INSIDERS, OUTSIDERS, AND INVOLUNTARY UNEMPLOYMENT:
SEXUAL HARASSMENT EXACERBATES GENDER INEQUALITY

DANIEL L. CHEN AND JASMIN SETHI*

Abstract   Is labor market gender inequality due to physiological differences, labor market choices, or discrimination? We examine the impact of sexual harassment on labor market choices using an insider-outsider model of involuntary unemployment and estimate its impact using legal precedents forbidding sexual harassment. Leveraging randomly assigned judges, we find that sexual harassment precedent favoring plaintiffs encouraged labor market entry of women. The effects were comparable to the Equal Employment Opportunity Act's impact on black employment share, greatest where male sexism was highest and in the heavily-litigated construction industry, and equivalent to one-third of the rise of sexual harassment human resources policies.

Keywords: Gender discrimination, micro-aggression, trauma, safe spaces, prejudice

JEL codes: J81, J83, K31, J31, J71

1 Introduction

Does forbidding sexual harassment ameliorate or exacerbate gender inequality? Sexual harassment is perceived to be a major impediment to female labor force participation. Policymakers in both developed and developing countries have taken steps to address this problem. For example, in India and Mexico, female-only trains and buses were introduced so women would face less harassment on their way to work. In the U.S., making the work environment friendlier to women has been one of the most dramatic labor market changes in the past half-century. Yet, the consequences of forbidding harassment on female labor force outcomes remain unknown. While sexual harassment has received the attention of one economic model–through the theory of compensating differentials (Basu 2003; Hersch 2011), it receives no mention in two recent summaries of the empirical literature on discrimination (Bertrand and Duflo 2016; Neumark 2016). This paper takes steps toward theoretically and empirically analyzing the impacts of sexual harassment.[1] Theoretically, the effects of forbidding sexual harassment are ambiguous. On the one hand, a large set of models and empirical analyses suggests that it would exacerbate gender inequality if the primary function of sexual harassment law is to mandate a benefit, requiring a compensating differential (Rosen 1974); the mandated benefit would impose costs on the targeted group through lower wages or lower employment (Acemoglu and Angrist 2001), act as a tax on labor demand (Summers 1989), and make women more costly to hire (Epstein 1995). On the other hand, a set of models previously applied to unions suggests that forbidding harassment can ameliorate gender inequality by opening job opportunities in previously harassing work environments. In this theory, insiders harass outsiders to lower their productivity, leading to involuntary unemployment (Lindbeck and Snower 1988). Our results support the second class of models, which has thus far relied on cross-sectional data (Lindbeck and Snower 2001) or lab experiments (Fehr and Fischbacher 2002). Our results may shed light on the economic consequences of micro-aggression, trauma, trigger warnings, safe spaces, and the #MeToo movement.

For researchers who study discrimination it is worth noting that much of U.S. policy surrounding labor market discrimination is carried out through the court system. Legal precedents include ruling on what constitutes sexual harassment, e.g., retaliation against a female employee for rejecting sexual advances of her boss. They also include shifting from a reasonable *person* to a reasonable *woman* standard for what constitutes sexual harassment, waiving the need for a plaintiff to prove emotional harm in court, and declaring that the firm can be subject to liability for hostile environment created

---

[1]Prior survey evidence indicate higher job satisfaction resulting from the presence of sexual harassment law (Newman et al. 2003).

by a supervisor. These precedents have great scope for large impacts on society. To estimate their effects, from a statistical perspective, the ideal experiment would be to randomize the court precedents. This would allow one to analyze the causal impact of a court precedent for or against sexual harassment plaintiffs. Since precedents are not random, we exploit the random assignment of judges, whose biographies predict their decisions in sexual harassment cases. Our research design exploits variation in the composition of judicial panels within-jurisdiction across time and within-year across jurisdictions. We collect and analyze the impacts of all cases developing sexual harassment precedent. We use the apex of the U.S. court system—the Courts of Appeals—whose decisions comprise almost the totality of court-made law in the U.S. We study the impact of diversity in the judiciary and the role of courts on social and economic outcomes.

We show that the composition of judicial panels appears to be random (uncorrelated with pre-trial characteristics) yet correlated with decisions in sexual harassment cases. Prior research has documented that politics and gender predict sexual harassment decisions (Farhang and Wawro 2004; Epstein 1995; Peresie 2005). But, it has been argued that judges vote more along party lines than along gender lines (Dixon 2010), that U.S. Presidents who appoint women candidates take the opportunity to appoint more ideologically extreme individuals (Asmussen 2011), and that female conservatives exhibit prejudice against females (Eisenman 1991). We find that female Republicans were 18 percentage points less supportive of sexual harassment plaintiffs, while male Democrats were 13 percentage points more supportive of sexual harassment plaintiffs.[2] We also collect data on judicial biographies to exploit the large variation in judicial decisions that is due to random combinations of biographical characteristics, and employ a sparse model, LASSO, to select among the combinatorial possible number of panel characteristics that predict sexual harassment decisions (Belloni et al. 2012).

