

# Judge Peer Effects in the Courthouse \*

Ozkan Eren

University of California, Riverside

Naci Mocan

Louisiana State University and NBER

August 2022

## Abstract

We investigate whether consequential decisions made by judges are impacted by the gender composition of these judges' peer group. We utilize the universe of decisions on juvenile defendants in each courthouse in Louisiana over fifteen years. Leveraging random assignment of cases to judges, and variations in judge peer composition generated by elections, retirements, deaths and resignations, we show that an increase in the proportion of female peers in the courthouse causes a rise in individual judges' propensity to incarcerate, and an increase in the assigned sentence length. This effect is driven by female judges. We also demonstrate that the impact of the proportion of female peers is not a proxy for other peer characteristics such as race, party affiliation, age or tenure. Further analysis suggests that this behavior of female judges is unlikely to be a reflection of an effort to conform to evolving norms of judicial stringency, measured by peers' harshness in sentencing, but that it is due to the sheer exposure to female colleagues.

---

\*Eren: Department of Economics, University of California, Riverside, Riverside, CA 92521 (e-mail: ozkane@ucr.edu). Mocan: E. J. Ourso College of Business, Department of Economics, Louisiana State University, Baton Rouge, LA 70803 (e-mail: mocan@lsu.edu). Data used in this study were provided by the Louisiana Office of Juvenile Justice. Access was provided by Louisiana State University. We thank Janet Currie, Randi Hjalmarsson, Kevin Lang, Adriana Lleras-Muney, Bentley Macleod, Marco Gonzalez-Navarro, Isaac Sorkin and the participants of the 6th Economics of Litigation Workshop in Granada, Spain, the Gender and Economics Workshop at the University of Luxembourg, and TOBB University of Economy and Technology for helpful comments.

# 1 Introduction

How does the behavior of an individual change as the attributes of his/her peers are altered? This is an important question both for scientific inquiry and for designing strategies to improve performance and productivity. The analysis of peer effects, however, is complicated because in most circumstances people are not assigned to a particular group of peers, nor do they choose their peers randomly. Rather, individuals associate themselves with their peers deliberately. Homophily, the tendency of individuals to choose their peers who share common attributes with them, creates the well-known selection problem (Manski 1993; Moffitt 2001; Sacerdote 2001). This makes it difficult to identify the influence of the group on an individual who is associated with that group.<sup>1</sup> Researchers tackled this issue by exploiting circumstances in which assignment to peer groups is arguably random. (e.g., Sacerdote 2001; Mas and Moretti 2009; Dahl et al. 2014, Murphy 2019).

While it is important to investigate whether an individual's own behavior, such as own effort and own productivity are influenced by the attributes of that person's peers, it is, equally important to analyze whether an individual's judgement on another person, or an individual's decision regarding a matter that impacts the well-being of another person is influenced by the characteristics of the decision-maker's peer group.

In this paper we investigate whether consequential decisions made by judges about defendants in a court-room are impacted by a particular attribute of these judges' peer group. Specifically, we focus on judicial decisions made about juvenile defendants, and we investigate how judges' incarceration and sentencing decisions are impacted by the gender composition of their peers in the courthouse. The decisions we analyze are solo-bench decisions; i.e., they are made by individual judges during bench trials in these judges' courtrooms.

## *Decisions of Panels vs. Individual Decisions*

A line of previous research has addressed a similar question, but this research has focused on decisions

---

<sup>1</sup>The other major identification issue, referred to as the endogeneity or the reflection problem, stems from the fact that it is often difficult to separate the impact of the peer group on the individual from the impact of the individual on the peer group (Manski 1993).

made by a panel of individuals and reported conflicting results. For example, it has been shown that an increase in the share of female evaluators on the panel has no impact on the promotion chances to Associate/Full Professorship of female candidates in Italy and Spain (Bagues et al. 2017). On the other hand, De Paola and Scoppa (2017) reported that promotion of females is less likely if the panel consists of exclusively males, but that the gender gap disappears if there is one female member of a five-member panel. Basques and Esteve-Volart (2010) analyzed the recruitment procedure to the four main Corps of the Spanish Judiciary and found that a female candidate is less likely to be hired if the committee's share of female evaluators gets larger. In contrast, Hoekstra and Street (2021) reported that in Florida own-gender jurors are significantly less likely to convict on drug charges.

Research specifically analyzing panel decisions of judges, and the challenges faced by this research require special mention. Previous work on judicial panels investigated if the race or gender composition of a panel of judges impacts case outcomes (Boyd et al. 2010; Kastellec 2013; Grossman et al. 2016, Holden et al. 2021). As summarized by Kastellec (2013), peer effects on panel decisions can emerge through three primary mechanisms. First, if a panel is composed of diverse set of individuals with different backgrounds and experiences, each one can approach the matter in front of them differently, which allows them to make different interpretations and reach different conclusions. Deliberations during this process can allow the numerical majority members of a panel to be persuaded by a member who is in numerical minority (e.g., two members of the panel being persuaded by the third member). Second, voting strategy on a panel is important because in most settings, and especially in judicial decisions, members of the panel may prefer to avoid casting dissenting votes. This "dissent aversion" can be used by the members of the panel to bargain and gain concession to sway the decision of the panel (Posner 1983). Third, exposure to a peer (the sheer presence of a peer with a particular attribute) can impact the behavior of other members of the panel.<sup>2</sup>

More formally, decisions that emerge from a panel of individuals are reflections of both contextual factors

---

<sup>2</sup>For example, as mentioned by Kastellec (2013), Supreme Court Justice Antonin Scalia said about Justice Thurgood Marshall, the first black member of the Supreme Court, that "Marshall could be a persuasive force just by sitting there. He wouldn't have to open his mouth to affect the nature of the conference and how seriously the conference would take the matters of race." (Liptak 2009).

(e.g., exogenous exposure to a male or female judge), and endogenous factors such as dissent aversion, which itself could be a function of the gender composition of the peers on the panel.<sup>3</sup> A detailed explanation of the difficulty in identifying the exogenous peer effect in a panel decision can be found in Moffitt (2001), and Holden et al. (2021) and Fischman (2015) provide good discussions of these complexities in the specifics of judicial panels.

Our design allows us to avoid the challenges posed by panel decisions. In our study judges make solo decisions. Therefore, our judges do not involve in deliberations about their cases with their peer judges, and they carry the sole responsibility of their decisions. This means that any peer effect we identify in our paper is not the result of deliberations with peers on the specifics of each case handled by individual judges. Similarly, peer effects stemming from voting strategy and dissent aversion are not relevant in solo decisions because strategy and dissent, by definition, involves the joint decisions of at least two judges. Thus, we isolate the effect of similarity/diversity, i.e., the impact of plausibly exogenous exposure to peers who are similar to, or different from the judge who makes his/her own decision on the case.

### *Gender Differences*

We contribute to a body of work analyzing gender differences in judicial decisions (e.g., Peresie 2005; Boyd et al. 2010; Choi et al. 2011; Knepper 2018; Eck and Crabtree 2020; Asmat and Kossuth 2021). The underlying conceptual justification that forms the basis of these studies is multifaceted. For example, Eckel and Grossman (1996) report that women are more likely to reciprocally punish unfair behavior in comparison to men. Croson and Buchan (1999) provide evidence for women's higher propensity for reciprocity. It has also been documented that women's decisions are less individually-oriented and more socially-oriented in comparison to those of men (Eckel and Grossman 2008). These results suggest that if female judges are more likely to consider the offenses of convicted defendants as transgressions towards society, female judges would

---

<sup>3</sup>Suppose that a panel consists of three members: Two males (M1, M2) and one female (F1). Now assume that M2 is exogenously replaced by a female member, F2, which altered the makeup of the panel to M1, F1, F2. M1's behavior may change because he now faces two female colleagues, rather one male and one female, but M1's behavior may also change because his propensity for dissent would change as he is now facing two females (F1, F2) with possibly different dissent aversions. Importantly, F1's behavior will also change when F2 replaces M2, both because F1 now faces one male and one female (as opposed to two males) and also because her dissent aversion may be altered as well. Thus, it is difficult to tease out the exogenous gender impact of replacing M2 with F2 on M1's and F1's behaviors from the endogenous effects generated by the same alteration.

be more likely to reciprocate on behalf of the society and impose harsher punishment. Equally important, female behavior is found to be more responsive to context (Eckel and Grosman 1996; Ben-Ner et al. 2004). As explained in detail by Croson and Gneezy (2009), women are more responsive to the environment in which they make decisions, and the responsiveness of women to the context explains some seemingly contradictory results about gender differences in experimental research. Thus, variations in the gender ratio of peers could differentially impact the behaviors of female vs. male judges.

Using the universe of judicial decisions made about juvenile defendants in Louisiana between 1998 and 2012, we find that an increase in the proportion of female colleagues raises the severity of punishment assigned by judges, and that this result is driven by female judges. We estimate the effect of peer composition off of the within-judge variation over time, and exploit movements of peers in and out of courthouse for identification. The composition of judges in a courthouse changes due to departure of judges for such reasons as retirement, resignation and death, and because of arrivals of newly elected judges.<sup>4</sup> Although random assignment of juvenile case files to judges (as discussed in detail below) alleviates concerns about confounding factors, it does not completely eliminate it because judges are not randomly assigned to their peers. Thus, we control for court and year fixed effects and court-specific trends in all specifications.

We demonstrate that the impact of the proportion of female peers is not a proxy for other peer characteristics such as the proportion of black peers, average age of judges, their tenure, or their party affiliation. We propose two mechanisms that may be responsible for our findings, and present results that are consistent with the critical mass hypothesis which postulates that individuals who are in the numerical minority adopt the behavioral norms of those who are in the majority, but that members of the minority group start making decisions that reflect their true tendencies as the proportion of minorities in the group rises. We provide extensive robustness checks supporting our findings. For example, we show that the results are not sensitive to excluding sex crimes, excluding the decisions made during the one-year window around major judicial elections, controlling for judge-specific vs. court-specific trends or dropping trends from the model, using

---

<sup>4</sup>There are only a few judges moving from one judicial district to another during their tenure and therefore, identification is obtained from movements of peers in and out of courthouse.

one-way or two-way clustered standard errors and alternative measures of gender peer composition, and many other sensitivity and falsification tests.

The remainder of the paper is organized as follows. Section 2 discusses the institutional settings. Section 3 presents the data. Section 4 describes the econometric methodology. Section 5 presents the results. Conclusions are provided in Section 6.

## 2 Institutional Background

The judicial branch in Louisiana includes courts with different jurisdictions, ranging from state supreme court to the courts of appeals, from district courts to juvenile courts. There are currently 42 judicial districts in Louisiana. Twenty-eight of these districts cover one single parish each, and judges handle cases in their respective courthouses in these parishes. There are 10 other judicial districts where each district covers multiple parishes, and judges in these districts rotate between the courthouses of these parishes. Finally, four judicial districts have their separate juvenile courts, where each of these juvenile courts serves one parish.<sup>5</sup>

Judges in district courts and juvenile courts are elected by voters in partisan elections, where the party affiliation of each candidate is shown on the ballot.<sup>6</sup> Candidates for judge positions compete in a primary against other candidates, including those from their own parties, and it is common to have multiple candidates from a given political party to run for the same judge position.<sup>7</sup> District and juvenile court judges serve six-year terms and they are eligible for re-election. Elections are spread throughout the year. For example, elections in 2018 took place in March and November, with the runoff election of the latter being held in December.

Judge turnover, and therefore a change in the composition of judges in a courthouse can be generated

---

<sup>5</sup>These parishes are Caddo, East Baton Rouge, Jefferson and Orleans.

<sup>6</sup>A candidate for the district/juvenile court must satisfy the following criteria to be eligible for judgeship: (i) licensed to practice law in the state for eight years, (ii) domiciled in the respective parish for at least one year, and (iii) be under the age of 70.

