants can expect to receive an eight percent shorter sentence relative to male defendants when they are assigned a female prosecutor. This finding suggests that judicial decision-making does not entirely undo in-group favoritism in charging.⁴⁵

This sentencing reduction appears to enter through the charging channel: when the defendant's base offense level is included as a control variable in column (7), the point estimate is attenuated and not statistically significant. Also as predicted, estimates of gender-based favoritism are statistically insignificant predictors of whether the defendant received a sentence below the recommended Guidelines range—an outcome that is largely within the judge's purview. Nor is gender match a statistically significant predictor of whether a defendant is released pending trial—a decision that the prosecutor can influence, but that is ultimately decided by a judge.

Finally, it is worth also highlighting that being a male defendant is significantly predictive of worse outcomes on every measure, consistent with the findings in Starr (2015) of large gender disparities at all stages of a criminal case. Relative to female defendants, a male defendant suffers: an increased likelihood of a mandatory minimum charge, a reduced probability of receiving a substantial assistance reduction, a reduced probability of being released pending trial, a longer sentence, and a reduced probability of being sentenced below the recommended Guidelines range. The estimated gender gap

⁴⁴ For example, Marianne Bertrand and Esther Duflo explain, “Implicit biases are more likely to drive behavior under conditions of ambiguity, high time pressures and cognitive loads, or inattentiveness to the task.” (Bertrand and Duflo 2017).

⁴⁵ This explanation is consistent with Joshua Fischman and Max Schanzenbach's finding that judicial discretion does not contribute to racial disparities and might mitigate racial disparity (Fischman and Schanzenbach 2012).

in sentence length reported in Table 4 is extremely close in magnitude to the analogous estimate in Starr (2015).⁴⁶

5.4. Robustness Check: Prosecutor and Defendant Characteristics Uncorrelated

The extent to which prosecutor gender is correlated with defendant gender (and other covariates) is likely to be heterogeneous across USAOs and USAO sections. In particular, some USAOs or USAO sections might assign cases using random assignment, or in ways that more closely approximate random assignment, than others. In this subsection, I examine the results in courthouse-sections in which prosecutor gender is uncorrelated with defendant gender and other defendant characteristics that are fixed at the time of case assignment.

To carry out this robustness check, I regress prosecutor gender on each of five defendant characteristics: age, and indicators for being black, Hispanic ethnicity, male, and a high school graduate. I do a separate regression for each characteristic in each courthouse-section. All regressions include year fixed effects. I then store the p-values associated with the coefficient estimate of the relationship between each defendant characteristic and prosecutor gender for each courthouse-section.

Appendix Table A.4 presents regression results when the data is restricted to courthouse-sections in which prosecutor gender is not related to *any* of the defendant characteristics at the one-, five-, and ten-percent levels. I find that the main regression results are highly robust to and virtually unchanged by these restrictions, even in the specification that removes courthouse-sections in which the coefficient estimate of prosecutor gender is significant at the ten-percent level for *any* of the five defendant characteristics (column (7)), a restriction that meaningfully reduces the size of the sample. These results suggest that the findings presented in the main analysis are not driven by non-random assignment of cases to prosecutors. The even-numbered columns report results without interaction terms and show that the coefficient estimate on the prosecutor gender variable (representing average gender differences between male and female prosecutors) is small and statistically insignificant in all specifications.

⁴⁶ The coefficient on defendant gender in the regression using the inverse hyperbolic sine of sentence length as a dependent variable is 0.781, which is equivalent to a roughly 54 percent increase in sentence length for male defendants relative to female defendants. This point estimate is very close to the finding in Starr (2015) of a roughly 60 percent gender gap in sentence length in federal criminal cases.

6. Heterogeneity

The previous section demonstrates that prosecutors appear to charge defendants more leniently when the defendant is the same gender as the prosecutor. In theory, there are many underlying mechanisms that could generate this finding. This section broadly considers two. First, gender-based leniency might derive from social preferences, or, taste-based discrimination (Becker 1957). I call this the *preference-based* explanation for the findings. Second, the prosecutorial process might generate gender-based leniency. Prosecutors might more effectively prosecute defendants with whom they have more in common. The paper refers to mechanisms in this category as *process-based* explanations for the findings. For example, if prosecutors are better able to evaluate the risk of recidivism among defendants of their gender than opposite-gender defendants, differential treatment could constitute statistical discrimination, along the lines of Cornell and Welch (1996). Conversely, defendant response to their prosecutor's gender could drive the results if, for example, defendants are more likely to trust and therefore cooperate with an own-gender prosecutor.

This section presents evidence of heterogeneity in the results. Heterogeneity is not only descriptively interesting in its own right, it also has the potential to shed light on the mechanisms responsible for in-group favoritism, which is the focus of this section. In particular, this section evaluates whether preference- or process-based mechanisms seem likely to be responsible for gender-based favoritism by examining how, if at all, the results are heterogeneous across many dimensions, including geography and sexism (subsection 5.1); offense type (subsection 5.2); and prosecutorial team composition (subsection 5.3). These subsample comparisons help discern possible explanations for the gender-based favoritism documented in Section 4. Subsection 5.4 further probes the process-based explanation by examining whether favoritism affects the defendant's level of cooperation, operationalized in several different ways.

Broadly speaking, I find substantially more support for the preference-based explanation of gender-based favoritism. As described in more detail below, the preference-based explanation is bolstered by findings that in-group favoritism is stronger in states with below-median measures of sexist attitudes (compared to states with above-median sexism), and in cases in which gender is more salient. I also find that favoritism

is stronger in same-gender prosecutorial teams (compared to mixed-gender teams or solo prosecutors), but I argue that depending on the nature of group decision-making, this result could be consistent with either preference- or process-based explanations (or both). I fail, however, to find any concrete evidence to support the process-based explanation: defendant-prosecutor gender match is not a significant predictor of whether the defendant receives a substantial assistance reduction, a safety valve reduction, an acceptance of responsibility reduction, or takes their case to trial.

6.1 Geography, Sexism, and Gender-Based Favoritism

In the United States, there is significant regional variation in the prevalence of sexist attitudes, and a locality's level of background sexism could affect the extent to which prosecutors demonstrate gender-based leniency, particularly if such favoritism is preference-based. On the other hand, because many aspects of the federal criminal process are standardized across geographic regions, one might not expect gender-based favoritism to vary across geographic regions if such favoritism is purely process-based.

A new working paper by Kerwin Kofi Charles, Jonathan Guryan, and Jessica Pan quantifies regional variation in sexism across the United States, and I exploit this state-level variation to explore the possibility that gender-based leniency in charging is related to the intensity of sexist attitudes in the locality.⁴⁷ I test whether the role of prosecutor-defendant gender match is moderated by the background level of sexism in the state in which the district is situated. It is worth emphasizing that sexism can manifest in many different ways. For example, sexism can be either conscious or implicit. No matter its form, however, one threshold prediction—which I verify below—is that female defendants will be treated more harshly relative to male defendants in states with higher levels of sexism than in states with less prevalent sexism.

In Table 5, I divide states in the sample by whether they exhibit either above- or below-median sexism according to Charles, Guryan, and Pan (2018).⁴⁸ As above, the defendant's base offense level is the outcome variable. As a preliminary matter, women

⁴⁷ Charles, Guryan, and Pan (2018) are interested in the labor market and familial status effects of sexism. They find that background sexism affects women's wages, labor force participation, marriage age, and childbearing.

⁴⁸ Six states and territories—Washington DC, Guam, Hawaii, Idaho, Nevada, and Puerto Rico—are removed from the sample because Charles, Guryan, and Pan (2019) do not report sexism scores for these states and territories.

appear to be charged more harshly relative to men in states with above-median sexism (row 5 in column (7)), providing validation that the sexism variable captures negative attitudes toward women.

The results in Table 5 also provide several important clues about the results presented in Section 4. First, as columns (1) and (2) demonstrate, gender-based favoritism in the data is largely driven by differential treatment of female defendants. For female defendants, the increased charging severity associated with having a male prosecutor (0.341 in the base offense level) is nearly three times larger than the decreased charging severity for male defendants assigned a male prosecutor (-0.137). In other words, female defendants earn a more important benefit from being assigned a gender-matching prosecutor than male defendants do.

This finding could be explained by the fact that gender is likely to be more salient for female than male defendants. Around 83 percent of criminal defendants are men, so a defendant's gender is likely to be more noteworthy when the defendant is female. Typically, in-group preferences are more pronounced for in-groups that have higher salience. (Everett, Faber, and Crockett 2015; McLeish and Oxoby 2011).

