NBER WORKING PAPER SERIES

JUDGE PEER EFFECTS IN THE COURTHOUSE

Ozkan Eren
Naci H. Mocan

Working Paper 27713
http://www.nber.org/papers/w27713

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020

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, 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, and the Gender and Economics Workshop at the University of Luxembourg for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Judge Peer Effects in the Courthouse
Ozkan Eren and Naci H. Mocan
NBER Working Paper No. 27713
August 2020
JEL No. D9,D91,J16,J71,K4,K41

ABSTRACT

Although there exists a large literature analyzing whether an individual's peers have an impact on that individual's own behavior and subsequent outcomes, there is paucity of research on whether peers influence a person's decisions and judgments regarding a third party. 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 between 1998 and 2012. 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 fully driven by female judges. Further analysis suggests that this behavior 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.

Ozkan Eren
Department of Economics
University of California, Riverside
Sproul Hall
Riverside, CA 92521
ozkan.eren@ucr.edu

Naci H. Mocan
Department of Economics
Louisiana State University
2439 BEC
Baton Rouge, LA 70803-6306
and NBER
mocan@lsu.edu

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; Moffit 2001; Sacerdote 2001). This makes it difficult to identify the influence of the group on an individual who is associated with that group.[1]

Researchers tackled this issue by exploiting circumstances in which assignment to peer groups is arguably random. The most well-studied peer group effect pertains to students because random assignment to peer groups is a fairly common occurrence in case of students (e.g., Sacerdote 2001; Whitmore 2005; Carell et al. 2013). Recent literature expanded the analysis of peer effects to other domains. For example, it has been investigated whether exposure to more productive co-workers has an impact on workers' own productivity (Falk and Ichino 2006; Mas and Moretti 2009; Bandiera et al. 2010; Cornelissen et al. 2017). Individuals are shown to be more likely to take paternity leave if a co-worker was exogenously induced to take up leave (Dahl et al. 2014). Soldiers in the U.S. Army engage in higher rates of misconduct when they are randomly assigned to peers who have criminal backgrounds (Murphy 2019).

It is important to investigate whether an individual's own behavior, such as effort and productivity, as well as the pursuing outcome (e.g., test scores and wages) are influenced by his/her peers. It is, however, equally important to analyze whether decisions that impact an outside party are influenced by the composition of decision-makers' peers. In this paper we investigate whether consequential decisions made by judges about defendants in a courtroom are impacted by a particular attribute of these judges' peer group. Specifically,

[1] 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).

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.[2]

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.[3] 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. To the extent that within-judge variation in the gender peer composition is not systematically related to time-varying unobserved factors affecting judicial decisions, the causal effect of gender peer composition in the courthouse can be uncovered. We provide several robustness checks and falsification tests supporting our identifying assumption.

This paper makes contributions to two areas of investigation. First, it contributes to the peer effect literature by analyzing how gender diversity in peer composition impacts individual decision-making when the decision has important consequences for another person. A line of previous research has addressed a similar question, but this research has focused on decisions made by a panel of individuals, and reported conflicting results (e.g., Bagues and Esteve-Volart 2010; Anwar et al. 2012 and 2017; Bagues et al. 2017; De Paola and Scoppa 2017). Instead of decisions made by a panel, our paper analyzes the decisions made by individual decision-makers.

The second contribution of the paper is to the literature on judicial behavior. Whether race or gender of the defendant causes bias in judicial decisions and whether judicial decisions are impacted by personal attributes of judges has received significant attention (e.g., Mustard 2001; Alesina and Ferrara 2014; Rehavi

[2] Put differently, there is no involvement of a jury, or of any other judge regarding the decisions made on the case.

[3] 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.

and Starr 2014; Depew et al. 2017). Along these lines, researchers investigated if the race and gender composition of a panel of judges impacts case outcomes (Boyd et al. 2010; Kastellec 2013; Grossman et al. 2016).

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. For example, 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).

It is a difficult task to disentangle these various components that constitute the primary mechanism(s) behind the peer effect on decisions made by a panel. In contrast, in our study judges make solo decisions. That is, they do not involve in deliberations about their case with their peer judges, and they carry the sole responsibility of their decisions. Similarly, peer effects stemming from voting strategy and dissent aversion are not relevant in solo decisions. Thus, we isolate the effect of similarity/diversity; i.e., the impact of exposure to peers "who are similar or different."