We then show that pro-plaintiff sexual harassment precedents increased female employment shares by an amount equivalent to what the Equal Employment Opportunity Act achieved for black employment shares (Chay 1998). To explore the mechanisms, three results emerge. First, the effects were greatest in the U.S. South, which scores highest in male sexism (Charles et al. 2010). and in the construction industry, which had the highest rate of sexual harassment claims. Second, forbidding harassment *increased* earnings per hour of female labor force participants, which is consistent with an insider-outsider model of involuntary unemployment outweighing the compensating differ-

---

[2]As shorthand, we will refer to judges appointed by Democratic presidents as "Democrats" and to judges appointed by Republican presidents as "Republicans."

entials mechanism underlying the dominant models of anti-discrimination law. Third, we show that pro-plaintiff sexual harassment led to the adoption of sexual harassment human resources policies (Dobbin and Kelly 2007); the impacts were equivalent to 37% of the adoption from 1982 to 1997.

Throughout our analysis, we verify robustness by changing our controls and specifications–adding controls for composition of the pool of judges available to be assigned to panels, Circuit-specific time trends, and state fixed effects; dropping 1 Circuit at a time, varying the lag structure, and collapsing the data to the Circuit-year level; as falsification, social outcomes are also unrelated to Circuit decisions before they were made and inferences robust to randomization inference and clustering standard errors in different ways. In a final section, we discuss an extension, where we incorporate exogenous variation in the decision to appeal using the random assignment of District court judges.

By documenting the endogeneity of labor market choices to discriminatory environments, our research blurs the boundary between different explanations for labor market gender inequality attributable to physiological differences, labor market choices, or discrimination (Summers 2005). Our research contributes to an empirical literature on sexual harassment, suggesting it can have economic consequences: McLaughlin et al. (2017) finds that sexual harassment increases financial stress, as many women leave their jobs to escape the harassing environment using interviews and survey data; Hersch (2011) reports that sexual harassment complaints filed with the EEOC (Equal Employment Opportunity Commission) correspond with economically measurable wage differentials. Our research also contributes to an empirical literature on sexual harassment *law*, which has primarily examined how labor lawyers and human resources consultants conveyed the risk of sexual harassment lawsuits after major Circuit decisions (Dobbin and Kelly 2007; Edelman 1992). This literature argues that millions of dollars were spent on training programs and establishing grievance procedures that aimed to reduce the risk of lawsuit more than to reduce the incidence of harassment,[3] and suggests that legal precedents can foment social change. Our findings also contribute to a broader literature on gender inequality (Blau and Kahn 2006; Card and DiNardo 2002; O'Neill 2003; Black and Strahan 2001), prejudice (Charles and Guryan 2008), and anti-discrimination law (Hellerstein et al. 2002; Eberts and Stone 1985; Chay 1998).

Our methods contribute to studies that employ the random assignment of judges to identify the impact of judicial decisions on the subsequent outcomes of litigants (Kling 2006; Maestas,

---

[3]Some firms mandated training akin to the diversity training implemented in response to the Civil Rights Act of 1964 to educate their workers about sexual harassment.

Mullen, and Strand 2013; Dahl, Kostøl, and Mogstad 2014). Because randomization occurs at a jurisdictional level, our analyses also capture general equilibrium effects (such as capital and labor migration across Circuits). Judge Richard Posner has lamented that, "[judicial] opinions lack the empirical support that is crucial to sound constitutional adjudication" (Posner 1998); similarly Justice Breyer remarked, "I believe that a[n] interpretive approach that undervalues consequences, by undervaluing related constitutional objectives, exacts a constitutional price that is too high" (Breyer 2006). Social scientists have long speculated on the relationship between the innovation of rights and socio-economic conditions, and scholarship to date has been unsatisfactory in exploring questions of causality vis-à-vis legal precedent. Methods to evaluate the impact of court-made law may help judges who are interested in the broader empirical consequences of their decisions.
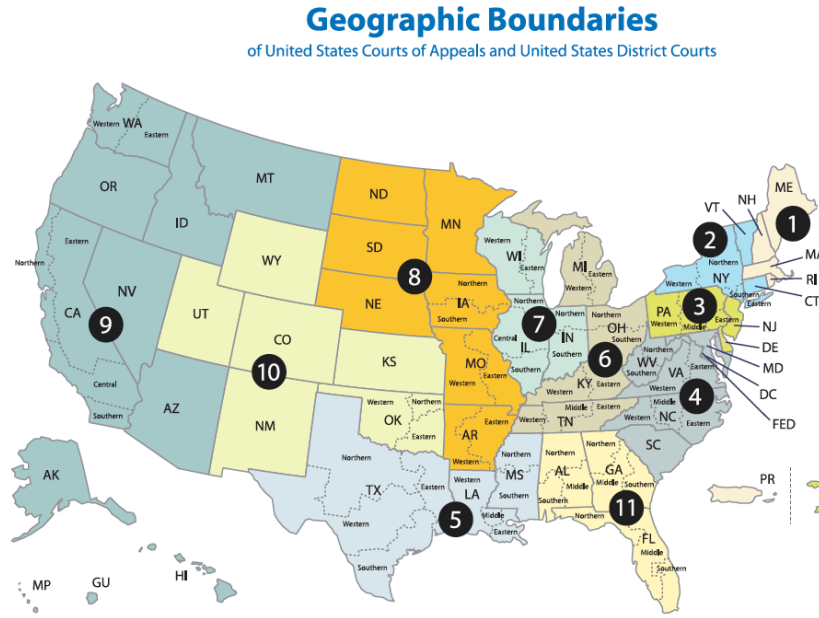
2 Background

