The Effect of Own-Gender Juries on Conviction Rates

Mark Hoekstra
Texas A&M University, NBER, and IZA

Brittany R. Street
Texas A&M University

May 2018

Abstract

This paper examines the extent to which criminal conviction rates are affected by the similarity in gender of the defendant and jury. To identify effects, we exploit the random variation in both the assignment to jury pools and the ordering of potential jurors. We do so using detailed administrative data on the juror selection process and trial proceedings for two large counties in Florida. Results indicate that own-gender juries result in significantly lower conviction rates on drug charges, though we find no evidence of effects for other charges. Estimates indicate that a one standard deviation increase in expected own-gender jurors (~10 percentage points) results in a 19 percentage point reduction in conviction rates on drug charges, which is highly significant even after adjusting for multiple comparisons. This results in a 13 percentage point decline in the likelihood of being sentenced to at least some jail time. 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.

Keywords: Crime, Jury Trials, Discrimination, Gender Bias

JEL Classification Codes: K41; K42, J16,

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), there is little evidence on whether modern juries favor own-gender defendants. The purpose of this paper is to test whether own-gender juries affect criminal conviction rates and sentencing outcomes.

The primary difficulty in doing so is that seated juries are the outcome of a nonrandom 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 juries 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 the randomization of the initial juror pool, and the random ordering of potential jurors within that pool, to predict the proportion of female jurors seated on the jury. This enables us to use only the variation in jury gender due to the fact women are (randomly) assigned to some jury pools more than others, and women are (randomly) assigned lower numbers in the ordering of some jury pools than in others. Because the seated jury consists of the first six or twelve ordered jurors who are not excluded by either a challenge for cause or a peremptory challenge (i.e., a challenge for which no reason must be given), this variation is orthogonal to other determinants of trial outcomes. We use this quasi-random variation in jury gender to identify own-gender effects by differencing out the impact of defendant and jury gender, similar to studies on racial bias (e.g., Price and Wolfers, 2010; Shayo and Zussman, 2011; West, 2017).

To implement this research design, 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 trial levels. Importantly, the data also include demographic information on potential jurors and the randomly-assigned ordering of each potential juror within the jury pool. Using this ordering and the empirical probabilities that jurors assigned a given number are seated on the jury, we predict the expected proportion of women on each jury, thereby isolating the as-good-as-random variation in the gender composition of seated juries. Importantly, we show that the predicted proportion of women on the jury is strongly predictive of the gender composition of the seated jury. We also show that this variation is uncorrelated with defendant and case characteristics, and with expected conviction rates of male and female defendants as predicted using exogenous characteristics.

Results provide strong evidence of own-gender juries on conviction rates for drug offenses. Estimates indicate that a one standard deviation increase in own-gender jurors (~10 percentage points) results in a 19 percentage point reduction in conviction rates on drug charges. Importantly, this effect is significant at the one percent level even after performing the multiple inference adjustment proposed by Anderson (2008). We also show that this change in jury gender composition leads to a 13 percentage point reduction in the likelihood of being sentenced to jail. In contrast, we find no evidence of effects for driving, property, or violent crime offenses. We hypothesize that the large effects for drug offenses are consistent with a model in which jurors are more likely to exhibit bias in cases where they have significant disagreements with U.S. law. However, we emphasize that we cannot rule out other explanations for the heterogeneity in own-gender effects. In addition, we present suggestive evidence that effects are driven largely by cases in which the jury reaches a verdict, as opposed to cases in which a plea deal is reached prior to the trial.

To our knowledge, this is the first paper to use random variation in own-gender juries 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.[1] In addition, this paper complements 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), who document the impact of having a black juror in the jury pool

[1] For example, see Abrevaya and Hamermesh (2012); Ayres and Siegelman (1995); Bagues and Esteve-Volart, 2010; Bagues, Sylos-Labini, and Zinovyeva (2017); Breda and Ly (2015); Dahl and Moretti (2008); De Paola and Scoppa, (2015); Goldin and Rouse (2000); Lavy (2008); Neumark, Bank, and Van Nort (1996); Moss-Racusin, Corinne, Dovidio, Brescoll, and Graham, and Handelsman (2012).

on felony conviction outcomes. It is also related to Anwar, Bayer, and Hjalmarsson (2016), 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] In assessing the role of jurors in affecting male versus female sentencing outcomes, this paper also complements a growing literature that documents and explains gender differences in sentencing (Bindler and Hjalmarsson, 2017; Butcher, Park, and Piehl, 2017). More generally, this study relates to a broader literature on the impact of judge gender (Johnson, 2014; Knepper, forthcoming; Schanzenbach, 2005; Steffensmeier and Hebert, 1999) and other judge and jury characteristics.[3] 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, forthcoming; Ouss and Philippe, forthcoming).

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, sizeable gender biases can still occur. In addition, we note that there is strong evidence that higher conviction and incarceration rates lead to increased recidivism and worsened labor market outcomes (Doyle and Aizer 2012; Mueller-Smith 2015). As a result, our findings suggest that drawing an opposite-gender jury can impose significant long-run costs on defendants.

1 Background and Data

1.1 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

[2] While we focus on the effect of own-gender juries in this paper, we also examine the effect of jury gender composition on overall conviction rates. Results are shown in Appendix Table A1, in which we regress an indicator for conviction on our measure of expected proportion women on the jury. Overall, we find no evidence that additional female jurors are more or less likely to convict overall.

[3] Examples include Anwar, Bayer, and Hjalmarsson, 2014; Anwar, Bayer, and Hjalmarsson (forthcoming); Mitchell, Haw, Pfeifer, and Meissner (2005); Cohen and Yang (forthcoming); Depew, Eren, and Mocan (2017); and George (2001).

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.

The potential jurors for a given case are then escorted to the courtroom 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 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.

1.2 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 August 2014 to August 2016. These are the third and fourth largest counties in Florida, respectively, each with a 2016 population of over 1.3 million people. 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.

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 genders for 93.8% 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 approach to inferring gender, we compare predicted gender to actual observed gender in Hillsborough County, and find that we accurately predict gender 98.63% 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 are able to compute empirical probabilities of being seated on the jury for each spot in the order in the jury panel.[4] To do so, we let the probability vary by size of the potential jury pool and the number of jurors being selected. Standard juries in Florida consist of six jurors, though the judge may decide to seat 12 jurors in some cases. Importantly, this decision is made prior to the assignment of the jury pool, and thus should not affect the internal validity of our approach.

These probabilities are shown in Figure 1, where panel a shows the probability of being seated on the jury for six-person juries, and panel b shows the same for twelve-person juries. For example, for six-person 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 20^{th}-ordered potential juror. By comparison, for 12-person juries selected from panels of 50 to 100 potential jurors, the probability of being seated ranges from 25 to 30 percent for the first 40 jurors to close to zero for the potential juror assigned last (e.g., 100^{th}) in the jury pool.

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. Summing this over the pool of potential jurors gives the expected number of females seated. Since trials in our data consist of both six and 12 person juries, we divide by the jury size to get the expected proportion of females. This enables us to make meaningful comparisons across jury panel sizes. In Section 3.1 we demonstrate that the predicted proportion of women on the jury is highly correlated with the actual proportion of women on the jury. In addition, in Section 3.2 we show

[4] 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 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.

that the expected proportion 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 our 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. 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 bias, since some cases settle after the prosecutor or defense attorney observes the composition of the potential jury pool or the actual seated 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 3.5 percent of all charges in our sample, and only 6.1 percent of drug charges. For the main analysis we treat this outcome as guilty, though in Table 7 we show that estimates are similar if we instead classify it as not guilty.

Our second outcome of interest is whether and for how long a defendant is sentenced to be incarcerated upon the conclusion of the trial. We define this outcome at the trial level, rather than charge level, since the sentences of individual charges are often served concurrently. In each case we observe the defendant's gender and race along with additional case characteristics including the severity of charges and the judge assigned to the case.

Finally, we note that because the purpose of this paper is to examine the effect of own-gender juries, we exclude cases linked to charges in which fewer than 10 percent of defendants are female. Consequently, we only consider cases that involve 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 very little variation in defendant gender.

Summary statistics are shown in Table 1, where Panel A shows characteristics at the trial level, and Panel B shows characteristics at the charge level. We have a total of 1,434 cases/defendants, representing 2,770 separate charges. 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). Across all cases, on average defendants are sentenced to 1,543 days in jail, though men are sentenced for significantly longer than women (1,795 versus 245 days).[5] Sixteen percent of our defendants are female, 48 percent are white, and the average age is 37.

