

Eren, Ozkan; Mocan, Naci

Working Paper

Judge Peer Effects in the Courthouse

IZA Discussion Papers, No. 13937

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Eren, Ozkan; Mocan, Naci (2020) : Judge Peer Effects in the Courthouse, IZA Discussion Papers, No. 13937, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/232689>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPER SERIES

IZA DP No. 13937

Judge Peer Effects in the Courthouse

Ozkan Eren
Naci Mocan

DECEMBER 2020

DISCUSSION PAPER SERIES

IZA DP No. 13937

Judge Peer Effects in the Courthouse

Ozkan Eren

University of California

Naci Mocan

Louisiana State University and IZA

DECEMBER 2020

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Judge Peer Effects in the Courthouse*

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. We also demonstrate that the impact of proportion of female peers is not a proxy for other peer characteristics such as race and age. 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.

JEL Classification: D9, K4

Keywords: peers, judicial, harshness, leniency, sentencing, judge, critical mass, juveniles, crime, court

Corresponding author:

Naci Mocan
Louisiana State University
Department of Economics
3039 Business Education Complex
Baton Rouge, LA 70803
USA
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, 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.

1 Introduction

The investigation of how an individual’s behavior changes as the attributes of his/her peers are altered is important both for scientific inquiry and for designing strategies to improve performance and productivity. The most well-studied peer group effect pertains to students because random assignment of individuals to peer groups is a fairly common occurrence in case of students (e.g., Sacerdote 2001; Whitmore 2005; Carrell et al. 2013; Anelli and Peri 2019). Recent literature expanded the analysis of peer effects to other domains ranging from worker productivity (Mas and Moretti 2009) to paternity leave decision (Dahl et al. 2014), to misconduct in the army (Murphy 2019).

While it is important to investigate whether an individual’s own behavior, such as effort and productivity, as well as the pursuing outcomes (e.g., test scores, choice of college major, wages) are influenced by his/her peers, it is equally important to analyze whether decisions that impact an outside party are influenced by the composition of the decision-makers’ peers. Recent research has addressed this question, but this research has focused on decisions made by a panel, rather than by individuals, and reported conflicting results (Basques and Esteve-Volart 2010; Anwar et al. 2012; Bagues et al. 2017; De Paola and Scoppa 2017).

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, deliberations can allow the numerical majority members of a panel to be persuaded by a member who is in numerical minority. Second, voting strategy on a panel is important because in most settings members of the panel may prefer to avoid casting dissenting votes (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.¹ 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 this paper we analyze the impact of peer composition on individual decision-makers. Specifically,

¹For 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).

we investigate whether incarceration and sentencing decisions made by juvenile court judges in their own courtrooms are impacted by the gender composition of their peers in the courthouse. These decisions do not involve deliberations with peer judges, and judges carry the sole responsibility of their decisions. Similarly, peer effects stemming from voting strategy and dissent aversion are not relevant in these solo decisions. Thus, we isolate the effect of exposure to peers “who are similar or different.”²

We estimate the effect of peer composition off of 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. Although random assignment of juvenile case files to judges 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 provide extensive robustness checks and falsification tests supporting our identifying assumption.

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 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 or average age of judges. 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.

These results are consistent with experimental evidence showing that women are more likely to reciprocally punish unfair behavior than men (Eckel and Grossman 1996; Croson and Buchan 1999) and that women’s decisions are sensitive to the context in which they are made (Croson and Gneezy 2009).

²Whether 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 and Starr 2014; Depew et al. 2017).

2 Institutional Background and the Data

There are 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 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.

Judges in district courts and juvenile courts are elected by voters in partisan elections, where candidates for judge positions compete in a primary election against other candidates, and it is common to have multiple candidates from a given political party to run for the same judge position. District and juvenile court judges serve six-year terms and they are eligible for re-election. Elections are spread throughout the year. Judge turnover 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. The state supreme court appoints a judge to the bench until a special election is held within a year.

Cases are randomly assigned to judges unless the juvenile was found guilty in a previous adjudication. Such repeat offenders are re-assigned to judges who handled the previous episode of the juvenile. If the judge finds the defendant not guilty, 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 incarcerated or they can be assigned a less severe arrangement (placement in residential non-secure treatment facility such as a group home or foster home.) Similarly, the judge can put the juvenile on probation. 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 their type of custody. There is no mandatory sentencing guidelines and judges exercise considerable discretion in sentencing.

