The Effect of Own-Gender Jurors on Conviction Rates

Mark Hoekstra * Brittany Street†

February 2021

Abstract

Despite concerns about gender bias in general and jury gender in particular, little is known about the effect of jury gender on conviction rates. We identify the effect of own-gender jurors by exploiting random variation in the assignment to and ordering of jury pools in two large Florida counties. Results indicate that own-gender jurors are significantly less likely to convict on drug charges, though we find no evidence of effects for other charges. Estimates indicate that adding one own-gender juror ($\sim$1.6 standard deviations) results in a 30 percentage point reduction in conviction rates on drug charges, which is highly significant even after adjusting for multiple comparisons. These findings highlight how drawing an opposite-gender jury can impose significant costs on defendants, and demonstrate that own-gender bias can occur even in settings where the importance of being impartial is actively pressed on participants.

JEL Classification: K41, K42, J16
Keywords: Crime, Jury Trials, Discrimination, Gender Bias

*Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843. Email: markhoekstra@tamu.edu.
†Department of Economics, University of Missouri, 118 Professional Bldg, Columbia, MO 65211. Email:streetb@missouri.edu.
We would like to thank Hillsborough and Palm Beach County Clerk's office for providing the data and officials in both counties for answering endless questions about how the jury assignment and selection processes work. In particular, we would like to especially thank Angela Gary, Keith Buss, and Tara Ramos. We would also like to thank Abigail Peralta, CarlyWill Sloan, and Meredee Tangvatcharapong, and conference participants at the Endogenous Rules, Risk and Legal Institutions Workshop and the annual meetings of the Association for Public Policy Analysis and Management, the Western Economic Association International, and the Southern Economic Association for helpful comments and suggestions. Any errors are our own.

1. Introduction

A central right of the accused in the U.S. criminal justice system is the right to a trial before an impartial jury. This right is enshrined in the 6th amendment of the Bill of Rights to the U.S. Constitution, and was inherited from the Magna Carta, which guaranteed that no man be punished without "the lawful judgment of his peers." There are ongoing concerns, however, about the actual impartiality of juries in general, and whether jurors favor those similar to themselves in particular. These concerns have resulted in court rulings that prohibit excluding potential jurors on the basis of race, ethnicity, or sex (Batson v. Kentucky 1986; J.E.B. v. Alabama 1994). However, while recent research has documented bias in favor of own-race defendants (Anwar, Bayer and Hjalmarsson 2012; Flanagan 2018), there is little evidence on whether modern juries favor own-gender defendants. The purpose of this paper is to test whether juries composed of more own-gender jurors affect criminal conviction rates and sentencing outcomes.

The primary difficulty in doing so is that seated juries are the outcome of a non-random jury selection process over which prosecutors, defense attorneys, and jurors have significant influence. As a result, it is difficult to distinguish the effect of own-gender jurors from confounding factors, such as defense attorney quality, that lead some cases to have more jurors of the same gender as the defendant. To overcome this selection problem, we use an approach similar to that of Anwar, Bayer and Hjalmarsson (2012), who examined the effect of at least one black juror being randomly assigned to the jury panel. Our analysis differs in that we are the first to exploit the random *ordering* of potential jurors within a given pool. This enables us to compute the expected number of women on the seated jury, since the seated jury consists of the first six ordered jurors who are not excluded by either a challenge for cause or a peremptory challenge (that is, a challenge for which no reason must be given). The advantage of this approach is that it accounts for the fact that some prospective jurors are much more likely to be seated than others.[1] We use this source of variation in jury composition to identify own-gender effects by differencing out the impact of defendant and jury gender, similar to studies on the impact of race (for example, Price and Wolfers 2010; Shayo and Zussman 2011; West 2017; Hoekstra and Sloan 2020).

To implement this approach, we use a new data set on juror characteristics and conviction and sentencing outcomes for Palm Beach and Hillsborough counties, which are the third and fourth most populous counties in Florida. These data include all felony and misdemeanor trials over a two year period, and contain detailed information on defendant characteristics as well as case characteristics measured at both the charge and case levels. The data also include the randomly-assigned ordering of

[1] For example, in our data the probability of being seated for the first potential juror in a panel of 30 is 38%, compared to only 2% for the individual in the 30th spot.

each potential juror within the jury pool. We use this ordering and the empirical probabilities that jurors assigned a given number are seated on the jury to predict the expected number of women on each jury. We also show this variation is uncorrelated with defendant and case characteristics, and with expected conviction rates of male and female defendants as predicted using exogenous characteristics.

Overall, we find no evidence that juror gender affects conviction rates or incarceration on average. Estimates rule out effects one-third the magnitude of a similar change in expected jury race, as estimated by Anwar, Bayer and Hjalmarsson (2012). In contrast, results provide strong evidence that own-gender jurors affect conviction rates for drug offenses. Estimates indicate that increasing the expected number of own-gender jurors by one (out of six) reduces conviction rates on drug charges by 30 percentage points. The effect seems largely driven by differential treatment of female defendants. Importantly, the effect is significant at the five percent level even after performing the multiple inference adjustment proposed by Anderson (2008). We find no evidence of effects for driving, property, or violent crime offenses. In addition, we show some evidence that the same change in jury gender leads to a 22 percentage point reduction in the likelihood of incarceration.

To our knowledge, this is the first paper to use random variation in jury gender to examine effects on convictions in modern criminal courts. In doing so, the paper contributes to two literatures. The first is the broad literature examining gender bias in education, labor, housing, and product markets (Abrevaya and Hamermesh 2012; Ayres and Siegelman 1995; Bagues and Esteve-Volart 2010; Bagues et al. 2017; Breda and Ly 2015; Dahl and Moretti 2008; De Paola and Scoppa 2015; Goldin and Rouse 2000; Lavy 2008; Neumark et al. 1996; Moss-Racusin et al. 2012). More specifically, this paper complements a growing literature that documents and explains gender differences in sentencing (Bindler and Hjalmarsson 2020; Butcher, Park and Piehl 2017). It also relates to a broader literature on the impact of judge gender (Johnson 2014; Knepper 2018; Schanzenbach 2005; Steffensmeier and Hebert 1999) and other judge and jury characteristics (Anwar et al. 2014; 2015; Mitchell et al. 2005; Cohen and Yang 2019; Depew et al. 2017; George 2001). Finally, in documenting how defendants who draw opposite-gender juries are more likely to be convicted and sentenced, this paper also complements recent papers documenting unfairness in conviction and sentencing based on other factors (Eren and Mocan 2018; Philippe and Ouss 2017).

This study is most closely related to a smaller body of research examining the impact of judge and jury characteristics on criminal trial outcomes. It is most similar to Anwar, Bayer and Hjalmarsson (2012) and Flanagan (2018), who show that having more own-race jurors in the jury panel affects felony conviction rates. It is also related to Anwar, Bayer and Hjalmarsson (2019), who show that while the introduction of women on English juries in 1919 had no effect on overall conviction rates, it resulted in

additional convictions for sex offenses and for violent crime cases with female versus male victims.[2] This paper differs from Anwar, Bayer and Hjalmarsson (2012) and Flanagan (2018) in that we focus on jury gender, rather than jury race. In addition, they study only felonies, while we observe both felony and misdemeanor charges. This paper differs from Anwar, Bayer and Hjalmarsson (2019) in that we focus on the effect of jury gender in a modern context in which effects might well be significantly different than in 1919. Our paper also differs from these three papers in that we also examine effects on sentencing, in addition to convictions. Finally, an additional contribution of our study to this literature is to compare our results to those that use average jury panel characteristics, independent of juror order, to predict the composition of the seated jury. We replicate this approach in results shown in the online appendix, which are similar to those that use the empirical probabilities of being seated conditional on order. While the similarity of the results surprised us, the upshot is that the less data-intensive approach used by Anwar, Bayer and Hjalmarsson (2012) and Flanagan (2018) is likely sufficient for future work.

The results of this study have important implications. First, they suggest that even in settings where participants are actively reminded of the importance and necessity of being fair and impartial, sizable gender biases can still occur. In addition, we note there is evidence that higher conviction and incarceration rates can lead to increased recidivism and worsened labor market outcomes (Aizer and Doyle Jr 2015; Mueller-Smith 2015). This suggests that differential treatment by juries on the basis of gender can lead to significant long-run differences in defendant outcomes.

2. Background and Data

2.1. *The Assignment of Potential Jurors to the Jury Pool and the Voir Dire Process*

As described above, a critical feature of our research design is the random assignment of residents to panels of potential jurors, and the random ordering of residents within each panel. In the Florida counties we study, county court offices randomly mail jury summons to residents who have a driver's license or identification card. Potential jurors arrive at the courthouse on the assigned day and enter their information into a computer system. Each potential juror is then randomly assigned to a case. In addition, within each case each potential juror is assigned a number. Jurors are then escorted into the courtroom by number and seated in order for the voir dire, or jury questioning, process.

As described by U.S. Supreme Court Justice Rehnquist, "Voir dire examination serves to protect [the right to an impartial jury] by exposing possible biases, both known and unknown, on the part of

[2]While we focus on the effect of own-gender jurors in this paper, we also examine the effect of female versus male jurors on overall conviction rates. Results are shown in Online Appendix Table OA2, in which we regress an indicator for conviction on our measure of expected number of women on the jury. Overall, we find no evidence that additional female jurors are more or less likely to convict overall.

the jurors. Demonstrated bias in the responses to questions on voir dire may result in a juror's being excused for cause; hints of bias not sufficient to warrant challenge for cause may assist parties in exercising their peremptory challenges" (McDonough Power Equipment, Inc. v. Greenwood 1984, page 464). Prosecutors and defense attorneys are allowed unlimited challenges for cause, though meeting the requirements for removing a potential juror is difficult, and such requests are not always granted by the judge. In Hillsborough and Palm Beach Counties, each side is typically allowed up to three peremptory challenges to remove jurors they believe unlikely to be favorable toward their side of the case. The final jury thus consists of the first six or twelve jurors not struck by either side, beginning with the potential juror assigned number one. Any remaining potential jurors are then excused or returned to jury services to be reassigned.

Jury verdicts are required to be unanimous in these counties. While there are several ways in which one could model how own-gender jurors impact outcomes when unanimous decisions are required, perhaps the simplest is to assume some fraction of potential jurors favor own-gender defendants. Consequently, in a model such as that proposed by Flanagan (2015) where anything less than a unanimous decision to convict is equivalent to an acquittal, increasing the number of own-gender jurors will result in more acquittals, on average.

2.2. *Data*

We obtained detailed administrative data for all misdemeanor and felony cases that were assigned potential juror pools in preparation for trial in Palm Beach and Hillsborough Counties from 2014 to 2016.[3] These are the third and fourth largest counties in Florida, respectively, each with a 2016 population of over 1.3 million people.[4] Importantly, these data include comprehensive information on the voir dire process along with case attributes. Specifically, we observe the pool of jurors randomly assigned to each case including name, seat number, and outcome of the selection process.[5]

Data from Hillsborough County also include the gender of potential jurors, as well as date of birth, race, and address. For Palm Beach County, we infer gender on the basis of the first name. We do so using an online application programming interface called genderize.io. The application predicts gender based on first name using a large dataset comprised of user profiles from several major social networks. Using this approach, we are able to predict probabilistic genders for 93% of potential jurors. For the names that we do not predict, we assign 0.5 to the female gender indicator variable under the assumption that the missing name is equally likely to be male or female. To verify the accuracy of this

[3]There are 32 cases in Palm Beach County and 1 case in Hillsborough County where there should be a jury panel but the information was not in the case file. Only two of these cases involve drug related charges.