2 Methods

In order to identify the effects of own-gender juries, we use a generalized difference-in-differences approach. Specifically, we estimate the following linear probability model:

$$(1)\ Convict_{ct} = \beta_1 DefFemale_t + \beta_2 E(PropFemale)_t + \beta_3 DefFemale_t * E(PropFemale)_t + X_t + County_t + CountyXCrime_{ct} + \varepsilon_{ct}$$

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, $E(Prop\ Female)_t$, the expected proportion of females seated on the jury for trial t, accounts for differences in the decision to convict due to the gender of jurors. The coefficient of interest, β_3, measures the effect of own-gender juries 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 gender. All specifications include county fixed effects along with county-by-crime fixed effects when considering more than one crime category. 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, because we also test for the presence of own-gender juries by crime severity (felony vs. misdemeanor), and by crime type (drug, driving, property, and violent), we also report False Discovery Rate (FDR) adjusted Q-values. 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.[6] These are

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

[6] While we also test for sentencing effects, the focus of the paper and therefore the multiple inference adjustment is on convictions. This is because in these counties, conviction is the only outcome over which juries have direct control. In these counties, sentencing decisions are made by judges based on those conviction outcomes.

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.

In addition, we also test whether the effect of own-gender juries on conviction translates to differences in sentencing, which is decided by a judge rather than the jury. For that reason, we focus primarily on the category of charges for which we find an effect on conviction. Due to the discrete nature of prison sentences, the presence of many zero observations, and the wide dispersion of sentence lengths, we estimate the effect of own-gender juries on the distribution of sentences using binary indicators. This is done at the trial level as the sentences for individual charges are often served concurrently. Formally we estimate the following ordinary least squares regression for each binary sentence length:

$$(2)\ At\ least\ X\ Days_t = \beta_1 DefFemale_t + \beta_2 E(PropFemale)_t + \beta_3 DefFemale_t * E(PropFemale)_t + X_t + County_t + CountyXCrime_t + \varepsilon_{ct}$$

where $At\ least\ X\ Days_t$ is a binary indicator for X days sentenced in trial t with X starting with at least 1 day and increasing by 6 month increments to 10 years. The covariates are defined as in the previous equation where β_3 is interpreted as the degree of own gender juries. We allow for correlation in errors among trials with the same defendant by clustering at the defendant level.

The intuition of this generalized difference-in-differences 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. Equivalently, we allow male defendants to be "more guilty" than female defendants, though we require that this difference in underlying guilt 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 juries. 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 juries. Similarly, if women are more likely than men to convict blacks, and if there is a higher proportion of black male defendants than black female defendants, then our estimated 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.

3 Results

3.1 Correlation between expected jury gender and actual jury gender

We begin by demonstrating that our measure of jury gender, which is the expected proportion of women on the jury based on the random potential juror assignments and orderings, 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 Figure 2, which graphs the actual proportion of women seated on the jury against the

expected proportion of women seated. It shows strong positive correlations for both 6-person juries and 12-person juries. In both cases the slope is close to one, suggesting that our (exogenous) measure of jury gender composition is strongly correlated with observed jury gender composition.

Regression results are shown in Table 2. Specifically, we estimate an equation of the same form as equation (1) above in that we regress the actual proportion of females on the predicted proportion of females, along with county-by-crime fixed effects. Results are consistent with Figure 2 in showing strong correlations between actual and expected gender composition. Column 1 shows a correlation of 0.949, significant at the 1 percent level, for all case types. 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 0.831 for driving cases to 1.031 for violent crime. All estimates are statistically significant at the 1 percent level. As a result, it is clear that the combination of 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.

3.2 Exogeneity tests of the measure of expected jury gender composition

The validity of our empirical approach depends in large part 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 proportion 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 proportion of women on the jury.

Results are shown in Table 3, with estimates at the trial level shown in Panel A, and at the charge level in Panel B. Overall, there is little evidence that these exogenous characteristics are correlated with our measure of expected jury gender composition. Of the 24 estimates shown, one is significant at the 10 percent level, and none are significant at the five percent level, which is

consistent with random chance. This contrasts with results from the same exercise using actual proportion 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 Appendix Table A2, eight of the 24 estimates are significant at the 10 percent level, and three are significant at the five percent level.[7] This reflects the fact that the actual proportion 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 3, 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 3a. The symbols represent local averages for charges against male and female defendants, and are grouped into 10 equal-sized bins. In addition, we fit separate lines to the underlying data for male and female defendants. Figure 3a 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 juries.

Results in Figure 3b show predicted conviction rates for drug charges, where we later show large effects of own-gender juries. Results are similar to Figure 3a in that while male defendants are predicted to have higher conviction rates than female defendants, this difference does not vary with expected jury gender. This is consistent with the identifying assumption, and suggests that any nonzero difference-in-difference estimate of the effect of jury gender will be due to the effect of jury gender, rather than some confounding factor.

[7] In cases where no jury is seated, we assign actual proportion female to be the expected proportion female. If we instead limit the sample to those trials in which jurors were seated, four estimates are significant at the 5 percent level.

3.3 Effect of own-gender juries 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 4 graphs the conviction rates of male and female defendants against the expected proportion of females on the jury. Results for all charges are shown in Figure 4a. It shows that the conviction rates of male defendants are relatively flat as the expected proportion of female jurors increases. By comparison, the conviction rates of females seem to decline somewhat as the expected proportion of female jurors increases, though the difference in slopes is relatively subtle.

Conviction rates for drug offenses are shown in panel b of Figure 4. Conviction rates for male defendants appear to increase somewhat as the expected proportion of female jurors increases. In contrast, conviction rates of female defendants decline sharply as the expected proportion of female jurors increases. The locally averaged conviction rates for female defendants facing juries with an expected proportion of females less than 0.5 range between 60 and 100 percent. By comparison, locally averaged conviction rates for juries expected to be more than half female range from 20 to 45 percent. In short, female defendants are much less likely to be convicted of a drug charge as the jury is more female, while if anything men are more likely to be convicted as the jury is more female.

Estimation results are shown in Table 4. All specifications control for the expected proportion 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 juries for all crimes. The coefficient is -0.228 and is not statistically significant. The magnitude of the coefficient implies that a 10 percentage point change in the expected gender of the jury is associated with a 2.88 percentage point reduction in the conviction rate.

Column 2 additionally controls for other defendant and case characteristics such as the defendant's age and race, judge gender, the number of charges in the case, and whether the defendant was also charged with a violent crime such as assault. Consistent with the identifying assumption, the coefficient changes little to -0.243 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. For example, if women convict at higher rates for all charges when the defendant is also charged with a violent crime, and if male defendants are more likely to be charged with violent crimes along with other crimes, then we can estimate a nonzero own-gender effect even if women apply this standard equally across all defendants. 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 proportion of female jurors. Specifically, we include interactions of the proportion of female jurors 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. This allows for the possibility that jurors are responding differentially to defendant characteristics that may be correlated with defendant gender.

Results from a specification that includes these pairwise interactions are shown in column 3 of Table 4. As shown there, the coefficient of interest becomes somewhat larger at -0.317, though is still statistically insignificant.

Columns 3 – 6 of Table 4 show results for felonies. Estimates range from -0.232 to -0.382, 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 -0.518 to -0.582 and none are statistically significant.

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 4. 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 0.816 to 0.33 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. Results are shown in Table 5. The format is similar to Table 4 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 juries affect convictions for driving, property, or violent crimes. In contrast, results in columns 10 - 12 indicate there is strong evidence of own-gender juries on conviction for drug charges. The estimate of -2.267 in column 10 suggests that a 10 percentage point change in the expected own-gender composition of the jury results in a 23 percentage point reduction in the conviction rate of defendants. Adding controls changes the estimate only slightly to -2.248, and further adding interaction controls results in an estimate of -1.926. All estimates are statistically significant at the one percent level. More importantly, FDR-adjusted Q-values are 0.001, 0.001, and 0.041, respectively. This indicates that even after accounting for the multiple statistical tests across the six major categories of crime charges in Tables 4 and 5, the coefficients in columns 10 – 12 of Table 5 are sufficiently extreme as to be unlikely to arise due to chance.