<sup>7</sup>This structure is sometimes referred to as the “jungle primary.” The candidate who receives the majority vote in the primary is elected. If no candidate receives the majority vote, a runoff election is held between the top two candidates in the general election.

not only by elections, but also by the newly-created judgeships or by a vacancy. Vacancies can result from a variety of reasons ranging from resignation to death, from suspension to retirement before the end of the term.<sup>8</sup> These vacant positions are filled by special elections within a year after the day the vacancy occurs. The supreme court appoints a judge to the bench until the special election is held.<sup>9</sup>

Cases are randomly assigned to judges unless the juvenile was found guilty in a previous adjudication.<sup>10</sup> Such repeat offenders are re-assigned to judges who handled the previous episode of the juvenile. At the adjudication hearing the judge may find the defendant not guilty and dismiss the case if the prosecutor is unable to provide evidence to find the youth delinquent. In this situation the juvenile is considered as not having entered the juvenile justice system, and the case is purged. If the judge finds the defendant guilty, the judge needs to make a disposition decision. Convicted defendants can be assigned by the judge to the custody of the Department of Public Safety and Corrections to be confined in secure placement (i.e., they are incarcerated). Instead of incarceration, the judge can assign a less severe arrangement where the juvenile is removed from his/her home and placed in a residential non-secure placement or treatment facility such as a group home, or foster home. Similarly, the judge can put the juvenile on probation.

In addition, the judge has to assign a sentence length to each convicted juvenile. This is true for those who are incarcerated and placed in secure custody, but it is also true for those who are placed on probation or in non-secure custody. That is, each convicted juvenile is assigned a sentence length regardless of the type of custody in which they are placed. Judges are responsible for weighing the severity of the offense committed and the prior offense history of the juvenile.<sup>11</sup> There is no mandatory sentencing guidelines and judges exercise considerable discretion in sentencing. Additional details can be found in Eren and Mocan (2021). Some of the defendants are not convicted by a judge. Rather, they plead guilty (or no contest) to the charge filed against them. In such cases, the only decision made by the judge was the determination of

---

<sup>8</sup>In Louisiana the mandatory retirement age for judges is 70.

<sup>9</sup>This person is ineligible as a candidate to be elected in the special election (Article V, Louisiana Constitution of 1974).

<sup>10</sup>Under the provisions of the Louisiana juvenile justice system, a computer generated random assignment (open to public) is implemented in each court by the Clerk's office for all case files (Rules for Louisiana District Courts, Chapter 14, Appendix 14.0A, various years).

<sup>11</sup>In general, the judge will impose the least restrictive disposition consistent with the circumstances of the case, the health and safety of the child, and the best interest of the society (Louisiana Children's Code CHC 683).

the sentencing. Estimating the impact of judges' female peers in this subsample provided similar results. Put differently, any concern about the group of sentenced offenders being a selected sample and thus leading to bias has no empirical validity in our context. We further discuss these details in Section 5.1.

### 3 Data

The data are obtained from two sources. Louisiana Department of Public Safety and Corrections, Youth Services, Office of Juvenile Justice (OJJ) provides the universe of case files from 1998 to 2012. Each case file contains information on the juvenile and the case, including the gender, race and the age of the juvenile, the statute offense committed, the date the juvenile was sentenced, sentence type (secure custody, probation, etc.), sentence length, the courthouse in which the disposition was held and the identifier of the judge.

The second source is the annual collections of the Supreme Court of Louisiana, the so-called *The Guide to Louisiana Courts*.<sup>12</sup> These annual periodicals feature statewide list of judges along with information on courthouse served and judicial turnovers. Specifically, the annual reports allow us to track movements into and out of courthouses (e.g., election, retirement and resignation) with exact dates. Using the Guide to Louisiana Courts, we construct a panel containing the universe of judges serving in each courthouse. We supplement this data by gathering information on judge attributes.<sup>13</sup>

We impose several restrictions to our research sample. First, to be able construct the gender peer measure, we use only those courthouses that have at least two judges in a given year. Second, we exclude judges from the benchmark model if they had handled fewer than 25 case files in a courthouse over 1998-2012. Finally, in order to avoid potential confounding effects that may arise from multiple offenses and/or criminal history of the juvenile, we limit our attention to first-time delinquents ages 10 through 17 who were convicted for only one statute offense. Recall that repeat offenders are assigned to the same judge who has handled the original case. This attribute of the juvenile justice system compromises the critical condition

---

<sup>12</sup>These annual collections are available at [http://www.lasc.org/press\\_room/annual\\_reports/default.asp](http://www.lasc.org/press_room/annual_reports/default.asp).

Although the universe of case files from OJJ dates back to 1996, we limit our attention to 1998 and onwards because annual collections are available online since 1998.

<sup>13</sup>We collect information on judges from variety of sources including online searches and Louisiana District Judges Association Periodicals (1956-2000).

of random assignment of defendants to judges. Thus, we exclude repeat offenders from the analysis, and focus on delinquents who had their first interaction with the juvenile justice system.<sup>14</sup> Having imposed these restrictions, we end up with a total of 20,244 juvenile case files handled by 138 judges in 59 courthouses.<sup>15</sup>

Table 1 presents the descriptive statistics. The incarceration rate is about 14 percent. Recall that all convicted juveniles are assigned a sentence length, irrespective of whether they are incarcerated. As shown in Panel A, the average sentence length is around 507 days. Black (62 percent) and white (36 percent) juveniles comprise approximately 98 percent of all offenders. About 24 percent of the sample is female. Age at first-conviction is almost 15, and 37 percent of juveniles are convicted of a felony crime.<sup>16</sup>

Panel B of Table 1 reveals that 24 percent of judges are female and that the mean age of judges when they made incarceration decisions is 53. There are on average 6 judges in a courthouse in a year.

## 4 Empirical Methodology

To evaluate the impact of judicial gender composition of the courthouse on the incarceration decision and the sentence length assigned by judges, we estimate the following equation

$$D_{ijct} = \beta_0 + \beta_1 F_{-j,ct} + X'_{ijct} \beta_2 + \lambda_j + \theta_t + \theta_c + \theta_{ct} + \epsilon_{ijct} \quad (1)$$

where  $D_{ijct}$  is the disposition of juvenile defendant  $i$  set by judge  $j$  in courthouse  $c$  at time  $t$ . This variable takes the value of one if juvenile had been incarcerated (placed in secure custody) following his/her conviction. If  $D_{ijct}$  is zero, this indicates that he/she was not incarcerated, but instead was placed on probation or held in non-secure custody. Alternatively,  $D_{ijct}$  stands for the sentence length assigned by judges.  $F_{-j,ct}$  is the proportion of female judges in the courthouse  $c$ , excluding judge  $j$ , at the beginning of year  $t$ .  $X'_{ijct}$  is a

---

<sup>14</sup>To avoid the influence of the outliers, we drop observations with sentence lengths shorter than 80 days (bottom 1 percent of the sentence length distribution) and longer than four years.

<sup>15</sup>Fifty-three of these are district courthouses, four are juvenile courts, and two of them are city courts. The locations of city courts are generally different from district courts. Cases handled in city courts typically involve juvenile and petty offenses, ordinance and traffic violations. In an overwhelming majority of city courts a single judge is in charge of adjudicating cases.

<sup>16</sup>Ungovernable (11 percent), simple battery (8 percent) and simple burglary (8 percent) are the most common offense types in the data.

vector of observed juvenile characteristics (i.e., gender, race, age and its square and offense type),  $\lambda_j$ ,  $\theta_t$  and,  $\theta_c$  stand for judge, year and court fixed effects, respectively,  $\theta_{ct}$  are court-specific trends to control for linearly trending unobserved court characteristics and  $\epsilon_{ijct}$  is the error term. Standard errors are clustered at the judge level.

Several comments are warranted about the gender composition measure. First, we use the universe of judges in the state to construct  $F_{-j,ct}$ , which allows us to measure the peer composition variable accurately. More specifically, in the creation of the peer measure we utilize all judges, even those who do not enter the analysis sample.<sup>17</sup> Second, recall that there are 10 judicial districts where each district covers multiple parishes, and judges in these districts rotate between the courthouses of different parishes. In these cases, ignoring judge mobility within the judicial district (i.e., ignoring the movement of judges between courthouses of different parishes) would prevent us from fully capturing the interactions between these judges. Thus, for these 10 judicial districts we calculate  $F_{-j,ct}$  at the district level, using all courthouses with which each judge is affiliated, although we also show that the results are insensitive to the inclusion/exclusion of these districts. Finally, using judge gender composition of the courthouse from the beginning-of-year may raise concerns for two related reasons, both of which may bias gender effect estimates towards zero. First, the peer composition measure described in equation (1) may not fully reflect the exposure of judges to their peers in the courthouse because it does not take into account possible judge turnover occurring throughout the year. Second, a change in judge behavior in response to a change in peer gender composition may take some time to manifest itself. Therefore, we also estimate variants of equation (1) by: (i) constructing gender peer measure at the year-by-quarter level, and (ii) including its lagged values.

We control for judge fixed effects in all specifications to circumvent any potential bias from selection into peer groups. Thus, we estimate the effect of peer gender composition faced by a particular judge off of the variation in that composition over time. By relying only on within-judge variation, identification is obtained

---

<sup>17</sup>For example, assume that a district courthouse consists of four judges: A, B, C and D. Assume further that only three of these judges (A, B, and C) deal with juvenile cases, while the fourth judge (D) in the same courthouse takes on other types of cases (e.g., civil cases and adult crimes). This fourth judge is included to calculate the proportion of female peers faced by A, B, and C. It is important, however, to note that the overwhelming majority (77 percent) of all judge-by-year observations involve juvenile cases. Consequently, gender peer effect estimates from equation (1) are similar when we construct the peer measure by ignoring judges who have not handled juvenile cases in that year (see Section 5.2 for these additional results).

from the movements of peers in and out of courthouse.<sup>18</sup>

Although random assignment of case files to judges mitigate concerns over omitted variable bias, it does not completely eliminate it because judges are not randomly assigned to their peers. It is conceivable that changes in the proportion of female judges are correlated with unobserved factors that also affect judicial decisions. For example, assume that state-wide awareness rises with regard to issues related to women's well-being (e.g., Me Too movement). Such awareness may motivate voters to elect female judges as opposed to male judges, and the same awareness may compel judges to be harsher against male defendants. In this case, a positive relationship may emerge between sentence severity and the proportion of female peers, but this would be an artifact of the change in the underlying state-wide sentiment towards females. To account for such confounders, we control for court and year fixed effects, as well as court-specific trends in all specifications.

The key identifying assumption underlying this framework is that time-varying unobservable variables affecting the outcome are not systematically related to within judge variation in the gender composition. To the extent that this assumption holds, the coefficient estimate  $\beta_1$  can be interpreted as the causal impact of gender peer effects on judicial decisions in the courthouse. Note that our identification strategy does not rule out all potential confounding factors. For example, consider a violent crime in a particular parish before a judicial election, in which the perpetrator is a man and the victim is a woman. This event can lead to the election of a female judge and also to harsher sentences imposed by all judges in that parish. In this scenario, the correlation between the behavioral response of judges and the change in gender peer composition is not causal, but it is driven by the event preceding the election in that location. This sort of a (court-specific) shock calls for conditioning on court-by-year fixed effects. Doing so, however, would result in almost no variation to exploit for identification.<sup>19</sup> We provide a number of robustness analyses which

---

<sup>18</sup>Judges, in principle, can switch peer groups if they relocate from one district to another. For example, a judge can resign from his post in a courthouse and can subsequently be elected to a position in another courthouse. This type of mobility, however, is extremely limited and comprises only about 2 percent of the sample. In conjunction with this, specifications that control for court-by-judge fixed effects provide estimates of gender peer effects that are almost identical to those presented in the paper. These results are available upon request.