[4]We acquired the data by contacting the five largest counties in Florida. We were unable to obtain records from Miami-Dade, Broward, and Orange County.

[5]We do not, however, have consistent data on why or by whom individual prospective jurors were struck.

approach to inferring gender, we compare predicted gender to actual observed gender in Hillsborough County, and find that we accurately predict gender 98.6% of the time. We then combine potential juror order and the gender of each potential juror to predict the number of women we would expect to serve, on average, for each trial.

From these data we construct our measure of expected jury gender. Specifically, we compute empirical probabilities of being seated on the jury for each spot in the order in the jury panel using a leave-one-out approach.[6] We do so based on the size of the potential jury pool and the number of jurors being seated.[7] Importantly, jury panel size and number of jurors are decided prior to the assignment of the jury pool, and thus should not affect the internal validity of our approach.

These probabilities are shown for panels of varying size in Figure 1.[8] For example, for jury trials with a panel size of 20 or less, the probability of being selected for the jury is around 40 percent for the first 10 or so potential jurors, and then declines to around 20 percent for the 20th-ordered potential juror. To predict the number of women that will be seated on the jury, we interact the estimated probabilities shown in Figure 1 with a gender indicator variable equal to one for females.[9] Summing this over the pool of potential jurors gives the expected number of females seated.[10] The distributions of expected and actual number of women seated on the jury are shown in Online Appendix Figure OA6.

In Section 3.1 we demonstrate that this measure of expected jury panel gender is highly correlated with the actual number of women on the jury. In addition, in Section 3.2 we show that the expected number of women on the jury is uncorrelated with case and charge characteristics, which is consistent with random assignment to panels and random ordering within panels.

For each case in the data, we observe the charges brought against the defendant and the outcome of each charge including verdict and sentencing. Our primary outcome of interest is an indicator for whether the defendant is convicted of the charge. In addition, as a robustness exercise we also perform the analysis at the case level, for which the outcome is defined as the proportion of guilty charges in the case. Importantly, our data include guilty and innocent verdicts issued for all cases for which a jury panel was assigned in preparation for trial. For example, we observe guilty pleas that arise after the jury pool was assigned as well as verdicts found by the jury. This precludes the possibility of selection

[6] In some cases, a second panel of potential jurors was used. Our understanding is this sometimes occurred because the first panel did not result in enough seated jurors, and sometimes because the judge chose not use the first panel at all for some reason. However, we still observe the first (and subsequent) juror panels in those cases, and we order the jurors accordingly. For example, if each of the first two panels had 50 potential jurors, we assign number 51 to the first ordered juror in the second panel, and number 100 to the last juror in that second panel. We do so even if no jurors from the first panel were seated on the jury.

[7] Standard juries in Florida consist of six jurors, though capital crimes consist of 12 jurors. We exclude all charges from these capital cases in our main analysis; previous versions of this paper include these charges and find similar effects.

[8] Each line is fit with a local linear polynomial at each panelist position using an epanechnikov kernel with varying Rule-of-Thumb (ROT) bandwidths. Figure 1, from smallest to largest panel size, uses one-sided bandwidths of 1, 1, 2, 2, and 10, respectively.

[9] The probability of a potential juror being female is used for panels in Palm Beach County.

[10] In addition, in a parallel analysis provided in the online appendix, we also show results using the proportion of the entire jury panel that is female as the measure of expected jury gender. See Online Appendix Figures OA1, OA2, OA3, and OA5, as well as Tables OA3, OA5, and OA6.

bias caused by some cases settling after the prosecutor or defense attorney observes the expected or actual composition of the jury. In addition, we note that for some charges in Florida, a verdict can be given in which adjudication is withheld. In that case the defendant is assigned a term of probation, and upon successful completion of that term is spared a conviction on his or her record. This is the outcome in only 3.75 percent of all charges in our sample, and only 4.28 percent of drug charges. For the main analysis we treat this outcome as guilty, though in Table 5 we show that estimates are similar if we instead classify it as not guilty.

Our second outcome of interest is whether a defendant is sentenced to be incarcerated upon the conclusion of the trial. We define this outcome at the case level, rather than charge level, since the sentences of individual charges are often served concurrently.

Finally, we note that because the purpose of this paper is to examine the effect of own-gender jurors, we exclude charges in which fewer than 10 percent of defendants are female. Consequently, we only consider charges related to a drug, driving, property, or violent crime. In addition, we limit violent crimes to domestic crimes, assaults, and robberies. This is due to the low number of female defendants in other violent crime categories, such as sexual assault and murder, which gives us little variation in defendant gender. In Online Appendix Table OA10 we show results are robust to the inclusion of these cases and to alternative classifications.

Summary statistics are shown at the charge level in Table 6, while Online Appendix Table OA1 shows characteristics at the case level. In total, we have data on 2,937 separate charges representing 1,481 cases. This is roughly twice as many cases as in Anwar, Bayer and Hjalmarsson (2012), who had 785 cases. Sixty-six percent of defendants are convicted of at least one charge, while men are convicted at somewhat higher rates than women (67 versus 63 percent).[11] Sixteen percent of our defendants are female, 48 percent are white, and the average age is 37.

3. Methods

In order to identify the effects of own-gender jurors, we use a generalized difference-in-differences approach that focuses on the interaction of jury and defendant gender.[12] Specifically, we estimate the following linear probability model:

$$\begin{aligned} Convict_{ct} =& \beta_1 DefFemale_t + \beta_2 E(NumFemale)_t + \beta_3 DefFemaleXE(NumFemale)_t \\ &+ X_t + County_t + CountyXCrime_{ct} + \varepsilon_{ct} \end{aligned} \tag{1}$$

[11] We assign a sentence of zero days to those defendants who are sentenced to time served.

[12] We note there could also be gender bias as it relates to the gender of the victim, as reported by Anwar, Bayer and Hjalmarsson (2019). We are unable to test for this type of gender bias because we do not observe victim gender.

where the outcome of interest $Convict_{ct}$ is a binary variable equal to one if the defendant is convicted guilty of charge c in trial t. $DefFemale_t$ an indicator variable equal to 1 if the defendant in trial t is female, controls for differences in conviction based on defendant gender. Similarly, $E(NumFemale)_t$ is the expected number of females seated on the jury for trial t. The coefficient of interest, β_3, measures the effect of own-gender jurors on the outcome. X_t is the set of control variables at the trial level including defendant's age and race, the total number of charges against the defendant, if the case involves a violent charge, the predicted age of the jury pool, and judge fixed effects. All specifications include county fixed effects along with county-by-crime fixed effects when considering more than one crime category.[13] Observations are weighted by the inverse of the total number of charges in a trial.

Robust standard errors are clustered at the defendant level to allow for errors to be correlated across charges and trials for a given defendant. In addition, for our main estimates of interest, we also report two-sided p-values from randomization inference. Specifically, we essentially replicate the data generating process used by the county in assigning prospective jurors by randomly re-assigning juror gender to each ordered slot for each case, assuming 50% of potential jurors are female and 50% are male. We then re-estimate our main results. We repeat this process 10,000 times, the result of which gives us an empirical distribution of t-statistics observed due to chance. We report the fraction of these 10,000 t-statistics that are more extreme than the absolute value of the t-statistic from our actual result, which we interpret as a two-sided p-value.

We also report False Discovery Rate (FDR) adjusted Q-values to adjust for the fact we test for effects by crime severity (felony vs. misdemeanor), and by crime type (drug, driving, property, and violent). These are computed using the method proposed by Anderson (2008), and adjust for the fact that we examine effects on conviction for six different categories of crime. These are interpreted similarly to p-values from a two-tailed test, and explicitly adjust for the increased likelihood of estimating extreme coefficients when making multiple comparisons.

We perform our main analysis of convictions at the charge level rather than the case level for several reasons. The first is that juries decide guilt on a per-charge basis, rather than at the case level, and so it seems sensible for the analysis to do the same. Second, there are several ways one can assign guilt at the case level, and we would prefer to avoid making decisions about which is best. These include whether there is at least one conviction in the case, the proportion of charges that result in conviction, the proportion of charges that result in conviction weighted by severity of the charge, or whether all charges are found guilty. Third, defining subgroups of crimes at the case level is somewhat arbitrary. For example, should a drug case be defined as one in which there is at least one drug charge, or where a drug charge is the most serious charge, or where all charges in the case are drug-related? For these

[13] Charges are grouped into four crime categories, namely drug, driving, property, or violent.

reasons, our analysis is performed at the charge level, though in Online Appendix Tables OA7 and OA8 we show similar results when looking at the proportion of charges that result in conviction.

The intuition of this approach is to compare the difference in how male and female defendants are judged by less-female juries to the difference in how male and female defendants are judged by more-female juries. This approach allows more-female juries to convict at different rates than more-male juries, so long as this difference is constant across male and female defendants in the absence of bias. Equivalently, we allow male defendants to be "more guilty" than female defendants, though we require this difference in underlying guilt to be similar for more-male and more-female juries.

The identifying assumption of this approach is that while male defendants may have different underlying likelihood of conviction than female defendants, in the absence of a treatment effect the difference in their conviction rates should be the same for more-male juries as for more-female juries. This assumption could be violated in a couple of different ways. The first is if our measure of jury gender is correlated with other factors that affect conviction rates. For example, if skilled defense attorneys are able to strike opposite-gender jurors at higher-than-average rates, then we might observe lower conviction rates when there are more same-sex jurors and falsely attribute it to own-gender jurors. To overcome this problem, we construct a measure of expected jury gender composition that is based on the random assignment of individuals to jury pools and the random ordering of individuals within the jury pool. We show that this measure of jury gender is both strongly correlated with the composition of the seated jury, and is orthogonal to other observed determinants of conviction rates such as defendant and case characteristics. We also show that the difference in the guilt propensity of male and female defendants, as predicted using all exogenous characteristics, does not vary with the gender composition of the jury.

The second way in which the identification assumption can fail is if female jurors tend to be more likely to convict defendants of certain crimes (or when certain other crimes are also being charged), and if those crimes are disproportionately committed by certain genders. For example, if women are more likely to convict on a theft charge when a violent crime was also committed at the same time, and if male defendants are more likely than female defendants to be charged with both theft and violent crime, this approach could overstate the effect of own-gender jurors. Similarly, if women are more likely than men to convict black defendants, and if there is a higher proportion of black male defendants than black female defendants, then our estimates could be biased. To address this possibility, we show the robustness of our estimates to the inclusion of controls that interact the (expected) gender composition of the jury with various case characteristics, such as race and whether the defendant is also being charged with a violent crime. In addition, we include controls that interact the gender composition of the jury with other defendant characteristics, such as race. If the inclusion of these interactions were to result in a decline in our estimate of interest, it suggests that at least some of the effect is due not

to own-gender bias, but to differential treatment of some other defendant characteristic correlated with defendant gender.

4. Results

4.1. *Correlation between expected jury gender and actual jury gender*

We begin by demonstrating that our measure of jury gender is predictive of actual jury composition. Note that in contrast to the main analysis, this exercise can only be performed for those cases in which a jury was seated for the trial. The underlying data are shown in panel (a) of Online Appendix Figure OA1, which graphs the actual number of women seated on the jury against the expected number of women seated on the jury. The slope is close to one, suggesting that our (exogenous) measures of jury gender composition are strongly correlated with observed jury gender composition.