To put these estimates in perspective, we note that Anwar, Bayer, and Hjalmarsson (2012) estimate that the impact of having one black potential juror in the jury pool (and thus likely less than a 10 percentage point increase in the expected proportion of jurors that are black) results in a 16 percentage point reduction in the conviction rates for black defendants.

3.4 Effects of own-gender juries on sentencing decisions

Next, we turn to the question of whether own-gender juries affect sentencing. While one may expect increased convictions to result in additional incarceration, we note that this link is *a priori* ambiguous for two reasons. The first is that the additional convictions may be for charges that do not result in incarceration. In addition, while juries make conviction decisions, in these counties judges decide sentencing. On the one hand, if judges treat all convictions similarly, we would expect to observe own-gender effects on sentencing for drug cases. On the other hand, if judges exercise discretion in sentencing based on either the facts of the case or even on the gender composition of the jury, we may not see evidence of own-gender effects in sentencing outcomes.

Results are shown in Figure 5, with panels a and b showing results for all cases and drug cases, respectively. Each panel shows estimates of the effect of own-gender juries in which the outcome of interest is whether the defendant was sentenced for at least one day, at least six months,

at least one year, at least 18 months, etc., up to at least 10 years. Results for all cases shown in Figure 5a indicate that while there is some evidence that own-gender juries resulted in reduced sentences – especially on the left-hand side of the distribution – none of the estimates are statistically significant.

Results in Figure 5b indicate there is a statistically significant decline in the likelihood of receiving a sentence of at least one day. Estimates for the effect on longer sentences are positive but not statistically significant. This suggests that juries are less likely to convict those own-gender defendants who might otherwise be convicted and sentenced to relatively short sentences.

These results are shown more formally in Table 6, which shows estimates of the effect of own-gender juries on the probability of being sentenced to at least some jail time. Consistent with Figure 5a, estimates in columns 1 – 3 for all charges are negative but not statistically significant. In contrast, estimates for cases that include at least one drug charge shown in columns 4 – 6 range from -1.363 to -1.631, and are all statistically significant at the five percent level. These estimates imply that a 10 percentage point change in the expected gender composition of the jury results in a 13 to 16 percentage point change in the likelihood of being sentenced to jail or prison.

These findings have several important implications. First, they suggest that own-gender juries do lead to differences in sentencing outcomes, even when sentencing decisions are made by judges. This mean that judges are either unwilling or unable to exercise discretion in an effort to offset the effect of jury gender composition on conviction decisions. In addition, the effects on sentencing imply that not only does drawing an opposite-gender jury lead to a criminal record, but it also leads to increased incarceration. Existing research on the effect of conviction and incarceration on recidivism and employment suggests that this results in significant long-term harm to defendants on drug charges (Doyle and Aizer 2012; Mueller-Smith 2015).

4 Robustness

As discussed earlier, a major threat to identification of own-gender jury effects 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 test for this by 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. Results in column 12 of Table 5 indicate our estimates are robust to the inclusion of these interactions, which provides evidence that the effects are due to the interaction of jury and defendant gender and not something else. 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 7, where column 1 replicates our main estimate for drug charges of -1.926 from column 12 of Table 5. Results in column 2 show that including these interactions increases the magnitude of the estimate to -2.058. 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.

In addition, we also test the robustness of our estimates to different specifications as well as to alternative ways of constructing our predicted jury gender measure. In column 3 of Table 7 we estimate the effect controlling for predicted juror age, which we only observe in Hillsborough County. The estimate is similar at -2.390, and is significant at the one percent level. Column 4 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 6.1 percent of the drug charges. The magnitude of the estimate is reduced slightly to -1.793, but is still statistically significant at the 5 percent level.

In columns 5 – 7 of Table 7, we estimate the own-gender jury effect when we classify the gender of potential jurors differently. Specifically, we classify jurors for whom we could not identify gender using genderize.io as either all female (column 5) or all male (column 6), respectively, rather than as having an equal likelihood of being male as female. Estimates are similar in magnitude and significance at -1.327 and -2.252, respectively. In addition, in column 7 we classify the gender of jurors based on the names and genders recorded in Florida by the Social Security Administration. The resulting estimate is -1.908, which is similar to the baseline estimate of -1.926.

Finally, in columns 8 – 10 of Table 7 we show that our estimate of own-gender juries is robust to alternative methods of predicting jury gender. In column 8 we estimate the effect when

we do not smooth the probability of being seated on the jury for a given jury and panel size using a local linear estimation with epanechnikov kernel, as we did for our main results. Instead, we use the raw probability that a juror assigned that number in a panel in a given range was seated on the jury. The estimate is -1.930 and is significant at the one percent level. The estimated effect is also similar if we use probit instead of local linear estimation, as shown by the estimate of -1.942 in column 9. The same is true when we use a local linear smoother but do not condition on jury panel size (-1.794), as shown in column 10.

In summary, we find no evidence that our estimated effect of own-gender juries 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 defining the outcome and predicting jury gender.

5 Discussion and Interpretation

There are several potential mechanisms through which own-gender juries could have such large effects on conviction and sentencing outcomes. The first is that seated jurors may exhibit own-gender bias when making conviction decisions on drug charges. 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. But under this interpretation, the results would be due to male and/or female jurors being either too lenient to own-gender defendants, being too tough (i.e., wrongfully convicting) on 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. However, we note that doing so would violate the legal

standard set by Batson v. Kentucky (1986) and J.E.B. v. Alabama (1994). In addition, the fact that predicted jury gender is so highly correlated with actual jury gender provides empirical evidence that the attorneys are unable to significantly offset random changes in expected jury gender.

Data limitations make it difficult for us to distinguish between these potential mechanisms with any certainty. However, to shed some light on this question, we estimate effects separately for cases that did and did not get to trial.[8] Results are shown in Appendix Table A3, which shows that both sets of point estimates are statistically significant at the one percent level. However, the magnitude of the effect for charges decided by jury is twice as large as the effect when the case was decided prior to the conclusion of the trial. We interpret this as suggestive evidence that effects are largely driven by changes during or after the trial, such as gender bias by juries. We note, however, that selection into whether a case goes to trial after the jury panel is assigned makes it difficult to interpret these differences with certainty.

A second question regarding the interpretation of this study's findings relates to the strength of the effects for drug charges compared to driving, property, and violent crime. Unfortunately, our data are not well-suited for explaining this difference across crime types. We speculate it is because even though Americans are supportive of existing and even stronger penalties for DUIs, violent crime, and property crime, Americans are critical of the prosecution of drug crimes. For example, recent surveys indicate that 40 percent of Americans believe the prison sentences for non-violent drug crimes are too harsh, and 64 percent support the full legalization of marijuana (YouGov/Huffington Post, 2015; Gallup News Service, 2017). Two-thirds of American adults believe the government should focus more on treatment for illegal users, compared to only 26 percent who believe more focus should be on prosecuting illegal users (Pew Research Center, 2014). A nontrivial proportion of Americans even disagree with the prosecution of "harder" drug crimes; 16 percent favor decriminalization of cocaine possession, and 9 percent favor legalization (Morning Consult, 2016). This shift in attitudes on drug laws is also reflected in recent state policy changes regarding drug possession.[9] These views are particularly relevant given the drug charges

[8] For this analysis we exclude the 41 cases representing 55 charges where the records did not indicate whether the case was decided by trial or prior to the start of the trial.

[9] The National Conference of State Legislatures (NCSL) reports that from 2011 to 2016, at least nine states have lowered some drug possession crimes from felonies to misdemeanors, and another nine have reduced mandatory sentences for some drug offenders (NCSL, 2016). In addition, as of 2018 over 20 states have decriminalized certain marijuana possession offenses (NORML, 2018). While Florida is not among the states making these changes, jurors

in our sample, over 65 percent of which are for possession of drugs or drug paraphernalia without intent to distribute.

In contrast, there is little to no public support for weakening the enforcement of non-drug laws, and significant support for even strengthening enforcement. While surveys of Americans' perceptions of non-drug offense prosecution are less common, what evidence there is contrasts sharply with views on drug crime enforcement. For example, only 11 and 1 percent of adults believe that the sentences typically given for non-violent property crimes and violent crimes, respectively, are too harsh (Huffington Post/YouGov, 2013). As a result, we interpret this study's findings as most consistent with a model in which jurors fairly enforce the laws with which they mostly agree, but disproportionately favor own-group defendants when deciding whether to enforce laws with which they might not agree. That is, while a juror may be willing and able to convict out-group defendants who break a law with which the juror disagrees, she is perhaps less willing to convict in-group defendants of the same crime. We emphasize, however, that there could be other explanations for the difference in results for across crime types.