Second, although matching with one's prosecutor on gender is important for female defendants, it is significantly *more* important in states with below-median sexism scores than in states with more prevalent sexism (columns (3)-(7)). This suggests that in-group favoritism might work against, rather than in tandem with, sexism. Overall, the results seem to reject the idea that sexism cultivates in-group favoritism. Instead, the results are consistent with gender favoritism that is generated by positive feelings toward women.

It is plausible that inter-group empathy bias is responsible for the findings. As others have shown, people are more likely to feel empathy towards in-group members (Cikara et. al 2011). And, perhaps most relevant to the prosecutorial setting,⁴⁹ in experimental settings, people are more likely to *help* in-group than out-group members (Everett, Faber, and Crockett 2015), and to feel pain and empathy when observing the *pain* of an in-group member compared to an out-group member (Xu et al. 2009, Gutsell

⁴⁹ Prosecutors are both witnesses to and agents of the pain felt and costs imposed on criminal defendants when they are prosecuted, convicted, and ultimately sentenced.

and Inzlicht 2012).⁵⁰ Others have found that people with higher implicit bias tend to have reduced empathy in response to out-group pain (Avenanti, Sirigu, & Algiuoti 2010). If sexism reduces in-group favoritism, it might break these empathies.

6.2 Heterogeneity by Offense Type

This subsection considers heterogeneity by offense type. Cases in the data are categorized into five discrete offense types: drug offenses, firearm offenses, regulatory and other offenses, violent offenses, and white collar offenses. A little over half of offenders are prosecuted for drug-related offenses. The least common offense type is violent crime, which accounts for only six percent of cases. If gender-based favoritism is preference-based, I predict that it will be the most intense among offense types for which gender is most salient.

Table 6 presents results of estimating equation (1) by offense type. The estimates of in-group favoritism are the strongest among firearm offenders. Among firearm offenders, only five percent are women—the lowest female representation in any offense type. Women are also rare among violent offenders, constituting just nine percent of defendants in this offense category. The estimate of gender-based favoritism for violent crime is noisily estimated due to the small number of federal defendants prosecuted for violent crimes, although the point estimate also suggests, if anything, stronger than average in-group favoritism. Gender might be especially salient for women in firearm and violent crime cases not only because female defendants are rarely charged in such cases, but also due to prevailing social norms that women do not engage in violent behavior (Russell 2014).

Conversely, the estimates of in-group favoritism are weakest for white collar offenders, in which one-third of defendants are female—the highest female representation in any offense type. These findings are consistent with the findings above that salience appears to impact the intensity of gender-based favoritism exhibited by prosecutors.

⁵⁰ While people also experience the opposite—pleasure in response to an out-group member’s adversity (a phenomenon called *schadenfreude*)—a review of the literature concluded that “positive social preferences of in-group love may play a stronger role than negative social preferences for outgroup derogation.” (Everett, Faber, and Crockett 2015).

The results in Table 6 also suggest that gender favoritism is heightened among more serious offenders, presenting a conflict with prior work that examines the labor market concerns of federal prosecutors. As described in section 2, Boylan and Long (2005) and Boylan (2005) find that better career outcomes in the private legal market are associated with more aggressive tactics by federal prosecutors—longer prison sentences and increased probabilities of taking a case to trial. Much labor literature suggests that male prosecutors are more likely to benefit from these future, private-sector rewards than female prosecutors (Goldin 2014, Wiswall and Zafar 2018). Yet, male prosecutors appear to show leniency in cases involving male defendants—who are, on average, significantly more serious offenders. This unexpected finding is consistent with a preference-based explanation because such leniency by male prosecutors seems to work against what one would predict would be in their self-interest.

An alternative explanation is that gender-based favoritism could be more intense for defendants of color, and these defendants tend to be over-represented in firearm and drug cases relative to regulatory and white collar cases. It is possible, of course, that the extent to which prosecutors behave with gender-based favoritism could vary depending on the defendant’s race and ethnicity. For example, Starr (2015) finds that gender disparity in federal criminal cases is significantly larger among Black than White defendants.

To examine this possibility more directly, Appendix Table A.5 presents results of estimating equation (1) by defendant race and ethnicity. The coefficient estimates of in-group favoritism are statistically significant at conventional levels in all racial groups except for Hispanic ethnicity defendants, who comprise a small portion of the sample. The point estimates are more extreme in the subsamples including minority defendants. However, three-way interactions between prosecutor gender, defendant gender, and defendant race or ethnicity are not statistically significant. I therefore cannot rule out the possibility of equal gender-based favoritism across racial groups.

6.3 Prosecution in Teams

So far the analysis has only considered the interaction between one defendant and one prosecutor—the first lead prosecutor assigned to each case. However, around one-third of the cases in the data involve more than one lead prosecutor. This subsection

compares levels of gender-based favoritism between cases prosecuted by solo prosecutors, mixed-gender prosecutorial teams, and same-gender prosecutorial teams. In this subsection, I first test whether prosecutors exhibit peer effects and, after finding that they do, attempt to untangle what this finding suggests about the mechanisms responsible for gender-based leniency.

To carry out this analysis, I compare estimates of in-group favoritism in three kinds of cases: (a) those prosecuted by solo prosecutors (63 percent of observations); (b) those prosecuted by same-gender teams of prosecutors (21 percent of observations); and (c) those prosecuted by mixed-gender teams of prosecutors (16 percent of observations). Although I use the word *team* to refer to the situation in which a case in the data includes more than one lead prosecutor, it is likely that at least some of these cases involve instances of reassignment, such as if an AUSA resigns from their job or takes temporary leave from work.

Table 7 presents results of estimating equation (1) among defendants prosecuted by solo prosecutors, those prosecuted by a same-gender team of prosecutors, and those prosecuted by a mixed-gender team of prosecutors. For reference, column (1) reports results from the full sample—it is identical to column (4) in Table 3. The results in Table 7 suggest that peer effects influence charging behavior. The estimates among same-gender teams of prosecutors (reported in column (3)) represent the effect of switching from an all-female team of prosecutors to an all-male team. The estimate of in-group favoritism (-1.144 in the base offense level) is statistically significant, and more than three times the magnitude of the estimate among solo prosecutors in column (2) (-0.365). Column (4) represents the effect of switching from an *any*-female team to an all-male team, and suggests that the presence of just one female prosecutor in the group matters (the point estimate is -0.759 and statistically significant), but does not generate as much favoritism as moving to an all-female team.

In contrast, the estimates of in-group favoritism among mixed-gender teams of prosecutors (reported in column (5)) are not statistically significant. In column (5), the point estimate is negative and the subsample includes many fewer observations than the full sample, suggesting that even these mixed-gender teams might not completely mitigate gender-based favoritism. Overall, however, the results are consistent with the

peer effects hypothesis that mixed-gender teams moderate the effects of favoritism, while same-gender teams heighten it.

Without knowing more about the group dynamics of prosecutorial teams, peer effects are consistent with both the preference- and process-based explanations. For example, if prosecutorial teams reach decisions by simply aggregating the preferences of the teams' individual members one would expect to find evidence of peer effects in gender-based favoritism. If gender-based leniency is preference-based, one might also expect same-gender teams to exhibit enhanced leniency towards own-gender defendants, akin to the now well-established finding that ideologically homogeneous panels of federal appellate judges exhibit stronger ideological preferences in group decision-making than heterogeneous panels (Kim 2009; Sunstein, Schkade and Ellman 2004).⁵¹

But peer effects could also be consistent with process-based explanations if one conceptualizes group decision-making by prosecutors as highly *deliberative*. If prosecutors are more skilled at prosecuting defendants with whom they have more in common, other prosecutors in the group might defer to the prosecutor who has the most commonality with the defendant, which would also generate results as in Table 7.

6.4 In-Group Favoritism, Bargaining, and Cooperation

Evidence contained in Table 4 and described in subsection 5.3 suggests that gender-based leniency is concentrated at charging, which itself challenges the process-based hypothesis. This subsection explores the process-based hypothesis more closely by examining whether prosecutor-defendant gender match is related to defendant cooperation or other aspects of the bargaining process.

Table 8 presents results of estimating equation (1) for six outcome variables that are meant to capture the extent to which the defendant places trust in and cooperates with their prosecutor. Columns (1) and(2) reconsider the substantial assistance reduction. A prosecutor can request a sentencing reduction for a defendant who provides substantial assistance to the government.⁵² Column (1) (reproduced from column (4) in Table 4) shows that across the whole defendant population, gender match between

⁵¹ Judges also appear to be influenced by their panel colleagues' genders in deciding cases in which gender is salient. (Boyd, Epstein, and Martin 2010; Farhang and Wawro 2004; Peresie 2005).