Regression results are shown in Table 1. Specifically, we estimate an equation of the same form as equation (1) above in that we regress the actual number of females on the predicted number of females, along with county-by-crime fixed effects. Results are consistent with Online Appendix Figure OA1 in showing strong correlations between actual and expected gender composition. Column 1 shows a correlation of 1.054 for all cases and is significant at the one percent level. The remaining columns show that this correlation remains strong for felonies, misdemeanors, and cases that include infractions related to drugs, driving, property crime, and violent crime. Correlations range from 0.932 for drug cases using average jury panel gender to 1.131 for misdemeanor cases. All estimates are statistically significant at the one percent level. As a result, it is clear that more women being assigned to a jury pool and being assigned earlier in the ordering leads to large subsequent differences in the actual gender composition of the seated jury.

4.2. *Exogeneity tests of the measure of expected jury gender composition*

The validity of our empirical approach depends on the assumption that predicted jury gender composition is uncorrelated with confounding factors. While we expect this assumption to hold based on our understanding of how potential jurors are assigned to and ordered within jury pools, we can also provide some empirical evidence. To do so, we regress exogenous defendant and case characteristics on the expected number of jurors who are female. These characteristics include jury panel size as well as defendant gender, race, age, the number of offenses, and whether the defendant is being charged with a felony, drug, driving, property, or violent crime. In addition, we also test whether average juror age (available only for Hillsborough County) or judge gender is correlated with our measure of the expected number of women on the jury.

Results are shown in Online Appendix Table OA3. Overall, there is little evidence that these exogenous characteristics are correlated with the expected number of female jurors. Of the 48 estimates shown, three are significant at the 10 percent level, and two are significant at the five percent level, which is consistent with random chance. This contrasts with results from the same exercise using actual number of women on the seated jury, rather than our measure of expected jury gender composition. In that exercise, the results of which are shown in Online Appendix Table OA4, nine of the 24 estimates are significant at the 10 percent level, and five are significant at the five percent level. This reflects the fact that the actual number of women seated for the jury is the outcome of the non-random jury selection process.

In addition, we also provide another test. The intuition of the test is to use all of the exogenous case and defendant characteristics shown in Table OA3, along with county-by-crime fixed effects, to predict conviction rate for each charge for each individual. This predicted conviction rate is thus a linear combination of all observable characteristics about that case and individual, where the weights are optimally chosen to best predict the likelihood of being convicted on that charge. We graph these predicted conviction rates for male and female defendants against our measure of expected jury gender composition. Our identifying assumption requires that the difference in the underlying propensity for guilt of male and female defendants be orthogonal to jury gender.

Results for all charges are shown in Figure 2a. The symbols represent local averages for charges against male and female defendants, and are grouped into 6 equal-sized bins. In addition, we fit separate lines to the underlying data for male and female defendants. Figure 2a shows that while male defendants are predicted to be found guilty more often than female defendants, this difference is constant across jury gender. This suggests that there is little reason, based on observable case and defendant characteristics, to expect a nonzero difference-in-differences estimate in the absence of an effect of own-gender jurors.

Results in Figure 3a show predicted conviction rates for drug charges, where we later show large effects of own-gender jurors. Results indicate that male and female defendants are predicted to have similar conviction rates regardless of expected jury gender. This is consistent with the identifying assumption that any nonzero difference-in-difference estimate of the effect of jury gender is due to the effect of jury gender, rather than some confounding factor.

4.3. *Effect of own-gender jurors on conviction rates*

Next, we turn to estimating the effect of jury gender on convictions. Before presenting formal estimates, we first show the raw data. Figure 2b graphs the conviction rates of male and female defendants across all charges. The figure indicates that the conviction rates of female defendants decline somewhat more than men's as the number of female jurors increases.

Conviction rates for drug offenses are shown in Figure 3b. Conviction rates for male defendants appear roughly flat (and linear) as the expected number of female jurors increases. In contrast, the conviction rates of female defendants decline sharply in a linear fashion as the expected number of female jurors increases. In short, female defendants are much less likely to be convicted of a drug charge as the jury is more female, while male conviction rates seem largely unaffected by jury gender.

Estimation results are shown in Table 2. All specifications control for the expected number of female jurors as well as an indicator for whether the defendant is female. In addition, all specifications control for county-by-crime fixed effects. Column 1 shows the estimate of own-gender jurors for all crimes. The coefficient is -0.038 and is not statistically significant. The magnitude of the coefficient implies that increasing the expected number of own-gender jurors by one (in a jury of six) is associated with a 3.8 percentage point reduction in the conviction rate. Importantly, the precision of this estimate enables us to reject effects that are one-third the estimated impact of a similar change in expected juror race reported by Anwar, Bayer and Hjalmarsson (2012).[14]

Column 2 additionally controls for other defendant and case characteristics such as the defendant's age and race, judge fixed effects, the number of charges in the case, and whether the defendant was also charged with a violent crime such as assault. The coefficient changes slightly to -0.016 and remains insignificant.

As discussed earlier, a major threat to identification is the possibility that more male or more female juries are responding not to defendant gender, but to a feature of the case or defendant that is systematically correlated with defendant gender. In order to address this concern, in the third column we examine the robustness to our estimate to the inclusion of controls that interact case characteristics with defendant gender and the expected number of female jurors. Specifically, we include interactions with defendant race, age, judge gender, number of charges in the case, whether the individual is being charged with a violent crime, and whether the defendant is being charged with a felony. Results from a specification that includes these pairwise interactions are shown in column 3 of Table 2. As shown there, the coefficient of interest changes slightly to -0.024, but is still statistically insignificant.

Columns 4 – 6 in of Table 2 show results for felonies, which are the subset of criminal cases studied in previous work on jury race. Estimates range from -0.014 to -0.064, though none are statistically significant at conventional levels. Similarly, results in columns 7 – 9 show results for misdemeanor charges. Again, all estimates are negative ranging from -0.071 to -0.080 and none are statistically significant.

[14] Anwar, Bayer and Hjalmarsson (2012) examine the impact of at least one black juror in the pool, which in their data translates to a 6.08 percentage point increase in the proportion black of the overall jury pool (1.66 potential jurors dived by an average pool size of 27.3). Assuming each potential juror is equally likely to be seated on a six-person jury, this translates to a 0.365 increase in the expected number of black jurors (0.0608*6). If we multiply our estimate and standard error in Column 1 of Table 4 by 0.365, we obtain an estimate of -0.01387 and a 95 percent confidence interval of [-0.050, 0.02]. By comparison, the estimated effect in Table IV of Anwar, Bayer and Hjalmarsson (2012) is -0.168.

Importantly, due to the fact that we report results for several different subcategories of crime, we also report False Discovery Rate (FDR) adjusted Q-values for each estimate in Table 2. These are computed using the method proposed by Anderson (2008), and adjust for the fact we examine a total of six subcategories of crime (felony, misdemeanor, drug, driving, property, and violent). The adjusted Q-values, which are interpreted similarly to two-sided p-values, range from 0.584 to 0.939 for the estimates in columns 4 – 9.

Next, we examine effects by category of the criminal charge. Specifically, we examine effects on conviction for driving, property, violent, and drug crime charges.[15] Results are shown in Table 3. The format is similar to Table 2 in that the first column for each category includes only county fixed effects, the second column adds controls for defendant and case characteristics, and the third column adds controls for interactions between jury gender and defendant and case characteristics.

Results in columns 1 – 9 suggest there is little evidence that own-gender jurors affect convictions for driving, property, or violent crimes. In contrast, results in columns 10 - 12 indicate there is strong evidence of own-gender jurors on conviction for drug charges. The estimate of -0.340 in column 10 suggests that adding one own-gender juror (in expectation) to a jury reduces the likelihood of conviction by 34 percent. Adding controls changes the estimate only slightly to -0.364. In addition, further adding interaction controls reduces estimates only slightly to -0.299. All three estimates are significant at the one percent level. We also report FDR-adjusted Q-values in columns 10 - 12, which are 0.003, 0.001, and 0.016, respectively. This indicates that even after accounting for the multiple statistical tests across the six major categories of crime charges in Tables 2 and 3, the coefficients in columns 10 – 12 of Table 3 are sufficiently extreme as to be unlikely to arise due to chance.

In addition to reporting robust standard errors clustered at the defendant level, we also report two-sided p-values from randomization inference in which we randomly assign males and females to jury panels in our data, and compare the t-statistics from our main coefficients to the distribution of 10,000 empirical t-statistics from the permutation exercise. Results in columns 10 - 12 indicate p-values of 0.017, 0.011, and 0.066, respectively.

Finally, we also perform the analysis at the case level. Tables OA7 and OA8 replicate Tables 2 and 3 except at the case level, rather than the charge level. We define the outcome as the proportion of charges that are guilty, and define subgroups as cases that have at least one charge within that subgroup. Results are similar in that we find no evidence of effects except in drug cases. For example, in our base specification in column 10 of Table OA8, we estimate that one more own-gender juror (in

[15] In Online Appendix Table OA10, we classify another 1,061 charges that were difficult to classify into groups into these four categories and re-estimate results. In addition, we show results just for the subgroup of charges we were unable to classify into the four categories in Table 3 ("All Other") and those that we were still unable to classify into a category in Table OA10 ("Still Other"). Results are similar.

expectation) results in a 27.3 percentage point decrease in conviction. That estimate is significant at the one percent level using conventional inference, has an FDR Q-value of 0.049, and a randomization-inference p-value of 0.042.

In summary, we have two primary findings. First, we find no effects of own-gender jurors on convictions overall for either misdemeanors or felonies, or for driving, violent, or property offenses. However, we show that increasing the number of expected own-gender jurors by one leads to a 30 - 36 percentage point reduction in convictions on drug charges, and that this effect is significant even after adjusting for multiple inference.

4.4. *Effect of own-gender jurors on sentencing*

Next, we ask whether increasing the expected number of own-gender jurors affects sentencing. In doing so, we note that while there is a clear link between drug charge convictions and sentencing, it is somewhat indirect. This is in part because sentencing is at the case level, and many cases with drug charges also include other charges. In addition, sentencing decisions are made by judges rather than juries. As a result, judges could potentially use their own assessment of cases to offset the impact of convictions generated by juror gender bias.

Results are shown in Table 4. The estimate in column 4 indicates that for cases with one or more drug charge, adding one (expected) own-gender juror to the jury reduces the likelihood of being sentenced to jail by around 22.5 percentage points. Adding case controls increases the estimate slightly to 27.6 percentage points. Both estimates are significant at the five percent level, though adding interactions reduces the estimate somewhat to a statistically insignificant but still economically large 16.8 percentage points.[16] Overall, these estimates provide some evidence that the marginal convictions due to jury gender do result in increased incarceration. Importantly, existing research on the effect of conviction and incarceration on recidivism and employment suggests increases in incarceration can result in significant long-term harm to defendants facing drug charges (Aizer and Doyle Jr 2015; Mueller-Smith 2015).

5. Robustness

We now test the robustness of our main finding that additional own-gender jurors leads to large reductions in conviction rates on drug charges. Results are shown in Table 5, where column 1 replicates our main estimate for drug charges of -0.364 from our preferred specification in column 11 of Table 3.

[16] In addition, we note this is the one result where there is a difference between using the measure of expected jury based on juror order and that based on overall jury pool gender, where the estimate for sentencing is still large but not statistically significant. See Online Appendix Figures OA4 and OA5.

We first return to a main threat to identification. As discussed earlier, that threat is the possibility that jurors of a given gender are responding not to the defendant's gender, but to some other defendant or case characteristic correlated with defendant gender. We tested for this in column 12 of Table 3 by showing results were unchanged when including interactions of jury gender with the number of charges in the case, whether there was a charge for a violent crime in the case, judge gender, and defendant race and age. However, one may also be concerned that jurors of different gender could respond differently to the type of drug charge in the case, which could be correlated with defendant gender. To test for this, we additionally include interactions of expected jury gender with indicators for marijuana possession, possession of other drugs, and possession of drug paraphernalia, where drug trafficking is the excluded group. Results are shown in column 2 of Table 5 and indicate that including these interactions slightly *increases* the magnitude of the estimate to -0.368. This provides further evidence that the effects shown are due to the interaction of defendant and jury gender, rather than the interaction of jury gender with some other characteristic correlated with defendant gender.