6 Conclusion

In this study, we test for the effect of own-gender juries on conviction and sentencing outcomes. To overcome potential bias due to nonrandom jury selection, we exploit the fact that potential jurors are randomly assigned to jury pools for each case, and are randomly ordered within each jury pool. This enables us to predict the gender composition of each jury for each case set to go to trial, thereby isolating the as-good-as-random variation in jury gender. We combine this variation with variation in defendant gender to estimate the effect of own-gender juries.

Results provide strong evidence that own-gender juries result in lower conviction rates for drug offenses. We estimate that a ten percentage point change in the expected own-gender composition of the jury results in a 19 percentage point decline in conviction rates on drug charges. A similar change in jury gender results in a 13 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

there are likely experiencing similar shifts in their views about drug laws.

prior research on the effect of juror race (Anwar, Bayer, and Hjalmarsson, 2012).[10]

We hypothesize that the reason we see such strong own-gender effects for drug charges but not others is because many Americans disapprove of the prosecution of drug crimes. We emphasize, however, that we cannot rule out other interpretations. Similarly, while we show evidence that effects are largest for cases that go to trial, it is difficult for us to determine which part of our effect is due to gender bias by jurors when deciding to convict, and what is due to changes in the offering or acceptance of plea deals based on perceptions of jury bias.

Our results are important for the debate over the use of peremptory challenges in selecting a jury. By documenting the significant harm that can arise to defendants who draw opposite-gender juries, we highlight the potential benefits to the prosecution of removing same-gender individuals from the jury pool. Similarly, defendants in drug cases stand to benefit greatly if their attorneys are able to successfully remove opposite-gender jurors from the jury pool. As a result, our results 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 documenting own-gender bias in decision-making. Our findings suggest that such bias can arise even in settings where the objective of impartiality is heavily emphasized and protected. Specifically, throughout the juror selection process the necessity of being impartial and fair is actively pressed on potential jurors. 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 has a significant effect on conviction and sentencing outcomes.

[10] They find that one black individual in the jury pool – and thus in expectation much less than one black juror on the seated jury – results in a 16 percentage point change in conviction rates.

References

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

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

Anderson, Michael (2008). 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*, 103(484): 1481-1495.

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

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

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

Ayres, Ian, and Peter Siegelman (1995). Race and Gender Discrimination in Bargaining for a New Car. *American Economic Review*, 85(3): 304-321.

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

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review*, 107(4): 1207-1238.

Batson v. Kentucky, 476 US 79 (1986).

Bindler, Anna, and Randi Hjalmarsson (2017). The Persistence of the Criminal Justice Gender

Gap: Evidence from 200 Years of Judicial Decisions. Working paper.

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

Butcher, Kristin F., Kyung H. Park, and Anne Morrison Piehl (2017). Comparing Apples to Oranges: Differences in Women's and Men's Incarceration and Sentencing Outcomes. *Journal of Labor Economics*, 35(1): 201-234.

Cohen, Alma, and Crystal S. Yang (forthcoming). Judicial Politics and Sentencing Decisions. Forthcoming in *American Economic Journal: Economic Policy*.

Dahl, Gordon B., and Enrico Moretti (2008). The Demand for Sons. *Review of Economic Studies*, 75: 1085-1120.

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

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

Eren, Ozkan, and Naci Mocan (forthcoming). Emotional Judges and Unlucky Juveniles. Forthcoming in *American Economic Journal: Applied Economics*.

Gallup News Service (2017). Record-High Support for Legalizing Marijuana use in U.S. Retrieved from http://news.gallup.com/poll/221018/record-high-support-legalizing-marijuana.aspx on March 8, 2018.

George, Tracey E (2001). Court Fixing. *Arizona Law Review*, 43(1): 9-62.

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

Huffington Post/YouGov (2013). Retrieved from http://big.assets.huffingtonpost.com/toplines_sentences_0813142013.pdf on March 10, 2018.

J.E.B. v. Alabama, 511 US 127 (1994).

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

Knepper, Matthew (forthcoming). When the Shadow is the Substance: Judge Gender and the Outcomes of Workplace Sex Discrimination Cases. Forthcoming in *Journal of Labor Economics*.

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

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

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

Morning Consult (2016). National Tracking Poll 160304. Retrieved from https://cdn0.vox-cdn.com/uploads/chorus_asset/file/6189021/Morning_Consult_Vox_drug_poll.0.pdf on March 10, 2018.

Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman (2012). Science Faculty's Subtle Gender Biases Favor Male Students. *Proceedings of the National Academy of Science* 109(41): 16474–79

National Conference of State Legislatures (2016). Drug Sentencing Trends. Retrieved from http://www.ncsl.org/research/civil-and-criminal-justice/drug-sentencing-trends.aspx on March 10, 2018.

National Organization for the Reform of Marijuana Laws (NORML) (2018). States That Have Decriminalized. Retrieved from http://norml.org/aboutmarijuana/item/states-that-have-decriminalized on March 10, 2018.

Neumark, David, Roy J. Bank, and Kyle D. Van Nort (1996). Sex Discrimination in Restaurant Hiring: An Audit Study. *Quarterly Journal of Economics*,111(3): 915-41.

Ouss, Aurelie, and Arnaud Philippe (forthcoming). No Hatred or Malice, Fear or Affection: Media and Sentencing. Forthcoming in *Journal of Political Economy*.

Pew Research Center (2014). America's New Drug Policy Landscape. Retrieved from http://www.people-press.org/2014/04/02/section-1-perceptions-of-drug-abuse-views-of-drug-policies/ on March 10, 2018.

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

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

Stanley, T.D., and Stephen B. Jarrell (1998). Gender Wage Discrimination Bias? A Meta-Regression Analysis. *Journal of Human Resources*, 33(4): 947-973.

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

YouGov/Huffington Post (2015). "Poll Results: Commuting Sentences." Retrieved from https://today.yougov.com/news/2015/04/06/poll-results-commuting-sentences/ on March 10, 2018.