The main data are obtained from two sources. Louisiana Department of Public Safety and Corrections, Youth Services, Office of Juvenile Justice 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 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 collections of the Supreme Court of Louisiana (*The Guide to Louisiana Courts*).³ These annual periodicals feature statewide list of judges along with information on courthouse served and judicial turnovers, and allows us to track movements in and out of courthouses (e.g., election, retirement and resignation) with exact dates. We supplement these data with information on judge attributes.

To construct the gender peer measure, we use those courthouses that have at least two judges in a given year. We exclude judges from the benchmark model if they had handled fewer than 25 case files in a courthouse over 1998- 2012, although as shown in the Appendix the results are not sensitive to this restriction. To avoid potential confounding that may arise from multiple offenses and/or criminal history of the juvenile, we focus on first-time delinquents ages 10 through 17 who were convicted for only one statute offense (repeat offenders are assigned to the same judge who handled the original case). The analysis sample contains 20,244 juvenile case files handled by 138 judges in 59 courthouses located in 42 judicial districts. Additional details on institutional structure and the data can be found in Eren and Mocan (2020).

Table 1 presents the descriptive statistics. The incarceration rate is about 14 percent. 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. 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.

³These annual collections are available at http://www.lasc.org/press_room/annual_reports/default.asp.

3 Empirical Methodology

We estimate the following equation

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

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 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.

We use the universe of judges in the state to construct $F_{-j,ct}$. More specifically, in the creation of the peer measure we utilize all judges, even those who do not enter the analysis sample.⁴ There are 10 judicial districts where each district covers multiple parishes, and judges in these districts rotate between the courthouses of different parishes. 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. 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. 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

⁴For 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.

within-judge variation, identification is obtained from the movements of peers in and out of courthouse.⁵ We also control for court and year fixed effects and court-specific trends in all specifications. Nevertheless, there could be court-specific shocks that may be correlated with peer composition, which call for conditioning on court-by-year fixed effects. Doing so, however, would result in almost no variation to exploit for identification.⁶ We provide robustness checks (e.g., replacing court trends with judge-specific linear trends) and falsification tests to provide evidence on the validity of our identification strategy.

We examine the validity of random assignment of case files to judges by running 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. The details of the randomization tests as well as the variation in the proportion of female peers are discussed in Section 1 of the Appendix.

4 Results

4.1 Baseline Results

Baseline estimates are presented in Table 2. 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. The results in Panel A demonstrate 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.,

⁵Judges, 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.

⁶To 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

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

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.

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 point 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).

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.

Some of the juveniles in the data 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 the judge made was the sentencing. Estimating the impact of judges’ female peers in this subsample 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.

If the proportion of female peers is correlated with some average judge attribute, any reaction to the variation in the proportion of female peers may reflect the impact of that attribute. 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. Similarly, female judges may be younger (or older) and peers’ average

age and not their gender composition may be related to judicial decisions. The analyses detailed in the Appendix refute this conjecture because adding to the model peers' race and their average age does not eliminate the gender peer effect (see Section 2 of the Appendix, in conjunction with Appendix Table A3).

We performed extensive robustness tests involving the calculation of the peer measure quarterly rather than annually, changing model specification in different ways, analyzing the impact on the lags and leads of the peer measure, exclusion of judicial districts, inclusion/exclusion of juveniles with multiple offenses and crime types (e.g., sex offenses, felony vs. non-felony offenses), analyzing whether the results are driven by a particular judge or by judges' case loads, assigning the actual peer values randomly to different years as a placebo test, and so on. These analyses are discussed in Section 2 of the Appendix.

4.2 Unbundling the Impact

To unbundle the overall gender peer effect, we investigate whether the estimated effect is sensitive to judge-defendant gender match. That is, we analyze 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 by estimating the following specification

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

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 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.⁷

⁷These specifications control for interactions of juvenile's gender with individual characteristics, offense and judge fixed effects

Table 3 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 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 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.

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 3. This magnitude implies that if the proportion of 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 4 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. Male judges do not react to the variation in the gender composition of their peers. On the other hand, this increase in the proportion of female peers in the courthouse 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 significant impact on sentence lengths

as well as interactions between judge's gender and individual characteristics and offense fixed effects.

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 3 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 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. As displayed at the bottom of Table 4, 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 harshness 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.