<sup>19</sup>To see this, consider a simplified version of equation (1)

$$D_{ijct} = \beta_0 + \beta_1 F_{-j,ct} + \beta_2 F_j + \theta_{ct} + \epsilon_{ijct}$$

indicate that such location-and-year specific shocks are not the drivers of our results. For example, we show that the results are not sensitive to controlling for the local (parish-level) juvenile and adult crime rates and violence against police in the parish of the courthouse. Similarly, the results are robust to the omission of observations within one-year window around major judicial elections, and the results are not driven by the punishment exerted by judges on sex crimes. An extensive array of other robustness and falsification exercises (e.g., replacing court trends with judge-specific trends, or dropping trends altogether, and so on) are discussed below.

We first examine the validity of random assignment of case files to judges. A typical test for this, in our context, is to run a series of regressions where the proportion of female peers in the courthouse is regressed on juvenile and case characteristics, while controlling for court and year fixed effects and court-specific trends. These results are reported in Table 2. Each cell represents a separate regression. The point estimates are all small in magnitude and none of them is statistically significant. We also run a single regression where we condition on all juvenile and case characteristics. The p-value for joint significance is 0.47 (reported in the last row of Table 2).<sup>20</sup>

Because our identification hinges on within-judge variation in the proportion of female peers over time, variation in this variable needs to be confirmed. Table 3 presents information related to variation in the proportion of female peers faced by each judge in each year. The mean and standard deviation in the proportion of female peers in the courthouse are 0.20 and 0.32, respectively. The standard deviation of the change in a judge's gender peer composition from one year to the next is 0.07, corresponding to around 23 percent of the overall variation. A simple variance decomposition exercise (sum of squares) shows that

---

where  $F_j$  is an indicator that takes the value one if judge  $j$  is female. As we demean the data at the court-by-year level, the equation transforms into

$$D_{ijct} - \bar{D}_{ct} = \beta_1(F_{-j,ct} - \bar{F}_{-j,ct}) + \beta_2(F_j - \bar{F}_{ct}) + (\epsilon_{ijct} - \bar{\epsilon}_{ct})$$

where  $\bar{F}_{-j,ct} = \bar{F}_{ct}$  and  $F_{-j,ct} - \bar{F}_{-j,ct} = \frac{-1}{N_{ct}-1}(F_j - \bar{F}_{ct})$  and  $N_{ct}$  is the number of judges in a courthouse. Substituting these equalities in the demeaned equation leads to

$$D_{ijct} - \bar{D}_{ct} = \beta_1 \frac{-1}{N_{ct}-1}(F_j - \bar{F}_{ct}) + \beta_2(F_j - \bar{F}_{ct}) + (\epsilon_{ijct} - \bar{\epsilon}_{ct})$$

It is evident that the first term varies independently only when the court size changes. See also Cornelissen et al. (2017) for an application of this identification strategy.

<sup>20</sup>We also experimented with similar randomization tests by dropping trends and controlling for court-by-year fixed effects. The point estimates from these alternative specifications yield the same conclusion about randomization.

within-judge variation accounts for around 7 percent of the total (within and between judge) variation in the proportion of female peers. Furthermore, the share of judge-year observations experiencing any judge turnover in courthouses is 17 percent. Thus, there is non-trivial variation in the composition of judges in courthouses to detect meaningful gender peer effects.

Figure 1 provides additional information about the extent of the variation in the proportion of female peers. The figure displays the distribution of the residuals from a regression where the proportion of female peers is regressed on judge, year, and court fixed effects, court-specific trends and defendant characteristics. Consistent with Table 3, non-negligible identifying variation is displayed in Figure 1, where the standard deviation of the residuals is 0.083. Figure A1 in the Appendix presents this information by judge gender.

Finally, Figure 2 presents the proportion of female peers, averaged across male and female judges over the years, and demonstrates the lack of any secular trend.<sup>21</sup> Figures A2 (a)-(c) in the Appendix plots several other measures related to variation in the gender peer composition.

## 5 Results

### 5.1 Baseline Results

Baseline estimates of the impact of gender composition in the courthouse on judicial decisions are presented in Table 4. Panel A reports the results of the models where the dependent variable is the incarceration decision of judges. The outcome in Panel B is sentence length. Standard errors are clustered at the judge level. Column (1) reports the results by controlling for judge, year, and court fixed-effects and court- specific trends. Column (2) adds juvenile characteristics and offense fixed-effects. The results in Panel A indicate a statistically significant point estimate of about 0.10, which implies that a 10 percentage point increase in the proportion of female judges (e.g., a change from being exposed to 1 female and 9 male colleagues in the courthouse to being exposed to 2 female and 8 male colleagues) leads to approximately 1.0 percentage

---

<sup>21</sup>We observe an increase in the average proportion of female peers for male judges beginning in 2009. This increase is attributable to the election that took place in November 2008. Excluding judicial decisions made after 2008 did not alter the conclusions. These results are reported in Section 5.2.

points increase in the probability of incarceration decision made by judges. Taking the incarceration rate of 0.137 from Table 1 as our benchmark, this estimated impact indicates an average increase of 7.3 percent. Panel B indicates that a 10 percentage point increase in the proportion of female judges in the courthouse generates an increase in sentence length by about seven days (or about 1.6% from the sample mean).<sup>22</sup>

As noted in Section 2, if the judge acquits the defendant, there is no sentencing decision to be made. In this case, the case file is purged and the juvenile defendant is treated as if he/she had no contact with the juvenile justice system. This means that the data contain only those individuals who are convicted. This particular limitation, due to institutional structure, is not a serious concern in the interpretation of our results to the extent that gender peer composition of the courthouse impacts the conviction/acquittal decision of judges in the same way as in sentencing decision. For example, if borderline cases (i.e., those with weak evidence) end up with guilty verdict rather than being dismissed when the proportion of female judges goes up in the courthouse, this would generate a sample which would include marginally guilty or “less guilty” defendants who are sentenced in courthouses with a higher percentage of female judges. Thus, the impact we identify could be an underestimate of the true judge peer effects.

To further explore this conjecture, we limit our attention to those juvenile offenders who plead guilty.<sup>23</sup> Given that the verdict on guilt vs. innocence is not a relevant margin for those who have plead guilty to the charge filed, estimating the impacts for the plead guilty sample provides valuable insights on the extent of a potential bias. The sample of juvenile offenders who plead guilty is small ( $n= 7,118$ ). In this sample the

---

<sup>22</sup>To put the estimates in perspective we compare our findings to those of previous work that examined the impact of factors unrelated to the merits of the case on judicial decisions. For example, Butcher et al. (2017) show that a one standard deviation change in judge stringency increases the probability of incarceration by 53 and 18 percent for female and male offenders, respectively. Philippe and Ouss (2018) find that facing a jury trial during a period after media coverage of crime incidents increases sentence lengths for juveniles by 7 percent. Finally, Hoekstra and Street (2021) find that adding one own-gender juror results in a 60 percent reduction in conviction rates on drug charges.

Analyses of decisions that are influenced by gender exposure include Battaglini et al. (2020) who find that a one-standard deviation increase in a U.S. appellate court judge’s exposure to female colleagues on judicial panels leads to a four percentage point increase in the probability that the judge hires a female clerk three years later. Washington (2008) finds that a U.S. congressperson’s propensity to vote liberally, especially on issues of productive rights, goes up with parenting female children. Each additional female child is associated with a 2-point (5 percent) increase in the score assigned to the congressperson by the National Association of Women (NOW).

<sup>23</sup>Prior to the adjudication hearing in which the judge makes a decision on guilt-vs-innocence, a petition hearing takes place. At this petition hearing the district attorney charges the juvenile with a crime. If the defendant pleads not guilty, the case goes to the trial, which takes place at a later date. In this case, the date of the adjudication hearing is later than the date of the petition hearing. If, on the other hand, the petition and adjudication dates are the same, this means that the judge has not made a guilty/not guilty decision; instead, the juvenile has pleaded guilty or no contest to the charge filed (Depew et al. 2017).

estimated coefficient of the gender peer composition on the propensity of incarceration is 0.096 (s.e=0.084) and it is 140.81(s.e=42.74) in the sentence length regression. These estimates are in line with those reported in Table 4. Note also that guilty vs. innocence margin is arguably less of a concern for juveniles convicted of a felony crime. Put differently, potential contamination due to selection bias is likely to be less pronounced for severe crimes. To this end, we examined gender peer effect by severity of crimes. The coefficients of the proportion of female peers in incarceration regressions are 0.124 (s.e=0.088) and 0.058 (s.e=0.043) for felony and non-felony offenses, respectively. The effects on sentence length are 86.113 (s.e=35.748) and 53.983 (s.e=49.712). The effect sizes on incarceration and sentence length for felony ad non-felony offenses are similar when they are benchmarked relative to their own sample means.

## 5.2 Robustness Checks

We implement several sensitivity checks to examine the robustness of our results. The first column of Table 5 presents the results from a specification where the measure of gender peer composition ( $F_{-j,ct}$ ) is constructed at the year-by-quarter level.<sup>24</sup> In this model, the proportion of female peers of each judge in a given courthouse is allowed to vary from quarter-to-quarter, but the coefficient estimates from this exercise for both the incarceration and the sentence length regressions are almost identical to those obtained from the baseline model of Table 4. Second, the turmoil during and after hurricanes Katrina and Rita may have impacted peer group composition and judicial decisions. To investigate this hypothesis, we exclude parishes that are known to be most affected from these natural disasters.<sup>25</sup> As shown in column (2) of Table 5, doing so provides a larger effect on incarceration, while the point estimate for sentence length is almost identical to those reported in Table 4. Third, recall that we calculate gender composition at the judicial district level in jurisdictions where judges rotate between different parish courthouses. Dropping these parishes in column (3) reduces the sample size to 12,755, and lowers the precision of the estimates, but it does not alter the point estimates.<sup>26</sup>

---

<sup>24</sup>We control for quarter fixed effects in this specification.

<sup>25</sup>These parishes are Jefferson, Lafourche, Orleans, Plaquemines, St. Bernard, St. Tammany, and Terrebonne.

<sup>26</sup>Pooling all parishes and running a fully interacted model produces similar results. The coefficient estimates for incarceration are 0.093 (s.e.=0.059) and 0.013 (s.e.=0.100) on the proportion of female judges and its interaction with rotating districts,

Fourth, recall that in the main analysis we exclude judges if they handled fewer than 25 case files in a given courthouse over the course of the analysis period. Including all judges with any number of dispositions, or restricting the sample to judges who handled at least 150 cases produce almost identical results. (columns 4 and 5 of Table 5, respectively). Fifth, adding first-time juvenile offenders with multiple convictions back to the sample reveal that the results are not very sensitive to this sample restriction either (column 6).

Sixth, we replace court trends with judge-specific linear trends under the identifying assumption that unobservable variables related to judicial outcomes do not deviate from an individual judge's trend when within-judge variation in the gender composition deviates from trend. The estimated effects, reported in column 7 of Table 5, are very similar in magnitude to those presented in Table 4.<sup>27</sup>

Seventh, recall that we constructed the peer composition measure using the universe of judges in the district courthouses. Our results remain similar when we define the peer measure only using female judges who handled juvenile cases (column 8). This is not surprising because, as noted, an overwhelming majority of the judges in these courthouses handled juvenile cases (77 percent of all judge-by-year observations). Finally, we drop all judicial decisions made after 2008 to circumvent concerns related to an increase in the trend of average proportion of female peers for male judges (Figure 2). Doing so does not largely alter our baseline findings (column 9).

