Female Judges and In-group Bias in Labor Courts

Raphael Corbi [*]
University of São Paulo

Rafael Ferreira [†]
University of São Paulo

Jaqueline Oliveira [‡]
Rhodes College

Danilo Souza [§]
Insper

August 12, 2021

Abstract

Does judge gender influence the outcome of sentences in labor courts? We address this question using data on judges, defendants (firms), and plaintiffs (workers) from labor court cases of São Paulo state, Brazil, spanning from 2006 to 2015. Exploiting the random assignment of judges to cases, we document that female judges are more likely to rule in favor of companies than male judges, and that this gap is larger when the plaintiff is female.

JEL Codes: J16, K40.

Keywords: Gender bias, Judicial outcomes.

[*]Email: rcorbi@usp.br
[†]Email: rafaelferreira@usp.br (corresponding author)
[‡]Email: oliveiraj@rhodes.edu
[§]Email: danilops2@insper.edu.br

1. Introduction

Is justice blind? A justice system is only objective and impartial if the identities of the parties involved in a case have no bearing on outcomes. We ask if that is the case in the labor courts of Brazil. More specifically, we investigate two issues. First, does the gender of the judge affect sentencing? And second, do we see evidence of gender in-group bias, that is, male judges favoring male plaintiffs and female judges favoring female plaintiffs?

To answer these questions, we build a dataset of 321,536 labor court cases spanning the years 2006 to 2015 and covering 18 jurisdictions within the state of São Paulo, Brazil's largest regional labor court. We focus on cases in which the defendants are formal sector companies and plaintiffs are employees.

The dataset contains information on court location, the names of judges, defendants and plaintiffs, and lawsuit outcome. The cases dataset is then merged onto an administrative employer-employee dataset that provides information on workers' characteristics such as age, gender, educational level, occupation, and wages, as well as the companies' sector, size, and location.

Following previous studies on related subjects, we rely on the random assignment of cases among judges within a given jurisdiction to draw causal inference from our empirical analysis.[1] We find that, conditional on jurisdiction-year, there is no significant difference in plaintiff and defendant attributes between cases presided by female and male judges. The random assignment ensures that any gender disparities in sentencing we find would not be driven by within-jurisdiction self-selection of judges into cases based on judges' preferences for plaintiffs' or companies' characteristics.

We then proceed to formally estimate the differences in the likelihood a plaintiff wins a case between female and male judges. Controlling for jurisdiction-year fixed effects, we find that plaintiffs are less likely to win a case randomly assigned to a female judge rather than to a male judge. Furthermore, the evidence indicates that this difference is larger for female plaintiffs. Our preferred model shows that having a case assigned to a female judge reduces a male plaintiffs' chance of winning by nearly 1.7 percentage points; the reduction for female plaintiffs is of 2.3 percentage points. A model that includes judge fixed effects confirms that female judges are harsher on female plaintiffs than male plaintiffs.

Our paper adds to a growing literature that examines the role of gender on judicial outcomes. The few previous studies that look at the judge-gender effect find mixed results. Steffensmeier and Hebert (1999) use data on sentencing outcomes in Pennsylvania and find evidence that female judges are more likely to incarcerate and impose longer sentences than male judges. Philippe (2017) and Schanzenbach (2005) find that female judges decrease the female-male gap in prison sentences - which usually favor females - in the French and U.S. federal criminal courts, respectively.[2] While these studies point to

[1] A non-exhaustive list of studies that rely on random assignment of cases for identification of causal effects include Shayo and Zussman (2011), Abrams et al. (2012), Anwar et al. (2012), Butcher et al. (2017), Lim et al. (2016), Knepper (2018), and Hoekstra and Street (2018).

[2] Indeed, that female defendants are treated with more lenience than male defendants is a pattern that emerges from previous findings. In a study of federal criminal cases in the U.S., Starr (2015) finds evidence that conditional on arrest offense, criminal history, and pre-charge observables, females receive lighter sentences; they are also more likely to avoid charges, convictions, and incarceration. Also looking at criminal

female judges being less favorable to female defendants, other studies reach different conclusions. Lim et al. (2016), for instance, find no effect of female judges on criminal sentencing decisions in Texas courts; Knepper (2018) shows that female judges are more favorable than male judges in workplace sex discrimination cases brought by female plaintiffs.