To gauge judge harshness, we calculated the incarceration rate and the sentence length at the judge-by-year level. Appendix Table A6 displays the means, weighted by total number of cases in each judge-by-year cell. As discussed in Section 3 of the Appendix, we find that male judges' assigned sentence length and the incarceration rate do not change appreciably regardless of whether they face low or high proportion of female peers. On the other hand, when female judges are exposed to a larger proportion of female peers, their average incarceration rate quadruples and their average assigned sentence length goes up by more than 100 days. Thus, the information presented in Appendix Table A6 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.⁸ 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 5 presents the results. 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 2. 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 rate of judge's peers is associated with a statistically insignificant 0.7 percentage point decrease in judges'

⁸The 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 3.

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 5 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. 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 that the influence of average peer harshness is smaller and indistinguishable from zero.

The coefficients reported in Table 5 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 5 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.⁹ Thus, the results

⁹The 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 3), 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

displayed in Table 5 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.

6 Conclusion

There is a large and growing body of research investigating peer effects in a variety of settings. This paper has aspects that are distinct from the existing peer effect literature. Most notably, while this literature is almost exclusively concerned with the effect of peers on one’s own behavior, our paper analyzes the impact of peers on decisions made about another person.

We use the universe of case files from Louisiana juvenile courts from 1998 to 2012. 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 solely by the reaction of female judges. Male judges do not respond to a change in peer gender composition. 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. We show that the proportion of female peers is not a proxy for the racial composition of peers or for average age of peers. Adding to the model the proportion of black judges and the average judge age in the courthouse 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 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 indicates that the

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.

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 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 sensitive to context (Eckel and Grosman 1996; Ben-Ner et al. 2004). As explained in detail by Croson and Gneezy (2009), women are sensitive 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 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.¹⁰ Notwithstanding, our results from a simulation exercise indicate an additional five juvenile incarcerations and 3,022 extra days of prison time imposed on juveniles per year.

¹⁰For example, while Aizer and Doyle (2015) report that the severity of juvenile punishment increases the propensity of criminal activity as an adult 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 as an adult, but that it increases the propensity of being convicted for a drug crime. Hjalmarsson (2009) reports that incarceration in juvenile facilities reduces recidivism in the state of Washington.

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 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: 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* (0.050)	0.098** (0.048)
Panel B: Sentence Length		
Proportion of Female Peers in the Court	62.530* (36.707)	67.802** (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 3: 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 (β_1)	0.026 (0.039)	65.030 (43.529)
Proportion of Female Peers in the Court*Juvenile is Female (β_2)	-0.003 (0.034)	-25.929 (35.814)
Proportion of Female Peers in the Court*Judge is Female (β_3)	0.214*** (0.064)	63.105 (61.028)
Proportion of Female Peers in the Court*Juvenile is Female*Judge is Female (β_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
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 4: 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 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 5: 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* (0.053)		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)
Panel B: Female Judge Sample (N=5,356)						
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)
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: 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%.

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.
- Anelli, M. and G. Peri (2019). The Effects of High School Peers’ Gender on College Major, College Performance and Income. *Economic Journal* 129(618), 553–602.
- Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The Impact of Jury Race in Criminal Trials. *Quarterly Journal of Economics* 127(2), 1017–1055.
- 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.
- Ben-Ner, A., L. Putterman, F. Kong, and D. Magan (2004). Reciprocity in a Two-Part Dictator Game. *Journal of Economic Behavior & Organization* 53(3), 333–352.
- 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.
- 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.
- 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.
- Eckel, C. C. and P. J. Grossman (1996). The Relative Price of Fairness: Gender Differences in a Punishment Game. *Journal of Economic Behavior & 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 (2019). Juvenile Punishment, High School Graduation and Adult Crime: Evidence from Idiosyncratic Judge Harshness. *Review of Economics and Statistics*, forthcoming.
- Eren, O. and N. Mocan (2020). Judge Peer Effects in the Courthouse. *NBER Working Paper 27713*.
- 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.
- 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).
- 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.
- 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.
- Posner, R. A. (1983). *The Economics of Justice*. Harvard University Press. Cambridge MA.
- Rehavi, M. M. and S. B. Starr (2014). 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).
- Whitmore, D. (2005). Resource and Peer Impacts on Girls’ Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review* 95(2), 199–203.

ONLINE APPENDIX:

1 Randomization Tests and Variation in the Peer Measure

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 A1. 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 A1).¹

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 A2 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 judges in courthouses to detect meaningful gender peer effects.

Figure A1 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 juvenile characteristics. Consistent with Table A2, non-negligible identifying variation is displayed in Figure A1, where the standard

¹We 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.

deviation of the residuals is 0.083. Figure A2 presents this information by judge gender.