Could the results be driven by local shocks in certain years that may have changed the sensibility of judges and voters? For example, if there is a brutal sex crime with a female victim before an election, voters may end up electing female judges, and judges may become harsher in their decisions because of this crime. To address this concern we added to the model property crimes and violent crimes committed by juveniles and adults in the parish of the courthouse, the number of police officers killed or assaulted in the parish, along with other parish attributes such as log of population, fraction of the county population with a high school degree or less and unemployment rate.<sup>28</sup> The results, reported in column (1) of Table 6, show that

---

respectively. The same coefficient estimates for sentence length are 53.999 (s.e.=45.860) and 54.458 (s.e.=59.884).

<sup>27</sup>We also experimented with our analysis by using the logarithm of sentence length as the dependent variable. The results from this exercise provided the same inference and they are available upon request.

<sup>28</sup>The data on local crime come from FBI's Uniform Crime Reporting (UCR) program. UCR provides annual county level property and violent crime statistics for both juveniles and adults, as well as information about officers who were feloniously

controlling for the extent of local crime and violence against police has no impact on the estimated coefficient of the proportion of female peers.

The prevalence of sex offenses may be correlated with judge harshness and the gender peer composition. Dropping from the sample sex offense cases had no impact on the estimated gender peer effect (column 2 of Table 6). To further analyze whether the results reflect the impact of a local event that took place before an election (which led to the election of female judges and an increase in harshness following the election) we re-estimated the model after deleting judicial decisions that took place during the year after the election of judges (column 3 of Table 6), and after dropping observations in the election year and the year after the election (column 4).<sup>29</sup> The results remained the same.

To investigate whether the results are driven by the decisions of a particular judge, we estimated equation (1) repeatedly, each time removing dispositions handed down by a different judge. Figures A3 and A4 in the Appendix plot the distribution of the coefficient estimates for the gender peer effect from a total of 138 regressions. The average of the coefficient estimates for incarceration is 0.098 (s.d=0.005), while it is 67.80 (s.d=3.13) for sentence length, indicating that the results are not driven by a particular judge.

It is plausible that judges were not responding to peers' gender but to some other judge characteristics that were systematically correlated with gender. For example, female judges may more likely be black (or white), and it could be the race of the peers rather than their gender that may be the driver of harshness. Column (1) of Table 7, however, shows that adding the proportion of black peers in the incarceration regression has no impact on the coefficient of female peers, and that the coefficient of black peers itself is not significantly different from zero. Conditioning on peers' political party affiliation does not change the results either. The point estimate of the proportion of Democrat judges in the courthouse is small and insignificant, while the impact of having more female peers remain intact (column 2). It can also be argued that female judges may be younger (or older), and that peers' average age and not their gender composition may be related to judicial decisions. Column (3) of Table 7 refutes this conjecture. The average age of peers

---

killed or assaulted while performing their duties. Other parish attributes are drawn from the U.S. Bureau of Labor Statistics.

<sup>29</sup>We considered the major elections of (November) 2002 and 2008. Column (3) drops pre-election years of 2002 and 2008. Column (4) drops both pre- and post-election years (2002, 2003, 2008, and 2009).

has no impact on the incarceration decision and adding average peer age to the model does not impact the coefficient of the proportion of female peers. Finally, column (4) demonstrates that inclusion of peers' race, party affiliation and age to the model jointly does not alter the inference. Columns (5) to (8) present the results of the same exercise, where the outcome is the sentence length. Once again, adding to the model peers' other attributes does not eliminate the gender peer effect.<sup>30</sup>

We analyzed the timing of the gender peer effects by augmenting the model with the lags and leads of the gender composition measure. Table 8 presents the results obtained from three different specifications. Columns (1), (2), (4) and (5) serve as falsification exercises as they investigate whether judicial decisions in a given year are influenced by the gender peer composition in the following year(s). Variations in the future values of the peer composition should not influence current decisions of judges, and as Table 8 reveals, the coefficient estimates of the leads of the proportion of female peers in courthouse are small and never statistically different from zero. This means that future values of the proportion of female peers have no impact on judges' current decisions on incarceration or sentence length. Columns (3), (4), (7) and (8) present the result from the models that investigate the existence of path-dependence in peer effects by regressing judicial decisions on current and lagged values of the gender peer composition measure. Overall, gender peer effects do not appear to exhibit a persistent pattern over time.

Finally, we estimate the models under placebo values of female peers. Specifically, we consider actual values of the proportion of female peers in each courthouse in each year (pertaining to both female and male judges) and randomly assign these values to different years for the same courthouse. We then run equation (1) and obtain the coefficient estimates of the proportion of female peers, and repeat this exercise for 1,000 times. Figures A5 and A6 in the Appendix display the distribution of the coefficient estimates obtained from this exercise. The vertical line depicts the actual point estimates from column (2) of Table 4. Only 3 of the 1,000 placebo regressions produce effects that are larger than the actual value in Figure A5 ( $p\text{-value}=0.00$ ), and only 14 placebo estimates in Figure A6 are greater than the corresponding estimate in

---

<sup>30</sup>We also entertained the possibility of the results being driven by judge experience. Controlling for total years of experience since graduating from law school did not change the inference. The results are available upon request

Table 4 (p-value=0.01).

### 5.3 Unbundling the Impact

In this section we investigate whether the peer effect results reported earlier mask a more complicated relationship regarding potentially heterogeneous gender-specific impacts. More specifically, to unbundle the overall gender peer effect we analyze whether the estimated effect is sensitive to judge-defendant gender match. To that end, we investigate whether variations in the proportion of female peers in the courthouse has a differential effect on the decisions made by male/female judges on male/female defendants. For this purpose, we estimate the following equation

$$\begin{aligned}
 D_{ijct} = & \beta_0 + \beta_1 F_{-j,ct} + \beta_2 Female\ Defendant * F_{-j,ct} + \beta_3 Female\ Judge * F_{-j,ct} \\
 & + \beta_4 Female\ Defendant * Female\ Judge + \beta_5 Female\ Defendant * Female\ Judge * F_{-j,ct} \quad (2) \\
 & + X'_{ijct} \beta_6 + \lambda_j + \theta_t + \theta_c + \theta_c t + \epsilon_{ijct}
 \end{aligned}$$

In equation (2), the impact on the outcome of the change in the proportion of female peers in case of male judge-male defendants is  $\beta_1$ . The impact is  $(\beta_1 + \beta_2)$  in case of male judge-female defendant pairs. If the judge is female and the defendant is male, the impact of a change in the proportion of female peers is  $(\beta_1 + \beta_3)$ , and the impact is  $(\beta_1 + \beta_2 + \beta_3 + \beta_5)$  for female judge-female defendant pairs.<sup>31</sup>

Table 9 presents the results. Estimated  $\beta_1$  and  $\beta_2$  are not significantly different from zero (in rows one and two) in either the incarceration or the sentence length regressions. Similarly, as shown at the bottom section of the table,  $(\beta_1 + \beta_2)$  is not significantly different from zero in either regression. This means that there is no compelling evidence of the influence of gender peer composition on male judges. More specifically, the incarceration decision and the assigned sentence length by male judges do not significantly react to a change in the proportion of their female peers in the courthouse, regardless of whether the defendant is male

---

<sup>31</sup>These specifications control for interactions of juvenile's gender with individual characteristics, offense and judge fixed effects as well as interactions between judge's gender and individual characteristics and offense fixed effects.

or female.

The gender peer effect on judicial decisions made by female judges on male defendants is  $(\beta_1 + \beta_3)$ , which is equal to 0.240 in the incarceration equation. It is 128.0 in case of the sentence length, and both sums are statistically significantly different from zero (with p-values of 0.00 and 0.00, respectively). This indicates that an increase in the proportion of female peers in the courthouse makes female judges more likely to incarcerate male defendants and also causes female judges to assign longer sentences on male defendants.<sup>32</sup>

The impact of an increase in the proportion of female peers on female judges' propensity to incarcerate female defendants is 0.173 ( $\beta_1 + \beta_2 + \beta_3 + \beta_5$ ), and it is significantly different from zero with a p-value of 0.00, as displayed at the bottom part of Table 9. This magnitude implies that if the proportion female peers faced by female judges goes up by 10 percentage points, this generates an increase in female judges' propensity to incarcerate female defendants by 1.7 percentage points. Finally, the impact on sentence length assigned by female judges on female defendants, induced by a change in gender peer composition in the courthouse, is small in magnitude (-33 days) and is not statistically different from zero.

Neither the point estimates nor their statistical significance change in a meaningful way under the modification to model specification. For example, if we exclude court-specific trends from the model,  $(\beta_1 + \beta_3)$  equals to 0.241 (p-value=0.00) in the incarceration regression and 121.33 (p-value=0.00) in the sentence length regression. Similarly,  $(\beta_1 + \beta_2 + \beta_3 + \beta_5)$  equals 0.167 (p-value=0.00) in the incarceration regression and it is -40.5 (p-value=0.82) in sentence length regression.<sup>33</sup> Along the same lines, clustering the standard errors at the courthouse level (rather than at the judge level) or two-way clustering at the courthouse and judge level do not affect the inference.<sup>34</sup>

Table 10 summarizes these results in the context of a courthouse consisting of 10 judges, two of whom are female. If the gender composition of judges changes so that the courthouse now has 3 female and 7

---

<sup>32</sup>This result is consistent with those reported by Liu (2021) who finds that female evaluators assign higher scores to female job applicants even when the identity of the applicants are masked.

<sup>33</sup>As a further robustness check, we control for quadratic court-specific trends. The results remain virtually identical and they are available upon request.

<sup>34</sup>For example, the p-value for the impact of judicial decisions made by female judges on the incarceration propensity of male defendants is 0.00 when the standard errors are clustered at the courthouse level, while the corresponding p-value for female defendants is 0.00.

male judges, this event increases the proportion of female peers for female judges by 0.11.<sup>35</sup> In response, female judges become 2.6 percentage points more likely to incarcerate male defendants and they assign two weeks longer sentences to males. The same increase in the proportion of female peers triggers an increase in female judges' proclivity to incarcerate female defendants as well (by 1.9 percentage points), but it has no significant impact on sentence lengths received by female juveniles.

To put gender peer effect in perspective, we provide some back-of-the-envelope calculations by randomly adding one more female judge in one-quarter of all courthouses in Louisiana. We repeat this exercise 1,000 times, each time calculating the difference between simulated and actual average gender peer measures. The mean of the differences from 1,000 simulations indicates a 5.5 percentage points increase in the average proportion of female peers. Using this change along with the estimates reported in Table 10 and the number of case files handled by female judges indicates that such an increase in female peers leads to five additional juveniles incarcerated and an additional 3,022 days of prison time imposed by female judges per year.

#### 5.4 Potential Mechanisms

Although it is not possible to determine with certainty the mechanism behind the peer effect identified in the paper, we propose two potential avenues, and test their validity. The analysis, described below, compels us to choose one of these mechanisms as the more relevant one over the other. Recall that incarceration rate is higher and sentence length is longer if cases are handled by female judges. As displayed at the bottom of Table 10, the incarceration rate of female judges is twice as high as that of male judges (0.205 vs. 0.104) and female judges assign sentences that are about one month longer on average (about 526 days vs. 498 days). To the extent that female judges are harsher, an increase in the proportion of female judges in a courthouse is likely associated with a rise in average strictness in judicial decisions in that courthouse. If female judges have the inclination to conform to the norms and customary standards of judicial decision-making, they would become harsher in their own judicial decisions as the average harshness goes up in their environment.

---

<sup>35</sup>Note that in a courthouse with 2 female and 8 male judges, the proportion of female peers is 0.22 (2/9) for each male judge, and the proportion of female judges is 0.11 (1/9) for each female judge. If a male judge of the courthouse is replaced by a female judge (e.g., through an election), then the courthouse contains 3 female and 7 male judges; and this particular event increases the proportion of female peers to 0.33 (3/9) for male judges and to 0.22 (2/9) for female judges.