Figures and Tables

Figure 1a: P(seated) for jury trials with 6 jurors

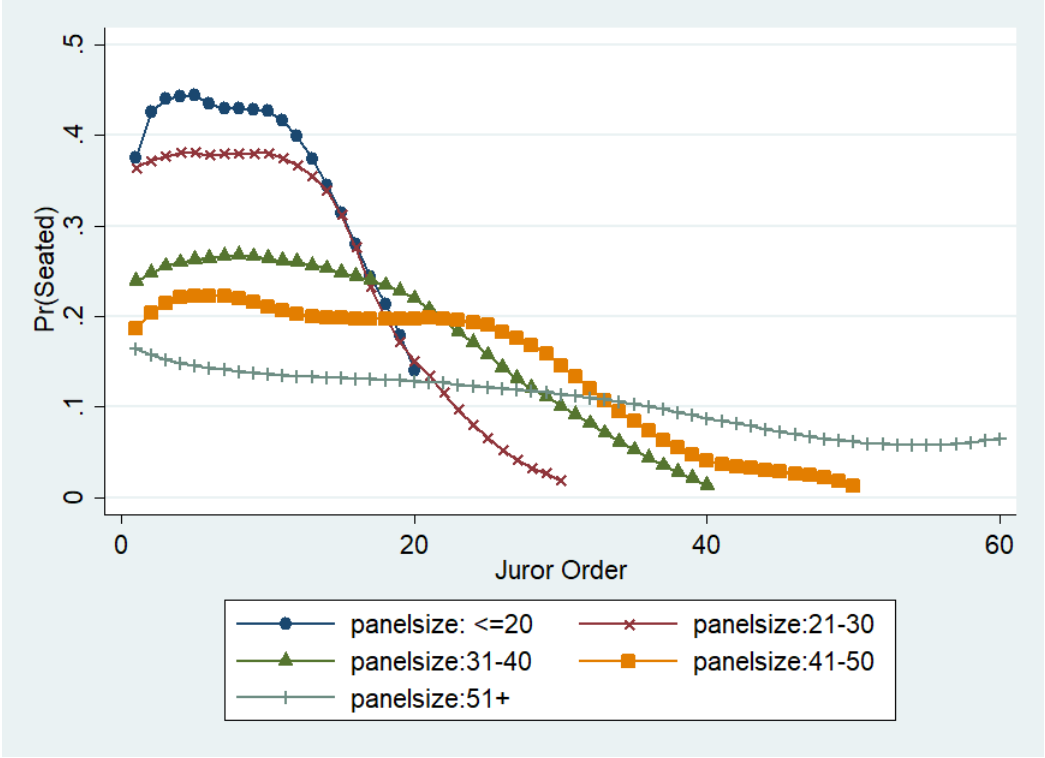
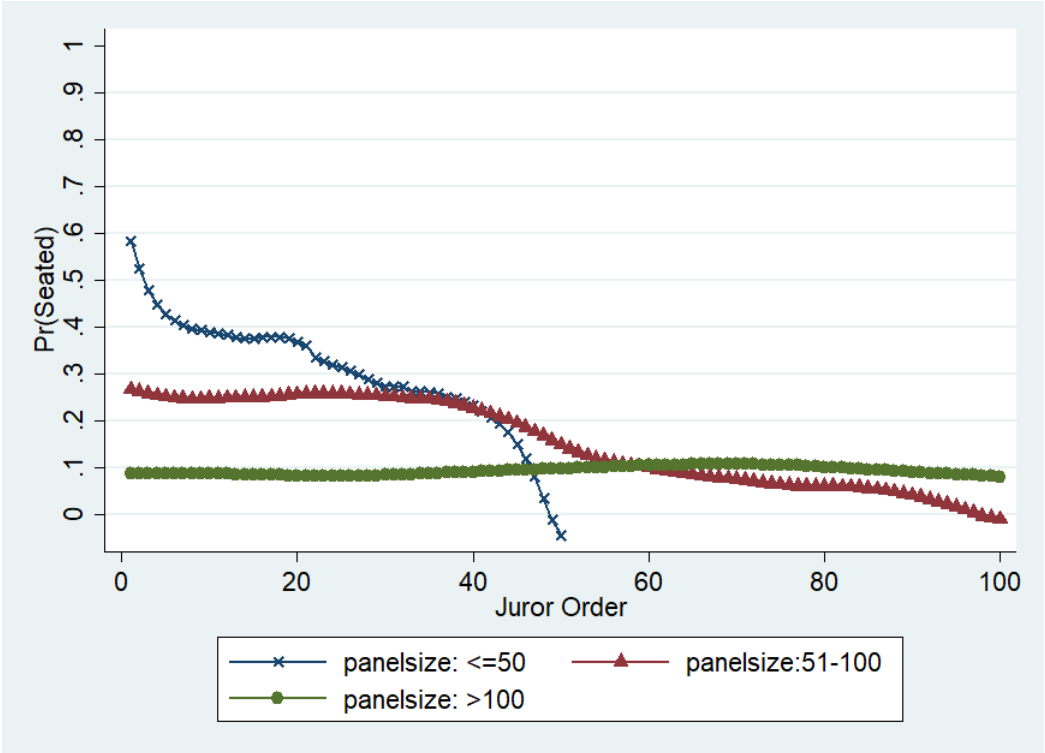


Figure 1b: P(seated) for jury trials with 12 jurors



Notes: Each line is fit with a local linear polynomial at each panelist position using a epanechnikov kernel with varying Rule-of-Thumb (ROT) bandwidths. Figure 1a from smallest to largest panelsize uses a one-sided bandwidth of 1,1,2,3, and 7. Figure 1b from smallest to largest panelsize uses a one-sided bandwidth of 4, 5, and 14.

Figure 2: Correlation between actual jury gender composition and expected gender composition

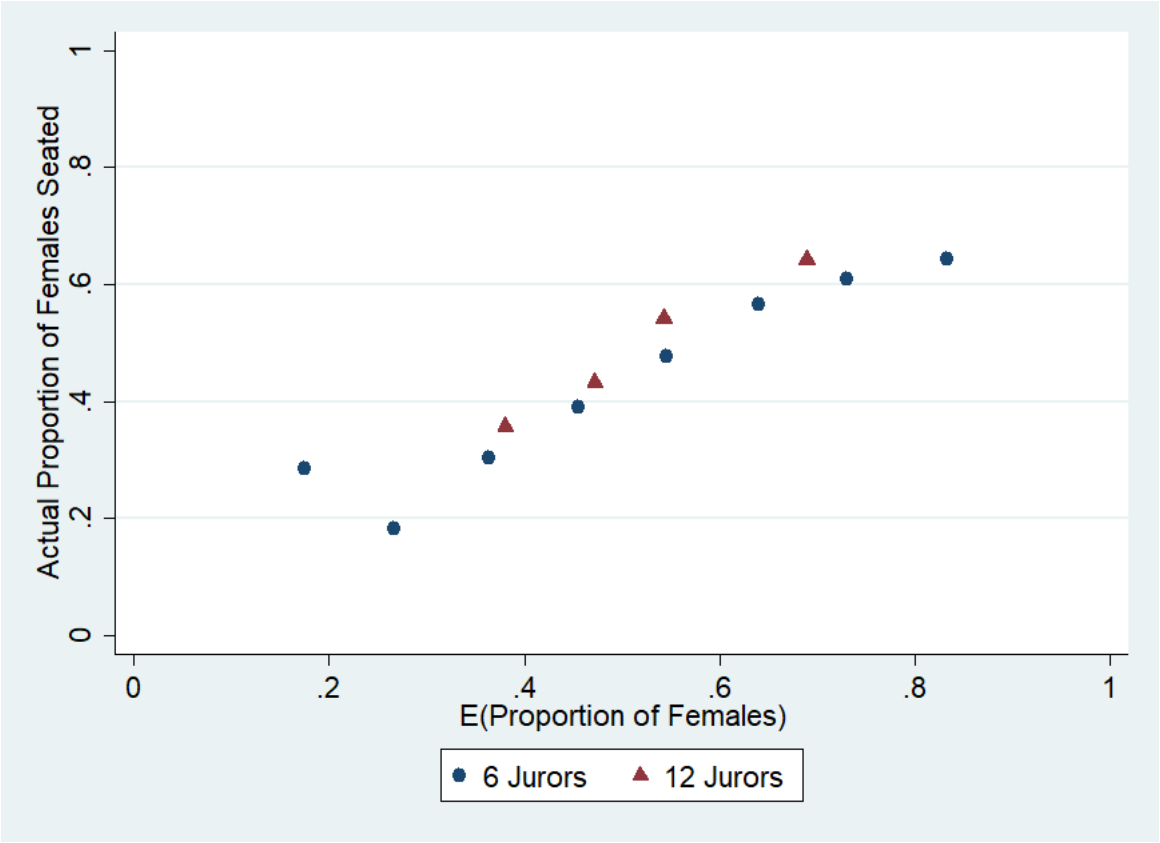


Figure 3a: Predicted conviction rates for male and female defendants - all charges