Similarly, we also ask whether our results are robust to the inclusion of controls for expected juror age, and expected juror age interacted with whether the defendant is female and the expected number of females on the jury. In addition, we also include interactions between the expected number of black jurors and whether the defendant is black, or female, and the expected number of females on the jury. We show results only for Hillsborough County, since we do not observe potential juror race and age in Palm Beach County. We show our preferred specification for Hillsborough County in column 3, which serves as a baseline. The estimate is larger and more significant for Hillsborough County (-0.451) than for Palm Beach (unreported estimate = -0.302), though both estimates imply economically large effects. Importantly, the estimate for Hillsborough County is similar (-0.406) when including the juror age and race interactions, as shown in column 4.

Column 5 of Table 5 shows the estimate from our main specification when we classify outcomes in which adjudication was withheld as not guilty rather than guilty, which occurs in 4 percent of the drug charges. The magnitude of the estimate is -0.342, which is similar to our baseline estimate of -0.364, and is still statistically significant at the one percent level.

In addition, we replicate our main set of results in Online Appendix Tables OA5 and OA6 where we instead conduct the analysis using the proportion of women in the entire pool, rather than using order to construct expected number of female jurors. Estimates are similar to those reported in Tables 2 and 3.[17]

Finally, in Online Appendix Table OA10 we show results for samples that include charges that were

[17] In Online Appendix Table OA9, we also show that estimates are robust to making alternative approaches to inferring jury gender as well as alternative ways of fitting the underlying data to predict the expected number of female jurors.

excluded in the primary analysis. Results are similar; the only subgroup for which effects are statistically significant is the drug category (in which we also include disorderly intoxication/conduct). We conclude that our main findings are robust to the inclusion or exclusion of charges that have few female defendants, or that are difficult to classify into a category.

In summary, we find no evidence that our estimated effect of own-gender jurors on convictions in drug cases is due to male or female jurors responding differentially to a characteristic correlated with defendant gender, rather than defendant gender itself. In addition, we find that this own-gender effect is robust to alternative ways of predicting jury gender, defining the treatment and outcome, and categorizing charges.

6. Discussion and Conclusion

There are several potential mechanisms through which own-gender jurors could cause these large effects on conviction outcomes for drug offenses. The first is that seated jurors may exhibit own-gender bias when making conviction decisions. Given that we do not observe true guilt, it is difficult for us to assess which jurors – male or female – are biased, and in what direction. Under the interpretation of juror bias, the results would be due to male and/or female jurors being either too lenient to own-gender defendants, being too tough on (that is, wrongfully convicting) opposite-gender defendants, or both.

Relatedly, effects could be due to the expectation of juror bias in criminal drug trials. For example, a defendant may be more likely to accept an otherwise unappealing plea deal if the expected jury composition is largely opposite-gender. It is also possible that prosecutors or defendants falsely believe jurors will engage in gender bias during the trial, resulting in a change in plea deal behavior prior to the start of the trial. Finally, an increase in the number of opposite-gender jurors could lead the defense to use their peremptory challenges on opposite-gender potential jurors. This would mean the attorney would have fewer peremptory challenges to use on other unfavorable jurors, thereby weakening the defendant's chances at acquittal.

Data limitations make it difficult for us to distinguish between these potential mechanisms with any certainty. What we know based on Figure 3b is that effects are driven by differential conviction rates of female defendants by juries of differing gender, rather than of male defendants. In addition, we find suggestive evidence that effects are somewhat larger for cases that go to trial compared to cases that do not.[18]

[18] Results are shown in Online Appendix Tables OA11 and OA12. In Table OA11, we first ask whether having a jury with more own-gender jurors affects the likelihood the case goes to trial. Estimates are economically small and statistically insignificant, suggesting there is no effect. This gives us some confidence that estimating effects separately by trial status is not confounded by selection into jury status. Estimated effects by trial status are shown in Table OA12, where columns 1 - 3 show estimates for cases that were not resolved via trial, and columns 4 - 6 show estimates for cases that were resolved via trial. Estimates for cases that were not resolved in trial range from -0.410 to -0.570, while estimates for cases that were resolved in trial are somewhat

Overall, we conclude that while there is little evidence that juror gender matters overall, there is strong evidence own-gender jurors reduce conviction rates for drug offenses. Estimates indicate that increasing the expected number of own-gender jurors by one results in a 30 percentage point decline in conviction rates on drug charges. Point estimates suggest that a similar change in jury gender results in a 22 percentage point reduction in the likelihood of being sentenced to at least some jail time. These are large effects, though we note this is consistent with prior research on the effect of juror race (Anwar, Bayer and Hjalmarsson 2012).

These results are important for the debate over the use of peremptory challenges in selecting a jury. In particular, our results for drug offenses provide support for recent court rulings that disallow prosecutors or defense attorneys to strike potential jurors from the jury pool on the basis of gender.

In addition, our results add evidence to a growing literature testing for own-gender bias in decision-making. On the one hand, our findings provide no evidence of gender bias overall, or for driving, property, or violent offenses. On the other hand, we show strong evidence of own-group bias for drug offenses. This finding suggests that bias can arise even in settings where the objective of impartiality is heavily emphasized and protected. In addition, the process explicitly allows for both sides to remove potential jurors from the jury if they are shown or believed to be unfair. We find that even in this process, the similarity in gender of the jury to the defendant can have a significant effect on convictions.

larger (-0.563 to -0.718). We note that for this analysis we exclude the 67 cases representing 135 charges where the records did not indicate whether the case was decided by trial or prior to the start of the trial.

References

Abrevaya, Jason and Daniel Hamermesh, "Charity and Favoritism in the Field: Are Female Economists Nicer (to Each Other)?," *Review of Economics and Statistics*, 2012, *94* (1), 202–207.

Aizer, Anna and Joseph Doyle Jr, "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges," *Quarterly Journal of Economics*, 2015, *130* (2), 759–803.

Anderson, Michael, "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, "The Impact of Jury Race in Criminal Trials," *Quarterly Journal of Economics*, 2012, *127* (2), 1017–1055.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, "The Role of Age in Jury Selection and Trial Outcomes," *Journal of Law and Economics*, 2014, *57* (4), 1001–1030.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, "Politics in the Courtroom: Political Ideology and Jury Decision Making," *Journal of the European Economic Association*, 2015.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson, "A Jury of Her Peers: The Impact of the First Female Jurors on Criminal Convictions," *The Economic Journal*, 2019.

Ayres, Ian and Peter Siegelman, "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, 1995, pp. 304–321.

Bagues, Manuel and Berta Esteve-Volart, "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment," *Review of Economic Studies*, 2010, *77* (4), 1301–1328.

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva, "Does the Gender Composition of Scientific Committees Matter?," *American Economic Review*, 2017, *107* (4), 1207–38.

Batson v. Kentucky, "476 US 79," 1986.

Bindler, Anna and Randi Hjalmarsson, "The Persistence of the Criminal Justice Gender Gap: Evidence from 200 Years of Judicial Decisions," *Journal of Law and Economics*, 2020, *63* (2), 297–339.

Breda, Thomas and Son Thierry Ly, "Professors in Core Science Fields Are Not Always Biased against Women: Evidence from France," *American Economic Journal: Applied Economics*, 2015, *7* (4), 53–75.

Butcher, Kristin, Kyung Park, and Anne Morrison Piehl, "Comparing Apples to Oranges: Differences in Women's and Men's Incarceration and Sentencing Outcomes," *Journal of Labor Economics*, 2017, *35* (S1), S201–S234.

Cohen, Alma and Crystal Yang, "Judicial Politics and Sentencing Decisions," *American Economic Journal: Economic Policy*, 2019.

Dahl, Gordon and Enrico Moretti, "The Demand for Sons," *Review of Economic Studies*, 2008, *75* (4), 1085–1120.

De Paola, Maria and Vincenzo Scoppa, "Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia," *Economica*, 2015, *82* (325), 162–188.

Depew, Briggs, Ozkan Eren, and Naci Mocan, "Judges, Juveniles, and In-Group Bias," *Journal of Law and Economics*, 2017, *60* (2), 209–239.

Eren, Ozkan and Naci Mocan, "Emotional Judges and Unlucky Juveniles," *American Economic Journal: Applied Economics*, 2018, *10* (3), 171–205.

Flanagan, Francis, "Race, Gender, and Juries: Evidence from North Carolina," *Journal of Law and Economics*, 2018, *61* (2), 189–214.

Flanagan, Francis X, "Peremptory challenges and jury selection," *The Journal of Law and Economics*, 2015, *58* (2), 385–416.

George, Tracey, "Court Fixing," *Arizona Law Review*, 2001, *43* (1), 9–62.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians," *American Economic Review*, 2000, *90* (4), 715–741.

Hoekstra, Mark and CarlyWill Sloan, "Does Race Matter for Police Use of Force? Evidence from 911 Calls," 2020. NBER Working paper 26774.

J.E.B. v. Alabama, "511 US 127," 1994.

Johnson, Brian, "Judges on Trial: A Reexamination of Judicial Race and Gender Effects across Modes of Conviction," *Criminal Justice Policy Review*, 2014, *25* (2), 159–184.

Knepper, Matthew, "When the Shadow Is the Substance: Judge Gender and the Outcomes of Workplace Sex Discrimination Cases," *Journal of Labor Economics*, 2018, *36* (3), 623–664.

Lavy, Victor, "Do Gender Stereotypes Reduce Girls' or Boys' Human Capital Outcomes? Evidence from a Natural Experiment," *Journal of Public Economics*, 2008, *92* (10-11), 2083–2105.

McDonough Power Equipment, Inc. v. Greenwood, "464 U.S. 548," 1984.

Mitchell, Tara, Ryann Haw, Jeffrey Pfeifer, and Christian Meissner, "Racial Bias in Mock Juror Decision-Making: A Meta-Analytic Review of Defendant Treatment," *Law and Human Behavior*, 2005, *29* (6), 621–637.

Moss-Racusin, Corinne, John Dovidio, Victoria Brescoll, Mark Graham, and Jo Handelsman, "Science Faculty's Subtle Gender Biases Favor Male Students," *Proceedings of the National Academy of Sciences*, 2012, *109* (41), 16474–16479.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," 2015. Working paper.

Neumark, David, Roy Bank, and Kyle Van Nort, "Sex Discrimination in Restaurant Hiring: An Audit Study," *Quarterly Journal of Economics*, 1996, *111* (3), 915–941.

Philippe, Arnaud and Aurélie Ouss, ""No Hatred or Malice, Fear or Affection": Media and Sentencing," *Journal of Political Economy*, 2017.

Price, Joseph and Justin Wolfers, "Racial Discrimination among NBA Referees," *Quarterly Journal of Economics*, 2010, *125* (4), 1859–1887.

Schanzenbach, Max, "Racial and Sex Disparities in Prison Sentences: The Effect of District-Level Judicial Demographics," *Journal of Legal Studies*, 2005, *34* (1), 57–92.

Shayo, Moses and Asaf Zussman, "Judicial Ingroup Bias in the Shadow of Terrorism," *Quarterly Journal of Economics*, 2011, *126* (3), 1447–1484.