Consequently, the first hypothesis is that female judges adjust their decisions to conform to evolving judicial stringency, generated by an increase in the proportion of female judges.<sup>36</sup>

The second hypothesis postulates that female judges do not try and adjust to the changing norms of stringency per se. Instead, female judges are influenced by the sheer presence of their female peers. In other words, female judges alter their behavior simply because they are exposed to more female peers. This second channel resembles the “critical mass” hypothesis, which posits that individuals who are in the minority of a group conform to the behavioral norms of the majority. As the share of minorities in the group goes up, members of the minority group start making decisions that reflect their true tendencies. Examples include females on corporate world, in politics, and in science (Kanter 1977, Dahlerup 1988, Etzkowitz et al. 1994). In our context because female judges are in the minority in a courthouse, they may feel pressure to adopt the incarceration and sentencing behavior of their male peers. When the number of female judges in the group (in the courthouse) goes up, the decision-making of female judges would start reflecting their true personal inclinations in incarceration and sentencing.

We calculated the incarceration rate and the sentence length at the judge-by-year level. Table 11 displays the means, weighted by total number of cases in each judge-by-year cell. Column (1) shows that average sentence length is about the same between female and male judges (550 days vs. 544 days), but that female judges are harsher than their male peers in incarceration tendency. The mean incarceration rate of female judges is almost twice that of male judges (0.22 vs. 0.13).<sup>37</sup> Columns (2) and (3) of Table 11, however, indicate that the averages reported in column (1) mask a more subtle picture. Specifically, we divided judges into two groups in each year: those judges who made decisions while facing a proportion of female peers which is lower than the average in that year, and judges whose female peers’ proportion was greater than the average of that year. Columns (2) and (3) of Table 11 present the means in judges’ incarceration rate and in assigned sentence length for male and female judges and by their exposure to their female peers.

---

<sup>36</sup>Using data from North Carolina, Abrams et al. (2021) find that judges adjust to the existing local norms of sentence harshness when they rotate between judicial districts.

<sup>37</sup>The significant difference in harshness in incarceration is confirmed by a regression where judge-by-year level incarceration rate is regressed on court-by-year fixed effects and an indicator for female judge. The estimated effect of female judge is a 4.3 percentage point increase in the incarceration rate (s.e=0.029).

Column (2) shows that when the proportion of female peers is low (below the sample mean), female judges are more lenient than male judges. The incarceration rate of female judges is 9 percent, in comparison to the 12 percent incarceration of male judges, and average sentence length of female judges is about 494 days, as opposed to the average of 560 days assigned by male judges. On the other hand, as shown in column (3), the picture is reversed when female judges face a high proportion (above average) of female peers. In this situation, female judges are significantly harsher than their male peers. The incarceration rate of female judges is almost two-and-a-half times higher (34 percent vs. 14 percent) and the average sentence length assigned by female judges is 92 days longer (602 days vs. 510 days) than their male counterparts. To put differently, Panel (B) and columns (2) and (3) show that male judges' assigned sentence length and their incarceration rate do not change appreciably regardless of whether they face low or high proportion of female peers. On the other hand, as shown in Panel (A) and columns (2) and (3), when female judges are exposed to a larger proportion of female peers, their average incarceration rate quadruples from about 9 percent to 34 percent, and their average assigned sentence length goes up from 494 days to 602 days.

The information presented in Table 11 is consistent with both of the hypotheses described above. To test the validity of these hypotheses, we calculated average peer harshness in incarceration and in sentencing for each judge, in addition to the proportion of their female peers.<sup>38</sup> We analyzed the extent to which an increase in average peer harshness and an increase in the proportion of female peers are related to a change in judicial decisions. Table 12 presents the results. Columns (1) to (3) pertain to incarceration decision and columns (4) to (6) display the results related to assigned sentence length. Column (1) shows that a 10 percentage point increase in the proportion of female judges increases the propensity to incarcerate by 0.9 percentage points, which is the same magnitude obtained from the benchmark sample of Table 4. Column (2), on the other hand, reveals that an increase in average incarceration rate of peers is not strongly related to the incarceration propensity of judges. Specifically, a 10 percentage point increase in average incarceration

---

<sup>38</sup>The calculation of peer harshness in incarceration and sentencing requires the courthouse to have at least two judges who handled juvenile cases. Thus, those judges who are the only ones in a courthouse who handle juveniles cases cannot be included in this particular analysis. Average peer harshness in incarceration and in sentence length are calculated in the same manner as the calculation of the proportion of female peers, as described in Section 4.

rate of judge's peers is associated with a statistically insignificant 0.7 percentage point *decrease* in judges' propensity to incarcerate. Column (3) presents the results obtained from the specification that accounts for both the proportion of female peers and average peer harshness in incarceration. The point estimate of the proportion of female peers is not impacted. The coefficient of average peer harshness switches from negative to positive but it is still negligible and statistically insignificant.

Column (4) of Table 12 shows that an increase in the proportion of female peers brings out a statistically significant increase in sentence lengths assigned by judges, with a magnitude similar to that reported in the benchmark model. In contrast, column (5) reveals that average sentence length assigned by peers is not associated with sentence lengths assigned by individual judges. Specifically, a 10 day increase in average sentence length assigned by the peers of the judge is associated with a statistically insignificant 1.4 day increase in judge's sentence assignment. Finally, column (6) shows that when peers' harshness in sentencing and the proportion of female peers are jointly included in the model, the impact of female peers remains about the same in magnitude and significance, and the influence of average peer harshness is smaller and indistinguishable from zero.

The coefficients reported in Table 12 should be interpreted with caution because of the reflection problem (Manski 1993), as average harshness of a judge's peers is likely to be endogenous because it could be impacted by the behavior of the judge. With this proviso, it is important to note that the coefficients of peer harshness variables are small in magnitude and never statistically significant. More importantly, inclusion of peer harshness has no discernable effect on the estimated coefficients of the proportion of female peers.

Panel B of Table 12 presents the same information using only female judges. The sample size goes down to 5,356 but the results are similar to those reported in Panel A, indicating that the effects are driven by female judges. Again, harshness of peers has no meaningful association with the incarceration and sentencing decisions of female judges. On the other hand, holding constant peers' harshness in judicial decisions, an increase in the proportion of female peers has a positive effect on female judges' decisions.<sup>39</sup> Thus, the

---

<sup>39</sup>The coefficient of female peers is not statistically significant in the sentence length regressions, although the magnitude (about 56 days) is still sizable. The loss of statistical significance is likely the result of both the reduced sample size and also the reflection that a rise in the proportion of female peers prompts female judges to assign harsher sentences in case of male

results displayed in Table 12 support the hypothesis that it is the exposure to female peers which has an impact on judicial decisions, and not exposure to the leniency/harshness of peers.<sup>40</sup>

## 6 Conclusion

In this paper we analyze the impact of peers on decisions made by judges. Because we analyze the decisions of individual judges, as opposed to decisions of a group (e.g., a panel), we isolate the influence of being exposed to those who are similar or different from the confounding effects of the group dynamics such as dissent aversion, and persuasion by the different viewpoints and arguments made by peers on the panel.

Using the universe of case files from Louisiana juvenile courts from 1998 to 2012, we investigate the impact of judges' gender peer composition on their judicial decisions. We exploit random assignment of defendants to judges, and leverage turnover of judges in courthouses generated by such events as retirement, resignation, or death of judges, and by departures and arrivals of judges to courthouses through elections. We find that an increase in the proportion of female peers faced by each judge generates an increase in the severity of punishment. That is, judges are more likely to incarcerate and they assign longer sentences when the proportion of their female peers in the courthouse goes up. This result is driven by the reaction of female judges. We show that the proportion of female peers is not a proxy for other peer characteristics such as race, party affiliation, age or tenure. Adding to the model the proportion of black judges, the average judge age in the courthouse, peers' party affiliation and tenure reveals that they have no influence on the estimated impact of peer gender composition.

Female judges, on average, are more likely to incarcerate than male judges, indicating that an increase in

---

defendants only (See Table 9), whereas the sample in Panel B includes both male and female defendants. When we estimate these models using female judges and male defendants, the sample size goes further down to 4,210. The estimated coefficients for female peers becomes 102.97 (s.e=80.18) in column (4), and the coefficient of average sentence length assigned by peers in column (5) is 25.98 (s.e=68.43). Inclusion of both variables (column 6) produces these coefficients as 102.84 (s.e=79.48) and 25.68 (s.e=67.22), respectively.

<sup>40</sup>This finding provides support to the critical mass hypothesis, although we do not take it literally. More specifically, we do not attempt to identify the location of a “critical proportion of female peers” where the leniency/harshness of female judges are turned on and off. This is because, the response of female judges could be gradual around a critical mass point, rather than exhibiting a discrete jump. Furthermore, even if such a unique threshold existed, its location could depend on the context (e.g., rural vs. urban courthouses and small vs. large courthouses) and it could also depend on judge attributes (e.g., younger judges vs. older judges).

the proportion of female judges in the courthouse is associated with increased average judicial harshness. This suggests that female judges might be reacting to the changing judicial norms of punishment by adopting to the enhanced severity of their environment. We present evidence, however, which suggests that the dominant force behind the increase in strictness of female judges is the sheer exposure to female colleagues, rather than a change in the overall leniency/harshness standards. This finding is consistent with the critical mass hypothesis, which would posit that female judges, who are in the numerical minority among their peers in a courthouse, would conform to the behavioral norms of the majority (male judges). As their share rises, female judges would start making decisions that reflect their true inclinations, which in this case translate into harsher punishment.

Eckel and Grossman (1996) report that women are more likely to reciprocally punish unfair behavior in comparison to men. Croson and Buchan (1999) provide evidence for women's higher propensity for reciprocity. It has also been documented that women's decisions are less individually-oriented and more socially-oriented in comparison to those of men (Eckel and Grossman 2008). These results suggest that if female judges are more likely to consider the offenses of convicted defendants as transgressions towards society, female judges would be more likely to reciprocate on behalf of the society and impose harsher punishment. Equally important, female behavior is found to be more responsive to context (Eckel and Grossman 1996; Ben-Ner et al. 2004). As explained in detail by Croson and Gneezy (2009), women are more responsive to the environment in which they make decisions, and the responsiveness of women to the context explains some seemingly contradictory results about gender differences in experimental research. Our results are consistent with this observation because we find that female judges' judicial decisions respond to the changing proportion of peers' gender in which these decisions are made.

The welfare implications of our findings are not straightforward for two reasons. First, it is unclear what the optimal level of punishment is. While we report that female judges are harsher than their male counterparts, this information in-and-of itself does not imply that male judges provide sub-optimal level of punishment or that female judges provide excessive punishment. Therefore, that the stringency of female judges goes up as they are exposed to more female peers may be beneficial or detrimental for social wel-

fare. Second, and related to the previous point, it is complicated to address all potential dimensions of social welfare that can be impacted by judicial decisions. Furthermore, judicial decisions regarding juvenile punishment have context-specific impacts in some of these dimensions.<sup>41</sup> Nevertheless, our results from a simulation exercise indicate an additional five juvenile incarcerations and 3,022 extra days of prion time imposed on juveniles per year.

---

<sup>41</sup>For example, while Aizer and Doyle (2015) report that the severity of juvenile punishment increases the propensity of further criminal activity as an adult (adult recidivism) in Cook County/Chicago, Eren and Mocan (2021) show that incarceration as a juvenile in Louisiana has no impact on the propensity to commit a violent crime, but that it increases the propensity of being convicted for a drug crime as an adult.