⁵² See note 26.

prosecutor and defendant is not a statistically or economically significant predictor of whether the defendant receives a substantial assistance reduction.

The substantial assistance reduction, however, is not equally important for all defendants. Column (3) restricts attention to defendants for whom the substantial assistance reduction is especially valuable—defendants facing a mandatory minimum who are ineligible for safety-valve relief.⁵³ For these defendants, a substantial assistance is especially beneficial because it is the only way that they can receive a sentence below their mandatory minimum. Defendants in this group comprise roughly 23 percent of the sample, but gender match is not a statistically significant predictor of whether the defendant receives a substantial assistance reduction among these defendants.

Column (3) estimates whether gender match is related to whether the defendant receives a safety valve reduction. Critically, to earn a safety valve reduction, a defendant must make a truthful proffer to the government of “all information and evidence the defendant has concerning the offense[s]”⁵⁴—an element that requires the defendant to place significant trust in their prosecutor. Columns (4) through (6) use other dependent variables that seek to capture the extent to which the defendant cooperated with the prosecution by pleading guilty rather than going to trial (column (4)) and by receiving an acceptance of responsibility reduction (columns (5) and (6)).

If gender match between prosecutor and defendant leads to increased cooperation, one would expect the coefficient estimates on the interaction terms to be greater than zero and statistically significant in all columns of Table 8. This is not borne out—none of the coefficient estimates are statistically significant and the confidence intervals suggest little influence of gender match on any of the outcome variables. This evidence thus provides more support for the idea that in-group favoritism is driven by automatic or implicit preferences, rather than process-based explanations.

7. Discussion

This paper uses novel federal criminal case data covering roughly 130,000 criminal defendants sentenced between 2002 and 2016 in 64 of the 94 federal district

⁵³ I define a defendant as likely to be ineligible for a safety-valve reduction if their case does not involve drugs, they are charged with a weapon, or they have more than one criminal history point.

⁵⁴ U.S. SENTENCING GUIDELINES MANUAL § 5C1.2 (2016).

courts in the United States. By exploiting the salutation field of the LIONS data, I obtain a variable that captures prosecutor gender. The paper finds that defendants receive more lenient charges when they match with their prosecutor on gender relative to when they are assigned a prosecutor of the opposite gender. Through this charging channel, prosecutor-defendant pairs that match on gender produce lower sentences relative to prosecutor-defendant pairs that mismatch. On the other hand, being assigned a prosecutor of the same gender does not affect the probability that a defendant will receive a sentence below the recommended Guidelines range—a decision that largely falls to the defendant’s sentencing judge.

By examining estimates in subsamples of the data, I suggest that the findings are more consistent with preference-based favoritism than process-based mechanisms. Specifically, gender-based favoritism is stronger in states with less prevalent sexist attitudes, in cases in which gender is likely to be more salient, and among same-gender prosecutorial teams. I argue that these findings suggest that gender-based leniency derives from prosecutor preferences, and, in particular, could be generated by female prosecutors feeling stronger empathic ties to female defendants than male prosecutors do. I find little evidence to support a process-based explanation for the results: across many alternative measures, defendants do not appear more likely to cooperate or place trust in a prosecutor with whom they match on gender.

This paper thus documents an important and perhaps troubling phenomenon. The U.S. Supreme Court has repeatedly emphasized that a central goal in federal sentencing law is to promote the uniformity of sentences across defendants.⁵⁵ The extent to which defendants are treated differently based on something they cannot control—whether they happen to be assigned a prosecutor of the same gender—conflicts with this essential idea.

Critically here, where there is no clear benchmark as to a defendant’s optimal sentence, one can not know whether the decisions produced by defendant-prosecutor gender matches are superior or inferior to decisions produced by mismatches. There are, however, a few reasons to believe the outcomes produced by gender *match* might produce

⁵⁵ See, e.g., *Booker*, 543 U.S. at 255 (“Congress enacted the sentencing statutes in major part to achieve greater uniformity in sentencing”); *Kimbrough v. United States*, 552 U.S. 85, 107 (2007) (“[I]t is unquestioned that uniformity remains an important goal of sentencing”).

better results. First, some might think that leniency is an inherently desirable outcome. There is a widespread view that the federal criminal system is too punitive. As evidence of this point, for example, nearly half of federal felony defendants are sentenced below their recommended Guidelines range while only two percent are sentenced above.

Second, if in-group leniency derives from empathy, as I suggest in subsection 6.1, gender-favoritism might be desirable because empathy is considered “prosocial,” or, beneficial for society (Cikara et. al 2011). Moreover, the fact that gender-based favoritism is more intense in states with less pervasive sexism also suggests that such favoritism could create desirable outcomes.

Third, it is plausible that prosecutors are incentivized to seek the highest charges possible—both because they are directed to do so by the Justice Manual⁵⁶ and because this might be a metric that is used for advancement. If so, leniency could represent thoughtful decision-making, which ought to be encouraged.

Fourth, in other contexts that examine in-group favoritism where there *is* an objective measure of outcome quality, prior work has found that in-group matches produce better results. For example, in the financial context, prior work has found that hedge fund managers are able to better predict the performance of firms led by CEOs of the manager’s gender (Jannati et al. 2016), and that loans perform better when the lender and borrower are culturally proximate (Fisman, Paravisini, and Vig 2017). Both of these papers suggest that gender match produces objectively better results. Of course, it is not clear that these results generalize to the criminal system.

Unfortunately, the lack of available data hinders understanding these important dynamics in the criminal setting. While scholars have made strides in understanding judicial decision-making thanks to rich and publicly-available data about judges,⁵⁷ there is little publicly-available data for researchers to use in analyzing decision-making by individual prosecutors. With more detailed data about prosecutors, researchers would be able to further explore other kinds of individual bias in prosecutorial decision-making.

⁵⁶ Justice Manual § 9-27-300.

⁵⁷ For example, researchers can easily download a database with detailed biographical information about all past and current federal judges from the Federal Judicial Center website: BIOGRAPHICAL DIRECTORY OF FEDERAL JUDGES: EXPORT, <https://www.fjc.gov/history/judges/biographical-directory-article-iii-federal-judges-export>.

With this suggestion in mind, the remainder of this paper discusses two potential policy changes that prosecutorial offices might implement.

Because gender-matching is associated with leniency, some might be tempted to embrace in-group favoritism and propose policies that promote explicit gender-matching in prosecutorial offices. Implementing an express policy of matching defendants to prosecutors on the basis of gender would undoubtedly raise both constitutional and practical concerns. Moreover, even if one wanted to match defendants to proximate prosecutors in an attempt to promote leniency, without having examined in-group favoritism on many other dimensions—like race—it is not clear that gender is the best variable on which to match.

Another class of solutions might attempt to address outgroup disfavoritism and therefore generate more equal treatment across defendants. Implicit bias training might be promising in this respect. In 2016, the U.S. Department of Justice announced that it would begin providing all federal law enforcement officers and prosecutors with implicit bias training, although President Donald Trump recently ordered a stop to all such trainings. If in-group favoritism stems from preference-based discrimination, one might be tempted to use implicit bias training to reduce discrimination. There is little evidence, however, about the extent to which such training programs can reduce bias in the criminal system and over the long-term. Such programs have some success in simulated settings in the short-term (Plant and Peruche 2005). In the employment setting more generally, however, many scholars have found that anti-bias and diversity trainings can be ineffective, and can even lead to decreased support for diversity among White participants (Plaut et al. 2011).

There are still many important questions about prosecutorial behavior that remain unanswered. For example, examining in-group favoritism on the basis of race is critical. Further work that could examine the scope of prosecutorial discretion and quantify disparities stemming from the assignment of prosecutors to defendants would also be valuable in defining the contours of prosecutorial discretion in the criminal system.