Finally, Figure A3 presents the proportion of female peers, averaged across male and female judges over the years, and demonstrates the lack of any secular trend.

2 Additional Estimations and Robustness Checks

It is plausible for the proportion of female peers to be correlated with some average judge attribute. Thus, any reaction to a variation in the proportion of female peers may reflect the impact of that attribute. 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 A3, 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. 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 (2) of Table A3 refutes this conjecture. The average age of peers 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 (3) demonstrates that inclusion of peers' race and age to the model jointly does not alter the inference. Columns (4) to (6) of Table A3 present the results of the same exercise, where the outcome is the sentence length. Once again, adding to the model peers' race and their average age does not eliminate the gender peer effect.

We implement several sensitivity checks to examine the robustness of our results. The first column of Table A4 presents the results from a specification where the measure of gender peer composition ($F_{-j,ct}$) is constructed at the year-by-quarter level.² 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 2. Second, the turmoil during and after hurricanes Katrina and Rita may have

²We control for quarter fixed effects in this specification.

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.³ As shown in column (2) of Table A4, doing so provides a larger effect on incarceration, while the point estimate for sentence length is almost identical to those reported in Table 2. 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 A4, 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 last column of Table A4, are very similar in magnitude to those presented in Table 2.⁴

The magnitude of the gender peer effect is similar between severe and less severe 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 and non-felony offenses are similar when they are benchmarked relative to their own sample means.

³These parishes are Jefferson, Lafourche, Orleans, Plaquemines, St. Bernard, St. Tammany, and Terrebonne.

⁴We 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.

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 A4 and A5 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 analyzed the timing of the gender peer effects by augmenting the model with the lags and leads of the gender composition measure. Table A5 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 A5 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 A6 and A7 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 2. Only 3 of the 1,000 placebo regressions produce effects that are larger than the actual value in Figure A6 (p-value=0.00), and only 14 placebo estimates in Figure A7 are greater than the corresponding estimate in Table 2 (p-value=0.01).

3 Judge Stringency by Gender

We calculated the incarceration rate and the sentence length at the judge-by-year level. Table A6 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).⁵

Columns (2) and (3) of Table A6, 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 A6 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. 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) of Table A6, 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 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

⁵The 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).

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.

Table A1: 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 (<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 A2: 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 A3: Estimates of Judge Peer Effects Controlling for Proportion of Black Judges and Average Age of Peers

	Incarceration			Sentence Length		
	(1)	(2)	(3)	(4)	(5)	(6)
	Coefficient (Standard Error)					
Proportion of Female Peers in the Court	0.089** (0.042)	0.090** (0.041)	0.084** (0.036)	64.365* (38.372)	70.256** (33.315)	66.911* (37.423)
Proportion of Black Peers in the Court	0.047 (0.074)		0.038 (0.075)	17.653 (59.275)		22.214 (63.407)
Average Age of Peers in the Court		-0.001 (0.002)	-0.001 (0.002)		0.352 (2.202)	0.492 (2.318)
Mean of the Proportion of Black Peers in the Court	0.136					
Mean of the Average Age of Peers	53.96					
Sample Size	20,244			20,244		
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: 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 A4: 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	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Panel A: Incarceration								
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)	0.108*** (0.052)	0.100** (0.045)	0.096* (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* (32.063)	66.854 (43.402)	55.314 (46.734)	67.679** (34.234)	66.958* (35.113)	48.489* (28.013)	77.931** (35.160)	65.569* (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 for 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 A5: 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** (0.055)	0.112* (0.064)	0.075 (0.051)	91.368* (49.075)	94.155* (50.349)	-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)	
Proportion of Female Peers in the Court ($t-1$)			-0.002 (0.044)			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	0.130	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: 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 A6: 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)
Panel A: Female Judges (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
Panel B: Male Judges (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.