Our paper is the first to rely on labor court data in a middle-income country setting. The outcomes of labor disputes may have relevant implications for the career prospects of females, especially in a context where the women's labor force participation is relatively low, and females make significantly less than their male counterparts.

2. Data

Our empirical analysis hinges on two data sources. The first source consists of electronic files with information on labor court cases extracted from judicial records[3] in the *Diário Oficial Eletrônico* (DOE), a daily publication by Brazil's *Tribunal Regional do Trabalho da* 2^a *Região* (TRT2), the country's largest labor tribunal. TRT2 handles all labor cases from the capital of São Paulo state and 45 of its neighboring municipalities, the greater São Paulo area. This region is divided into several smaller areas, called jurisdictions. A jurisdiction is a geographic area – often comprised of one or more[4] municipalities – that share the same pool of judges.

The electronic files of TRT2 cases are publicly available and contain data on the names of all parties involved in a case – plaintiff, defendant, lawyers, and judge – as well as court number, filling date, and decision date. It also has the judges' categorical ruling: pro-worker or pro-firm. We focus on lower labor court cases[5] in which workers are the plaintiffs and firms are the defendants, spanning from 2006 to 2015. Our data contains many of the active cases that reach a judicial verdict, comprising 153 courts across all 18 jurisdictions in the greater São Paulo area that have at least two courts. Out of the 528 judges in our sample, 57% are women.

The second source of data is *Relação Anual de Informações Sociais* (RAIS), a yearly matched employer-employee administrative database with information on the universe of Brazilian labor market contracts. This dataset includes workers' age, gender, education level, race, occupation and wage, as well as firms' sector, size and location. Using workers, judges and firms names, we merged the labor court and the employer-employee datasets, which resulted in a lawsuit-level dataset with 321,536 cases and data on the sociodemographic characteristics of plaintiffs (workers) and judges, the sector and size of the defen-

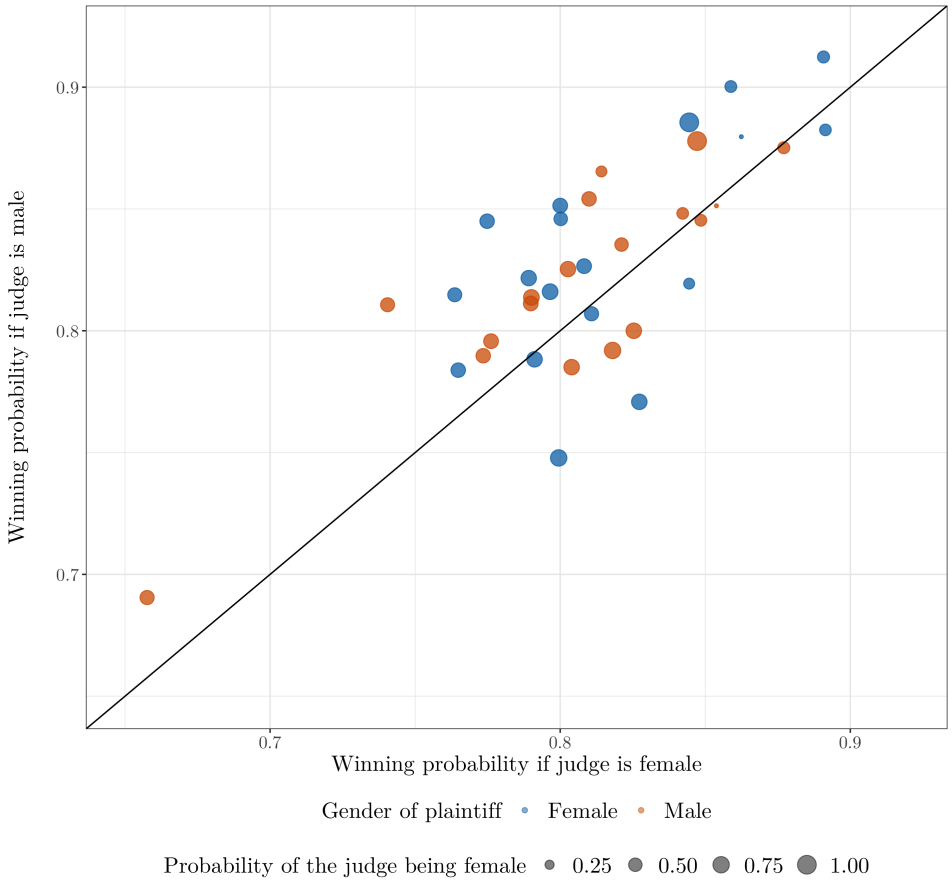
courts in the U.S., Butcher et al. (2017) show evidence of favorable treatment of women defendants that cannot be explained by observed case characteristics, but find that harsher judges increase the likelihood of female incarceration. The historical study conducted by Bindler and Hjalmarsson (2020) finds similar results for London criminal courts, indicating the females are less likely to be found guilty and receive lighter sentences for observationally similar offenses.