Steffensmeier, Darrell and Chris Hebert, "Women and Men Policymakers: Does the Judge's Gender Affect the Sentencing of Criminal Defendants?," *Social Forces*, 1999, *77* (3), 1163–1196.

West, Jeremy, "Racial Bias in Police Investigations," 2017. Working paper.

Tables and Figures

Table 1: Correlation between actual jury gender composition and expected gender composition

	All	Felony	Misdemeanor	Driving	Property	Violent	Drug
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
E(Num Fem | Order)	1.054**	0.992**	1.161**	1.083**	0.999**	1.055**	0.932**
	(0.042)	(0.053)	(0.069)	(0.080)	(0.090)	(0.064)	(0.100)
Observations	1313	848	465	372	313	595	204
F stat	67	38	30	92	63	137	50

Source. Criminal cases slated for jury trials in Hillsborough and Palm Beach counties between 2014 and 2016.
Note. Each column represents a separate regression. Columns 2 - 4 restrict the sample to cases with at least one charge in that category. All regressions include county fixed effects and columns 1-3 include county-by-crime fixed effects. Robust standard errors are in parentheses.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table 2: Effect of own-gender jurors on conviction rates, by severity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All Charges			Felony Charges			Misdemeanor Charges		
E(Num Fem | Order)xDef Fem	-0.038	-0.016	-0.024	-0.064	-0.014	-0.037	-0.079	-0.071	-0.080
	(0.050)	(0.051)	(0.051)	(0.071)	(0.072)	(0.071)	(0.063)	(0.064)	(0.065)
FDR Adjusted q-values				[0.584]	[0.939]	[0.883]	[0.584]	[0.80]	[0.653]
Observations	2937	2937	2937	1640	1640	1640	1297	1297	1297
Mean Dependent Variable	0.53	0.53	0.53	0.55	0.55	0.55	0.48	0.48	0.48
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes

Source. Criminal cases slated for jury trials in Hillsborough and Palm Beach counties between 2014 and 2016.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table 3: Effect of own-gender jurors on conviction rates, by crime type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Driving Charges			Property Charges			Violent Charges			Drug Charges		
E(Num Fem | Order)xDef Fem	0.018	0.006	-0.013	0.088	0.064	0.080	-0.055	0.006	0.015	-0.340**	-0.364**	-0.299**
	(0.087)	(0.080)	(0.087)	(0.105)	(0.115)	(0.110)	(0.078)	(0.082)	(0.081)	(0.094)	(0.092)	(0.098)
FDR Adjusted q-values	[0.834]	[0.939]	[0.883]	[0.584]	[0.939]	[0.883]	[0.584]	[0.939]	[0.883]	[0.003]	[0.001]	[0.016]
Permutation based p-values										{0.017}	{0.011}	{0.066}
Observations	775	775	775	723	723	723	981	981	981	467	467	467
Mean Dependent Variable	0.53	0.53	0.53	0.56	0.56	0.56	0.49	0.49	0.49	0.59	0.59	0.59
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Source. Criminal cases slated for jury trials in Hillsborough and Palm Beach counties between 2014 and 2016.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table 4: Effect of own-gender jurors on sentencing, at the case level

	All Cases			Cases with One or More Drug Charge		
	(1)	(2)	(3)	(4)	(5)	(6)
E(Num Fem | Order)	-0.015	-0.022	0.054	0.005	0.017	0.273
	(0.021)	(0.020)	(0.063)	(0.052)	(0.056)	(0.227)
Def female	-0.044	-0.088	-0.277	0.506	0.690	-0.656
	(0.157)	(0.163)	(0.193)	(0.357)	(0.418)	(0.650)
E(Num Fem | Order)xDef Fem	-0.037	0.001	-0.008	-0.225*	-0.276*	-0.168
	(0.049)	(0.050)	(0.051)	(0.114)	(0.130)	(0.116)
Observations	1480	1480	1480	244	244	244
Mean Dependent Variable	0.42	0.42	0.42	0.50	0.50	0.50
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes

Source. Criminal cases slated for jury trials in Hillsborough and Palm Beach counties between 2014 and 2016.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge fixed effects are not included in these interactions.
Standard errors are in parentheses.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table 5: Robustness of estimates of own-gender jurors on conviction rates - drug charges only

	(1)	(2)	(3)	(4)	(5)
E(Num Fem | Order)	0.038	0.038	0.058	-0.681	0.041
	(0.048)	(0.077)	(0.049)	(0.802)	(0.047)
Defendant female	1.086**	0.993**	1.254**	1.754	0.987**
	(0.288)	(0.346)	(0.287)	(1.304)	(0.304)
E(Num Fem | Order)xDef Fem	-0.364**	-0.368**	-0.451**	-0.406**	-0.342**
	(0.092)	(0.097)	(0.090)	(0.111)	(0.095)
Observations	467	467	312	303	467
Mean Dependent Variable	0.59	0.59	0.59	0.59	0.59
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
County	Both	Both	Hills	Hills	Both
Drug Type Controls & Interactions	No	Yes	No	No	No
Juror Race & Age Controls & Interaction	No	No	No	Yes	No
Adjudication Withheld=Not Guilty	No	No	No	No	Yes

Source. Criminal cases slated for jury trials in Hillsborough and Palm Beach counties between 2014 and 2016.
Note. Standard errors are in parentheses and are clustered at the defendant level. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Drug types are marijuana related charges, possession related charges of any drug, and drug paraphernalia charges which are interacted with defendant gender and the expected gender of the jury panel. Notably, we only observe potential jurors' race and age in Hillsborough County, so we replicate Column 1 for Hillsborough in Column 3. In Column 4, we add controls for the expected juror age and interact it with defendant gender and the expected number of female jurors. We also control for the expected number of black jurors and interact it with defendant race, defendant gender, and the expected number of female jurors. Column 5 redefines adjudication withheld as not guilty.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Figure 1: Probability seated

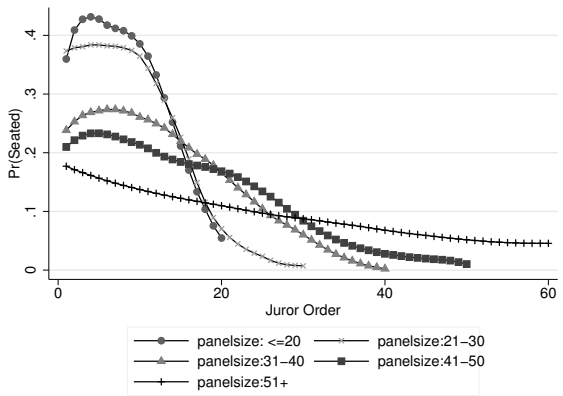
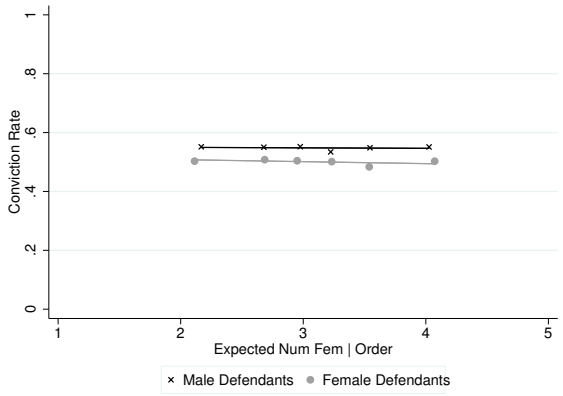


Figure 2: Predicted and actual conviction rates for male and female defendants, all charges

(a) predicted



(b) actual

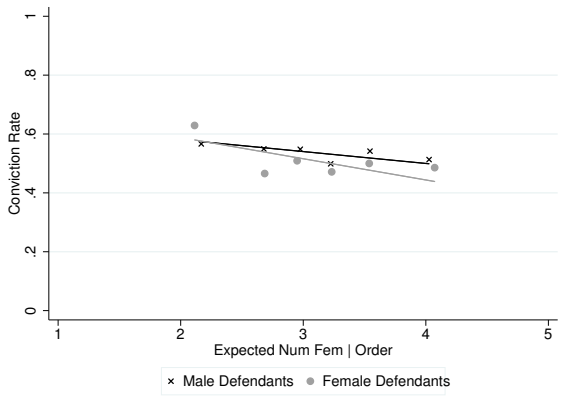
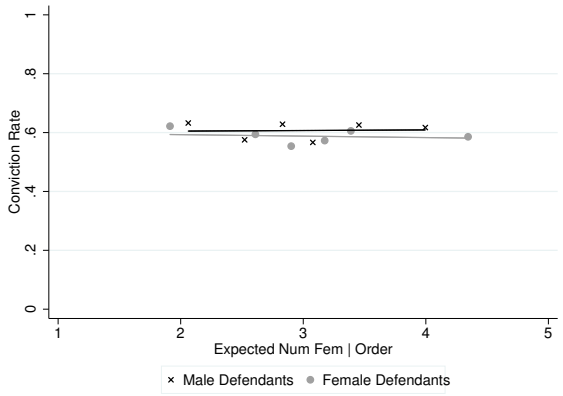


Figure 3: Predicted and actual conviction rates for male and female defendants, drug charges

(a) predicted



(b) actual

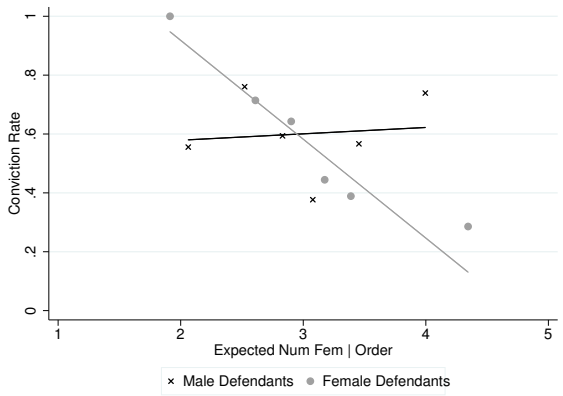


Table 6: Summary statistics for charges

	All	Male	Female	Felony	Misdem.	Driving	Property	Violent	Drug
Outcomes									
Guilty	0.53	0.54	0.51	0.55	0.48	0.53	0.56	0.49	0.59
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
Case Characteristics									
Defendant female	0.13	0.00	1.00	0.11	0.20	0.17	0.12	0.12	0.13
	(0.34)	(0.00)	(0.00)	(0.31)	(0.40)	(0.37)	(0.32)	(0.33)	(0.34)
Defendant white	0.50	0.49	0.59	0.43	0.66	0.71	0.45	0.39	0.47
	(0.50)	(0.50)	(0.49)	(0.50)	(0.47)	(0.45)	(0.50)	(0.49)	(0.50)
Defendant age	37.33	37.45	36.56	37.01	38.10	37.97	37.78	35.58	39.35
	(13.13)	(13.44)	(10.92)	(13.48)	(12.26)	(12.36)	(15.17)	(12.31)	(12.21)
Number of charges	4.15	4.42	2.44	4.85	2.51	3.27	5.81	3.17	5.04
	(4.90)	(5.19)	(1.39)	(5.64)	(1.40)	(2.44)	(6.52)	(3.59)	(6.34)
Violent charge in case	0.41	0.41	0.38	0.51	0.18	0.09	0.16	1.00	0.06
	(0.49)	(0.49)	(0.49)	(0.50)	(0.38)	(0.28)	(0.37)	(0.00)	(0.24)
Felony charge in case	0.70	0.72	0.56	1.00	0.00	0.26	0.88	0.85	0.84
	(0.46)	(0.45)	(0.50)	(0.00)	(0.00)	(0.44)	(0.33)	(0.35)	(0.37)
Judge female	0.32	0.31	0.41	0.34	0.29	0.31	0.27	0.38	0.31
	(0.47)	(0.46)	(0.49)	(0.47)	(0.46)	(0.46)	(0.44)	(0.49)	(0.46)
Jury Characteristics									
Actual number female	2.69	2.69	2.69	2.59	2.92	2.88	2.79	2.55	2.50
	(1.48)	(1.48)	(1.49)	(1.49)	(1.44)	(1.39)	(1.47)	(1.50)	(1.55)
E(num fem | pool)	3.10	3.10	3.11	3.09	3.12	3.13	3.14	3.08	3.03
	(0.55)	(0.54)	(0.61)	(0.52)	(0.60)	(0.57)	(0.51)	(0.55)	(0.56)
E(num fem | order)	3.10	3.10	3.12	3.09	3.12	3.11	3.18	3.09	2.99
	(0.62)	(0.62)	(0.65)	(0.61)	(0.65)	(0.64)	(0.57)	(0.62)	(0.66)
Average juror age in pool	45.00	45.05	44.68	45.01	44.93	44.68	44.78	45.11	45.20
	(3.20)	(3.19)	(3.25)	(3.22)	(3.12)	(3.18)	(3.08)	(3.23)	(3.31)
Observations	2937	2541	396	2061	876	775	723	981	467

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Summary statistics are at the charge level representing a total of 1,481 cases. Summary statistics at the case level are available in the online appendix, Table OA1.