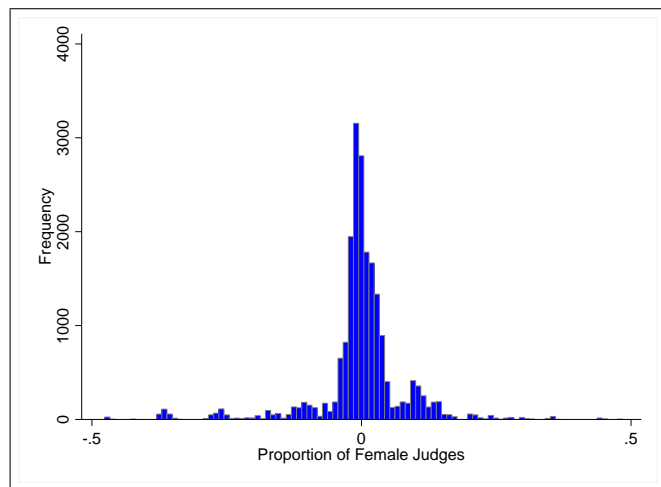
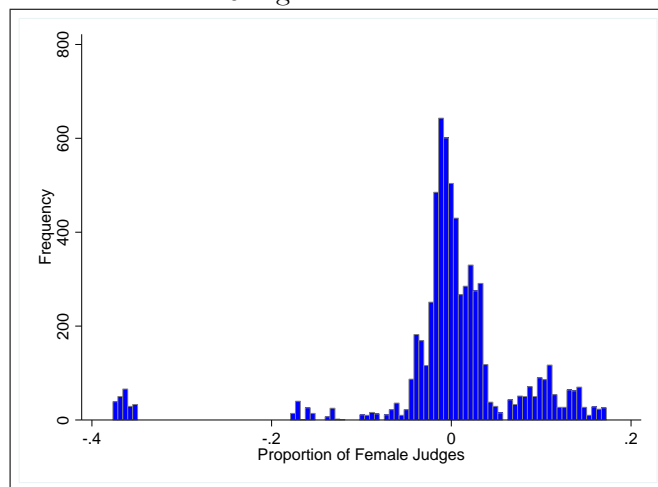


Figure A1: Distribution of Proportion of Female Peers-Residualized

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

Panel A: Female Judges



Panel B: Male Judges

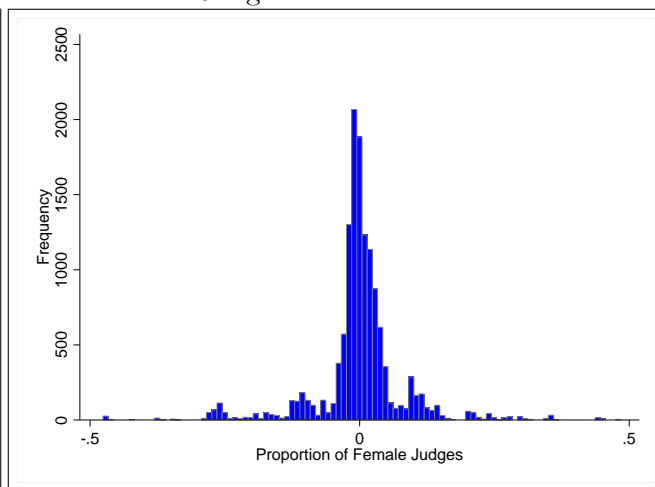


Figure A2: Distribution of Proportion of Female Peers-by Judge's Race-Residualized

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.