[3]Available in <https://aplicacoes1.trtsp.jus.br/ConsultaDOE/doe/completo.jsp>.

[4]Jurisdictions with more than one municipality often contain too small towns to have jurisdictions of their own. The city of São Paulo is the only municipality in our sample with more than one jurisdiction: its territory is divided into three jurisdictions.

[5]In Brazilian labor courts, most claims are related to notice of contract termination, severance payments, overtime, damages, unpaid wages, benefits, or unused vacations (Tribunal Superior do Trabalho, 2019).

Figure 1: Winning probabilities across jurisdiction, by judge and plaintiff gender



dants (firms), as well as judicial ruling. We coded outcome as an indicator variable that takes on value one if the case was *accepted* or *partially accepted* (meaning that the worker got some of their requests accepted but not all), and zero if rejected.

Figure 1 displays the proportion of cases within a jurisdiction that was favorable to the plaintiff, by plaintiff and judge gender. A few patterns emerge. With the exception of one jurisdiction, plaintiffs won over 80% of the cases, suggesting that labor courts are usually "pro-workers". The proportion of female judges (represented by circle size) vary across jurisdictions, and it is not uncommon for women to make up the majority of the judges within a jurisdiction. Also, plaintiff winning probabilities are higher when cases are assigned to male judges as most circles are above the 45-degree line, with such judge gender gap pattern in rulings slightly more pronounced for female plaintiffs (represented by blue circles).

3. Random Judge Assignment

Within jurisdictions with more than one judge, the assignment of cases to judges is randomized by the court's internal software. Once a lawyer files a petition in a labor court-

house, a random draw determines which judge from the pool of judges from that jurisdiction will decide on the case. Assignment probabilities should not depend on any case characteristics other than geographic location.[6]

In order to investigate the extent to which cases are in fact randomly assigned to judges within the same jurisdiction, we test whether characteristics of plaintiffs and defendants correlate with the gender of the assigned judge. Formally, we estimate

$$X_{ijct} = \alpha_0 + \alpha_1 \text{Female judge}_j + \gamma_{ct} + u_{ijct},$$

where X_{ijct} is a characteristic of case i assigned to judge j in jurisdiction c in year t, Female judge_j indicates the gender of the judge, and γ_{ct} is a jurisdiction-year fixed effects. Jurisdiction fixed effects are included because the case assignment is randomized at the jurisdiction level. These effects are allowed to differ by year to account for possible changes in the gender composition of the labor courts over time. If assignment of cases to judges within a jurisdiction is random, we should expect the estimates of α_1 to be small and insignificant once we add jurisdiction-year fixed effects.

Columns (1) and (2) of Table 1 display the mean of plaintiff (Panel A) and defendant (Panel B) characteristics according to judge gender. Columns (3) and (4) from Table 1 present estimates of α_1 obtained by comparing judges across and within jurisdictions in a given year, respectively. Each point estimate, and the corresponding standard error in brackets, pertains to a separate OLS regression. The estimates from column (3) represent the raw mean difference in case characteristics between female and male judges.

While most of those differences are statistically significant at 5%, the magnitudes are small. More specifically, female judges are slightly more likely to rule on cases involving females, white plaintiffs with college degree, and employed in the manufacturing sector. Once we include jurisdiction-year fixed effects in column (4), all differences become insignificant, which is expected under random assignment.