We employ the universe of judicial decisions made about juvenile defendants in Louisiana between 1998 and 2012. Having constructed a panel containing all judges serving in each courthouse, we analyze whether

and to what extent the incarceration and sentence length decisions are influenced by the proportion of female peers to whom judges are exposed in the courthouse. 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 propose two mechanisms that may be responsible for this outcome and present results that provide suggestive evidence favoring one mechanism over the other. More precisely, our results 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.

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

2.1 Judicial Branch in Louisiana

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. District courts are the backbone of the state's justice system and they have original jurisdiction over all civil and criminal cases (Article V, Louisiana Constitution of 1974). 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.[4]

Judges in district courts and juvenile courts are elected by voters in partisan elections, where the party

[4] These parishes are Caddo, East Baton Rouge, Jefferson and Orleans.

affiliation of each candidate is shown on the ballot.[5] 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.[6] 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. 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 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.[7] 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.[8]

City courts constitute another component of the court system. 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.

2.2 Juvenile Justice System

Youth in Louisiana may enter the juvenile justice system when a formal complaint against them is filed by a citizen, parent, school official, or by the arresting police officer. Having received the complaint, the District Attorney's Office makes a decision about how to proceed with the case. They may choose to dismiss the case because of lack of sufficient evidence, they may refer the case to the Families in Needs of Services program, or

[5] 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.

[6] This structure is sometimes referred to as the "jungle primary."

[7] In Louisiana the mandatory retirement age for judges is 70.

[8] This person is ineligible as a candidate to be elected in the special election (Article V, Louisiana Constitution of 1974).

they may engage in informal adjustment arrangements with the juvenile and his/her parents.[9] Alternatively, prosecutors may proceed with a petition to the court. In this situation the case file is brought to the court. Cases are randomly assigned to judges unless the juvenile was found guilty in a previous adjudication.[10] 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.[11] 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.[12] There is no mandatory sentencing guidelines and judges exercise considerable discretion in sentencing. Additional details can be found in Eren and Mocan (2019).

[9] This could entail the juvenile participating in community service, restitution, or treatment and complying with certain behavioral requirements such as satisfactory school attendance (Louisiana Children's Code CHC 631).

[10] 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).

[11] The potential threat of sample selection bias stemming from this data restriction is discussed in Section 5.1.

[12] 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).

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*.[13] 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 judges' gender.[14]

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 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. Having imposed these restrictions, we end up with a total of 20,244 juvenile case files handled by 138 judges in 59 courthouses.[15]

[13] These annual collections are available at 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.

[14] We collect information on judges from variety of sources including online searches and Louisiana District Judges Association Periodicals (1956-2000).

[15] Fifty-three of these are district courthouses, two of them are city courts, and four are juvenile courts.

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.[16]

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

$$D_{ijct} = \beta_0 + \beta_1 F_{-j,ct} + X'_{ijct}\beta_2 + \lambda_j + \theta_t + \theta_c + \theta_c t + \epsilon_{ijct} \tag{1}$$

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 vector of observed juvenile characteristics (i.e., gender, race, age and its square and offense type), λ_j, θ_t and, θ_c stand for judge, year and court fixed effects, respectively, $\theta_c t$ are court-specific trends to control for linearly trending unobserved court characteristics and ϵ_{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

[16] Ungovernable (11 percent), simple battery (8 percent) and simple burglary (8 percent) are the most common type of offense types in the data.

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.[17] 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 peer 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 from the movements of peers in and out of courthouse.[18]

It is also conceivable that changes in the proportion of female peers are correlated with unobserved factors that also affect judicial decisions. Although random assignment of case files to judges mitigate

[17] For example, assume that a district courthouse consists of four judges, but only one of these four judges deals with juvenile cases, while the other three take on other types cases (e.g., commercial cases, civil cases, and adult crimes). In this case, the judge who deals with juvenile cases is included in the analysis and because he is exposed to his peers in the courthouse, the remaining three judges in the same courthouse are used to calculate the proportion of female peers in the courthouse.