2.1 Institution   The U.S. has a common law system where decisions become precedent for future cases in the same jurisdiction. We focus on the appellate courts, referred to as Circuit Courts. Only 2-3% of Circuit cases are appealed again to the U.S. Supreme Court. There are 12 Circuit Courts, each in charge of a geographic region comprising 4 to 9 U.S. states (see Figure 1). Cases originate from the District Courts, numbering 1 to 4 per state. The EEOC first issued guidelines with the term "sexual harassment" in 1980. Before 1982, there were very few sexual harassment cases. We follow the data collection and coding method in the Chicago Judges Project (Sunstein et al. 2006), We collect data on all sexual harassment cases in Circuit and District Courts back to 1982 to examine the totality of sexual harassment cases brought in an employment context. Treated Circuit-years would be ones that received a pro-plaintiff decision and the control would be one that received a pro-defendant decision. Since decisions are not randomly assigned, we instrument for the decisions using the random assignment of judges.[4] A number of papers have documented the effect of judges' demographic background on sexual harassment cases (Farhang and Wawro 2004; Epstein 1995; Peresie 2005).

Figure 2 plots the growth in sexual harassment cases, number of pro-plaintiff decisions, and number of pro-defendant decisions during this time period. The coding of pro-plaintiff is if the plaintiff was afforded any relief. Table I indicates that on average, there were roughly 1 sexual harassment precedent per Circuit-year for a total of 251.[5] Roughly two-thirds of the decisions

---

[4]For a randomization check, we employ Boyd et al. (2010)'s coding of case characteristics for a large set of discrimination cases from the Chicago Judges Project.

[5]252 Circuit-years is not a multiple of 12 because Circuit 11 was founded in 1981, so Circuit 11 has 6 fewer

FIGURE 1.—

Geographic Boundaries
of United States Courts of Appeals and United States District Courts

Source: US Government, http://www.uscourts.gov/uscourts/images/CircuitMap.pdf

were pro-plaintiff. We also collected 3,754 cases between 1982 and 2002 in District Courts.[6] To implement our research design, we also collect judicial background characteristics from the Appeals Court Attribute Data, District Court Attribute Data,[7] Federal Judicial Center, and newspapers.[8] We filled in missing data by searching transcripts of Congressional confirmation hearings and other official or news publications on Lexis.

While there were 18.50 judges available, on average across Circuit-years, for assignment to panels, the typical judge saw very few sexual harassment cases.[9] Thus we focus on shared biographical characteristics. For example, one-third of seats would be assigned a male Democrat in expectation, providing enough data to estimate attitudes towards sexual harassment.[10] Additional summary

---

observations than the other Circuits. Circuit 11 was created by splitting it off from Circuit 5; Circuit 5's decisions before this split are considered binding precedent in Circuit 11. We account for this split in our analyses by assigning pre-1981 precedent in Circuit 5 to observations in Circuit 11.

[6]We searched Westlaw using "((SEX! +2 DISCRIMINATION) (GENDER +2 DISCRIMINATION)) & (SEX! +2 HARASSMENT)".

[7]http://www.cas.sc.edu/poli/juri/attributes.html

[8]Variables include: geographic history, education, occupational history, governmental positions, military service, religion, race, gender, and political affiliations. Some data on religion come from Goldman (1999). Sisk's data are available at http://courseweb.stthomas.edu/gcsisk/religion.study.data/cover.htm.

[9]There are 180 life-tenured judges and 251 cases.

[10]Each Circuit Court decides many thousands of cases per year, so the composition of judicial panels in other legal areas is uncorrelated with the composition in sexual harassment panels.