4. Gender disparities in sentencing outcome

Given the evidence that case assignment is random, we move to document the impact of judge gender on judicial outcomes of female and male plaintiffs. We estimate the following regression model:

$$\begin{aligned} Y_{ijct} = \beta_0 + \beta_1 \text{Female plaintiff}_j + \beta_2 \text{Female judge}_j + \\ \beta_3 \text{Female plaintiff} \times \text{Female judge}_j + \Phi X_{it} + \lambda_{ct} + \lambda_j + \epsilon_{ijct}, \end{aligned}$$

where Y_{ijct} is a dummy indicating that the plaintiff from case i judged by judge j in jurisdiction c in year t was victorious. X_{it} include case-specific controls for plaintiff race and education, firm sector and firm size measured by the number of workers. λ_{ct} is jurisdiction-year fixed effects; these are included because the random assignment of cases

[6] By law, labor cases must be filed in the jurisdiction where the work took place.

Table 1: Random assignment check

	Mean		Male-female judge gap	
Dependent variables	Female judge (1)	Male judge (2)	Across jurisdiction (3)	Within jurisdiction (4)
Panel A: Plaintiff variables				
Female plaintiff	0.334	0.331	0.003	-0.003
			[0.002]	[0.002]
Plaintiff with college degree	0.178	0.173	0.004***	-0.003*
			[0.001]	[0.001]
White plaintiff	0.581	0.574	0.007***	-0.001
			[0.002]	[0.002]
Black plaintiff	0.344	0.352	-0.008***	0.001
			[0.002]	[0.002]
Panel B: Defendant variables				
Firm sector: industry	0.171	0.159	0.012***	-0.000
			[0.001]	[0.001]
Firm sector: construction	0.076	0.081	-0.005***	0.001
			[0.001]	[0.001]
Firm sector: retail	0.176	0.175	0.001	0.002
			[0.001]	[0.001]
Firm sector: transportation	0.119	0.130	-0.011***	0.000
			[0.001]	[0.001]
Firm sector: administration	0.180	0.181	-0.001	-0.001
			[0.001]	[0.001]
Firm sector: others	0.278	0.274	0.004**	-0.002
			[0.002]	[0.002]
Firm size (workers)	773.701	743.681	30.019***	-12.834
			[10.934]	[11.232]

Notes: Columns (1) and (2) show the mean of each variable conditional on the judge gender. Column (3) shows the estimated coefficient of a OLS regression in which an indicator variable for female judge is the independent variable. In column (4) we replicate the column (3) regression including jurisdiction-year fixed effects. In both cases, standard errors are reported in brackets. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

to judges happens within jurisdictions. Some versions of the model also include judge fixed effects, λ_j, which use within-judge variation in plaintiff gender to estimate β_3.[7]

[7]It is worth noting that once judge fixed effects are included, the coefficient β_2 is not identified.

The coefficient β_1 captures the plaintiff-gender difference in the likelihood of winning cases presided by male judges. This coefficient may reflect judicial bias in favor or against female plaintiffs, as well as unobservable differences in the quality of cases brought about by females.

The coefficients that reveal gender differences in labor court sentencing are β_2 and β_3. If it is the case that gender does not influence how a judge rules in labor courts, both coefficients would be insignificant. If female judges hold more "pro-worker" views than their male counterparts, regardless of plaintiff's gender, we would expect β_2 to be positive and significant and β_3 to be statistically insignificant. A coefficient β_3 different from zero implies that the gender of the judge plays a role in any existing gender gap in the likelihood a worker wins a labor court case.

Table 2: Baseline results

	Dependent variable: Y = 1 if plaintiff wins				
	(1)	(2)	(3)	(4)	(5)
Female plaintiff	0.0006	0.0046	0.0060	-0.0018	-0.0020
	[0.0109]	[0.0126]	[0.0101]	[0.0060]	[0.0059]
Female judge	-0.0216***	-0.0193***	-0.0170***	-0.0171***	
	[0.0042]	[0.0041]	[0.0025]	[0.0025]	
Female plaintiff × Female judge		-0.0068*	-0.0062**	-0.0063**	-0.0057*
		[0.0039]	[0.0028]	[0.0028]	[0.0030]
Jurisdiction × Year FE	N	N	Y	Y	Y
Case controls	N	N	N	Y	Y
Judge FE	N	N	N	N	Y
Dep. variable mean	0.787	0.787	0.787	0.787	0.787
Observations	321,536	321,536	321,536	321,536	321,536
Adjusted R2	0.001	0.001	0.009	0.018	0.042