[18] 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 subsequent 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.

concerns over omitted variable bias, it does not completely eliminate it because judges are not randomly assigned to their peers. Thus, we control for court and year fixed effects, as well as court-specific linear 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 β_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 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. 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 [19] We provide robustness checks (e.g., replacing court trends with judge-specific linear trends) and falsification tests in Section 5.1 to provide evidence on the validity of our identification strategy.

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

[19] To see this, consider a simplified version of equation (1)

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

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

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

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

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

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.

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).[20]

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 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 the courthouse to detect meaningful gender peer effects.

5 Results

5.1 Baseline Results and Robustness Checks

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

[20] We also experimented with similar randomization tests by controlling for court-by-year fixed effects. The point estimates from this exercise yield the same conclusion about randomization.

the courthouse to being exposed to 2 female and 8 male colleagues) leads to approximately 1.0 percentage 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).[21]

Recall that if the judge finds the defendant not guilty and acquits him/her, 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.[22]

[21] 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, Anwar et al. (2012) find that adding 1 to 2 black jurors to a pool of almost 30 potential jurors decreases conviction rates of black defendants more than 10 percent and increases it by a comparable amount for white defendants. 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.

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).

[22] 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). Given that the verdict on guilt vs. innocence is not a relevant margin for those who have plead guilty (or no contest) to the charge filed, estimating the impacts of incarceration for the plead-guilty sample provides insights in the extent of a potential bias. The sample of juvenile offenders who plead guilty is small (n= 7,118). In this sample the 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.

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.[23] 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.[24] 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 12,500, and lowers the precision of the estimates, but it does not alter the point estimates.

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, recognizing that juvenile sex offenders may be treated differently by judges and that the change in the gender peer composition may impact judicial decisions on sex offenders differently in comparison to other juvenile defendants, we dropped these cases from the sample, but the results remained intact (column 7).

Finally, 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 the

[23] We control for quarter fixed effects in this specification.

[24] These parishes are Jefferson, Lafourche, Orleans, Plaquemines, St. Bernard, St. Tammany, and Terrebonne.

last column of Table 5, are very similar in magnitude to those presented in Table 4.[25]

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 A1 and A2 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.

We also analyzed the timing of the gender peer effects by augmenting the model with the lags and leads of the gender composition measure. Table 6 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 6 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) and (6) 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 court house 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 A3 and A4 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 A3

[25] 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.

(p-value=0.00), and only 14 placebo estimates in Figure A4 are greater than the corresponding estimate in Table 4 (p-value=0.01).

5.2 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

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

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 β_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.[26]

Table 7 presents the results. Estimated β_1 and β_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 male judges are not influenced by the gender peer composition. More specifically, the incarceration decision

[26] 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.

and the assigned sentence length by male judges do not react to a change in the proportion of their female peers in the courthouse, regardless of whether the defendant is male 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 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.

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 7. 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.

Table 8 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 male judges, this event increases the proportion of female peers for female judges by 0.11.[27] Male judges do not react to the variation in the gender composition of their peers. This is true regardless of the gender of the defendant and both for the incarceration decision and the sentence length decision. On the other hand, an increase in the proportion of female peers in the courthouse by 0.11 prompts female judges to be tougher. They 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

[27] 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.

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 7 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.

5.3 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 female judges are harsher than their male counterparts. As displayed at the bottom of Table 8, 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). Thus, 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. 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.

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 divided judicial decisions into two groups: those made by judges when the proportion of female peers is less than 0.20 (which is the sample mean), and the decisions made when the proportion of female peers is greater than or equal to 0.20. Appendix Table A1 presents the probability of incarceration and the sentence length for both male and female judges in both circumstances. The incarceration rate and sentence length are similar between female and male judges when the proportion of female peers is low. On the other hand, the incarceration rate of female judges is three times higher than that of male judges (0.33 vs. 0.10) when judges are exposed to a larger proportion of female peers. Similarly, female judges assign sentences that are about 100 days longer in comparison to male judges when female judges have more female peers.