FIGURE 2.— Sexual Harassment Precedents, 1982-2002
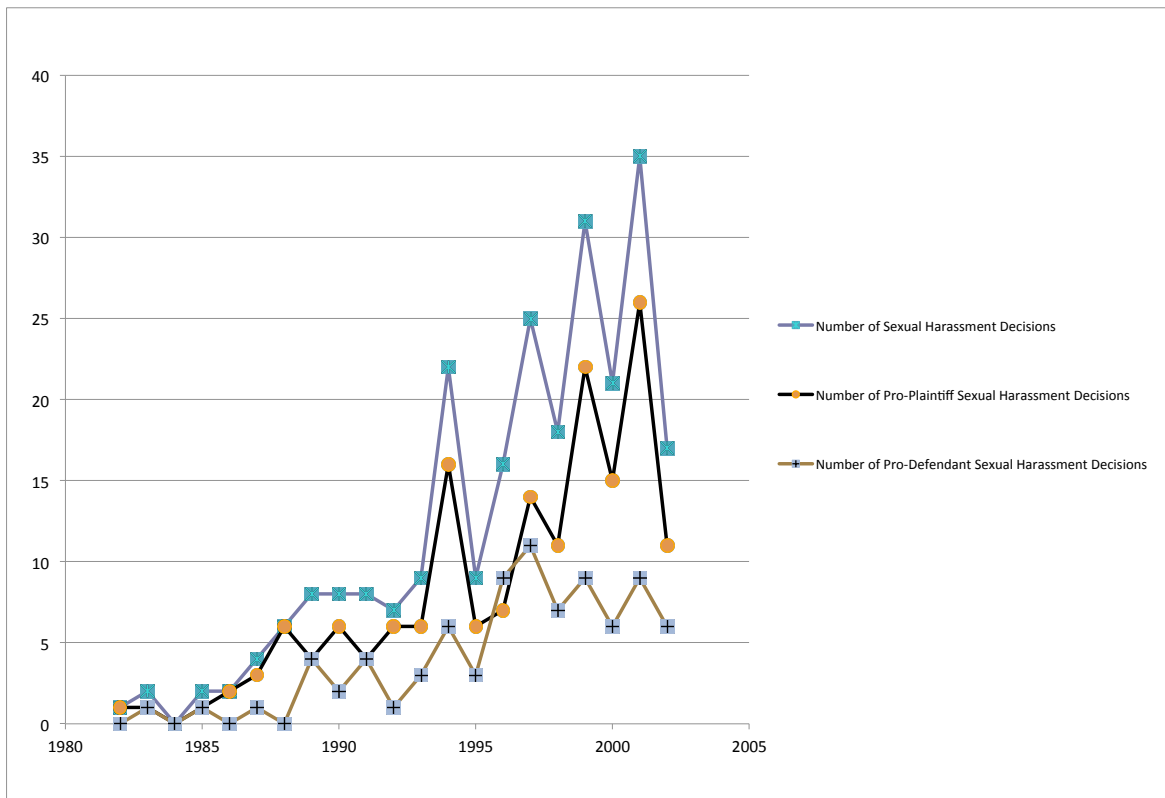
Notes: X-axis is year and Y-axis is number of cases.

TABLE I

Summary Statistics

Circuit-Year level	Mean	Sd
Number of Judges	18.504	7.356
Number of Panels	0.996	1.471
Expected % of Females	0.117	0.081
Expected % of Democrats	0.407	0.121
Expected % of Female Republicans	67,350	0.040
Expected % of Male Democrats	0.326	0.119
% Pro-Plaintiff	67%	
Total (Circuit-years)	252	

Notes: Column 1 presents the average and Column 2 the standard deviation of summary statistics on sexual harassment cases.

statistics are displayed in Table I. To identify the judges in District Court cases, we use administrative data from the Administrative Office of the U.S. Courts (AOC) and PACER filings.[11]

2.2 Theories of Sexual Harassment Whether forbidding sexual harassment exacerbates gender inequality is a priori ambiguous and this motivates our empirical investigation. Broadly speaking, law and economics models of anti-discrimination law posit that these laws act as an unfunded mandated benefit, making it more costly to hire women, lowering their wages or employment relative to men (Epstein 1995). In addition, compensating differentials may lower female wages by making it more pleasant for women who would be willing to work for lower wages in work environments that previously allowed harassment (Basu 2003). A collection of empirical findings support this class of models.[12] There are, however, reasons to think that employers are less likely to experience sexual harassment law as an unfunded mandated benefit since it may be difficult to know in advance who is going to be a sexual harassment plaintiff and because men could bear some of the cost of the mandated benefit as potential harassers.[13]

We present an alternative theory building on a class of insider-outsider models, previously adapted to unionization (Lindbeck and Snower 1988). In brief, while harassment is allowed, outsiders are unable to find jobs even though they are prepared to work for less than the prevailing wages of incumbent workers (insiders). The outsiders cannot underbid insiders; if they did and were to become

---

[11]Sixteen years of Public Access to Court Electronic Records are available on open source sites for 33 Districts.

[12]The Americans with Disabilities Act (ADA) and maternity mandates had detrimental effects on the groups they were intended to protect and undoing, in part, the redistributive goals of these policies (ADA: Acemoglu and Angrist 2001; Employment protection: Autor et al. 2006; Maternity mandates: Gruber 1994).

[13]For example, with ADA, disabilities are often visible to employers and, therefore, the unfunded mandate of accommodations may have led to calculated decisions to not hire particular disabled workers whereas employers could not as easily make the same calculated decisions vis-a-vis women. In addition, unlike the cost of complying with the ADA or the federal requirement of providing maternity mandates, the cost of compliance with sexual harassment law could be reduced by not hiring either the group being harassed or the group doing the harassing.

new employees, insiders would make the work experience of these entrants unpleasant. Insiders would harass the entrants, thereby reducing their productivity. Firms, therefore, find it costly to substitute outsiders for insiders. These harassment and labor turnover costs create economic rents, which the insiders capture via wage setting, and as a result, involuntary unemployment arises. The insider-outsider theory predicts that forbidding harassment allows outsider women to enter the labor force as they can now compete for jobs previously dominated by men and the insider women who tolerated sexual harassment. Economic rents captured by insiders are dissipated when they can no longer harass the outsiders, and on the margin, some insiders may leave. In addition, some females—those who previously obtained the insider rents—may see less benefit in their employment outcomes, especially as they face increased competition from previously outsider females.