## References

- Abrams, D., R. Galbiati, E. Henry, and A. Philippe (2021). When in Rome... On Local Norms and Sentencing Decisions. *Journal of the European Economic Association, forthcoming*.
- Aizer, A. and J. J. Doyle Jr. (2015). Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges. *Quarterly Journal of Economics 130*(2), 759–803.
- Asmat, R. and L. Kossuth (2021). Gender Differences in Judicial Decisions under Incomplete Information: Evidence from Child Support Case. *Working Paper*.
- Bagues, M., M. Sylos-Labini, and N. Zinovyeva (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review 107*(4), 1207–1238.
- Bagues, M. F. and B. Esteve-Volart (2010). Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment. *Review of Economic Studies 77*(4), 1301–1328.
- Battaglini, M., J. M. Harris, and E. Patacchini (2020). Professional Interactions and Hiring Decisions: Evidence from the Federal Judiciary. *NBER Working Paper 26726*.
- Ben-Ner, A., L. Putterman, F. Kong, and D. Magan (2004). Reciprocity in a Two-Part Dictator Game. *Journal of Economic Behavior and Organization 53*(3), 333–352.
- Boyd, C. L., L. Epstein, and A. D. Martin (2010). Untangling the Causal Effects of Sex on Judging. *American Journal of Political Science 54*(2), 389–411.
- Butcher, K. F., K. H. Park, and A. M. 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.
- Choi, S. J., M. Gulati, M. Holman, and E. A. Posner (2011). Judging Women. *Journal of Empirical Legal Studies 8*(3), 504–532.
- Croson, R. and N. Buchan (1999). Gender and Culture: International Experimental Evidence from Trust Games. *American Economic Review 89*(2), 386–391.
- Croson, R. and U. Gneezy (2009). Gender Differences in Preferences. *Journal of Economic Literature 47*(2), 448–474.
- Dahl, G. B., K. V. Løken, and M. Mogstad (2014). Peer Effects in Program Participation. *American Economic Review 104*(7), 2049–2074.
- Dahlerup, D. (1988). From a Small to a Large Minority: Women in Scandinavian Politics. *Scandinavian Political Studies 11*(4), 275–298.
- De Paola, M. and V. Scoppa (2015). Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia. *Economica 82*(325), 162–188.

- Depew, B., O. Eren, and N. Mocan (2017). Judges, Juveniles, and In-Group Bias. *Journal of Law and Economics* 62(2), 209–239.
- Eck, K. and C. Crabtree (2020). Gender Differences in the Prosecution of Police Assault: Evidence from a Natural Experiment in Sweden. *PLOS ONE* 15(7), 1–16.
- Eckel, C. C. and P. J. Grossman (1996). The Relative Price of Fairness: Gender Differences in a Punishment Game. *Journal of Economic Behavior and Organization* 30(2), 143–158.
- Eckel, C. C. and P. J. Grossman (2008). Differences in the Economic Decisions of Men and Women: Experimental Evidence. In *Handbook of Experimental Economics Results*, eds. Charles R. Plott and Vernon L. Smith, Volume 1, Chapter 57, pp. 509–519. Elsevier.
- Eren, O. and N. Mocan (2021). Juvenile Punishment, High School Graduation and Adult Crime: Evidence from Idiosyncratic Judge Harshness. *Review of Economics and Statistics* 103, 34–47.
- Etzkowitz, H., C. Kemelgor, M. Neuschatz, B. Uzzi, and J. Alonso (1994). The Paradox of Critical Mass for Women in Science. *Science* 266(5182), 51 – 54.
- Fischman, J. B. (2015). Interpreting Circuit Court Voting Patterns: A Social Interactions Framework. *Journal of Law, Economics, and Organization* 31(4), 808–842.
- Grossman, G., O. Gazal-Ayal, S. D. Pimentel, and J. M. Weinstein (2016). Descriptive Representation and Judicial Outcomes in Multiethnic Societies. *American Journal of Political Science* 60(1), 44–69.
- Hoekstra, M. and B. Street (2021). The Effect of Own-Gender Jurors on Conviction Rates. *The Journal of Law and Economics* 64(3), 513–537.
- Holden, R., M. Keane, and M. Lilley (2021). Peer Effects on the United States Supreme Court. *Quantitative Economics* 12(3), 981–1019.
- Kanter, R. M. (1977). Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women. *American Journal of Sociology* 82(5), 965–990.
- Kastellec, J. P. (2013). Racial Diversity and Judicial Influence on Appellate Courts. *American Journal of Political Science* 57(1), 167–183.
- Knepper, M. (2018). When the Shadow Is the Substance: Judge Gender and the Outcomes of Workplace Sex Discrimination Cases. *Journal of Labor Economics* 36(3), 623–664.
- Liptak, A. (2009). The Waves Minority Judges Always Make. *New York Times* (May 31).
- Liu, H. H. (2021). Masked Evaluations: The Role of Gender Homophily. *The Journal of Legal Studies* 50(2), 303–330.
- Louisiana Children's Code (2012). Louisiana State Legislature-Various Chapters.
- Louisiana District Judges Association Periodicals (2000). Guide to the Louisiana Judiciary (1956-2000).

- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60(3), 531–542.
- Markussen, S. and K. Røed (2017). The Gender Gap in Entrepreneurship: The Role of Peer Effects. *Journal of Economic Behavior and Organization* 134, 356–373.
- Mas, A. and E. Moretti (2009). Peers at Work. *American Economic Review* 99(1), 112–145.
- Moffitt, R. (2001). Policy Interventions, Low-Level Equilibria, and Social Interactions. In *Social Dynamics*, eds. Steven N. Durlauf and H. Peyton Young, pp. 45–82. Cambridge, MA: MIT Press.
- Murphy, F. X. (2019). Does Increased Exposure to Peers with Adverse Characteristics Reduce Workplace Performance? Evidence from a Natural Experiment in the US Army. *Journal of Labor Economics* 37(2), 435–466.
- Peresie, J. L. (2005). Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts. *Yale Law Journal* 114(7), 1759–1790.
- Philippe, A. and A. Ouss (2018). No Hatred or Malice, Fear or Affection: Media and Sentencing. *Journal of Political Economy* 126(5), 2134–2178.
- Posner, R. A. (1983). *The Economics of Justice*. Harvard University Press. Cambridge MA.
- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics* 116(2), 681–704.
- The Supreme Court of Louisiana (2012). Annual Reports (1998-2012).

**Table 1: Summary Statistics for Juveniles and Judges**

	Mean	SD
<b>Panel A: Juvenile Characteristics</b>		
Incarceration (Secure Custody)	0.137	0.344
Sentence Length	507.16	298.01
Black	0.621	0.485
White	0.363	0.481
Female	0.236	0.425
Age	14.75	1.45
Committed a Felony	0.369	0.483
Sample Size	20,244	
<b>Panel B: Judge Characteristics</b>		
Female	0.239	0.428
Age at Disposition	53.09	8.63
Average Number of Judges in the Court	5.57	3.32
Number of Judges	138	

NOTES: The statistics above reflect our research sample, which consists of first-time juvenile offenders over the period from 1998 to 2012. The sample is restricted to juveniles whose disposition decisions were made in courts where there were at least two regular judges in the beginning-of-year.

**Table 2: Randomization Tests for Judge Peer Effects**

<b>Dependent Variable:</b> <b>(Proportion of Female Peers in the Courthouse)</b>	Coefficient (Standard Error)
Female	0.003 (0.003)
White	0.000 (0.002)
Juvenile Age	-0.000 (0.001)
<b>Offense Types:</b>	
Violent	-0.007 (0.005)
Property	-0.002 (0.004)
Drug	0.000 (0.003)
Felony	-0.004 (0.004)
Joint Significance ( <i>p-value</i> )	0.47
Sample Size	20,244

NOTES: Standard errors are clustered at the court level. The sample consists of courts where there were at least two judges in the beginning-of-year (1998-2012). Each cell represents a separate regression of the proportion of female peers on juvenile characteristics and offense types. Randomization regressions control for court and year fixed effects and court-specific trends. See text for further details.

**Table 3: Variation Analysis for Identification**

	Value
<b>Variation in Peer Measure</b>	
(Unit of Analysis: Judge-by-Year)	
Mean of the Proportion of Female Peers in the Courthouse	0.204
Standard Deviation of the Proportion of Female Peers in the Courthouse	0.324
Standard Deviation Change of the Proportion of Female Peers in the Courthouse from $t-1$ to $t$	0.074
Sum of Squares within Judges of the Proportion of Female Peers in the Courthouse (%)	6.89
Sum of Squares across Judges of the Proportion of Female Peers in the Courthouse (%)	93.11
Share of Judge-Year Observations in Courts with Turnover (%)	17.05

NOTES: The statistics above describe the variation in the proportion of female peers in the courthouse which we exploit in subsequent estimations.

**Table 4: Estimates of Judge Peer Effects on Incarceration and Sentence Length Imposed by Judges**

	Coefficient (Standard Error)	
	(1)	(2)
<b>Panel A: Incarceration</b>		
Proportion of Female Peers in the Court	0.087* (0.050)	0.098** (0.048)
<b>Panel B: Sentence Length</b>		
Proportion of Female Peers in the Court	62.530* (36.707)	67.802** (33.707)
Sample Size	20,244	20,244
<b>Controls:</b>		
Court and Year Fixed Effects	Yes	Yes
Court-Specific Trends	Yes	Yes
Judge Fixed Effects	Yes	Yes
Juvenile Characteristics	No	Yes
Offense Fixed Effects	No	Yes

NOTES: The sample consists of courts where there were at least two judges in the beginning-of-year (1998-2012). Standard errors are reported in parentheses and are clustered at the judge level. Juvenile controls include indicators for juvenile's gender and race and juvenile's age and its square. Offense fixed effects include indicators for type of offense a juvenile was convicted for (violent, property, drug-related and other offenses).

\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 5: Robustness Checks-Judge Peer Effects**

	Alternative Peer Measure Using Quarterly Variation	Drop Hurricane Katrina/Rita Regions	Judicial Districts Serving Single Parishes (Courthouses)	All Judges (Case Files Judge Handled>0)	Case Files Judge Handled>=150 in the Same Court
	Coefficient (Standard Error)				
	(1)	(2)	(3)	(4)	(5)
<b>Panel A: Incarceration</b>					
Proportion of Female Peers in the Court	0.098** (0.050)	0.160*** (0.049)	0.095 (0.063)	0.094** (0.045)	0.096* (0.050)
Mean of Outcome	0.137	0.122	0.168	0.137	0.141
<b>Panel B: Sentence Length</b>					
Proportion of Female Peers in the Court	58.173* (32.063)	66.854 (43.402)	45.314 (46.734)	67.679** (34.234)	66.958* (35.113)
Mean of Outcome	507.07	494.36	516.51	506.14	510.08
Sample Size	20,216	15,212	12,755	20,647	18,020
<b>Controls:</b>					
Court and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes	Yes

**Table 5 cont.**

	<b>Juvenile Offenders with Multiple Convictions Added</b>	<b>Control Judge-Specific Trends</b>	<b>Alternative Peer Measure-Using Judges Dealing with Juv. Cases</b>	<b>Case Disposition Year&lt;=2008</b>
	Coefficient (Standard Error)			
	(6)	(7)	(8)	(9)
<b>Panel A: Incarceration</b>				
Proportion of Female Peers in the Court	0.108** (0.052)	0.096* (0.055)	0.126*** (0.049)	0.137** (0.069)
Mean of Outcome	0.148	0.137	0.137	0.150
<b>Panel B: Sentence Length</b>				
Proportion of Female Peers in the Court	49.447* (28.075)	65.569* (38.830)	52.310 (40.132)	138.386*** (37.013)
Mean of Outcome	517.62	507.16	507.16	525.48
Sample Size	23,015	20,244	20,244	15,114
<b>Controls:</b>				
Court and Year Fixed Effects	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes

NOTES: Standard errors are reported in parentheses and are clustered at the judge level. Peer measure in Column 1 is constructed using quarterly variation in the proportion of female peers. Column 2 excludes parishes which were affected from the Hurricane Katrina/Rita (Orleans, Jefferson, St. Tammany, Plaquemines, St. Bernard, Lafourche, and Terrebonne). Column 3 excludes judicial districts presiding over multiple parishes (courthouses). Column 4 extends the effective sample to include juveniles without imposing any restrictions on the number of dispositions made by judges, while Column 5 limits the effective sample to include juveniles whose dispositions were made by judges with at least 150 cases in the same courthouse over the period from 1998 to 2012. Column 6 includes first-time offenders who were convicted for more than one statute offenses. Column 7 replaces court trends with judge-specific linear trends while Column 8 constructs the peer measure using judges dealing with juvenile cases. The last column excludes dispositions made after 2008.