The information presented in Appendix Table A1 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.[28] 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 9 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 of Table 9, on the other hand, reveals that an increase in average incarceration rate of peers is

[28] 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.

not strongly related to the incarceration propensity of judges. Specifically, a 10 percentage point increase in average incarceration 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 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 9 should be interpreted with caution because of the reflection problem (Manski 1993), as average harshness of a judge's peers is likely 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 9 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.[29] Thus, the results displayed in Table 9 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.[30]

6 Conclusion

There is a large and growing body of research on peer effects, investigating in a variety of settings the impact of peers on one's own behavior or own outcomes. Examples include the impact of peers on student achievement, the effect of fellow employees on worker productivity, the influence of colleagues on paternity leave take-up, and the impact of peers in the army on soldiers' propensity to misconduct. In this paper we focus on judges who made decisions on juvenile defendants, and analyze the extent to which variation in the gender composition of judges in a courthouse has an impact on judicial decisions regarding incarceration and sentence length of defendants. The paper has aspects that are distinct from the existing literature. Most notably, while the majority of the peer effect literature is concerned with the effect of peers on one's own behavior, our paper analyzes the impact of peers on decisions made about another person (juvenile defendants).

A related, but separate line of research has investigated how peer interactions in a group-setting influence the decisions made by that group. More specifically, this strand of research has focused on the impact of race- or gender composition of a group of individuals on the decisions made by that group. Examples of this genre include the investigation of how the collective actions of a panel of judges, a group of jurors, or hiring

[29] 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 defendants only (See Table 8), 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.

[30] 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).

and promotion committees change in response to variations in racial or gender composition of the members of the group. In contrast, the judicial decisions we analyze are not the result of a group deliberation. Rather, they are solo-decisions made by individual judges in their own courtrooms. All judges, however, are exposed to their peers in the courthouses with which they are affiliated.

We use the universe of all case files from Louisiana juvenile courts from 1998 to 2012. We exploit random assignment of defendant case files to judges, and leverage turnover of judges in courthouses generated by such events as retirement, resignation, or death of judges, and by departure and arrival of judges to courthouses through elections. Our specifications control for judge, court and year fixed effects, as well as court-specific trends. Under the assumption that within-judge variation in the gender peer composition is not systematically related to time-varying unobserved factors affecting individual judicial decisions, one can uncover the causal effect of gender peer composition in the courthouse. Several robustness checks and falsification tests support our identification strategy.

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 sentence terms when the proportion of their female peers in the courthouse goes up. Further analyses to unbundle this effect reveal that it is driven solely by the reaction of female judges. Put differently, male judges do not respond to a change in the proportion of their female peers. Female judges, on the other hand, increase their stringency of punishment in response to an increase in the proportion of their female peers in the courthouse.

Female judges, on average, are more likely to incarcerate and they assign longer sentences than male judges, indicating that an increase in 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 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 all judges 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.

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 welfare. 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.[31] Notwithstanding, 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.

[31] 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 (2019) 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. Hjalmarsson (2009) analyzes data from the state of Washington and reports that incarceration in juvenile facilities reduces recidivism.

References

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.

Alesina, A. and E. La Ferrara (2014). A Test of Racial Bias in Capital Sentencing. *American Economic Review 104*(11), 3397–3433.

Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The Impact of Jury Race in Criminal Trials. *Quarterly Journal of Economics 127*(2), 1017–1055.

Anwar, S., P. Bayer, and R. Hjalmarsson (2017). A Jury of Her Peers: The Impact of the First Female Jurors on Criminal Convictions. *Economic Journal 129*(618), 603–650.

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.

Bandiera, O., I. Barankay, and I. Rasul (2010). Social Incentives in the Workplace. *Review of Economic Studies 77*(2), 417–458.

Battaglini, M., J. M. Harris, and E. Patacchini (2020). Professional Interactions and Hiring Decisions: Evidence from the Federal Judiciary. *NBER Working Paper 26726*.

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.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica 81*(3), 855–882.

Cornelissen, T., C. Dustmann, and U. Schönberg (2017). Peer Effects in the Workplace. *American Economic Review 107*(2), 425–456.

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.

Eren, O. and N. Mocan (2019). Juvenile Punishment, High School Graduation and Adult Crime: Evidence from Idiosyncratic Judge Harshness. *Review of Economics and Statistics, forthcoming*.

Etzkowitz, H., C. Kemelgor, M. Neuschatz, B. Uzzi, and J. Alonzo (1994). The Paradox of Critical Mass for Women in Science. *Science 266*(5182), 51 – 54.

Falk, A. and A. Ichino (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics 24*(1), 39–57.