References

- Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies* 347-383.
- Aizer, Anna, and Joseph J., Jr. Doyle. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *Quarterly Journal of Economics* 130 (2): 759-803.
- Antonovics, Kate L., and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91 (1): 163-175.
- Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics* 127 (2): 1017-1055.
- Barkow, Rachel E. 2013. "Prosecutorial Administration: Prosecutor Bias and the Department of Justice." *Virginia Law Review* 99: 271-342.
- Barkow, Rachel E. 2009. "Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law." *Stanford Law Review* 61: 869-921.
- Becker, Gary S. 1957. *The Economics of Discrimination*. University of Chicago Press.
- Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." In *Handbook of Field Experiments*, by Abhijit Banerjee and Esther Duflo. North Holland.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119: 249-.
- Bibas, Stephanos. 2009. "Prosecutorial Regulations versus Prosecutorial Accountability." *University of Pennsylvania Law Review* 157: 959-1016.
- Bjerk, David. 2005. "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing." *The Journal of Law and Economics* 48: 591-625.
- 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.
- Boylan, Richard T. 2005. "What Do Prosecutors Maximize? Evidence from the Careers of U.S. Attorneys." *American Law and Economics Review* 7 (2): 379-402.

- Boylan, Richard T., and Cheryl X. Long. 2005. "Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors." *Journal of Law and Economics* 48 (2).
- Brown, Robert A., and James Frank. 2006. "Race and Officer Decision Making: Examining Differences in Arrest Outcomes between Black and White Officers." *Justice Quarterly* 23 (1): 96-126.
- Charles, Kerwin Kofi, Jonathan Guryan, and Jessica Pan. 2018. "The Effects of Sexism on American Women: The Role of Norms vs. Discrimination." *NBER Working Paper 24904*.
- Chen, Daniel L. 2017. "Implicit Egoism in Sentencing Decisions: First Letter Name Effects with Randomly Assigned Defendants." *TSE Working Paper 16-726*.
- Cikara, Mina, Emile G. Bruneau, and Rebecca R. Saxe. 2011. "Us and Them: Intergroup Failures of Empathy." *Current Directions in Psychological Science* 20 (3): 149-153.
- Cornell, Bradford, and Ivo Welch. 1996. "Culture, Information, and Screening Discrimination." *Journal of Political Economy* 104: 542-.
- Cox, Adam B., and Thomas J. Miles. 2008. "Judging the Voting Rights Act." *Columbia Law Review* 108 (1): 1-54.
- Cross, Frank B. 2003. "Decisionmaking in the U.S. Circuit Courts of Appeals." *California Law Review* 91 (6): 1457-1515.
- Depew, Briggs, Ozkan Eren, and Naci Mocan. 2017. "Judges, Juveniles, and In-Group Bias." *The Journal of Law and Economics* 60 (2): 209-239.
- Di Tella, Rafael, and Ernesto Schargrotsky. 2013. "Criminal Recidivism after Prison and Electronic Monitoring." *Journal of Political Economy* 121 (1): 28-73.
- Didwania, Stephanie Holmes. 2020 (forthcoming). "The Immediate Consequences of Federal Pretrial Detention." *American Law and Economics Review*.
- Doleac, Jennifer L., and Benjamin Hansen. 2017. "The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes when Criminal Histories are Hidden." *Working Paper*. http://jenniferdoleac.com/wp-content/uploads/2015/03/Doleac_Hansen_BanTheBox.pdf.
- Donohue III, John J., and Steven D. Levitt. 2001. "The Impact of Race on Policing and Arrests." *Journal of Law and Economics* 44 (2): 367-394.

- Everett, Jim A.C, Nadira S Faber, and Molly Crockett. 2015. "Preferences and Beliefs in Ingroup Favoritism." *Frontiers in Behavioral Neuroscience* 9 (15): 1-15.
- 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, and Organization* 20 (2): 299-330.
- Fischman, Joshua B., and Max M. Schanzenbach. 2012. "Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums." *Journal of Legal Studies* 9 (2): 729-764.
- Fiske, Susan T., Amy J.C Cuddy, Peter Glick, and Jun Xu. 2002. "A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition." *Journal of Personality and Social Psychology* 82 (6): 878-902.
- Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2017. "Cultural Proximity and Loan Outcomes." *American Economic Review* 107 (2): 457-92.
- Flanagan, Francis X. 2018. "Race, Gender, and Juries: Evidence from North Carolina." *Journal of Law and Economics* 61 (2): 189-214.
- Fryer Jr., Roland G. 2016. "An Empirical Analysis of Racial Differences in Police Use of Force." *NBER Working Paper 22399*.
- Gazal-Ayal, Oren, and Raanan Sulitzeanu-Kenan. 2010. "Let My People Go: Ethnic In-Group Bias in Judicial Decisions-Evidence from a Randomized Natural Experiment." *Journal of Empirical Legal Studies* 7 (3): 403-428.
- Geller, Amanda, Irwin Garfinkel, and Bruce Western. 2011. "Paternal Incarceration and Support for Children in Fragile Families." *Demography* 48 (1): 25-47.
- Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091-1119.
- Gorman, Elizabeth H. 2005. "Gender Stereotypes, Same-Gender Preferences, and Organizational Variation in the Hiring of Women: Evidence from Law Firms." *American Sociological Review* 70 (4): 702-728.
- Greenwald, Anthony G, and Thomas F Pettigrew. 2014. "With Malice Toward None and Charity for Some: Ingroup Favoritism Enables Discrimination." *American Psychologist* 69 (7): 669-84.

- Gutsell, Jennifer N., and Michael Inzlicht. 2012. "Intergroup Differences in the Sharing of Emotive States: Neural Evidence of an Empathy Gap." *Social Cognition and Affective Neuroscience* 7 (5): 596-603.
- Hoyt, Crystal L., Stefanie Simon, and Lindsey Reid. 2009. "Choosing the Best (Wo)Man for the Job: The Effects of Mortality Salience, Sex, and Gender Stereotypes on Leader Evaluations." *The Leadership Quarterly* 20 (2): 233-246.
- Huang, Bert I. 2011. "Lightened Scrutiny." *Harvard Law Review* 124 (5): 1109-1152.
- Jackson, Robert H. 1940-1941. "The Federal Prosecutor." *American Institute of Criminal Law and Criminology* 41: 3-6.
- Jannati, Sima, Alok Kumar, Alexandra Niessen-Ruenzi, and Justin Wolfers. 2016. "In-Group Bias in Financial Markets." *Working Paper*. December 24.
- Kennedy, Jessica, Mary-Hunter McDonnell, and Nicole Stephens. 2017. "Do Women Face a Higher Ethical Bar? Exploring Discrimination in the Punishment of Ethical Violations." *Working Paper*. July 13.
- Kim, Pauline T. 2009. "Deliberation and Strategy on the United States Courts of Appeals: An Empirical Exploration of Panel Effects." *University of Pennsylvania Law Review* 157: 1319-1381.
- Krieger, Linda Hamilton. 1998. "Civil Rights Perestroika: Intergroup Relations After Affirmative Action." *California Law Review* 86: 1251-.
- Kutateladze, Besiki L, Nancy R Andiloro, Brian D Johnson, and Cassia C Spohn. 2014. "Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing." *Criminology* 52 (3): 514-551.
- Kutateladze, Besiki Luka, Andiloro, Nancy R., Johnson, Brian D. 2014. "Opening Pandora's Box: How Does Defendant Race Influence Plea Bargaining?" *Justice Quarterly* 33 (3): 398-426.
- Lantigua-Williams, Juleyka. 2016. "Are Prosecutors the Key to Justice Reform?" *The Atlantic*, May 18. <https://www.theatlantic.com/politics/archive/2016/05/are-prosecutors-the-key-to-justice-reform/483252/>.
- Leonetti, Carrie. 2012. "When the Emperor Has No Clothes III: Personnel Policies and Conflicts of Interest in Prosecutors' Offices." *Cornell Journal of Law and Public Policy* 22: 53-.

- Lewis, Amy C., and Steven J. Sherman. 2010. "Perceived Entitativity and the Black-Sheep Effect: When Will We Denigrate Negative Ingroup Members." *Journal of Social Psychology* 150: 211-.
- Lim, Claire S.H., Bernardo S. Silveria, and James M. Snyder. 2016. "Do Judges' Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts." *American Law and Economics Review* 18 (2): 302-357.
- Lindeman, Marjaana, and Lilli Sundvik. 1995. "Evaluative Bias and Self-Enhancement Among Gender Groups." *European Journal of Social Psychology* 25 (3): 269-280.
- Marques, Jose M., Vincent Y. Yzerbyt, and Jacques-Philippe Leyens. 1988. "The 'Black Sheep Effect': Extremity of Judgments Towards In-group Members as a Function of Group Identification." *European Journal of Social Psychology* 18: 1-.
- McAdams, Richard H. 1998. "Race and Selective Prosecution: Discovering the Pitfalls of Armstrong." *Chicago-Kent Law Review* 73: 605-.
- McLeish, Kendra N., and Robert J. Oxoby. 2011. "Social Interactions and the Salience of Social Identity." *Journal of Economic Psychology* 32 (1): 172-178.
- Meares, Tracy L. 1995. "Rewards for Good Behavior: Influencing Prosecutorial Discretion and Conduct with Financial Incentives." *Fordham Law Review* 64: 851-919.
- Miles, Thomas J. 2014. "Does the 'Community Prosecution' Strategy Reduce Crime? A Test of Chicago's Experience." *American Law and Economics Review* 16 (1): 117-43.
- Motivans, Mark. 2017. "Federal Justice Statistics, 2014 -- Statistical Tables." *Bureau of Justice Statistics*. March 23. Accessed November 21, 2017. <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=5873>.
- Mueller-Smith, Michael. 2015. "The Criminal and Labor Market Impacts of Incarceration." *Working Paper*.
- Ouziel, Lauren M. 2017. "Ambition and Fruition in Federal Criminal Law: A Case Study." *Virginia Law Review* 103: 1077-.
- Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108 (5): 937-975.
- Peresie, Jennifer L. 2005. "Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts." *Yale Law Journal* 114: 1759-.