**Table 6: Robustness Checks cont.-Judge Peer Effects**

	<b>Include Parish Level Crime/Other Controls</b>	<b>Drop Sex Offenses</b>	<b>Drop Years After the Major Elections</b>	<b>Drop Major Election Years and the Years After Elections</b>
	Coefficient (Standard Error)			
	(1)	(2)	(3)	(4)
<b>Panel A: Incarceration</b>				
Proportion of Female Peers in the Court	0.111*** (0.042)	0.100** (0.045)	0.099* (0.052)	0.085** (0.042)
Total Juvenile Violent Crimes in the Parish/100	-0.008 (0.022)			
Total Juvenile Property Crimes in the Parish/100	0.020*** (0.004)			
Total Officers Killed/Assaulted in the Parish/100	-0.003 (0.003)			
Mean of Outcome	0.137	0.133	0.139	0.137
<b>Panel B: Sentence Length</b>				
Proportion of Female Peers in the Court	63.163* (34.186)	77.931** (35.160)	81.287** (34.891)	71.187*** (26.926)
Total Juvenile Violent Crimes in the Parish/100	-21.624 (19.386)			
Total Juvenile Property Crimes in the Parish/100	6.885 (6.916)			
Total Officers Killed/Assaulted in the Parish/100	8.979** (4.433)			
Mean of Outcome	507.16	497.38	509.40	508.07
Mean of Juvenile Property Crimes	111.56			
Mean of Officers Killed/Assaulted	45.46			
Sample Size	20,244	19,561	17,197	14,491
<b>Controls:</b>				
Court and Year Fixed Effects	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes

NOTES: Standard errors are reported in parentheses and are clustered at the judge level. Column 1 includes the following parish level controls: number of adult property and violent crimes, number of juvenile property and violent crimes, number of officers feloniously killed or assaulted, log of county population, unemployment rate and fraction of the county population with a high school degree or less. Column 2 drops sex crime convictions. Column 3 drops judicial decisions that took place during the year after the major elections (2002 and 2008), while the last column drops observations in the election year and the year after the elections.

\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 7: Estimates of Judge Peer Effects-Controlling for Other Peer Attributes**

	Incarceration				Sentence Length			
	(1)	(2)	(3)	(4)	Coefficient (Standard Error)	(5)	(6)	(7)
Proportion of Female Peers in the Court	0.089*** (0.042)	0.098*** (0.048)	0.094*** (0.042)	0.087*** (0.037)	64.993* (38.428)	68.538** (34.611)	79.132** (32.374)	72.793* (39.674)
Proportion of Black Peers in the Court	0.046 (0.074)				0.043 (0.076)	18.342 (59.391)		41.457 (63.034)
Proportion of Democrat Peers in the Court		-0.005 (0.058)		-0.007 (0.057)		-65.251** (32.531)		-66.651** (31.653)
Average Age of Peers in the Court			-0.001 (0.002)	-0.000 (0.002)		1.536 (2.058)	1.536 (2.174)	1.793 (2.058)
Sample Size	20,113				20,113			
<b>Controls:</b>								
Court and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES: Standard errors are reported in parentheses and are clustered at the judge level. See Tables 4 and the text for further details.  
 \*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 8: Timing of Judge Peer Effects and Falsification Test**

	Incarceration				Sentence Length			
	(1)	(2)	(3)	(4)	Coefficient (Standard Error)	(5)	(6)	(7)
Proportion of Female Peers in the Court	0.117** (0.055)	0.112* (0.064)	0.070 (0.045)	0.056 (0.051)	91.368* (49.075)	94.155* (50.349)	-12.496 (35.546)	-41.037 (37.149)
Proportion of Female Peers in the Court ( $t+1$ )	0.017 (0.023)	-0.024 (0.030)			30.939 (47.385)	27.314 (41.351)		
Proportion of Female Peers in the Court ( $t+2$ )		0.027 (0.037)			20.128 (61.751)	20.128 (61.751)		
Proportion of Female Peers in the Court ( $t-1$ )			-0.002 (0.036)	-0.002 (0.044)		97.085** (48.958)	108.54** (50.610)	
Proportion of Female Peers in the Court ( $t-2$ )				0.014 (0.038)			-25.957 (42.982)	
Mean of Outcome	0.140	0.143	0.134	0.130	512.15	518.31	507.30	505.54
Sample Size	19,049	17,638	18,159	15,931	19,049	17,638	18,159	15,931
<b>Controls:</b>								
Court and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTES: Standard errors are reported in parentheses and are clustered at the judge level. See Table 4 and the text for further details.

\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 9: Estimates of Judge Peer Effects-Judge and Defendant Gender Interactions**

	Incarceration	Sentence Length
	Coefficient (Standard Error)	
	(1)	(2)
Proportion of Female Peers in the Court ( $\beta_1$ )	0.026 (0.039)	65.030 (43.529)
Proportion of Female Peers in the Court*Juvenile is Female ( $\beta_2$ )	-0.003 (0.034)	-25.929 (35.814)
Proportion of Female Peers in the Court*Judge is Female ( $\beta_3$ )	0.214*** (0.064)	63.105 (61.028)
Proportion of Female Peers in the Court*Juvenile is Female*Judge is Female ( $\beta_5$ )	-0.064 (0.087)	-136.117 (153.763)
<i>p-value</i> ( $\beta_1 + \beta_2$ )	0.62	0.38
<i>p-value</i> ( $\beta_1 + \beta_3$ )	0.00	0.00
<i>p-value</i> ( $\beta_1 + \beta_2 + \beta_3 + \beta_5$ )	0.00	0.85
Sample Size	20,244	20,244
<b>Controls:</b>		
Court and Year Fixed Effects	Yes	Yes
Court-Specific Trends	Yes	Yes
Judge Fixed Effects	Yes	Yes
Juvenile Characteristics	Yes	Yes
Offense Fixed Effects	Yes	Yes

NOTES: Standard errors are reported in parentheses and are clustered at the judge level. For each juvenile control, offense and judge fixed effects, the level term and its interaction with the focal variable (juvenile and judge's gender) are included in the specifications.

\*significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 10: Simulation Exercise-The Impact of Replacing One Male Judge with One Female Judge in a Courthouse of Two Female and Eight Male Judges (10 Judges in total)**

<b>Judge-Juvenile Defendant:</b>	<b>The Impact on Incarceration</b>	<b>The Impact on Sentence Length</b>
	(1)	(2)
Female Judge-Female Juvenile Defendant (N=1,555)	1.9 pp. (9.3%)	insignificant
Female Judge-Male Juvenile Defendant (N=5,068)	2.6 pp. (12.7%)	14 days (2.68 %)
Male Judge-Female Juvenile Defendant (N=3,230)	insignificant	insignificant
Male Judge-Male Juvenile Defendant (N=10,391)	insignificant	insignificant
Female Judge's Incarceration Rate/ Average Sentencing	0.205	525.72
Male Judge's Incarceration Rate/Average Sentencing	0.104	498.13

NOTES: The percentages in parentheses represent changes relative to the average judge-gender specific dispositions. N represents the sample sizes. See Table 8 the text for further details.

**Table 11: Distribution of Judge Harshness in Incarceration and Sentence Length by Judge's Gender**

	Full Sample	Proportion of Female Peers <=Sample Mean	Proportion of Female Peers>Sample Mean
	(1)	(2)	(3)
<b>Panel A: Female Judges</b> (Judge-by-Year)			
Mean of Judge Harshness in Incarceration	0.220	0.090	0.343
Mean of Judge Harshness in Sentence Length	549.88	494.16	602.63
<b>Panel B: Male Judges</b> (Judge-by-Year)			
Mean of Judge Harshness in Incarceration	0.129	0.122	0.141
Mean of Judge Harshness in Sentence Length	544.43	560.58	510.03

NOTES: The entries represent weighted means, where the weights are based on judges' total number of dispositions in each year. The mean proportion of females that are used in columns (2) and (3) are calculated separately for each year in the sample.

**Table 12: Estimates of Judge Peer Effects-Controlling for Average Peer Harshness of Judges Handling Juvenile Cases**

	Incarceration		Sentence Length			
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Full Sample (N=17,395)</b>						
Proportion of Female Peers in the Court	0.090* (0.053)		0.092* (0.056)	75.053** (34.991)	72.960** (34.924)	
Average Judge Peer Harshness in Incarceration		-0.071 (0.141)	0.033 (0.172)			
Average Judge Peer Harshness in Sentence Length/100			13.667 (20.368)	11.001 (17.087)		
<b>Panel B: Female Judge Sample (N=5,356)</b>						
Proportion of Female Peers in the Court	0.392*** (0.043)		0.394*** (0.042)	55.852 (86.286)	56.047 (86.393)	
Average Judge Peer Harshness in Incarceration		-0.043 (0.219)	-0.157 (0.261)			
Average Judge Peer Harshness in Sentence Length/100				14.590 (64.952)	14.983 (64.295)	
<b>Controls:</b>						
Court and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes	Yes	Yes
Judge Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Juvenile Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Offense Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

NOTES: Standard errors are reported in parentheses re clustered at the judge level. N represents the sample sizes. See Tables 4 and 7 and the text for further details.  
 \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

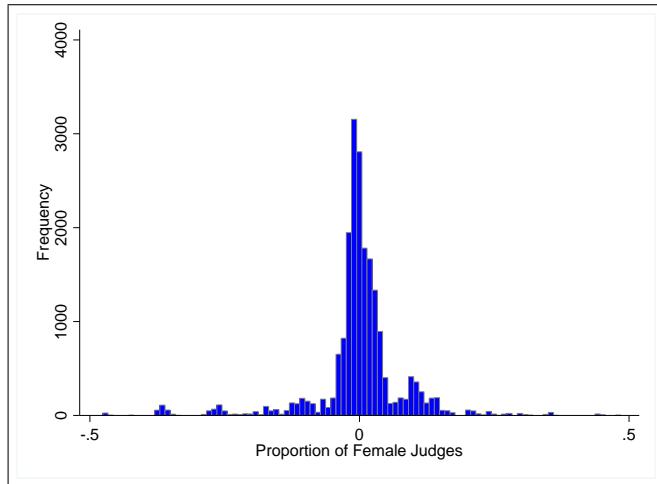


Figure 1: Distribution of Proportion of Female Peers-Residualized

NOTES: The residuals are obtained from a regression of proportion of female peers on judge, year and court fixed effects, court-specific trends and defendant characteristics.