These two broad views of anti-discrimination law are distinguished by the overall impact of sexual harassment law on gender inequality. However, there are a few alternative theories of sexual harassment law, such as a redistribution from females to males or tax on the hiring of men consistent with a reduction in gender inequality. We can be distinguish the insider-outsider theory from these theories by examining the impact on insider women. We can infer insider status by contrasting the experience of wage-earning women in the labor force with the experience of women in the general population. Redistributional theories would predict that insider male outcomes fall relative to insider female outcomes. However, forbidding harassment raises productivity of females (and possibly that of males). A final view of sexual harassment law is simply an elimination of masculine culture prevalent before females were hired in substantial numbers. However, wages and productivity should be unrelated to the elimination of machismo.

The insider-outsider model of involuntary unemployment assumes that firms do not forbid harassment on their own. First, there may be managerial unawareness of a link between harassment and productivity. Second, an insider cooperates with entrants if his gains, a share of the additional profit resulting from his cooperation, exceed his losses in market power as his wage falls towards his reservation wage. However, this only happens if the firm relinquishes a share of gross profit, something that may make it a net loser compared to other firms. Agency and transaction costs may prevent Coasian bargaining between insiders and the firm. Then the firm has no incentive to implement the new contract. Third, insider employees may be risk averse; forbidding harassment could change the insider profit-sharing scheme and thereby impose additional risk on insider employees, who then suffer a utility loss. The firm may be unable to compensate them for this loss. Fourth,

there may be additional sources of labor turnover costs preventing firms from simply replacing all the insiders with outsiders. Entry of firms that hire the outsiders may not occur due to setup costs, capital market imperfections, scarcity of entrepreneurial skills, and reduction of product prices. However, a legal regime equalizes the playing field across all firms when no firm by itself would have the incentive to forbid harassment.

3 Design of Study

3.1 Specification In order to measure the effect of sexual harassment legal precedent on gender inequality we use regressions of the form:

(1) $$Y_{ict} = \theta_c + \theta_t + \sum_{n=0}^{L} \beta_{1t-n} Law_{ct-n} + \sum_{n=0}^{L} \beta_{2t-n} 1\left[M_{ct-n} > 0\right] +$$
$$\sum_{n=0}^{L} \beta_{3t-n} Law_{ct-n} * F_{ict} + \sum_{n=0}^{L} \beta_{4t-n} 1\left[M_{ct-n} > 0\right] * F_{ict} + \eta X_{ict} + \varepsilon_{ict}$$

where

- Y_{ict} is employment status, hours worked, earnings, and management status for individual i in circuit c and year t.
- Law_{ct} is pro-plaintiff sexual harassment precedent. Typically $Law_{c(t-n)}$ is 1 (100% pro-plaintiff) or 0 (100% pro-defendant) (See Table I). We focus on five years of lags and one lead ($n = -1$ to 5) and report the average coefficient. We vary the lag structure for robustness.
- F_{ict} is an indicator for female. For the insider-outsider theory, we expect $\beta_3 > 0$ for both wages and employment. For other theories, we expect $\beta_3 < 0$ for at least one dimension of wages or employment.
- $1\left[M_{ct-n} > 0\right]$ is an indicator for the presence of a decision. M is the number of cases, which is typically 0 or 1. When there are no cases, we encode Law_{ct} as 0.
- θ_c are indicators for Circuit and θ_t are indicators for year.
- X_{ict} is a set of control variables, such as age, gender, educational attainment, and race (these enter as dummies with the exception of age).

In robustness checks, we also include Circuit-specific time trends and time-varying Circuit-level controls, such as the characteristics of the pool of judges available to be assigned in Circuit c and time $t - n$.

Since random assignment is at the Circuit-year level, we expect to see similar results whether clustering standard errors at the Circuit or Circuit-year level.[14] Our results are also unaffected if we

---

[14]Barrios et al. (2012) show that random assignment of treatment addresses serial and spatial correlation across

cluster at the state level, collapse our data to one observation per Circuit-year, use randomization inference assigning the legal variation to another Circuit, and use wild bootstrap. The coefficients on the leads serve as an omnibus falsification check for spurious significance.

Our identification strategy is illustrated in Figures 3A and 3B. Let N_{ct} be a judicial characteristic. The jagged line displays $\frac{N_{ct}}{M_{ct}}$, variation in actual composition of judicial panels, and the smooth line displays $\mathbf{E}(\frac{N_{ct}}{M_{ct}})$, variation in expected composition of judicial panels. More formally, let $p_{ct} = \frac{N_{ct}}{M_{ct}} * \mathbf{1}\,[M_{ct} > 0]$, i.e., defined to be 0 when $\mathbf{1}\,[M_{ct} > 0] = 0$. Then: $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct}))\varepsilon_{ict}] = \mathbf{Pr}[M_{ct} > 0]\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct}))\varepsilon_{ict}|M_{ct} > 0] + \mathbf{Pr}[M_{ct} = 0]\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct}))\varepsilon_{ict}|M_{ct} = 0] = 0$. Next, $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct}))\varepsilon_{ict}] = \mathbf{E}(p_{ct}\varepsilon_{ict}) - \mathbf{E}[\mathbf{E}(p_{ct})\varepsilon_{ict}] = \mathbf{E}(p_{ct}\varepsilon_{ict}) - \mathbf{E}(p_{ct})\mathbf{E}(\varepsilon_{ict}) = \mathbf{E}[p_{ct}\varepsilon_{ict}]$. Thus, p_{ct} and $p_{ct} - \mathbf{E}(p_{ct})$ can serve as valid instruments.[15] All lags and leads of Law_{ct} are instrumented. As standard practice, we also interact the instruments as we lag and interact Law_{ct} with the female indicator.