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.

Hjalmarsson, R. (2009). Juvenile Jails: A Path to the Straight and Narrow or to Hardened Criminality? *Journal of Law and Economics 52*(4), 779–809.

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.

Liptak, A. (2009). The Waves Minority Judges Always Make. *New York Times* (May 31).

Louisiana Children's Code (2012). Louisiana State Legislature-Various Chapters (1998-2012).

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.

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 *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.

Mustard, D. B. (2001). Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts. *Journal of Law and Economics 44*(1), 285–314.

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.

Rehavi, M. M. and S. B. Starr (2014, dec). Racial Disparity in Federal Criminal Sentences. *Journal of Political Economy 122*(6), 1320–1354.

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).

Washington, E. L. (2008). Female Socialization: How Daughters Affect Their Legislator Fathers. *American Economic Review 98*(1), 311–332.

Whitmore, D. (2005). Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review 95*(2), 199–203.

Table 1: Summary Statistics for Juveniles and Judges

	Mean	SD
Panel A: Juvenile Characteristics		
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	
Panel B: Judge Characteristics		
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 a period from 1998 to 2012. The sample is further 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

Dependent Variable: (Proportion of Female Peers in the Courthouse)	Coefficient (Standard Error)
Female	0.003
	(0.003)
White	0.000
	(0.002)
Juvenile Age	-0.000
	(0.001)
Offense Types:	
Violent	-0.007
	(0.005)
Property	-0.002
	(0.004)
Drug	0.000
	(0.003)
Felony	-0.004
	(0.004)
Joint Significance (*p-value*)	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 in the courthouse on the juvenile characteristics and offense type. Randomization regressions control for court and year fixed effects and court-specific trends. See text for further details.
*significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3: Variation Analysis for Identification

	Value
Variation in Peer Measure	
(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)
Panel A: Incarceration		
Proportion of Female Peers in the Court	0.087*	0.098**
	(0.050)	(0.048)
Panel B: Sentence Length		
Proportion of Female Peers in the Court	62.530*	67.802**
	(36.707)	(33.707)
Sample Size	20,244	20,244
Controls:		
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	Juvenile Offenders with Multiple Convictions Added	Drop Sex Offenses	Control Judge-Specific Trends
	Coefficient (Standard Error)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Incarceration								
Proportion of Female Peers in the Court	0.098**	0.160***	0.095	0.094**	0.096*	0.108**	0.100**	0.096*
	(0.050)	(0.049)	(0.063)	(0.045)	(0.050)	(0.052)	(0.045)	(0.055)
Mean of Outcome	0.137	0.122	0.168	0.137	0.141	0.148	0.133	0.137
Panel B: Sentence Length								
Proportion of Female Peers in the Court	58.173*	66.854	55.314	67.679**	66.958*	48.489*	77.931**	65.569*
	(32.063)	(43.402)	(46.734)	(34.234)	(35.113)	(28.013)	(35.160)	(38.830)
Mean of Outcome	507.07	494.36	516.51	506.14	510.08	517.63	497.38	507.16
Sample Size	20,216	15,212	12,755	20,647	18,020	23,063	19,561	20,244
Controls:								
Court and Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Court-Specific Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
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. 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 more than one statute offenses and Column 7 excludes sex offenses. The last column replaces court trends with judge-specific linear trends.

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

Table 6: Timing of Judge Peer Effects and Falsification Test

	Incarceration			Sentence Length		
	Coefficient (Standard Error)					
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion of Female Peers in the Court	0.117**	0.112*	0.075	91.368*	94.155*	-41.037
	(0.055)	(0.064)	(0.051)	(49.075)	(50.349)	(37.149)
Proportion of Female Peers in the Court *(t+1)*	0.017	-0.024		30.939	27.314	
	(0.023)	(0.030)		(47.385)	(41.351)	
Proportion of Female Peers in the Court *(t+2)*		0.027			20.128	
		(0.037)			(61.751)	
Proportion of Female Peers in the Court *(t-1)*			-0.002			108.54**
			(0.044)			(50.610)
Proportion of Female Peers in the Court *(t-2)*			0.014			-25.957
			(0.038)			(42.982)
Mean of Outcome	0.140	0.143	0.130	512.15	518.31	505.54
Sample Size	19,049	17,638	15,931	19,049	17,638	15,931
Controls:						
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: The sample consists of courts where there were at least two judges in the beginning-of-year (1998 to 2012). Standard errors are reported in parentheses and are clustered at the judge level. See Table 4 and text for further details.