Notes: In brackets, standard errors are clustered at the jurisdiction level. Case controls include an indicator variable for plaintiff with college degree, dummies for white and black plaintiffs, dummies for the 5 firm sectors with more lawsuits in our dataset, and the firm size measured by the number of workers. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

Table 2 shows the results. Columns (1) and (2) reveal that the gender of the judge affects the likelihood a worker wins a case in Brazilian labor courts. More specifically, when a presiding judge is female, plaintiffs of both genders are 2.16 percentage points *less* likely to win, suggesting that female judges are less "pro-workers" than male judges. Furthermore, we find that female judges are even less pro-worker than male judges in cases when the plaintiff is also a female. This is in contrast with in-group bias findings in Shayo and Zussman (2011), as there is no evidence that a female judge presiding over a case will confer an advantage to a female plaintiff relative to a male judge presiding.

Table 3: Robustness check

	Dependent variable: Y = 1 if plaintiff wins					
	(1)	(2)	(3)	(4)	(5)	(6)
Female plaintiff	-0.0023	-0.0073	-0.0056	-0.0038	0.0153	-0.0240*
	[0.0094]	[0.0076]	[0.0036]	[0.0040]	[0.0203]	[0.0130]
Female judge	-0.0170***		-0.0163***		-0.0161***	
	[0.0025]		[0.0026]		[0.0025]	
Female plaintiff × Female judge	-0.0065**	-0.0055*	-0.0067**	-0.0057*	-0.0068**	-0.0060*
	[0.0028]	[0.0029]	[0.0030]	[0.0029]	[0.0030]	[0.0033]
Female plaintiff × Judge age	0.0000	0.0001			-0.0006	0.0005
	[0.0001]	[0.0001]			[0.0005]	[0.0003]
Female plaintiff × Judge tenure			0.0004	0.0002	-0.0003	0.0027
			[0.0004]	[0.0003]	[0.0012]	[0.0024]
Female plaintiff × Judge age × Judge tenure					0.0000	-0.0001
					[0.0000]	[0.0000]
Jurisdiction × Year FE	Y	Y	Y	Y	Y	Y
Case controls	Y	Y	Y	Y	Y	Y
Judge FE	N	Y	N	Y	N	Y
Dep. variable mean	0.787	0.787	0.787	0.787	0.787	0.787
Observations	321,536	321,536	321,536	321,536	321,536	321,536
Adjusted R2	0.018	0.043	0.018	0.043	0.018	0.043

Notes: In brackets, standard errors are clustered at the jurisdiction level. Case controls include an indicator variable for plaintiff with college degree, dummies for white and black plaintiffs, dummies for the 5 firm sectors with more lawsuits in our dataset, and the firm size measured by the number of workers. We also include judge age as another case control in columns (1), (2), (5) and (6), and judge tenure as case control in columns (3), (4), (5) and (6). Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

As randomization in case assignment occurs at the jurisdiction level, column (3) adds jurisdiction-year fixed effects in order to restrict the identification to within-jurisdiction-year variation. The estimated effect of a female judge when plaintiff is male is slightly reduced to -0.0170. Column (4) adds case controls. As expected if treatment is random, the coefficients for female judge and its interaction with female plaintiff are virtually unchanged. Also, including judge fixed effects has little effect on the estimates as shown in column (5). Overall, female judges are less likely to grant victory to plaintiffs, especially female workers.

The random assignment of cases to judges implies we can estimate the coefficients on judge gender and its interaction with plaintiff's gender without bias. It is possible, however, that the interaction with judge gender is reflecting other judge's characteristics potentially correlated with gender. If female judges are younger and have shorter tenure than male judges, and younger and/or less experienced judges hold more pro-firm views, our results could be driven by age and tenure and not gender. To check for that, in Table 3 we re-estimate the main model adding interactions with judge's age (columns (1), (2),

(5), and (6)) and tenure (columns (3), (4), (5), and (6)). We draw two conclusions from the new estimates. First, age and tenure of judges do not seem to play a role in sentencing outcomes. And second, the estimates of the coefficient β_2 are robust to the inclusion of interactions with other judge's characteristics, changing little when judge fixed effects are added.