- Pfaff, John F. 2017. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform*. New York: Basic Books.
- Plant, E. Ashby, and B. Michelle Peruche. 2005. "The Consequences of Race for Police Officers' Responses to Criminal Suspects." *Psychological Science* 16: 180-.
- Plaut, Victoria C., Flannery G. Garnett, Laura E. Buffardi, and Jeffrey Sanchez-Burks. 2011. "What About Me? Perceptions of Exclusion and Whites' Reactions to Multiculturalism." *Journal of Personality and Social Psychology* 101 (2): 337-53.
- Price, Joseph, and Justin Wolfers. 2010. "Racial Discrimination among NBA Referees." *Quarterly Journal of Economics* 1859-1887.
- Raggi, Reena. 1989. "Prosecutors' Offices: Where Gender is Irrelevant." *Fordham Law Review* 57 (6): 975-979.
- Reedt, Lou, Courtney Semisch, and Kevin Blackwell. 2013. *Effective Use of Federal Sentencing Data*. Presentation to the American Society of Criminology, United States Sentencing Commission.
- Rehavi, M. Marit, and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6): 1320-1354.
- Revesz, Richard L. 1997. "Environmental Regulation, Ideology, and the D.C. Circuit." *Virginia Law Review* 83 (8): 1717-1772.
- Robertson, Christopher, Shima Baradaran Baughman, and Megan S. Wright. 2019. "Race and Class: A Randomized Experiment with Prosecutors." *Journal of Empirical Legal Studies* 16 (4): 807-847.
- Rosenblum, Darren. 2000. "'Trapped' in Sing Sing: Transgendered Prisoners Caught in the Gender Binarism." *Michigan Journal of Gender and Law* 6: 499.
- Ross, Jannell. 2016. "A Look Inside the Training the White House Thinks Could Stop Police Shooting, Protest, Damning DOJ Report Cycle." *Washington Post*, August 16. https://www.washingtonpost.com/news/the-fix/wp/2016/08/10/a-look-inside-the-training-the-white-house-thinks-could-stop-the-police-shootingprotestdamning-doj-report-cycle/?utm_term=.dea997f9ddc2.
- Rowe, Brian. 2009. "Gender Bias in the Enforcement of Traffic Laws: Evidence Based on a New Empirical Test." *Unpublished Manuscript*. September.
- Rudman, Laurie A., and Peter Glick. 1999. "Feminized Management and Backlash Toward Agentic Women: The Hidden Costs to Women of a Kinder, Gentler Image

- of Middle Managers." *Journal of Personality and Social Psychology* 77 (5): 1004-1010.
- Rudman, Laurie A., and Stephanie A. Goodwin. 2004. "Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More than Men Like Men?" *Journal of Personality and Social Psychology* 87 (4): 494-509.
- Russell, Brenda L. 2014. *Perceptions of Female Offenders: How Stereotypes and Social Norms Affect Criminal Justice Responses*. New York: Springer.
- Sanga, Sarath. 2014. "Does Officer Race Matter?" *American Law and Economics Review* 16: 403-.
- Schanzenbach, Max M. 2005. "Racial and Gender Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics." *Journal of Legal Studies* 57-92.
- Schanzenbach, Max M. 2015. "Racial Disparities, Judge Characteristics, and Standards of Review in Sentencing." *Journal of Institutional and Theoretical Economics* 27-47.
- Shayo, Moses, and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *The Quarterly Journal of Economics* 126 (3): 1447-1484.
- Sklansky. 2017. "The Nature and Function of Prosecutorial Power." *Journal of Criminal Law and Criminology* 106: 473.
- Sloan, CarlyWill. 2019. "Racial Bias by Prosecutors: Evidence from Random Assignment." *Working Paper*.
- Spohn, Cassia. 1990. "The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities." *Law & Society Review* 24 (5): 1197-1216.
- Starr, Sonja B., and M. Marit Rehavi. 2013. "Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effects of Booker." *Yale Law Journal* 2-80.
- Starr, Sonja. 2015. "Estimating Gender Disparities in Federal Criminal Cases." *American Law and Economics Review* 17 (1): 127-159.
- Steffensmeier, Darrell, and Chris Hebert. 1999. "Women and Men Policymakers: Does the Judge's Gender Affect the Sentencing of Criminal Defendants?" *Social Forces* 77 (3): 1163-1196.

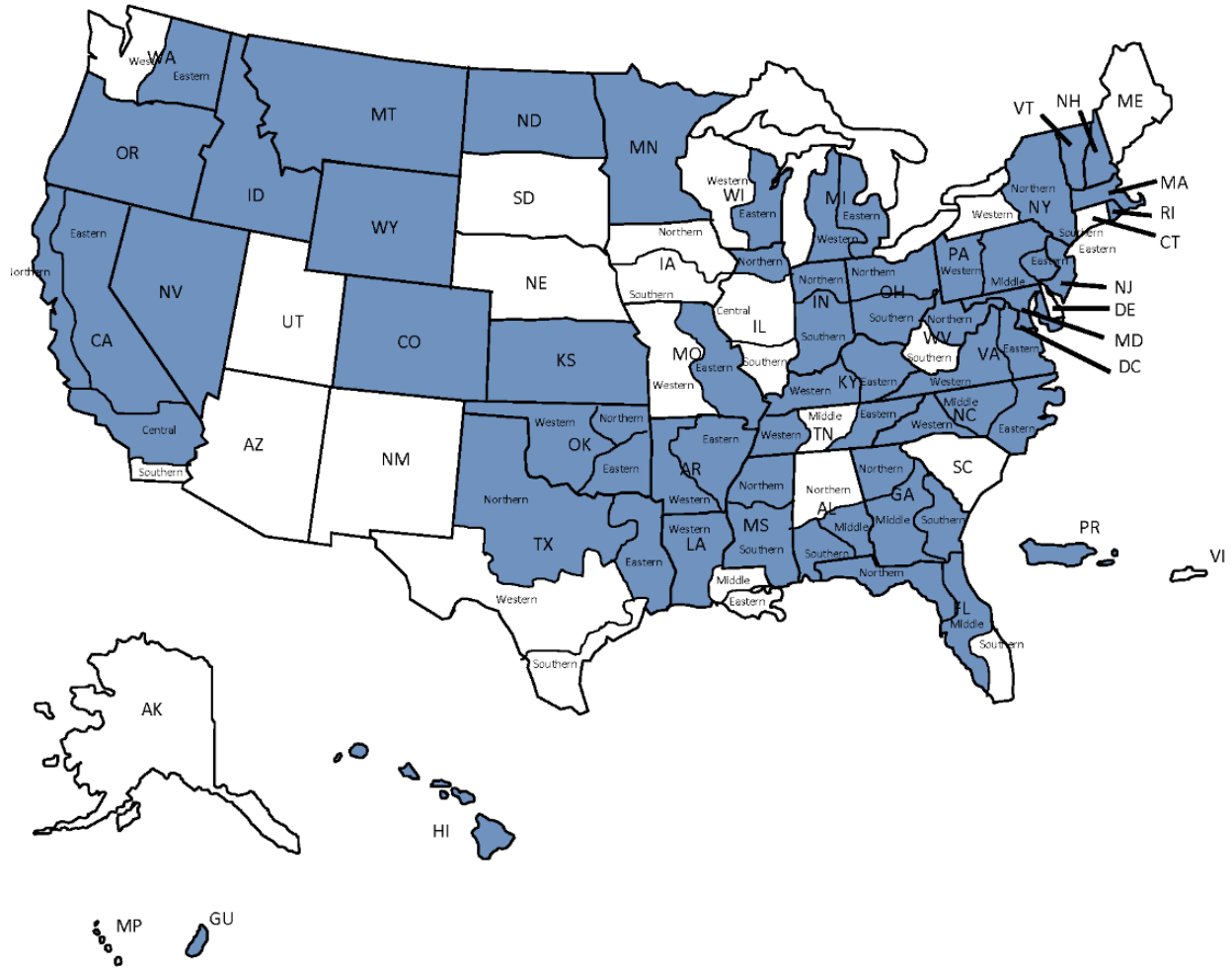
- Sunstein, Cass R., David Schkade, and Lisa Michelle Ellman. 2004. "Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation." *Virginia Law Review* 90 (1): 301-354.
- Uhlman, Thomas M. 1978. "Black Elite Decision Making: The Case of Trial Judges." *American Journal of Political Science* 22 (4): 884-895.
- Vial, Andrea C, Victoria L Brescoll, Jaime L Napier, John F Dovidio, and Tom R Tyler. 2017. "Differential Support for Female Supervisors Among Men and Women." *Journal of Applied Psychology* 103 (2): 215-227.
- Ward, Geoff, Amy Farrell, and Danielle Rousseau. 2009. "Does Racial Balance in Workforce Representation Yield Equal Justice? Race Relations of Sentencing in Federal Court Organizations ." *Law and Society Review* 757-806.
- Welch, Susan, Michael Combs, and John Gruhl. 1988. "Do Black Judges Make a Difference?" *American Journal of Political Science* 32 (1): 126-136.
- Western, Bruce. 2002. "The Impact of Incarceration on Wage Mobility and Inequality." *American Sociological Review* 67 (4): 526-546.
- Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender." *The Quarterly Journal of Economics* 133 (1): 457-507.
- Xu, Xiaojing, Xiangyu Zuo, Xiaoying Wang, and Shihui Han. 2009. "Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses." *The Journal of Neuroscience* 29 (26): 8525-8529.
- Yang, Crystal S. 2015. "Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing." *Journal of Legal Studies* 44 (1): 75-111.
- Yang, Crystal S. 2014. "Have Interjudge Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker." *New York University Law Review* 89 (4): 1268-1342.