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

Table 7: Estimates of Judge Peer Effects-Triple Interaction Models

	Incarceration	Sentence Length
	Coefficient (Standard Error)	
	(1)	(2)
Proportion of Female Peers in the Court (β_1)	0.026	65.030
	(0.039)	(43.529)
Proportion of Female Peers in the Court*Juvenile is Female (β_2)	-0.003	-25.929
	(0.034)	(35.814)
Proportion of Female Peers in the Court*Judge is Female (β_3)	0.214***	63.105
	(0.064)	(61.028)
Proportion of Female Peers in the Court*Juvenile is Female*Judge is Female (β_5)	-0.064	-136.117
	(0.087)	(153.763)
p-value ($\beta_1 + \beta_2$)	0.62	0.38
p-value ($\beta_1 + \beta_3$)	0.00	0.00
p-value ($\beta_1 + \beta_2 + \beta_3 + \beta_5$)	0.00	0.85
Sample Size	20,244	20,244
Controls:		
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 8: 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)

Judge-Juvenile Defendant:	The Impact on Incarceration	The Impact on Sentence Length
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 sample consists of courts where there were at least two judges in the beginning-of-year (1998 to 2012). The numbers in parentheses are relative to average judge gender-specific dispositions. N represents the sample sizes. See Table 7 and the text for further details.

Table 9: Estimates of Judge Peer Effects-Controlling for Average Peer Harshness of Judges Handling Juvenile Cases

	Incarceration			Sentence Length		
	Coefficient (Standard Error)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Sample (N=17,395)						
Proportion of Female Peers in the Court	0.090*		0.092*	75.053**		72.960**
	(0.053)		(0.056)	(34.991)		(34.924)
Average Judge Peer Harshness in Incarceration		-0.071	0.033			
		(0.141)	(0.172)			
Average Judge Peer Harshness in Sentence Length/100					13.667	11.001
					(20.368)	(17.087)
Mean of the Proportion of Female Peers in the Courthouse	0.311					
Mean of the Average Judge Peer Harshness in Incarceration	0.167					
Mean of the Average Judge Peer Harshness in Sentence Length	538.04					
Panel B: Female Judge Sample (N=5,356)						
Proportion of Female Peers in the Court	0.392***		0.394***	55.852		56.047
	(0.043)		(0.042)	(86.286)		(86.393)
Average Judge Peer Harshness in Incarceration		-0.043	-0.157			
		(0.219)	(0.261)			
Average Judge Peer Harshness in Sentence Length/100					14.590	14.983
					(64.952)	(64.295)
Mean of the Proportion of Female Peers in the Courthouse	0.537					
Mean of the Average Judge Peer Harshness in Incarceration	0.254					
Mean of the Average Judge Peer Harshness in Sentence Length	561.80					
Controls:						
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: The average peer harshness measures are calculated using judges handling juvenile cases and are based on the dispositions of the first-time delinquents who were convicted for only one statute offense. The analysis sample is further restricted to courts where there were at least two judges handling juvenile cases. Standard errors are reported in parentheses are 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%.

Appendix:

Table A1: Distribution of Incarceration Probability and Sentence Length by the Proportion of Female Peers in the Courthouse and Judge's Gender

	Proportion of Female Peers <=Sample Mean	Proportion of Female Peers>Sample Mean
	(1)	(2)
Panel A: Female Judges		
Incarceration	0.079	0.328
Sentence Length	475.01	575.20
Panel B: Male Judges		
Incarceration	0.106	0.098
Sentence Length	509.47	471.04