We surveyed a number of courts of appeal and evaluated measures taken by them to ensure that the assignment of judges to panels is random. In one court, two to three weeks before the oral argument, a computer program is used to randomly assign available judges, including any visiting judges, to panels that will hear cases. The program used is an in-house creation. There is a mechanism in the program that ensures the same judges are not sitting together on panels. This is also checked manually, although the clerk could not remember ever having manually to change judicial assignments for this reason. There is no specialization among judges; the cases are "all over the map" in regard to subject matter. Senior judges tell the clerk how often they are willing to sit and hear cases, and they are added to the program for randomized assignment in accordance with their schedules. There is an administrative office that sets the baseline number of cases senior judges must hear per term.

In another court, random assignment of panels occurs before the random assignment of cases. Panels of judges are organized to hear cases on a yearly basis, randomly assigned together by computer program and given dates for hearings. There are "holes" left in some of the panels by

---

treatment units, since "if the covariate of interest is randomly assigned at the cluster level, only accounting for non-zero covariances at the cluster level, and ignoring correlations between clusters, leads to valid standard errors and confidence intervals." A prior draft clusters standard errors at the Circuit-year level and this draft clusters at the Circuit level upon request of a previous referee.

[15]This draft presents estimates using the following identification assumption (i.e., moment condition) for causal inference: $\mathbf{E}[\frac{N_{ct}}{M_{ct}}\varepsilon_{ict}|\mathbf{E}(\frac{N_{ct}}{M_{ct}}), \mathbf{1}\,[M_{ct} > 0]] = 0$. Previous drafts obtained similar results using $\mathbf{E}[N_{ct}\varepsilon_{ict}|\mathbf{E}(\frac{N_{ct}}{M_{ct}}), \mathbf{1}\,[M_{ct} > 0], M_{ct}] = 0$ (using the number of pro-plaintiff decisions controlling for the number of decisions) and $\mathbf{E}[N_{ct}\varepsilon_{ict}|\mathbf{E}(\frac{N_{ct}}{M_{ct}}), \mathbf{1}\,[M_{ct} > 0], Q_{ct}] = 0$ (controlling for the size of the court docket and checking if pro-plaintiff vs. pro-defendant decisions had opposite-signed effects).

FIGURE 3.—

Panel A
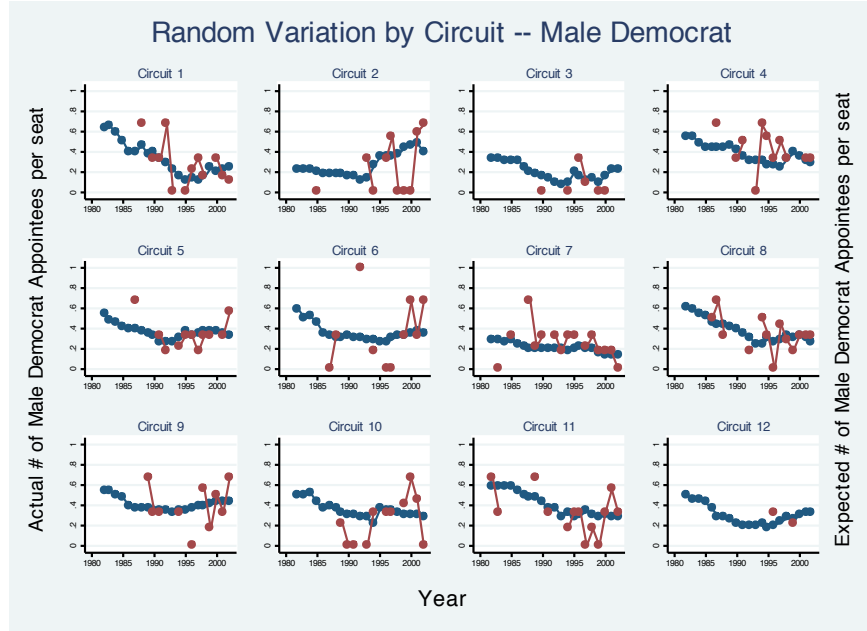
Panel B
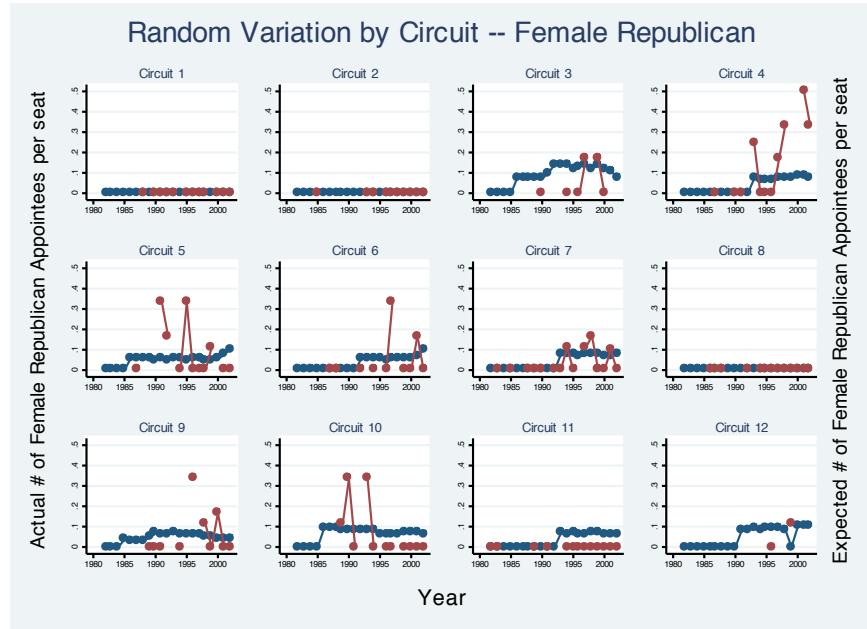
Notes: X-axis is year. Red dots correspond to p_{ct} and the left-hand Y-axis. Blue dots correspond to $\mathbf{E}(p_{ct})$ and the right-hand Y-axis. Panel A presents variation in sexual harassment panels with male Democrats and Panel B presents the same for female Republicans. Each Circuit is displayed separately.