Appendix A: Gender Coding

This Appendix lists all of the salutation codes in the LIONS data and indicates how they were coded in the paper. F indicates female; M indicates male; and U indicates unknown.

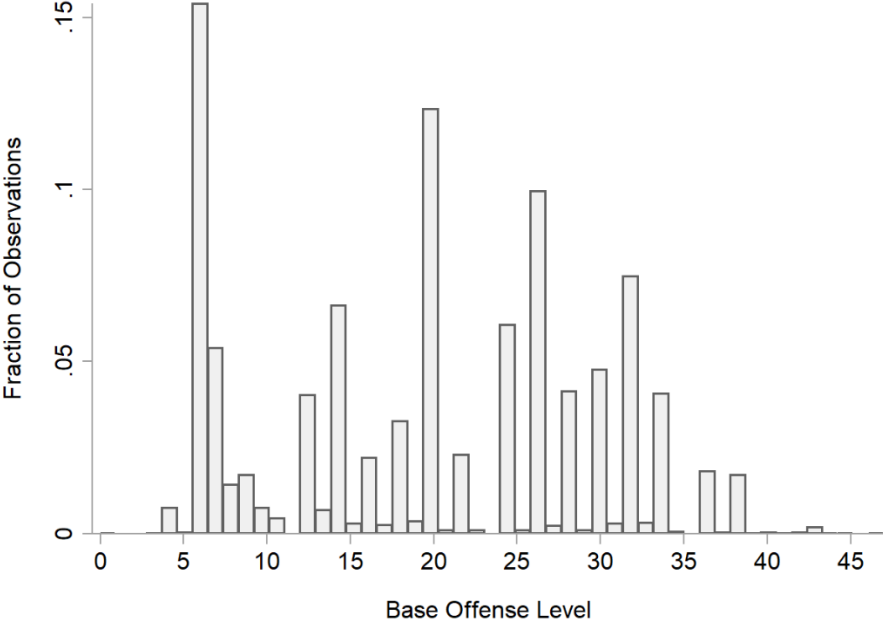
Salutation	Frequency (pct)	Coding	Continued...	Frequency (pct)	Coding
AUS	0.33	U	SAUSA	0.01	U
AUSA	1.98	U	UFAP	0.04	U
Al	0.01	M	USA	0.00	U
Amy	0.01	F	Vaughn	0.02	M
Ausa	0.00	U	a	0.02	U
CAPT	0.00	U	aa	0.00	U
Capt	0.01	U	ch	0.00	U
Capt.	0.02	U	d	0.00	U
Captain	0.02	U	db	0.04	U
Captian	0.00	U	ep	0.00	U
Caption	0.00	U	gh	0.00	U
Cpt	0.01	U	js	0.00	U
Cpt.	0.01	U	kp	0.00	U
Diana	0.00	F	mR.	0.05	M
Edie	0.00	U	mS.	0.00	F
FAUSA	0.01	U	mr	0.16	M
Gentry	0.02	U	mr.	0.01	M
Hello	0.04	U	ms.	0.10	F
Hon.	0.04	U	s	0.05	U
Janet	0.01	F	ss	0.04	U
Jeff	0.02	M	wn	0.08	U
Jim	0.01	M			
Joe	0.00	M			
John	0.02	M			
Lt.	0.00	U			
M	0.02	U			
MISS	0.00	F			
MR	1.34	M			
MR.	2.77	M			
MRS	0.04	F			
MRS.	0.05	F			
MS	0.89	F			
MS.	1.45	F			
Miss	0.07	F			
Mr	4.65	M			
Mr,	0.00	M			
Mr.	59.10	M			
Mr.s	0.00	F			
Mrs	0.13	F			
Mrs.	2.52	F			
Ms	2.10	F			
Ms,	0.00	F			
Ms.	21.71	F			
Ns.	0.00	U			
Richard	0.00	M			

Figure 1. Federal Judicial Districts Appearing in the Data



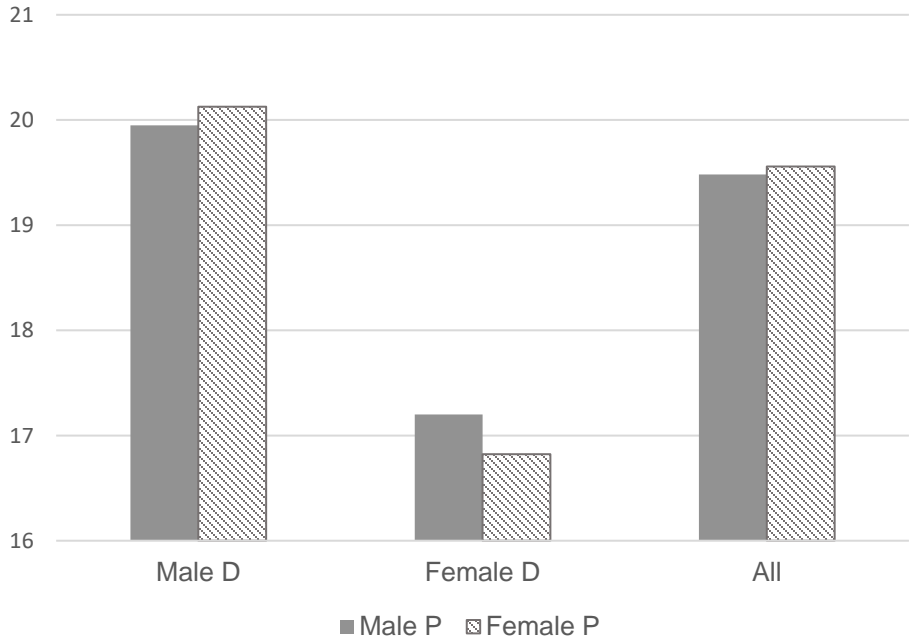
Notes: Shaded districts are those that are represented in the paper.

Figure 2. Histogram of Base Offense Level



Notes: This figure plots the distribution of base offense levels in the data ($N=131,462$).

Figure 3. Base Offense Level Averages by Defendant and Prosecutor Gender



Notes: Average base offense level de-meaned by courthouse.