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

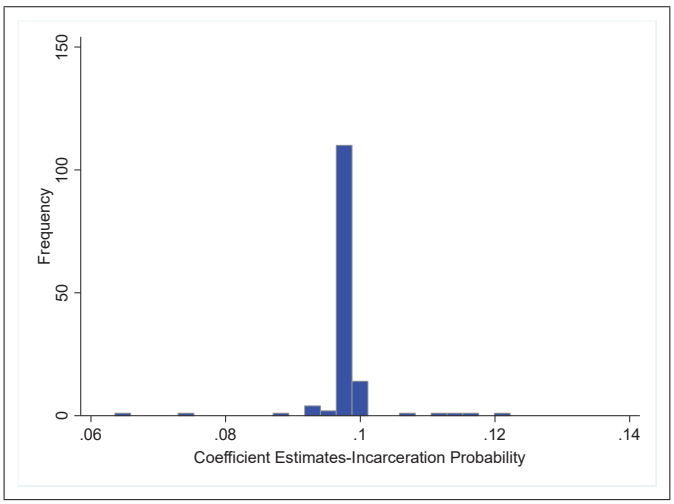

Figure A1: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Incarceration

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.

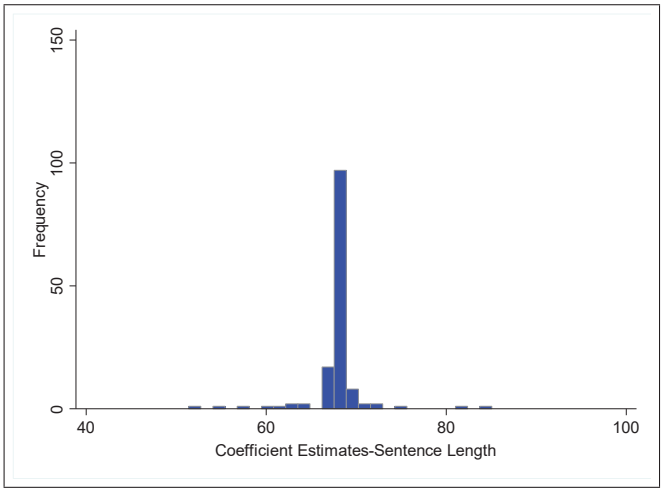

Figure A2: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Sentence Length

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.

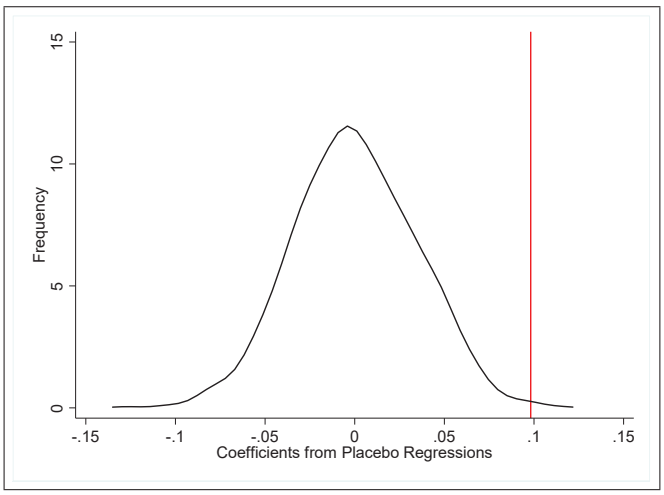

Figure A3: Placebo Coefficients of the Proportion of Female Peers in Incarceration Regression

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.

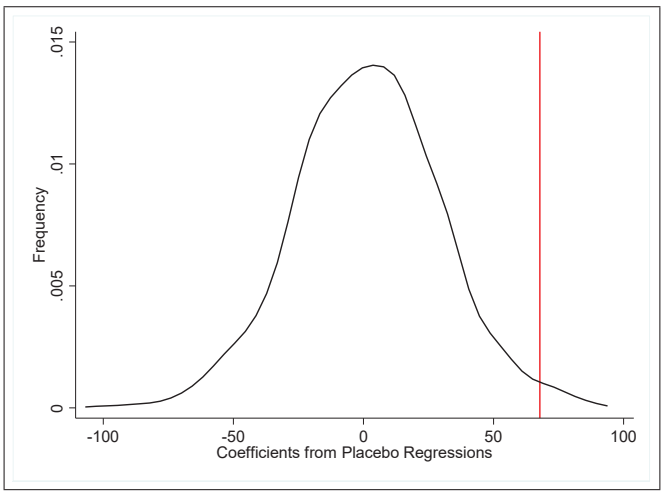

Figure A4: Placebo Coefficients of the Proportion of Female Peers in Sentence Length Regression

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.