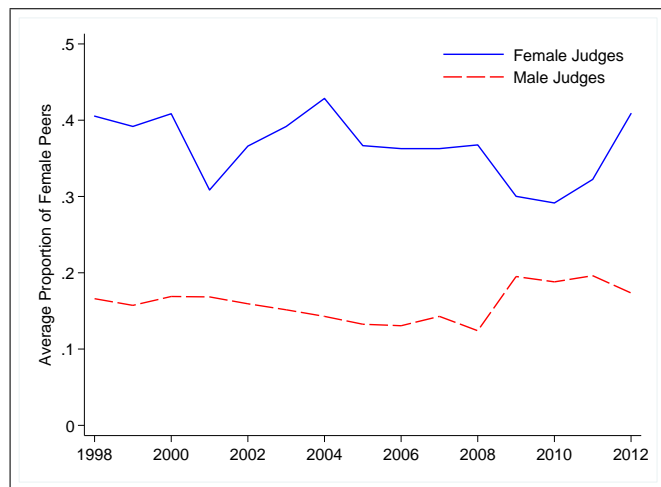


Figure A3: Average Proportion of Female Peers by Judge's Gender

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

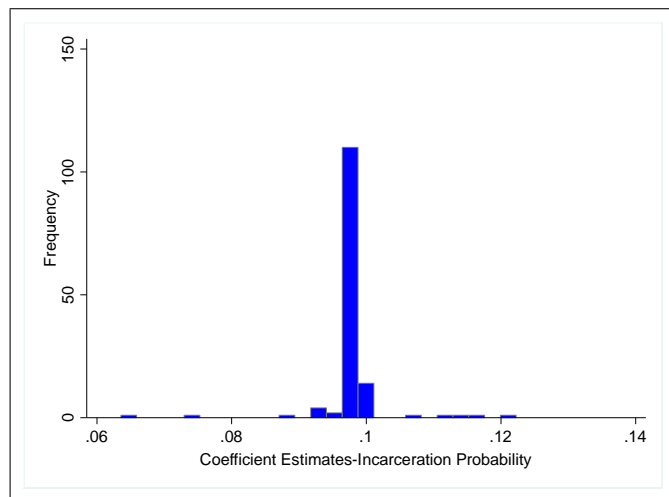


Figure A4: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Incarceration

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.

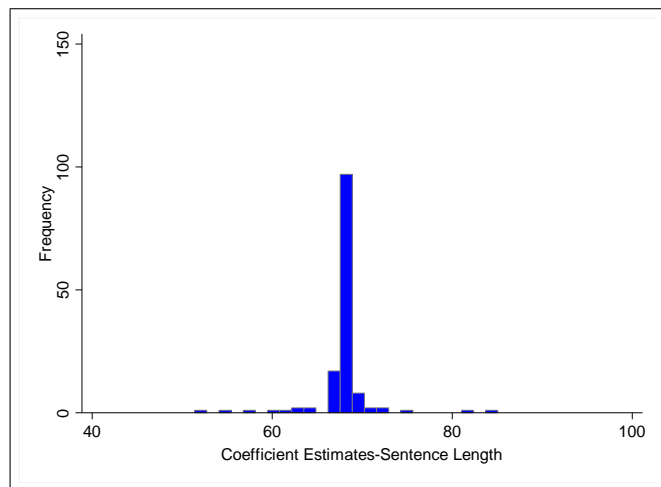


Figure A5: Distribution of Leave-One Judge Out Estimates of the Gender Peer Effects in the Courthouse on Sentence Length

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.

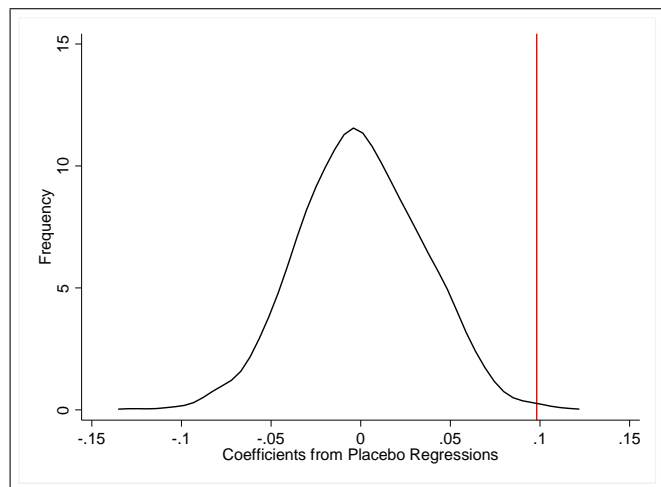


Figure A6: Placebo Coefficients of the Proportion of Female Peers in Incarceration Regression

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.

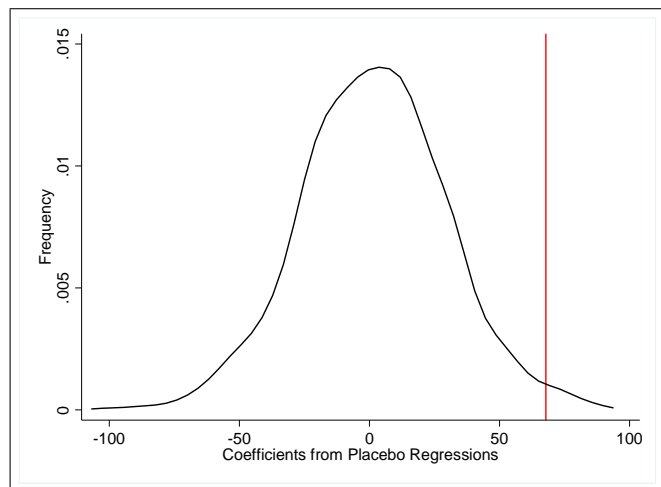


Figure A7: Placebo Coefficients of the Proportion of Female Peers in Sentence Length Regression

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.