Table 1. Summary Statistics

	Female Defendants	Male Defendants	All Defendants	Min	Max
Defendant Characteristics					
Male	-	-	0.833	0	1
Black	0.321	0.440	0.420	0	1
Hispanic	0.100	0.124	0.120	0	1
White	0.521	0.392	0.414	0	1
Another Race	0.059	0.045	0.047	0	1
Less than High School	0.240	0.325	0.311	0	1
High School Only	0.378	0.421	0.413	0	1
Some College	0.306	0.192	0.211	0	1
College Graduate	0.076	0.062	0.065	0	1
Age (years)	37.0	35.9	36.1	16	92
Number of Dependents	1.329	1.506	1.479	0	28
Criminal History Points	1.90	4.831	4.34	0	82
Case Characteristics					
Lead Male Prosecutor	0.716	0.726	0.724	0	1
Defendants	3.62	3.41	3.43	1	62
Prosecutors	1.65	1.71	1.70	1	15
Base Offense Level	16.0	20.7	19.9	0	47
Final Offense Level	17.0	21.5	20.8	1	43
Case Start Year	2007.3	2007.2	2007.2	1999	2016
Sentencing Year	2008.8	2008.6	2008.7	2001	2016
Intermediate Outcomes					
Mandatory Minimum (0/1)	0.225	0.366	0.342	0	1
Substantial Assistance Reduction	0.265	0.218	0.225	0	1
Mean Guidelines Range	49.4	93.7	86.3	0	470
Case Outcomes					
Any Incarceration	0.700	0.902	0.868	0	1
Sentence (months) ⁵⁸	27.5	68.3	61.5	0	470
Sentence / Mean Guidelines Range	0.466	0.713	0.672	0	16
Below-Guidelines (0/1)	0.607	0.478	0.499	0	1
Above-Guidelines (0/1)	0.013	0.024	0.022	0	1
In-Range (0/1)	0.380	0.498	0.478	0	1
Observations	21,995	109,467	131,462	131,462	131,462

⁵⁸ Not including alternative confinement, such as house arrest. Sentences are capped at 470 months—the Commission’s value for life sentences.

Table 2. Regressing Male Prosecutor on Case Characteristics

	All Cases (1)	All Cases (2)	Female Def Cases (3)	Male Def Cases (4)
Male (fraction)	0.019** (0.008)	0.021** (0.005)	-	-
Black (fraction)	-0.035** (0.015)	-0.0003 (0.007)	-0.007 (0.012)	-0.0001 (0.006)
Hispanic (fraction)	-0.072*** (0.023)	-0.009 (0.008)	-0.021 (0.017)	-0.008 (0.008)
Another Race (fraction)	-0.058 (0.044)	-0.011 (0.014)	-0.049* (0.027)	-0.011 (0.013)
Age (years)	-0.0003 (0.0003)	-0.0002 (0.0002)	0.00006 (0.0004)	-0.0002 (0.0002)
Dependents	-0.002 (0.002)	0.0003 (0.0008)	-0.0010 (0.003)	0.00006 (0.0008)
HS Only (fraction)	-0.002 (0.006)	-0.003 (0.003)	-0.005 (0.010)	-0.003 (0.003)
Some College (fraction)	0.010 (0.007)	0.007 (0.005)	0.004 (0.012)	0.007 (0.005)
College Grad (fraction)	0.019* (0.011)	0.012 (0.008)	0.002 (0.016)	0.014* (0.009)
Defendants	0.0009 (0.002)	-0.0005 (0.001)	-0.00009 (0.012)	0.0003 (0.003)
Prosecutors	-0.015 (0.008)	-0.005 (0.004)	-0.002 (0.009)	-0.007 (0.004)
Prosecutor Workload (cases/yr)	-0.0005 (0.001)	-0.0002 (0.0008)	-0.0004 (0.009)	-0.0002 (0.0008)
White Collar Offense	0.015 (0.028)	-0.0008 (0.009)	0.001 (0.023)	0.005 (0.009)
Drug Offense	0.026 (0.027)	-0.005 (0.009)	0.006 (0.024)	-0.005 (0.009)
Regulatory Offense	-0.007 (0.024)	-0.012 (0.009)	-0.002 (0.024)	-0.012 (0.009)
Violent Offense	-0.004 (0.028)	-0.008 (0.010)	0.029 (0.030)	-0.008 (0.010)
F-Stat: Crim Hist Category	0.79 (p=0.56)	0.44 (p=0.82)	0.78 (p=.57)	0.49 (p=0.78)
Month and Year FEs		X	X	X
Courthouse-Section FEs		X	X	X
Observations	94,395	94,395	12,245	79,929

OLS regressions of the gender of the defendant's first lead prosecutor (1=male; 0=female) on case characteristics. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the courthouse-section level. Month and year fixed effects include sentencing year, Guidelines version used, and month and year of case initiation. Omitted racial category is white, omitted education category is less than a high school completion, and omitted offense type is firearm offenses.

Table 3. Main Results (Base Offense Level)

	(1)	(2)	(3)	(4)	(5)	(6)
Male D * Male P	-0.691** (0.345)	-0.510*** (0.141)	-0.505*** (0.141)	-0.492*** (0.136)	-0.456*** (0.136)	-
Male Defendant	5.249*** (0.295)	1.697*** (0.123)	1.656*** (0.123)	1.661*** (0.119)	1.611*** (0.117)	1.309*** (0.062)
Male Prosecutor	0.571 (0.594)	0.387** (0.194)	0.390** (0.194)	0.329** (0.163)	-	-0.079 (0.096)
Demographic Controls		X	X	X	X	X
Month and Year FEs			X	X	X	X
Courthouse Section FEs				X	X	-
Prosecutor FEs						X
Observations	131,462	131,462	131,462	131,462	131,462	131,462

OLS regressions of the defendant's base offense level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the prosecutor level. Demographic controls include race, Hispanic ethnicity, age at sentencing, number of dependents, prosecutor annual caseload, the number of defendants, the number of prosecutors, and indicators for the defendant's educational attainment and criminal history category. Month and year fixed effects include sentencing year, Guidelines version used, and month and year of case initiation.

Table 4. Additional Outcomes

Panel A: Prosecutor-Based Outcomes					
	Final Offense Level (1)	Facing MM (0/1) (2)	GL Range (3)	Sub. Assist. (0/1) (4)	
Male D *	-0.302**	-0.016**	-0.043**	-0.0004	
Male P	(0.143)	(0.008)	(0.020)	(0.009)	
Male Defendant	2.088*** (0.125)	0.075*** (0.007)	0.312*** (0.017)	-0.055*** (0.007)	
Observations	131,462	131,324	131,461	131,462	
Panel B: Judge- and Prosecutor-Based Outcomes					
	Released Pretrial (0/1) (5)	Sentence (IHS) (6)	Sentence (IHS) (control for BOL) (7)	Sentence (% of GL Mean) (8)	Below-GL Sentence (0/1) (9)
Male D *	0.010	-0.077**	-0.027	-0.003	-0.006
Male P	(0.008)	(0.039)	(0.034)	(0.009)	(0.010)
Male Defendant	-0.124*** (0.007)	0.781*** (0.032)	0.606*** (0.029)	0.151*** (0.008)	-0.088*** (0.008)
Observations	128,022	131,462	131,462	131,460	131,461

OLS regressions. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and prosecutor fixed effects.

Table 5. Gender Favoritism and Sexism

	Male Defs	Female Defs	Male Defs	Female Defs	Low Sexism States	High Sexism States	All Defs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Male Defendant	-	-	-	-	2.126*** (0.191)	1.238*** (0.195)	2.041*** (0.169)
Male Prosecutor	-0.137 (0.127)	0.341* (0.186)	-0.017 (0.158)	0.816*** (0.224)	0.623** (0.241)	0.125 (0.270)	0.883*** (0.241)
High Sexism State (0/1)	-	-	-0.075 (0.208)	0.476 (0.302)	-	-	0.563* (0.326)
Male D * Male P	-	-	-	-	-0.876*** (0.200)	-0.234 (0.176)	-0.911*** (0.198)
Male D * High Sexism	-	-	-	-	-	-	-0.661*** (0.243)
Male P * High Sexism	-	-	-0.231 (0.251)	-0.897** (0.357)	-	-	-0.890** (0.385)
Male D * Male P * High Sexism	-	-	-	-	-	-	0.669** (0.287)
Observations	103,511	20,850	103,511	20,850	60,291	64,070	124,361

OLS regressions of the the defendant's base offense level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects.