the program, and visiting judges are plugged into those spots by the chief judge. This program also ensures that the same judges are not seated together repeatedly. Thus, the judges know at the beginning of the year which days they will be hearing cases and the compositions of the panels on which they will sit.

Once all the briefing is completed, a case is put into a pool of cases "ready to calendar." If a panel of judges has previously looked at a case, it will be sent back to them (for example, if it was remanded to resolve one issue). Otherwise, a different program randomly assigns cases to these pre-established panels and dates. About eight weeks before the scheduled argument, a preliminary calendar is sent out and the judges review it for recusal. If a judge must recuse himself, the case is taken off the calendar and placed back in the pool for reassignment. Senior judges decide how many days and which months they will work, and this information is entered into the program for random assignment. Before the advent of computer programs, one judge did all of the panel assignments by hand, and the clerks randomly assigned the cases by hand. For more information about random assignment of cases at the Circuit level, see Brown Jr. and Lee (2000).[16]

Other variations from random assignment include: en banc cases that are heard by the entire pool of judges (or a significant fraction in Circuit 9). We do not use these cases, which are also relatively infrequent. Judges can also take sick leave or go on vacation, but this is determined far in advance. Not accounting for vacation, sick leave, senior status, en banc, remand, and recusal can lead to the inference that judges are not randomly assigned. Our identification strategy assumes that these kinds of deviations from random assignment are ignorable. Even a gold-standard random process — the roll of a die — has a deterministic element. If known with precision, the force and torque applied to the die, the subtle air currents, the hardness of the surface, etc., might allow us (or a physicist) to determine with certainty the outcome of these "random" rolls. Despite this obvious non-randomness, we would still have faith in the outcome of a trial with treatment assignments based on die rolls because we are certain that the factors affecting the assignment have no impact on the outcome of interest and hence are ignorable.

To check for randomization, we use data from Boyd et al. (2010), which codes some case characteristics for a subset of 415 gender discrimination cases in the Chicago Judges Projects data (Sunstein et al. 2006). We regress case characteristics on male Democratic (female Republican) judges per seat and find that most characteristics are not correlated with the judicial panel composition. Table

---

[16]See also, http://law.du.edu/images/uploads/neutral-assignment/Neutral_assignment_links.pdf.

II shows that of 19 case characteristics, one is correlated with male Democrats per seat and one is correlated with female Republicans per seat. Both correlations are statistically significant at the 10% level. Given the number of tests, it may be expected that 10% would be statistically significant at the 10% level. In the appendix, we present interviews and another randomization check on how similar the string of actual panel assignments is to a random string.

3.2 Economic Outcomes We examine labor market outcomes using the Merged Outgoing Rotation Groups (MORG) Current Population Survey (CPS).[17]

- Employment status: We create an indicator to distinguish between no-employment (including non-labor force participants) vs. part- or full-time employment.[18] Non-labor force participants include discouraged workers, the potential outsiders in an insider-outsider theory of involuntary unemployment.

- Weekly earnings: Earnings are adjusted to be in 2000 real terms. Logs are taken of 1+earnings. We recode log earnings as zero for individuals who are not in the labor force or not employed.

- Amount of time worked: We use hours worked last week instead of usual weekly hours because usual weekly hours are not consistently available. We recode the number of hours worked as zero for individuals who are not in the labor force or not employed.

- Management status: We use an indicator for whether an individual has an administrator, official, public administration, executive, or other management-related occupation.[19]

- Demographic controls: We consider age, sex, race, marital status, educational attainment, and the geographic location of the individual, which allows us to match the individual's state of residence to the Circuit having legal jurisdiction.