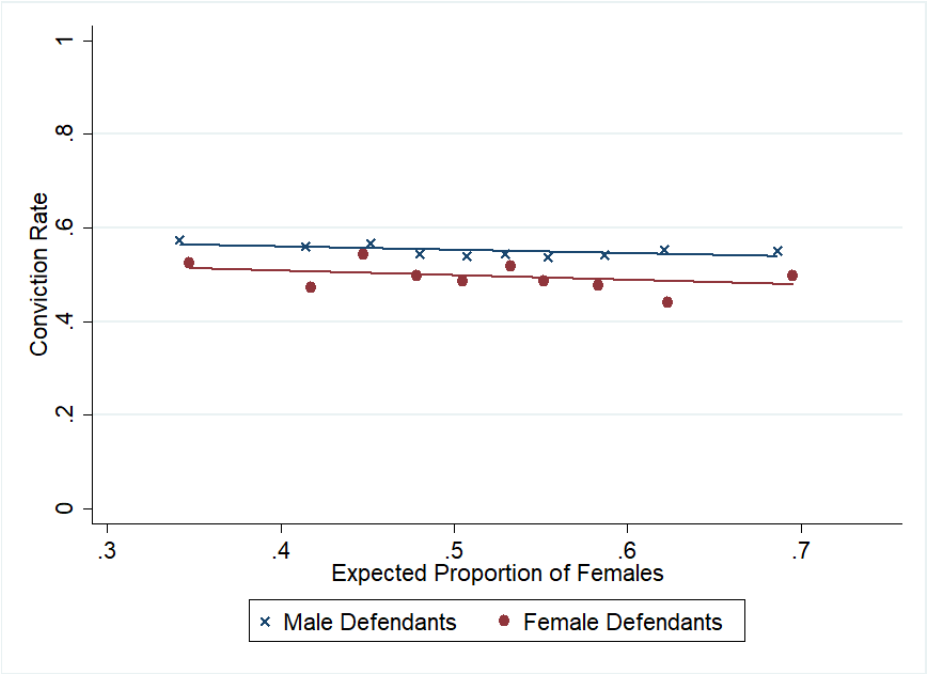
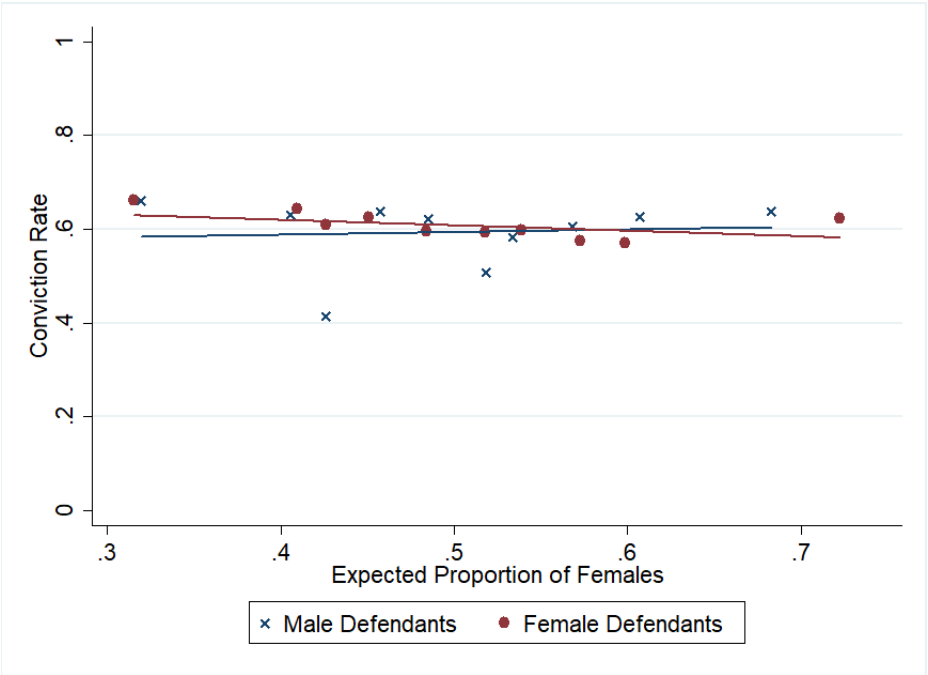


Figure 3b: Predicted conviction rates for male and female defendants - drug charges only



Notes: For each charge, we predict the probability of conviction using all observable characteristics. The line represents a linear fit across all predicted conviction rates.

Figure 4a: Actual conviction rates for male and female defendants - all charges

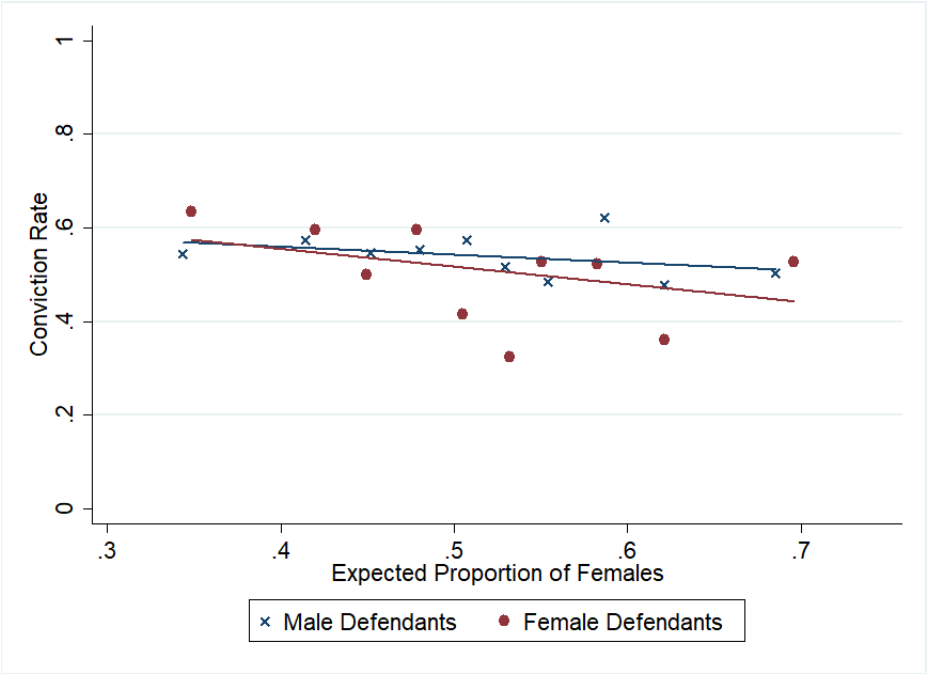
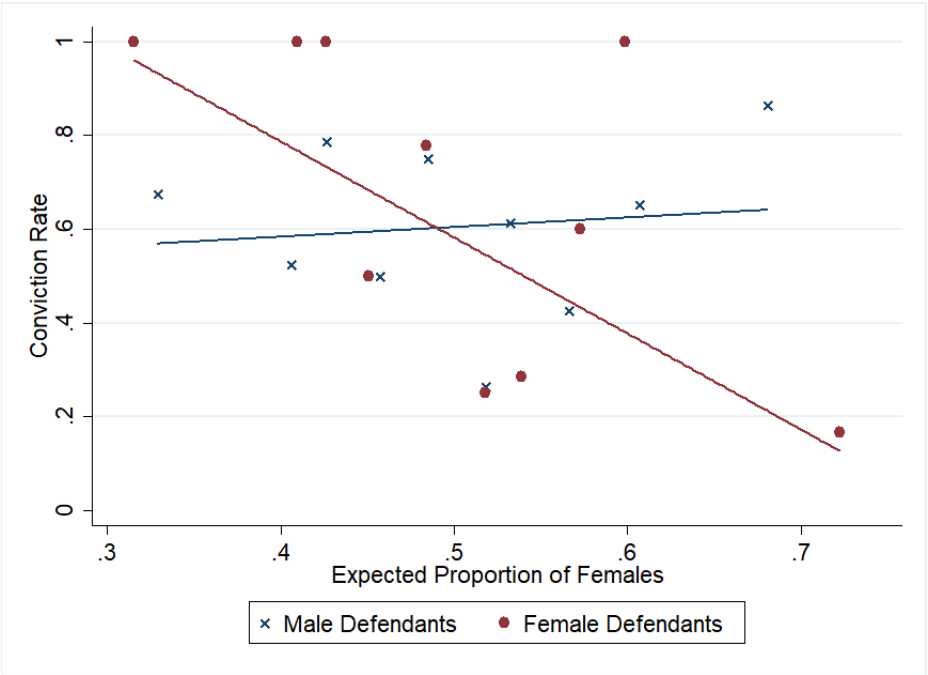


Figure 4b: Actual conviction rates for male and female defendants - drug charges only



Notes: Each figure graphs the actual conviction rates for male and female defendants against the expected gender composition of the jury. Observations are grouped such that each circle represents an equal number of charges.