Table 6. Gender Favoritism by Offense Type

	Drugs	Firearms	Regulatory	Violent	White Collar
<i>Dependent Variable: Base Offense Level</i>					
Male D * Male P	-0.486** (0.245)	-1.286*** (0.364)	-0.442* (0.259)	-0.724 (0.618)	-0.191 (0.135)
Male Defendant	1.402*** (0.248)	2.862*** (0.261)	2.486*** (0.241)	1.143** (0.533)	0.911*** (0.107)
Male Prosecutor	0.189 (0.309)	1.347*** (0.387)	0.028 (0.278)	0.817 (0.676)	0.397*** (0.117)
Male Defendants (share)	0.86	0.95	0.76	0.91	0.66
Male Prosecutors (share)	0.73	0.72	0.71	0.71	0.73
Observations	66,608	20,009	14,871	8,279	21,695

OLS regressions of the the defendant's base offense level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the courthouse-section level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and courthouse-section fixed effects.

Table 7. Gender Favoritism in Prosecutorial Teams

	All Cases	Solo Prosecutors	Multiple Ps of Same Gender	Multiple Ps: No Female	Ps of Mixed Genders
	(1)	(2)	(3)	(4)	(5)
Male D * Male P	-0.492*** (0.136)	-0.365** (0.154)	-1.144*** (0.301)	-0.759*** (0.182)	-0.267 (0.298)
Male Defendant	1.661*** (0.119)	1.431*** (0.134)	2.279*** (0.274)	1.924*** (0.135)	1.955*** (0.223)
Male Prosecutor	0.329** (0.163)	0.174 (0.178)	0.928*** (0.313)	0.573** (0.195)	0.222 (0.340)
Observations	131,462	83,047	27,888	48,415	20,538

OLS regressions of the the defendant's base offense level. ***: p<0.01; **: p<0.05; *: p<0.10. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and courthouse-section fixed effects.

Table 8. Alternative Measures of Cooperation

	Substantial Assistance	Sub Assist: MM & Safety Inelig.	Safety Valve	Pled Guilty (No Trial)	Any Acceptance of Responsibility	Max Acceptance of Responsibility
	(1)	(2)	(3)	(4)	(5)	(6)
Male D * Male P	-0.0004 (0.009)	0.010 (0.023)	-0.009 (0.010)	-0.0002 (0.003)	0.001 (0.005)	0.0005 (0.005)
Male Defendant	-0.055*** (0.007)	-0.157*** (0.019)	-0.064*** (0.009)	-0.015*** (0.003)	-0.012*** (0.004)	-0.018*** (0.004)
Observations	131,462	33,435	66,053	131,461	131,462	131,462

OLS regressions. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and prosecutor fixed effects. Column (2) includes defendants who were facing a mandatory minimum greater than twelve months and are ineligible for a safety-valve reduction, either because their cases did not involve a drug offense, they were charged with a weapon, or they had more than one criminal history point.

Table A.1. Representativeness of the Sample

	Data in this paper (1)	Matched Defendants (3)	Commission Defendants (3)
Defendant Characteristics			
Male	0.83	0.83	0.84
Black	0.42	0.39	0.39
Hispanic	0.12	0.12	0.13
White	0.41	0.44	0.43
Another Race	0.05	0.05	0.05
Less than High School	0.31	0.29	0.30
High School Only	0.41	0.39	0.39
Some College	0.21	0.23	0.22
College Graduate	0.07	0.08	0.08
Age (years)	36.1	36.9	36.7
Number of Dependents	1.48	1.4	1.4
Criminal History Points ⁵⁹	4.3	4.0	4.0
Case Characteristics			
Lead Male Prosecutor Defendants	0.72	-	-
Prosecutors	3.43	-	-
Base Offense Level	1.70	-	-
Final Offense Level	19.9	19.1	19.3
Case Start Year	20.8	20.6	20.6
Sentencing Year	2007	-	-
	2008	2009	2009
Intermediate Outcomes			
Facing Mandatory Minimum	0.35	0.33	0.34
Substantial Assistance Reduction	0.23	0.20	0.20
Mean Guidelines Range	86.3	84.6	87.6
Case Outcomes			
Any Incarceration	0.87	0.84	0.83
Sentence (months) ⁶⁰	61.5	60.8	62.9
Sentence / Mean Guidelines Range	0.67	0.65	0.64
Below-Guidelines (0/1)	0.50	0.50	0.49
Above-Guidelines (0/1)	0.02	0.02	0.02
In-Range (0/1)	0.48	0.48	0.48
Observations	131,462	409,078	533,757

⁵⁹ Total criminal history points are unadjusted.

⁶⁰ Not including alternative confinement, such as house arrest. Sentences were capped at 470 months—the Commission's value for life sentences.

Table A.2. Gender Favoritism and Selection into the Sample

	(1)	(2)	(3)	(4)
<i>Dependent Variable: Acquittal/Dismissal/Declination</i>				
Interaction	0.003 (0.006)	0.004 (0.006)	0.002 (0.006)	-
Male Defendant	-0.014** (0.006)	-0.015** (0.006)	-0.015** (0.006)	-0.013*** (0.003)
Male Prosecutor	-0.002 (0.008)	-0.003 (0.008)	-0.006 (0.007)	-0.004 (0.003)
Demographic Controls	X	X	X	X
Month and Year FEs		X	X	X
Courthouse-Section FEs			X	X
Observations	89,117	89,117	87,183	87,183

OLS regressions of whether a defendant's case was acquitted, declined, or dismissed (0/1). ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the courthouse-section level. Month and year fixed effects include sentencing year and month and year of case initiation. Demographic controls are indicators for whether the defendant is male, whether the case involved multiple defendants, whether the case had multiple prosecutors, and offense type.

Table A.3. Alternative Measures of Charging Severity

	Base Offense Level (1)	Stat Min (IHS) (2)	Average Sentence (IHS) (3)
Male D * Male P	-0.456*** (0.136)	-0.080** (0.040)	-0.033** (0.016)
Male Defendant	1.611*** (0.117)	0.390*** (0.034)	0.124*** (0.014)
Observations	131,462	131,324	129,606

OLS regressions. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and prosecutor fixed effects. The statutory minimum is the total statutory minimum prison term for all counts of conviction transformed by its inverse hyperbolic sine.⁶¹ The average sentence is the leave-one-out-mean sentence for defendants convicted under the same Chapter Two Guideline in the defendant's courthouse, transformed by its inverse hyperbolic sine.

⁶¹ The statutory minimum is capped at 470 months; 410 defendants had statutory minimums above this value.

Table A.4. Courthouse-Sections in which Prosecutor Gender Uncorrelated with Five Characteristics

	Full Sample (1)	p>0.01 (2)	p>0.05 (3)	p>0.10 (4)
Male D * Male P	-0.456*** (0.136)	-0.692*** (0.173)	-0.643*** (0.206)	-0.482* (0.265)
Male Defendant	1.611*** (0.117)	1.693*** (0.144)	1.595*** (0.166)	1.642*** (0.202)
Observations	131,462	72,609	43,584	22,430

OLS regressions of the defendant's base offense level. ***: p<0.01; **: p<0.05; *: p<0.10. Standard errors are clustered at the prosecutor level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and prosecutor fixed effects.

Table A.5. Results by Defendant Race/Ethnicity

	All Cases	Black Defendants	Non-Black Defendants	Hispanic Defendants	Non-Hispanic Defendants	White Defendants	Non-White Defendants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Dependent Variable: Base Offense Level</i>							
Male D * Male P	-0.292* (0.174)	-0.538*** (0.199)	-0.379** (0.159)	-0.353 (0.369)	-0.437*** (0.140)	-0.337* (0.174)	-0.472** (0.186)
Male Defendant	0.946*** (0.182)	2.252*** (0.189)	1.281*** (0.178)	1.966*** (0.382)	1.521*** (0.152)	1.069*** (0.179)	2.153*** (0.179)
Male Prosecutor	0.056 (0.247)	0.359* (0.201)	0.234 (0.206)	0.430 (0.371)	0.245 (0.190)	0.170 (0.233)	0.342* (0.189)
Male D * Male P * Black	-0.227 (0.257)	-	-	-	-	-	-
Male D * Male P * Hispanic	-0.077 (0.407)	-	-	-	-	-	-
Observations	125,357	55,316	70,041	15,695	109,662	54,346	71,011

OLS regressions of the the defendant's base offense level. ***: p<0.01; **: p<0.05; *: p<0.10. Standard errors are clustered at the courthouse-section level. All regressions include demographic controls listed in Table 3, month and year fixed effects, and courthouse-section fixed effects. Regression coefficients for race, race*prosecutor gender and race*defendant gender are unreported but included in the regression reported in column (1).