Finally, it is possible that female and male plaintiffs sort into different types of firms according to the firms' unobservable characteristics. For example, women may be more adversely affected than men by a firm breaking the rules and could, therefore, self-select to work for firms with a proven track record of abiding by labor regulations. Also, men may be more likely than women to work for larger and wealthier corporations, and these corporations may be better equipped to divert resources to win labor court cases. We went back and re-estimated the models after including firm fixed effects, which then allowed us to compare the differences in winning probabilities across cases with the same (firm) defendant. The results[8] are presented in Table 4. As we can see, the main conclusions do not change with the addition of firm fixed effects. This indicates that firm unobservable characteristics, and the sorting of workers of different genders on the basis of these characteristics, do not bias the estimates of judge gender effects because of the random assignment of judges to cases.

Table 4: Baseline with firm fixed effects

	Dependent variable: Y = 1 if plaintiff wins					
	(1)	(2)	(3)	(4)	(5)	(6)
Female plaintiff	0.0049	-0.0047	-0.0050	-0.0114***	-0.0121***	-0.0125***
	[0.0105]	[0.0060]	[0.0060]	[0.0030]	[0.0030]	[0.0029]
Female judge	-0.0161***	-0.0162***		-0.0171***	-0.0171***	
	[0.0027]	[0.0027]		[0.0030]	[0.0030]	
Female plaintiff × Female judge	-0.0066***	-0.0067***	-0.0059**	-0.0057**	-0.0056**	-0.0049**
	[0.0023]	[0.0022]	[0.0022]	[0.0023]	[0.0023]	[0.0018]
Jurisdiction × Year FE	Y	Y	Y	Y	Y	Y
Case controls	N	Y	Y	N	Y	Y
Judge FE	N	N	Y	N	N	Y
Firm FE	N	N	N	Y	Y	Y
Dep. variable mean	0.789	0.789	0.789	0.789	0.789	0.789
Observations	282,535	282,535	282,535	282,535	282,535	282,535
Adjusted R2	0.010	0.021	0.047	0.144	0.144	0.170

Notes: In brackets, standard errors are clustered at the jurisdiction level. Case controls include an indicator variable for plaintiff with college degree, dummies for white and black plaintiffs, dummies for the 5 firm sectors with more lawsuits in our dataset, and the firm size measured by the number of workers. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$.

[8]The sample size is smaller because we can only use firms that were defendants in multiple cases.

5. Concluding remarks

This paper aims to investigate the existence of gender in-group bias in labor court rulings in Brazil. Our identification strategy exploits the random assignment of cases to judges within a particular jurisdiction. We find that plaintiffs are less likely to win labor cases presided by female judges. Furthermore, we also find evidence of negative gender in-group bias, with female workers being less likely to win a case when the judge is also female.

References

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do Judges Vary in Their Treatment of Race? *The Journal of Legal Studies 41*(2), 347–383.

Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The impact of jury race in criminal trials. *Quarterly Journal of Economics 127*(2), 1017–1055.

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

Butcher, K. F., K. H. Park, and A. M. Piehl (2017). Comparing Apples to Oranges: Differences in Women's and Men's Incarceration and Sentencing Outcomes. *Journal of Labor Economics 35*(S1), S201–S234.

Hoekstra, M. and B. Street (2018). The Effect of Own-Gender Juries on Conviction Rates. Working Paper No. 25013, National Bureau of Economic Research.

Knepper, M. (2018). When the shadow is the substance: judge gender and the outcomes of workplace sex discrimination cases. *Journal of Labor Economics 36*(3), 623–664.

Lim, C. S., B. S. Silveira, and J. M. Snyder (2016). Do judges' characteristics matter? ethnicity, gender, and partisanship in texas state trial courts. *American Law and Economics Review 18*(2), 302–357.

Philippe, A. (2017). Gender disparities in criminal justice. TSE Working Paper No. 17-762, Université Toulouse 1 Capitole.

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

Shayo, M. and A. Zussman (2011). Judicial ingroup bias in the shadow of terrorism. *Quarterly Journal of Economics 126*(3), 1447–1484.

Starr, S. B. (2015). Estimating gender disparities in federal criminal cases. *American Law and Economics Review 17*(1), 127–159.

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

Tribunal Superior do Trabalho (2019). Relatório Geral da Justiça do Trabalho. Technical report.