Figure 5a: Estimated effects of own-gender juries on sentencing - all cases

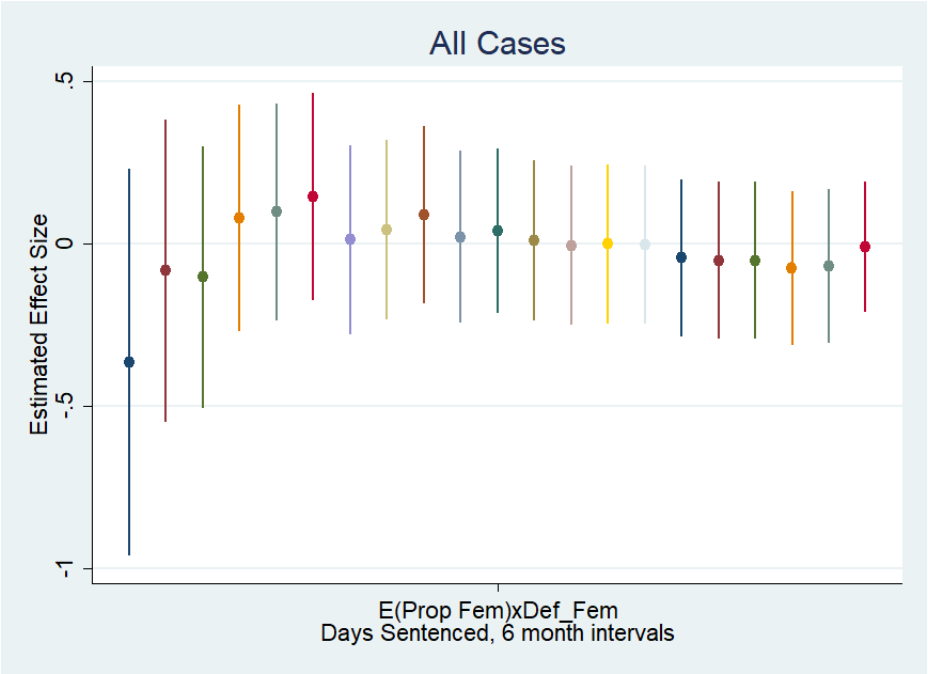
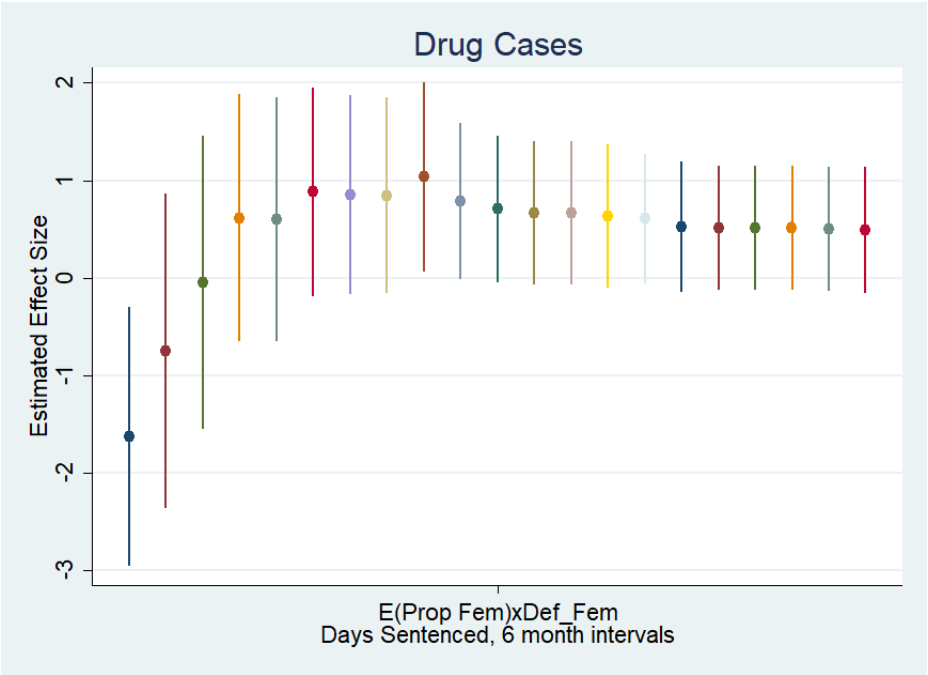


Figure 5b: Estimated effects of own-gender juries on sentencing - drug-related cases only



Notes: Each estimate shown represents the effect of own-gender juries on total sentencing in the case. 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. Figure 4a includes all drug, driving, property, and violent crime cases. Figure 4b restricts to cases with at least one drug charge.

Table 1: Summary Statistics

Panel A: By Case

	All	Male	Female	Felony	Misdem.	Driving	Property	Violent	Drug
Outcomes									
Conviction Rate	0.66	0.67	0.63	0.67	0.65	0.75	0.75	0.59	0.76
Total days sentenced (including zeros)	1542.56 (4932.62)	1785.44 (5320.29)	244.72 (1055.47)	2184.88 (5806.43)	105.77 (612.22)	452.37 (3229.67)	2311.49 (5644.33)	2502.13 (6805.12)	866.60 (1806.30)
P(sentenced $\geq$ 1 days)	0.42	0.45	0.28	0.49	0.26	0.35	0.54	0.42	0.51
P(sentenced $\geq$ 1 years)	0.26	0.29	0.08	0.36	0.03	0.10	0.40	0.29	0.35
P(sentenced $\geq$ 5 years)	0.16	0.18	0.04	0.22	0.01	0.03	0.25	0.21	0.14
Case Characteristics									
Defendant female	0.16	0.00	1.00	0.13	0.21	0.19	0.17	0.13	0.13
Defendant white	0.48	0.46	0.57	0.43	0.57	0.66	0.42	0.41	0.37
Defendant's age	36.86 (12.52)	37.16 (12.79)	35.23 (10.80)	36.56 (12.70)	37.52 (12.10)	37.69 (12.13)	36.32 (13.34)	36.09 (12.72)	37.35 (11.40)
Number of Charges	2.34 (2.12)	2.43 (2.25)	1.88 (1.11)	2.49 (2.22)	2.03 (1.83)	2.52 (2.18)	2.95 (2.94)	2.26 (1.86)	2.90 (2.70)
Violent charge in case	0.47	0.49	0.38	0.56	0.26	0.07	0.26	1.00	0.11
Felony charge in case	0.69	0.71	0.58	1.00	0.00	0.39	0.83	0.82	0.74
Judge female	0.33	0.32	0.38	0.37	0.24	0.29	0.30	0.38	0.35
Jury Characteristics									
Actual Prop Female	0.45 (0.24)	0.44 (0.24)	0.48 (0.24)	0.44 (0.25)	0.46 (0.23)	0.46 (0.23)	0.44 (0.25)	0.44 (0.24)	0.43 (0.26)
E(Proportion Female)	0.52 (0.10)	0.51 (0.10)	0.52 (0.11)	0.51 (0.10)	0.52 (0.10)	0.52 (0.11)	0.52 (0.09)	0.51 (0.10)	0.51 (0.11)
Predicted Average Juror Age (only in Hillsborough)	44.99 (3.51)	45.09 (3.52)	44.41 (3.43)	45.01 (3.58)	44.90 (3.23)	44.55 (3.39)	45.15 (3.25)	45.05 (3.56)	44.84 (3.62)
Observations	1434	1210	224	987	447	392	351	659	240

Panel B: By Charges

Outcomes									
Conviction Rate	0.54	0.54	0.51	0.56	0.47	0.52	0.58	0.49	0.60
Case Characteristics									
Defendant female	0.14	0.00	1.00	0.12	0.20	0.18	0.12	0.13	0.14
Defendant white	0.50	0.48	0.61	0.43	0.66	0.70	0.44	0.39	0.46
Defendant's age	37.17 (13.14)	37.27 (13.49)	36.57 (10.76)	36.41 (13.25)	38.89 (12.73)	38.13 (12.44)	36.76 (14.62)	35.55 (12.35)	39.59 (13.03)
Number of Charges	4.09 (4.70)	4.35 (4.99)	2.52 (1.41)	4.25 (4.62)	3.73 (4.87)	3.52 (3.11)	5.42 (6.06)	3.08 (3.33)	5.06 (6.13)
Violent charge in case	0.41	0.42	0.37	0.52	0.17	0.09	0.18	1.00	0.08
Felony charge in case	0.69	0.71	0.57	1.00	0.00	0.29	0.89	0.86	0.74
Judge Female	0.32	0.31	0.41	0.34	0.27	0.29	0.27	0.37	0.33
Jury Characteristics									
Actual Prop Female	0.45 (0.24)	0.45 (0.24)	0.45 (0.24)	0.44 (0.25)	0.46 (0.22)	0.46 (0.22)	0.46 (0.24)	0.43 (0.24)	0.42 (0.27)
E(Proportion Female)	0.52 (0.10)	0.52 (0.10)	0.52 (0.10)	0.51 (0.10)	0.52 (0.10)	0.52 (0.10)	0.52 (0.09)	0.52 (0.10)	0.50 (0.10)
Predicted Average Juror Age (only in Hillsborough)	44.85 (3.48)	44.92 (3.49)	44.40 (3.43)	44.85 (3.52)	44.83 (3.23)	44.55 (3.27)	44.63 (3.62)	45.08 (3.40)	44.81 (3.58)
Observations	2770	2378	392	1919	851	753	660	920	450