A. Online Appendix

Table OA1: Summary statistics for cases

	All	Male	Female	Felony	Misdem.	Driving	Property	Violent	Drug
Outcomes									
Convict any	0.66	0.67	0.63	0.67	0.65	0.75	0.75	0.59	0.74
	(0.47)	(0.47)	(0.48)	(0.47)	(0.48)	(0.43)	(0.44)	(0.49)	(0.44)
Proportion guilty	0.54	0.54	0.50	0.55	0.50	0.55	0.61	0.48	0.61
	(0.44)	(0.44)	(0.44)	(0.44)	(0.43)	(0.40)	(0.42)	(0.45)	(0.42)
Total days sentenced, including zeros	1326	1527	241	1986	48	390	1951	2104	845
	(4712)	(5085)	(1016)	(5690)	(191)	(2879)	(5085)	(66245)	(1783)
P(sentenced $\geq$ 1 days), including zeros	0.42	0.44	0.28	0.49	0.26	0.35	0.54	0.41	0.50
	(0.49)	(0.50)	(0.45)	(0.50)	(0.44)	(0.48)	(0.50)	(0.49)	(0.50)
P(sentenced $\geq$ 1 years), including zeros	0.25	0.28	0.09	0.37	0.02	0.10	0.39	0.28	0.35
	(0.43)	(0.45)	(0.28)	(0.48)	(0.13)	(0.30)	(0.49)	(0.45)	(0.48)
P(sentenced $\geq$ 5 years), including zeros	0.14	0.16	0.03	0.22	0.00	0.03	0.23	0.20	0.14
	(0.35)	(0.37)	(0.18)	(0.41)	(0.06)	(0.18)	(0.42)	(0.40)	(0.34)
Case Characteristics									
Defendant female	0.16	0.00	1.00	0.13	0.21	0.18	0.17	0.13	0.14
	(0.36)	(0.00)	(0.00)	(0.33)	(0.41)	(0.39)	(0.38)	(0.34)	(0.34)
Defendant white	0.48	0.46	0.55	0.42	0.60	0.67	0.42	0.41	0.38
	(0.50)	(0.50)	(0.50)	(0.49)	(0.49)	(0.47)	(0.49)	(0.49)	(0.49)
Defendant age	36.95	37.26	35.28	36.57	37.69	37.74	36.47	36.16	37.20
	(12.48)	(12.73)	(10.93)	(12.58)	(12.27)	(12.13)	(13.43)	(12.75)	(10.97)
Number of charges	2.31	2.39	1.83	2.54	1.85	2.43	2.91	2.23	2.77
	(2.11)	(2.23)	(1.08)	(2.43)	(1.12)	(1.95)	(2.96)	(1.88)	(2.59)
Violent charge in case	0.46	0.47	0.39	0.56	0.26	0.07	0.25	1.00	0.10
	(0.50)	(0.50)	(0.49)	(0.50)	(0.44)	(0.25)	(0.43)	(0.00)	(0.30)
Felony charge in case	0.66	0.68	0.54	1.00	0.00	0.31	0.79	0.81	0.76
	(0.47)	(0.47)	(0.50)	(0.00)	(0.00)	(0.46)	(0.40)	(0.40)	(0.43)
Judge female	0.33	0.32	0.38	0.38	0.24	0.29	0.30	0.39	0.35
	(0.47)	(0.47)	(0.49)	(0.48)	(0.43)	(0.45)	(0.46)	(0.49)	(0.48)
Actual number female	2.69	2.66	2.86	2.60	2.86	2.85	2.68	2.62	2.56
	(1.49)	(1.49)	(1.46)	(1.50)	(1.45)	(1.44)	(1.52)	(1.48)	(1.58)
E(num fem | pool)	3.09	3.09	3.12	3.09	3.10	3.11	3.13	3.08	3.05
	(0.58)	(0.57)	(0.63)	(0.56)	(0.62)	(0.59)	(0.53)	(0.58)	(0.59)
E(num fem | order)	3.08	3.08	3.11	3.07	3.10	3.10	3.15	3.07	3.03
	(0.65)	(0.64)	(0.67)	(0.64)	(0.66)	(0.65)	(0.61)	(0.64)	(0.68)
Average juror age in pool	45.06	45.15	44.58	45.12	44.93	44.68	45.13	45.09	45.05
	(3.31)	(3.31)	(3.28)	(3.39)	(3.09)	(3.29)	(3.02)	(3.39)	(3.33)
Observations	1481	1250	231	977	504	404	365	668	244

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.

Table OA2: Effect of gender jurors on conviction rates

	All	Felony	Misdemeanor	Driving	Property	Violent	Drug
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
E(Num Fem | Order)	-0.022	-0.017	-0.017	0.016	-0.053	-0.015	-0.007
	(0.018)	(0.024)	(0.024)	(0.031)	(0.038)	(0.029)	(0.045)
Observations	2937	1640	1297	775	723	981	467

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Each column represents a separate regression. Columns 1 - 3 include county-by-crime fixed effects, and columns 4 - 7 include county fixed effects. Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA3: Exogeneity Tests

Panel A: Case-Level								Case has at least one charge that is classified as:				
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
Panel A.1: Expected Gender Composition from Overall Pool												
E(Num Fem | Pool)	0.013	0.030	-0.508	-0.320	-0.351	-0.008	-0.009	0.005	-0.006	0.031+	-0.006	-0.012
	(0.018)	(0.022)	(0.570)	(0.205)	(0.494)	(0.021)	(0.068)	(0.019)	(0.020)	(0.018)	(0.023)	(0.017)
Observations	1481	1481	1481	823	1481	1481	1481	1481	1481	1481	1481	1481
Panel A.2: Expected Gender Composition from Ordering of Pool												
E(Num Fem | Order)	0.008	0.024	-0.257	-0.172	0.051	-0.010	0.008	-0.003	-0.006	0.042*	-0.008	-0.013
	(0.015)	(0.019)	(0.519)	(0.178)	(0.446)	(0.019)	(0.061)	(0.016)	(0.018)	(0.017)	(0.020)	(0.015)
Observations	1481	1481	1481	823	1481	1481	1481	1481	1481	1481	1481	1481
Panel B: Charge-Level												
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
Panel B.1: Expected Gender Composition from Overall Pool												
E(Num Fem | Pool)	0.015	0.030	-0.573	-0.299	-0.449	-0.007	-0.007	0.001	-0.006	0.027	-0.006	-0.013
	(0.018)	(0.022)	(0.571)	(0.204)	(0.488)	(0.021)	(0.071)	(0.018)	(0.019)	(0.017)	(0.022)	(0.015)
Observations	2937	2937	2937	1470	2937	2937	2937	2937	2937	2937	2937	2937
Panel B.2: Expected Gender Composition from Ordering of Pool												
E(Num Fem | Order)	0.009	0.024	-0.314	-0.155	-0.026	-0.009	0.022	-0.009	-0.007	0.032*	-0.010	-0.012
	(0.015)	(0.019)	(0.520)	(0.178)	(0.434)	(0.019)	(0.065)	(0.016)	(0.017)	(0.015)	(0.019)	(0.014)
Observations	2937	2937	2937	1470	2937	2937	2937	2937	2937	2937	2937	2937

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Each column in each panel reports estimates from a separate regression in which we regress observable characteristics on the expected number of females on the jury based on order. Columns 1 - 7 include county-by-crime fixed effects, and columns 8 - 12 include county fixed effects. The first three columns show results for defendant characteristics. Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA4: Exogeneity tests with actual number of female jurors

Panel A: Case-Level								Case has at least one charge that is classified as:				
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
Actual Num Fem	0.016*	0.020+	-0.178	-0.130	-0.110	0.010	-0.025	-0.008	0.009	0.018+	-0.028*	0.002
	(0.008)	(0.011)	(0.276)	(0.102)	(0.231)	(0.010)	(0.036)	(0.009)	(0.010)	(0.009)	(0.011)	(0.008)
Observations	1481	1481	1481	823	1481	1481	1481	1481	1481	1481	1481	1481
Panel B: Charge-Level												
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
Actual Num Fem	0.012+	0.018*	-0.034	0.013	-1.392**	-0.003	-0.063*	-0.010	0.013+	0.002	-0.009	-0.008
	(0.006)	(0.008)	(0.215)	(0.078)	(0.363)	(0.008)	(0.029)	(0.007)	(0.007)	(0.007)	(0.008)	(0.006)
Observations	2937	2937	2937	1470	2937	2937	2937	2937	2937	2937	2937	2937

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Each column in each panel reports estimates from a separate regression in which we regress observable characteristics on the actual number of females on the seated jury. If a jury is not seated, the expected number of females seated is used. Columns 1 - 7 include county-by-crime fixed effects, and columns 8 - 12 include county fixed effects. The first three columns show results for defendant characteristics. Standard errors are in parentheses and are clustered at the defendant level. Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA5: Effect of own-gender jurors on conviction rates, expected number female in overall pool, by severity

	All Charges			Felony Charges			Misdemeanor Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E(Num Fem | Pool)	-0.015	-0.014	0.071	-0.008	-0.006	0.115	-0.011	-0.015	0.049
	(0.022)	(0.022)	(0.075)	(0.030)	(0.030)	(0.091)	(0.031)	(0.031)	(0.102)
Def female	0.145	0.123	-0.008	0.235	0.114	-0.056	0.190	0.219	0.160
	(0.160)	(0.164)	(0.220)	(0.234)	(0.239)	(0.286)	(0.204)	(0.211)	(0.293)
E(Num Fem | Pool)xDef Fem	-0.064	-0.052	-0.060	-0.101	-0.057	-0.076	-0.069	-0.073	-0.081
	(0.051)	(0.052)	(0.052)	(0.073)	(0.074)	(0.071)	(0.066)	(0.068)	(0.069)
FDR Adjusted q-values				[0.388]	[0.453]	[0.427]	[0.40]	[0.53]	[0.427]
Observations	2937	2937	2937	1640	1640	1640	1297	1297	1297
Mean Dependent Variable	0.53	0.53	0.53	0.55	0.55	0.55	0.48	0.48	0.48
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CountyXCrime Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.