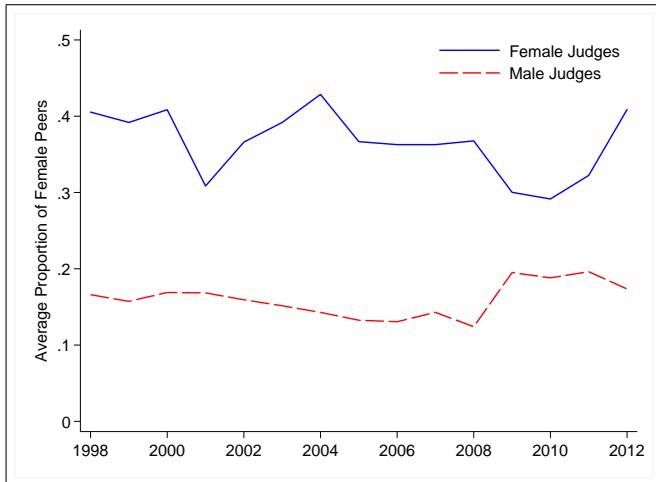
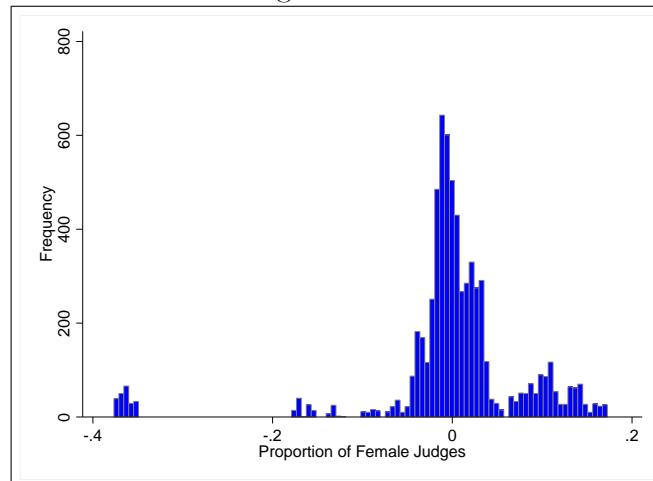


Figure 2: Average Proportion of Female Peers by Judge's Gender

NOTES: The analysis sample consists of courthouses where there were at least two judges in the beginning-of-year (1998 to 2012).

## Appendix:

Panel A: Female Judges



Panel B: Male Judges

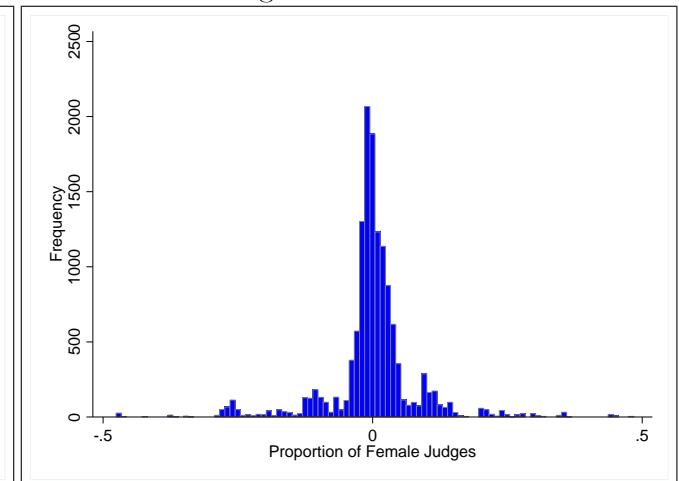


Figure A1: Distribution of Proportion of Female Peers-by Judge's Race-Residualized

NOTES: The residuals are obtained from a regression of proportion of female peers on judge, year and court fixed effects, court-specific trends and juvenile characteristics.

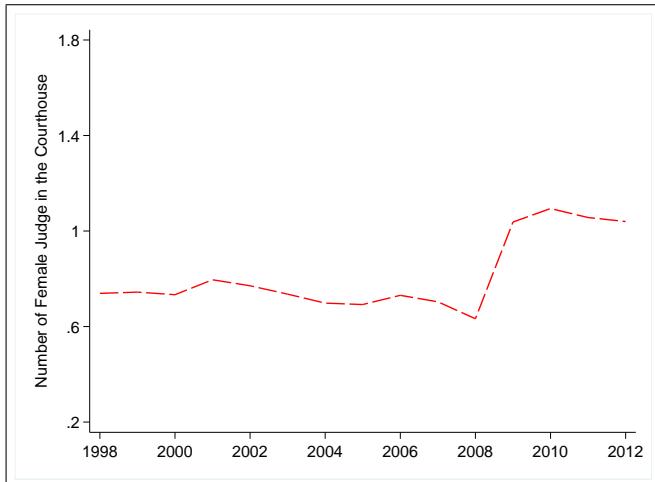


Figure A2(a): Number of Female Judges in the Courthouse

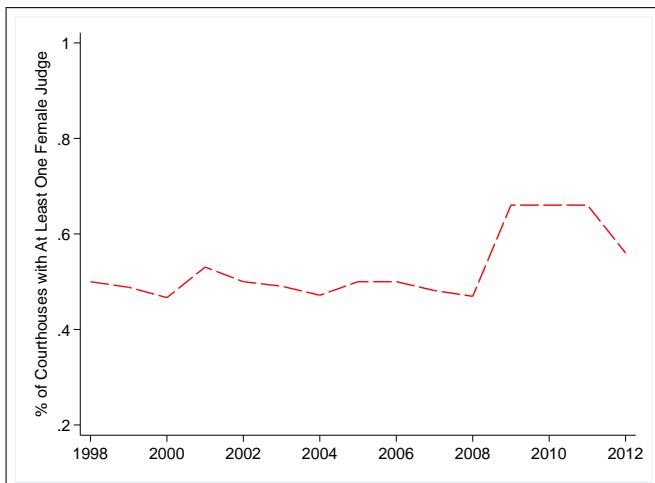


Figure A2(b): Fraction of Courthouses with at Least One Female Judge

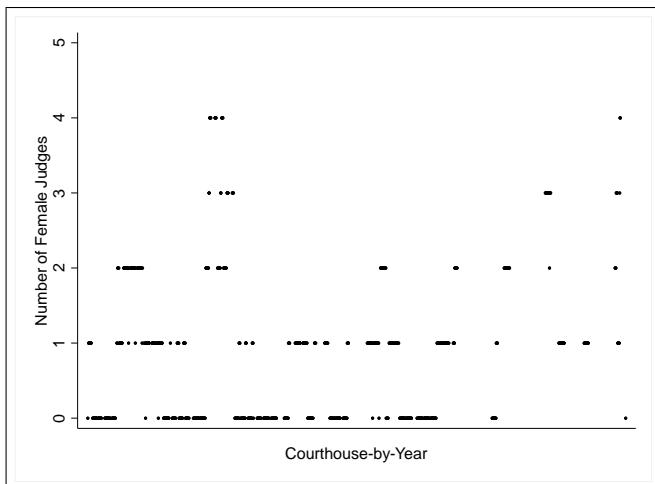


Figure A2(c): Number of Female Judges-Courthouse-by-Year

NOTES: The analysis sample consists of courthouses where there were at least two judges in the beginning-of-year (1998 to 2012).

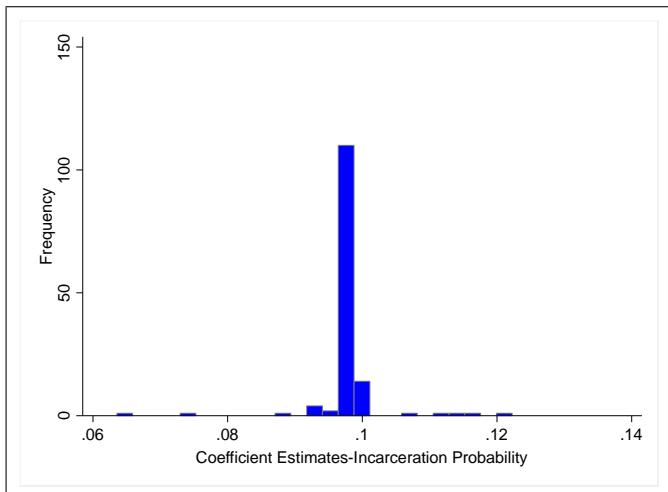


Figure A3: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Incarceration

NOTES: The distribution of the coefficient estimates of the proportion of female peers in the courthouse on incarceration decision set by the judges. The baseline specification is estimated repeatedly, each time removing dispositions set by a different judge. There are 138 judges in the effective sample.

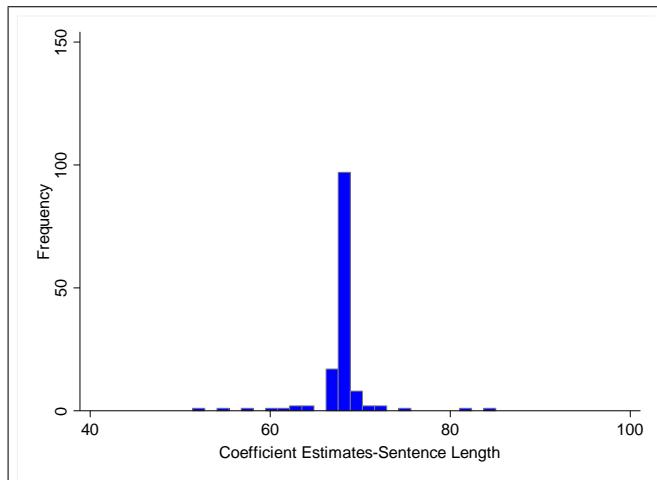


Figure A4: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Sentence Length

NOTES: The distribution of the coefficient estimates of the proportion of female peers in the courthouse on sentence length set by the judges. The baseline specification is estimated repeatedly, each time removing dispositions set by a different judge. There are 138 judges in the effective sample.

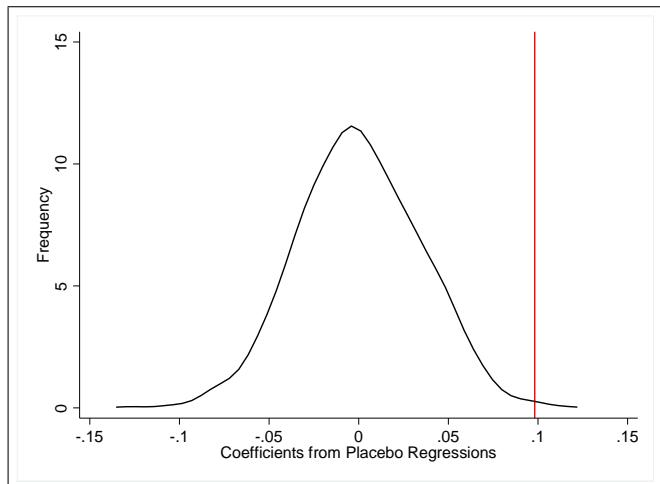


Figure A5: Placebo Coefficients of the Proportion of Female Peers in Incarceration Regression

NOTES: The figure displays the distribution of placebo coefficients of the proportion of female peers, where the proportions of female peers of a courthouse are randomly assigned to different years of the same courthouse. The vertical line represents the actual point estimate reported in Column 2 of Table 4.

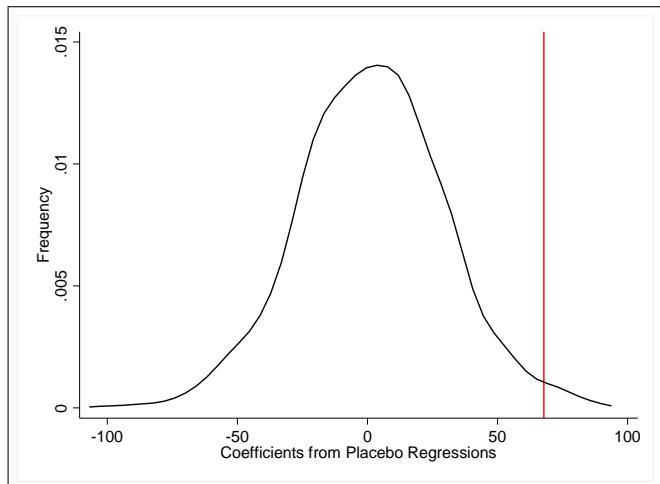


Figure A6: Placebo Coefficients of the Proportion of Female Peers in Sentence Length Regression

NOTES: The figure displays the distribution of placebo coefficients of the proportion of female peers, where the proportions of female peers of a courthouse are randomly assigned to different years of the same courthouse. The vertical line represents the actual point estimate reported in Column 2 of Table 4.