Table 2: 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(Proportion Female)	0.949***	0.921***	1.010***	0.830***	0.892***	1.031***	0.937***
	(0.059)	(0.076)	(0.091)	(0.114)	(0.149)	(0.082)	(0.154)
Observations	1434	987	409	392	351	659	240
F stat	29	16	22	31	22	79	23

Notes: 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.
*p<0.10, **p<0.05, ***p<0.01

Table 3: Exogeneity Tests

Panel A: Trial-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
E(Prop Fem)	0.124	0.135	-3.547	-1.300	-0.703	-0.057	-0.039	-0.030	0.021	0.179*	-0.046	-0.123
	(0.100)	(0.125)	(3.218)	(1.076)	(3.359)	(0.119)	(0.402)	(0.117)	(0.116)	(0.105)	(0.130)	(0.099)
Observations	1434	1434	1434	831	1434	1434	1434	1434	1434	1434	1434	1434
Panel B: Charge-Level												
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
E(Prop Fem)	0.122	0.138	-3.149	-1.106	-0.719	-0.023	0.079	-0.054	0.026	0.151	-0.026	-0.132
	(0.101)	(0.126)	(3.243)	(1.083)	(3.412)	(0.121)	(0.427)	(0.119)	(0.109)	(0.094)	(0.127)	(0.089)
Observations	2770	2770	2770	1445	2770	2770	2770	2770	2770	2770	2770	2770

Notes: Each column in each panel reports estimates from a separate regression in which we regress observable characteristics on the expected proportion of females on the jury. 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.
*p<0.10, **p<0.05, ***p<0.01

Table 4: Effect of own-gender juries on conviction rates, by severity

	All Charges			Felony Charges			Misdemeanor Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
E(Prop Fem)xDef_Fem	-0.228	-0.243	-0.317	-0.232	-0.255	-0.382	-0.572	-0.582	-0.518
	(0.310)	(0.310)	(0.311)	(0.426)	(0.426)	(0.412)	(0.411)	(0.412)	(0.424)
Adjusted FDR-Q values				[0.705]	[0.816]	[0.532]	[0.33]	[0.351]	[0.445]
Observations	2770	2770	2770	1538	1538	1538	1232	1232	1232
Mean Dependent Variable	0.54	0.54	0.54	0.54	0.54	0.54	0.52	0.52	0.52
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

Notes: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 age, the number of charges in the case, and indicators for defendant's race, judge's gender, and whether there was charge for a violent crime in the case. Interactions include controls for each of those characteristics interacted with the expected proportion of female jurors.
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.
*p<0.10, **p<0.05, ***p<0.01

Table 5: Effect of own-gender juries on conviction rates, by crime type

	Driving Charges			Property Charges			Violent Charges			Drug Charges		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
E(Prop Fem)xDef_Fem	0.085	0.013	-0.053	0.975	0.916	0.863	-0.300	-0.212	-0.254	-2.267***	-2.248***	-1.926***
	(0.509)	(0.504)	(0.538)	(0.667)	(0.675)	(0.642)	(0.507)	(0.514)	(0.503)	(0.577)	(0.572)	(0.704)
Adjusted FDR-Q values	[0.868]	[0.98]	[0.922]	[0.33]	[0.351]	[0.445]	[0.705]	[0.816]	[0.737]	[0.001]	[0.001]	[0.041]
Observations	753	753	753	660	660	660	920	920	920	450	450	450
Mean Dependant Variable	0.55	0.55	0.55	0.53	0.53	0.53	0.50	0.50	0.50	0.64	0.64	0.64
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

Notes: All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant age, the number of charges in the case, and indicators for defendant's race, judge's gender, and whether there was charge for a violent crime in the case. Interactions include controls for each of those characteristics interacted with the expected proportion of female jurors.

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.

*p<0.10, **p<0.05, ***p<0.01

Table 6: Effect of own-gender juries on being sentenced to jail

	All Charges			Drug Charges		
	(1)	(2)	(3)	(4)	(5)	(6)
E(Prop Fem)xDef_Fem	-0.365	-0.297	-0.354	-1.631**	-1.493**	-1.363**
	(0.304)	(0.296)	(0.298)	(0.674)	(0.747)	(0.645)
Observations	1421	1421	1421	237	237	237
Mean Dependant Variable	0.41	0.41	0.41	0.49	0.49	0.49
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

Notes: All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant age, the number of charges in the case, and indicators for defendant's race, judge's gender, and whether there was charge for a violent crime in the case. Interactions include controls for each of those characteristics interacted with the expected proportion of female jurors and defendant's gender.
Standard errors are in parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table 7: Robustness of estimates of own-gender juries on conviction rates - drug charges only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
E(Proportion Females)	-1.926***	-2.058***	-2.390***	-1.793**	-1.327**	-2.252***	-1.908***	-1.930***	-1.942***	-1.794**
	(0.704)	(0.751)	(0.885)	(0.753)	(0.650)	(0.722)	(0.691)	(0.697)	(0.682)	(0.694)
Observations	450	450	291	450	450	450	450	450	450	450
Mean Dependant Variable	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64
Def & Jury Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drug type Interactions	No	Yes	No	No	No	No	No	No	No	No
Juror Age Control & Interaction	No	No	Yes	No	No	No	No	No	No	No
Adjudication Withheld=Not Guilty	No	No	No	Yes	No	No	No	No	No	No
Missing genders	half	half	half	half	female	male	half	half	half	half
Predicted genders	API	API	API	API	API	API	SS	API	API	API
Pr(Seated)	LL	LL	LL	LL	LL	LL	LL	Raw	Probit	LL
Pr(Seated—panelsize)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Notes: Standard errors are in parentheses and are clustered at the defendant level.
*p<0.10, **p<0.05, ***p<0.01

Appendix (For Online Publication)

Table A1: Effect of own-gender juries on conviction rates

	All	Felony	Misdemeanor	Drug	Driving	Property	Violent
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
E(Prop Fem)	-0.069	-0.122	0.058	0.201	-0.376	-0.088	0.138
	(0.115)	(0.148)	(0.162)	(0.197)	(0.252)	(0.184)	(0.285)
Observations	2770	1538	1232	753	660	920	450

Notes: 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.
*p<0.10, **p<0.05, ***p<0.01

Table A2: Exogeneity tests with actual proportion of female jurors

Panel A: Trial-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 Proportion Female	0.121**	0.140**	-1.307	-1.123*	7.677	0.028	0.034	-0.002	0.048	0.060	-0.135*	-0.002
	(0.050)	(0.069)	(1.890)	(0.610)	(6.649)	(0.065)	(0.288)	(0.065)	(0.062)	(0.058)	(0.071)	(0.053)
Observations	1434	1434	1434	831	1434	1434	1434	1434	1434	1434	1434	1434
Panel B: Charge-Level												
	female	white	age	avg juror age	panel size	judge female	number charges	felony	driving	property	violent	drug
Actual Prop Female	0.121**	0.136*	-0.971	-1.098*	7.495	0.031	-0.036	0.004	0.045	0.094*	-0.112	-0.025
	(0.050)	(0.069)	(1.898)	(0.606)	(6.833)	(0.065)	(0.291)	(0.065)	(0.060)	(0.052)	(0.069)	(0.047)
Observations	2770	2770	2770	1445	2770	2770	2770	2770	2770	2770	2770	2770

Notes: Each column in each panel reports estimates from a separate regression in which we regress observable characteristics on the actual proportion of females on the seated jury. 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.
*p<0.10, **p<0.05, ***p<0.01

Table A3: Effect of own-gender juries on conviction rates, by jury trial status

	Non-Trial			Jury Trial		
	(1)	(2)	(3)	(4)	(5)	(6)
E(Prop Fem)xDef_Fem	-2.588***	-2.716***	-3.519***	-5.735***	-6.006***	-4.932**
	(0.723)	(0.726)	(0.941)	(1.270)	(1.262)	(2.184)
Observations	161	161	161	152	152	152
Mean Dependant Variable	0.64	0.64	0.64	0.64	0.64	0.64
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

Notes: All specifications include controls for defendant gender and expected gender composition of the jury, as well as county fixed effects. Additional controls include defendant age, the number of charges in the case, and indicators for defendant's race, judge's gender, and whether there was charge for a violent crime in the case. Interactions include controls for each of those characteristics interacted with the expected proportion of female jurors.
Standard errors are in parentheses and are clustered at the defendant level.
*p<0.10, **p<0.05, ***p<0.01