Note. Each column replicates columns in Table 2 using the number of female jurors based on the overall pool for comparison to the measure used in Anwar et al. (2012). All specifications include controls for defendant gender and expected gender composition of the jury, as well as county-by-crime fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge fixed effects are not included in these interactions.

Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.

Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA6: Effect of own-gender jurors on conviction rates, expected number female in overall pool, by crime type

	Driving Charges			Property Charges			Violent Charges			Drug Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
E(Num Fem | Pool)	0.007	0.004	-0.113	-0.110*	-0.106*	-0.015	0.003	-0.012	-0.008	0.030	0.057	0.171
	(0.039)	(0.038)	(0.138)	(0.049)	(0.050)	(0.173)	(0.035)	(0.034)	(0.104)	(0.057)	(0.058)	(0.232)
Def female	0.223	0.244	0.126	-0.559	-0.436	-0.858+	0.163	-0.004	-0.466	1.063**	1.234**	1.084+
	(0.262)	(0.263)	(0.508)	(0.361)	(0.403)	(0.506)	(0.281)	(0.286)	(0.335)	(0.292)	(0.317)	(0.612)
E(Num Fem | Pool)xDef Fem	-0.081	-0.076	-0.110	0.151	0.105	0.145	-0.061	-0.006	0.010	-0.362**	-0.414**	-0.360**
	(0.084)	(0.085)	(0.092)	(0.116)	(0.128)	(0.117)	(0.086)	(0.087)	(0.084)	(0.098)	(0.102)	(0.110)
FDR Adjusted q-values	[0.40]	[0.53]	[0.427]	[0.388]	[0.53]	[0.573]	[0.477]	[0.942]	[0.902]	[0.002]	[0.001]	[0.008]
Permutation based p-values										{0.032}	{0.013}	{0.034}
Observations	775	775	775	723	723	723	981	981	981	467	467	467
Mean Dependent Variable	0.53	0.53	0.53	0.56	0.56	0.56	0.49	0.49	0.49	0.59	0.59	0.59
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Each column replicates columns in Table 3 using the number of female jurors based on the overall pool for comparison to the measure used in Anwar et al. (2012). All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA7: Effect of own-gender jurors on conviction rates at the case level, by severity

	All Charges			Felony Charges			Misdemeanor Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E(Num Fem | Order)	-0.022	-0.026	0.048	-0.025	-0.022	0.067	-0.011	0.098	0.098
	(0.019)	(0.019)	(0.065)	(0.024)	(0.024)	(0.072)	(0.031)	(0.113)	(0.113)
Def female	0.043	-0.008	-0.151	0.065	-0.035	-0.176	0.053	0.062	0.062
	(0.155)	(0.162)	(0.212)	(0.208)	(0.212)	(0.260)	(0.234)	(0.349)	(0.349)
E(Num Fem | Order)xDef Fem	-0.030	-0.008	-0.021	-0.046	-0.008	-0.026	-0.019	-0.009	-0.009
	(0.048)	(0.050)	(0.050)	(0.065)	(0.065)	(0.064)	(0.074)	(0.076)	(0.076)
FDR Adjusted q-values				[0.715]	[0.929]	[0.993]	[0.799]	[0.929]	[0.933]
Observations	1481	1481	1481	977	977	977	504	504	504
Mean Dependent Variable	0.54	0.54	0.54	0.55	0.55	0.55	0.50	0.50	0.50
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CountyXCrime Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.

Note. Each column is replicated for Table 2 at the case level. The outcome is the proportion of guilty charges in the case. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county-by-crime fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.

Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.

Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA8: Effect of own-gender jurors on conviction rates at the case level, by crime type

	Driving Charges			Property Charges			Violent Charges			Drug Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
E(Num Fem | Order)	-0.014	-0.013	0.034	-0.093*	-0.093*	0.086	-0.018	-0.040	-0.119	0.049	0.056	0.214
	(0.032)	(0.035)	(0.120)	(0.039)	(0.041)	(0.137)	(0.030)	(0.029)	(0.091)	(0.045)	(0.046)	(0.191)
Def female	-0.141	-0.063	0.001	-0.475	-0.410	-0.499	0.122	0.003	-0.451	0.798**	1.062**	0.830
	(0.275)	(0.254)	(0.505)	(0.325)	(0.384)	(0.426)	(0.240)	(0.261)	(0.313)	(0.301)	(0.307)	(0.550)
E(Num Fem | Order)xDef Fem	0.036	0.024	-0.001	0.132	0.096	0.063	-0.054	-0.007	-0.026	-0.273**	-0.347**	-0.279*
	(0.085)	(0.081)	(0.087)	(0.103)	(0.121)	(0.109)	(0.074)	(0.079)	(0.078)	(0.102)	(0.094)	(0.108)
FDR Adjusted q-values	[0.799]	[0.929]	[0.933]	[0.601]	[0.929]	[0.933]	[0.715]	[0.929]	[0.993]	[0.049]	[0.002]	[0.064]
Permutation based p-values										{0.042}	{0.007}	{0.059}
Observations	404	404	404	365	365	365	668	668	668	244	244	244
Mean Dependent Variable	0.55	0.55	0.55	0.61	0.61	0.61	0.48	0.48	0.48	0.61	0.61	0.61
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Each column is replicated for Table 3 at the case-level. The outcome is the proportion of guilty charges in the case. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level. False discovery rate (FDR) adjusted Q-values adjust for multiple inference given the six subcategories of crime examined. They are constructed using the method proposed by Anderson (2008) and are interpreted as two-sided p-values.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA9: Robustness of estimates of own-gender jurors on conviction rates - drug charges only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
E(Num Fem — Order) x Def Fem	-0.364**	-0.316**	-0.394**	-0.348**	-0.352**	-0.393**	-0.369**
	(0.092)	(0.093)	(0.090)	(0.093)	(0.092)	(0.092)	(0.100)
Observations	467	467	467	467	467	467	467
Mean Dependent Variable	0.59	0.59	0.59	0.59	0.59	0.59	0.59
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing Genders	half	female	male	half	half	half	half
Predicted Genders	API	API	API	SS	API	API	API
Pr(Seated)	LL	LL	LL	LL	Raw	Probit	LL
Pr(Seated—Panelsize)	Yes	Yes	Yes	Yes	Yes	Yes	No

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. Standard errors are in parentheses and are clustered at the defendant level. Column 1 replicates the estimate from column 11 of Table 3. The next two columns classify jurors for whom we could not identify gender using genderize.io as either all female (Column 2) or all male (Column 3). In column 4 we classify juror gender using names and genders recorded in Florida by the Social Security Administration. In columns 5 - 7 we show our estimate of own-gender jurors is robust to alternative methods of predicting jury gender, including using the raw probability rather than smoothing with epanechnikov kernel (column 5), probit (column 6), and a local linear smoother that does not condition on jury panel size (column 7).
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA10: Effect of own-gender jurors on conviction, other charges

	All	Felony	Misdemeanor	All Other	Still Other	Driving	Property	Violent	Drug
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E(Num Fem | Order)	-0.018	-0.005	-0.020	0.022	0.108	0.020	-0.074+	-0.034	0.043
	(0.018)	(0.023)	(0.024)	(0.035)	(0.166)	(0.034)	(0.043)	(0.025)	(0.046)
Def female	0.023	0.004	0.184	-0.115	1.057	0.011	-0.260	-0.151	1.092**
	(0.158)	(0.226)	(0.194)	(0.501)	(2.575)	(0.258)	(0.355)	(0.255)	(0.284)
E(Num Fem | Order)xDef Fem	-0.016	-0.020	-0.062	0.002	-0.351	0.000	0.062	0.037	-0.370**
	(0.049)	(0.070)	(0.061)	(0.155)	(0.855)	(0.081)	(0.113)	(0.078)	(0.091)
Observations	3998	2330	1668	1061	89	791	753	1954	475
Mean Dependent Variable	0.53	0.55	0.48	0.69	0.65	0.53	0.56	0.49	0.59
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.

Note. Each column replicates the second specification from Tables 2 and 3 for the listed sample including charges that are not clearly classified into one of the four primary crime categories. Each column includes controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects.

Other charges added in each category:

Drug: disorderly intoxication/conduct.

Driving: fleeing police.

Property: littering, corruption, official misconduct, possession of snook, second hand metal recyclers, falsely reporting a crime, falsely impersonating a person, perjury, prostitution, racketeering.

Violent: murder, manslaughter, sexual assault, sexual activity with a minor, molestation, shooting into a building, improper exhibition of firearm/weapon, felon in possession of firearm, carrying concealed firearm/weapon, violation of domestic injunction, resist/obstruct/oppose police officer, stalking, harassing telephone calls, child abuse, abuse of elderly, animal abuse, dangerous dog causing injury.

Charges still not in a crime category: failure of sexual offender to register, escape from jail, tampering with witness, false name to law enforcement, traveling to meet a minor, loitering, interference with custody, failure of caregiver to ensure school attendance, unlawful misuse of 911 system, violation of pretrial conditions, indecent exposure, lewd conduct, transmission of harmful material to minor, rico violation, disrupting school functions, abuse of dead human body, accessory after the fact.

Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA11: Effect of own-gender jurors on jury trial status, for drug charges only

	Outcome: Non-Trial			Outcome: Jury Trial		
	(1)	(2)	(3)	(4)	(5)	(6)
E(Num Fem | Order)	-0.019	-0.021	-0.381*	0.005	0.009	0.223
	(0.046)	(0.048)	(0.189)	(0.048)	(0.050)	(0.227)
Defendant female	-0.155	-0.147	0.671	0.232	0.197	-0.611
	(0.446)	(0.453)	(0.707)	(0.460)	(0.473)	(0.728)
E(Num Fem | Order)xDef Fem	0.065	0.068	0.045	-0.097	-0.088	-0.084
	(0.141)	(0.151)	(0.122)	(0.144)	(0.156)	(0.136)
Observations	467	467	467	467	467	467
Mean Dependent Variable	0.31	0.31	0.31	0.42	0.42	0.42
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects.Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Table OA12: Effect of own-gender jurors on conviction rates, by jury trial status, for drug charges only

	Non-Trial			Jury Trial		
	(1)	(2)	(3)	(4)	(5)	(6)
E(Num Fem | Order)	0.094	0.023	-0.312	0.036	0.095	0.238
	(0.082)	(0.079)	(0.614)	(0.068)	(0.075)	(0.397)
Def female	1.270**	1.172**	3.642+	1.514**	1.968+	1.409
	(0.373)	(0.439)	(1.912)	(0.454)	(0.993)	(1.328)
E(Num Fem | Order)xDef Fem	-0.440**	-0.410*	-0.570+	-0.563**	-0.718*	-0.671
	(0.123)	(0.156)	(0.313)	(0.139)	(0.347)	(0.437)
Observations	146	146	146	195	195	195
Mean Dependent Variable	0.66	0.66	0.66	0.58	0.58	0.58
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects.
Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

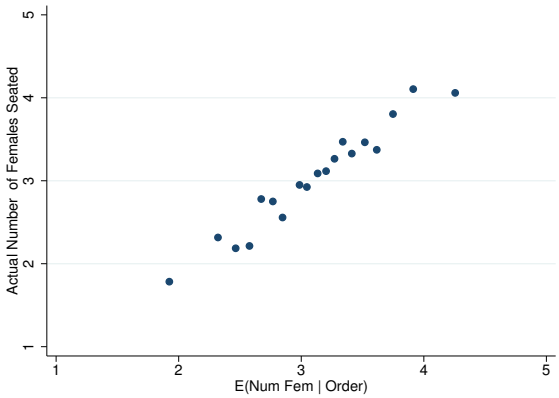
Table OA13: Effect of own-gender jurors on convictions, by drug charge type

	Marijuana or Paraphernalia			Possession or Paraphernalia			Other Drug Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E(Num Fem | Order)	0.051	0.043	0.208	0.119+	0.087	0.512	-0.054	-0.068	0.174
	(0.070)	(0.063)	(0.328)	(0.062)	(0.069)	(0.316)	(0.077)	(0.089)	(0.407)
Def female	0.910*	0.869+	-0.384	1.291**	1.331**	1.358+	0.318	0.152	-0.558
	(0.417)	(0.445)	(0.845)	(0.314)	(0.378)	(0.806)	(0.642)	(0.770)	(1.127)
E(Num Fem | Order)xDef Fem	-0.329*	-0.332*	-0.271*	-0.475**	-0.475**	-0.569**	-0.056	-0.006	0.097
	(0.142)	(0.144)	(0.137)	(0.101)	(0.115)	(0.152)	(0.225)	(0.257)	(0.321)
Observations	163	163	163	197	197	197	188	188	188
Mean Dependent Variable	0.64	0.64	0.64	0.64	0.64	0.64	0.57	0.57	0.57
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interactions	No	No	Yes	No	No	Yes	No	No	Yes

Sources. Hillsborough County and PalmBeach County Clerk of Courts, criminal cases spanning 2014-2016 slated for a jury trial.
Note. All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant race and age, the number of charges in the case, and indicators for whether there was charge for a violent crime or a felony charge in the case, and judge fixed effects. Interactions include controls for each of those characteristics interacted with the expected number of female jurors and the defendant's gender; judge gender is included in these interactions in place of judge fixed effects. Columns 1-3 estimate effects for all marijuana-related and paraphernalia charges. Columns 4-6 estimate the effect for possession and paraphernalia charges (not including marijuana possession). The last three columns restrict to drug charges that are not marijuana related, any type of possession, or drug paraphernalia.
Standard errors are in parentheses and are clustered at the defendant level.
Significant at the + $p<0.10$, * $p<0.05$, and ** $p<0.01$ level.

Figure OA1: Correlation between actual jury gender composition and expected gender composition

(a) proportion female in pool



(b) expected proportion female

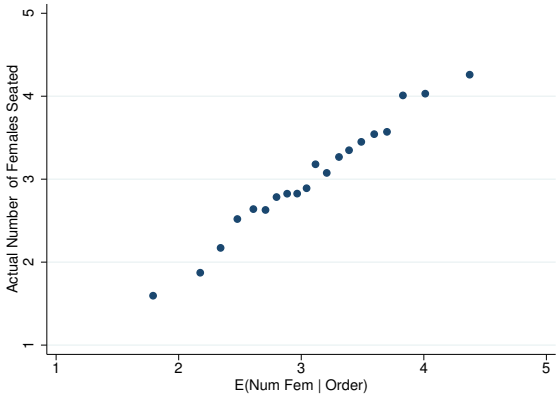
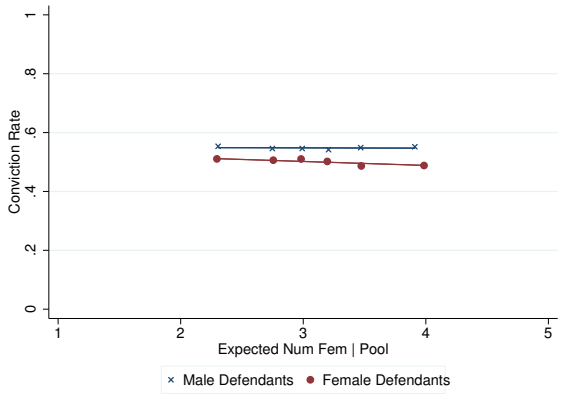
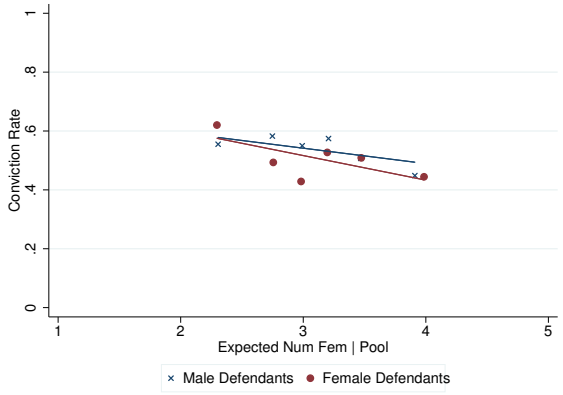


Figure OA2: Predicted and actual conviction rates for male and female defendants, all charges

(a) predicted



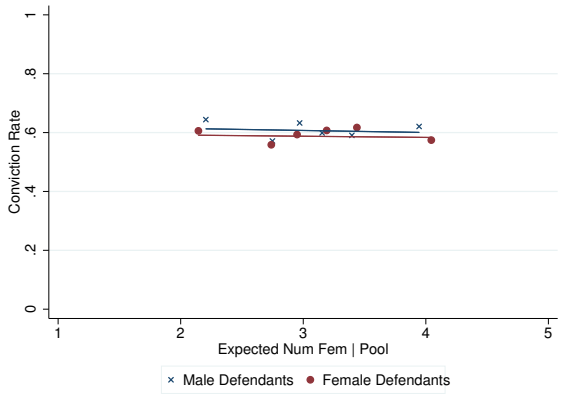
(b) actual



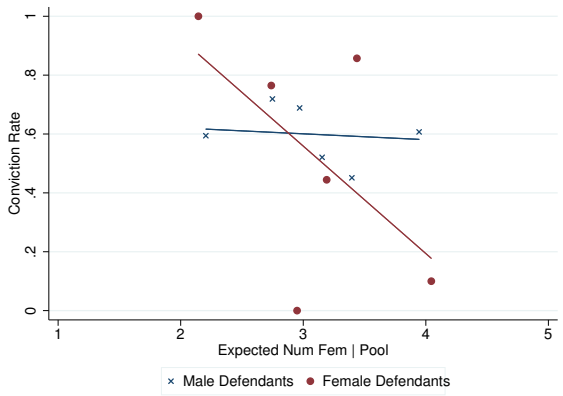
Note. For each charge, we predict the probability of conviction using all observable characteristics in Panel A. The line represents a linear fit across all predicted conviction rates. Panel B graphs the actual conviction rates for male and female defendants against the expected gender composition of the jury.

Figure OA3: Predicted and actual conviction rates for male and female defendants, drug charges

(a) predicted



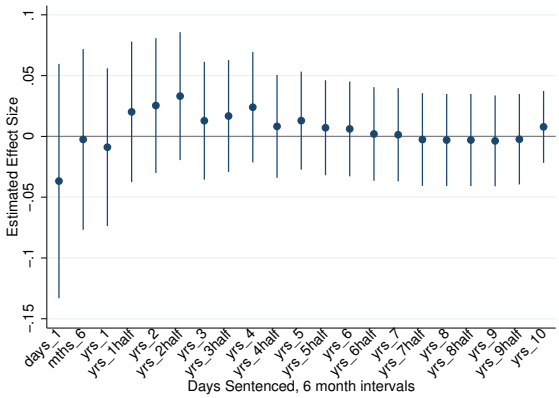
(b) actual



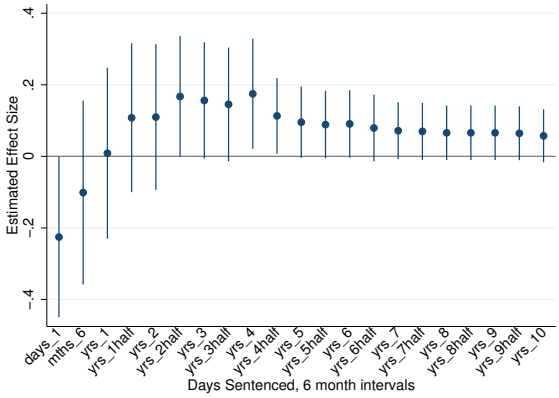
Note. For each drug charge, we predict the probability of conviction using all observable characteristics in Panel A. The line represents a linear fit across all predicted conviction rates. Panel B graphs the actual conviction rates for male and female defendants against the expected gender composition of the jury.

Figure OA4: Estimated effects of own-gender jurors on sentencing

(a) all charges



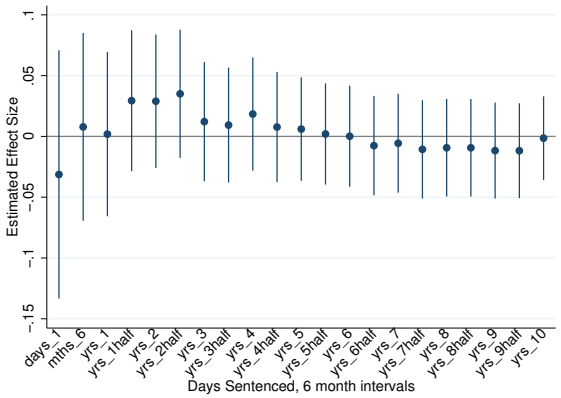
(b) drug charges



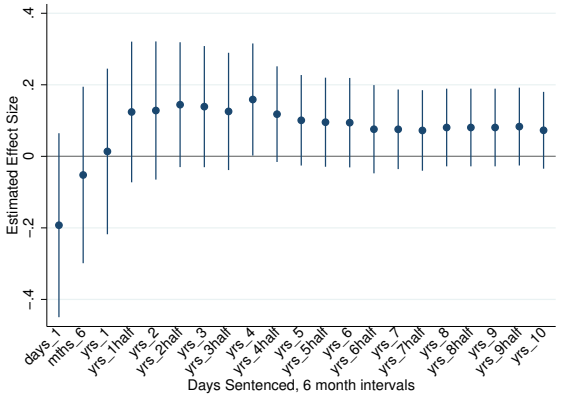
Note. Each estimate shown represents the effect of own-gender jurors on total sentencing in the case controlling for county by crime and judge fixed effects. The outcomes of interest, from left to right, are a set of indicators for sentenced to at least one day, sentenced to at least six months, 1 year, 1.5 years, 2 years, etc., up to at least 10 years. In Panel A, the sample includes all drug, driving, property, and violent crime cases. In Panel B, the sample restricts to cases with at least one drug charge.

Figure OA5: Estimated effects of own-gender jurors in *overall pool* on sentencing

(a) all charges



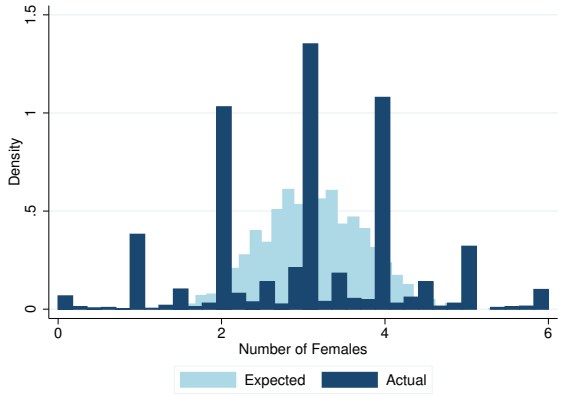
(b) drug charges



Note. Each estimate shown represents the effect of own-gender jurors on total sentencing in the case controlling for county by crime and judge fixed effects. The outcomes of interest, from left to right, are a set of indicators for sentenced to at least one day, sentenced to at least six months, 1 year, 1.5 years, 2 years, etc., up to at least 10 years. In Panel A, the sample includes all drug, driving, property, and violent crime cases. In Panel B, the sample restricts to cases with at least one drug charge.

Figure OA6: Distribution of actual and expected female jurors

(a) all charges



(b) drug charges