Some studies drop non-labor force participants and find a smaller differences in labor market outcomes between men and women. We later restrict the sample to labor force participants for the analysis of insiders.

We employ a dataset on human resources policies (Dobbin and Kelly 2007) to investigate the adoption of sexual harassment grievance procedures and policies in response to sexual harassment law. We measure the presence of firm-level sexual harassment policies in a national sample of 389 workplaces interviewed in 1997 on the history of human resources practices dating back to 1965.

---

[17]We weight our analysis with CPS-provided weights.

[18]According to the Bureau of Labor Statistics, "Persons who are neither employed nor unemployed are not in the labor force. This category includes retired persons, students, those taking care of children or other family members and others who are neither working nor seeking work."

[19]Occupation is available for about 90% of the unemployed and 33% of those not in the labor force, about 10% of which are managerial. Respondents may interpret this question as being about their previous job.

TABLE II

Randomization Check: Orthogonality with Case Characteristics as Determined by Lower Court

Case Characteristics as Determined by Lower Court	Male Democrat (1)	Female Republican (2)
Direction of Lower Court Decision	0.0115	-0.171
	(0.0856)	(0.187)
Plaintiff claims employer acted in retaliation	-0.102	0.184
	(0.0936)	(0.205)
All plaintiffs are female	0.0126	-0.0920
	(0.0747)	(0.164)
Title IX claim	0.0415	-0.0558
	(0.0252)	(0.0553)
Section 1983 claim	0.0533	-0.0474
	(0.0500)	(0.110)
Constructive discharge from employment	0.00764	0.0726
	(0.0559)	(0.122)
Procedural issues dominate	0.0167	0.163
	(0.0586)	(0.128)
Plaintiff suing under state law	0.0677	-0.283
	(0.0830)	(0.181)
Plaintiff claims illegally denied promotion	-0.0591	-0.0465
	(0.0755)	(0.165)
Plaintiff claims illegally not being hired	-0.0909+	0.105
	(0.0529)	(0.116)
Plaintiff claims illegally fired	0.0460	-0.159
	(0.0961)	(0.210)
Plaintiff claims unequal pay	-0.0235	-0.0868
	(0.0675)	(0.148)
Plaintiff sued under 14th Amendment	0.0606	-0.167+
	(0.0429)	(0.0938)
Plaintiff sued under 1st Amendment	0.0574	-0.0503
	(0.0353)	(0.0775)
Damages major point of contention	0.0765	0.166
	(0.0669)	(0.147)
Contains Section 1981 claim	0.0295	-0.0818
	(0.0585)	(0.128)
Contains age discrimination claim	0.0368	-0.241
	(0.0695)	(0.152)
Contains pregnancy discrimination claim	0.0232	0.0911
	(0.0484)	(0.106)
Contains emotional distress claim	-0.0781	0.0432
	(0.0530)	(0.116)

Notes: Significant at +10%, *5%, **1%. Heteroskedasticity-robust standard errors are in parentheses. Each coefficient represents a separate regression of a distinct case characteristic on the fraction of the panel comprising of male Democrats (respectively, female Republicans).

There is a response rate of 56% but no bias in survey response along observable characteristics: establishment size, organization size, subsidiary status, branch/headquarter status, region, and female chief executive.[20] Our firm-level analysis uses the same controls as Dobbin and Kelly (2007): number of employees, percent of women at location, sex of top executive named, percent women among named executives, age of establishment, dummies for manufacturing, service, and trucking, regional unemployment rate, and percent female in industry.

4 The Effect of Judge Identity on Court Decisions

Judges vote more along party lines than along gender lines (Dixon 2010), and presidents who appoint nontraditional (e.g., women) candidates take the opportunity to appoint more ideologically extreme individuals (Asmussen 2011). Female conservatives exhibit prejudice against females (Eisenman 1991), and female Republicans were 18 percentage points less likely to vote pro-plaintiff in our data. Male Democrats were 13 percentage points more likely to vote pro-plaintiff.[21] Table III Column 1 includes both characteristics together. Column 2 examines the case-level. Since the judicial characteristics are now included as $\frac{N_{ct}}{M_{ct}}$ (a per-seat measure), coefficients need to be divided by three to interpret the effect of an additional female Republican on a three-judge panel. Column 3 reports estimates at the Circuit-year level. Coefficients at the Circuit-year level can differ from the case level due to uneven bunching across Circuit-years.[22] Column 4 presents the first-stage for the individual-level analysis. Column 5 presents the first-stage for the firm analysis. Coefficients differ since the number of firms per Circuit-year is not constant. In each of the last three columns, the F-statistics are strong.

Table IV presents a falsification test of the instrument and shows that the proportion of pro-plaintiff precedents is not related to the number of male Democrats per seat or the number of female Republicans per seat in the one or two years before and after the true instrument. This is expected since Circuit Courts only handle cases that present new legal precedent. This result assures us that our instrument is not picking up general societal trends.

We also employed LASSO to instrument for Law_{ct} (Belloni et al. 2012).[23] The joint F statistic

---

[20]Unfortunately, there is no variable indicating whether the firm had offices in multiple Circuits. However, to the extent that firms make HR decisions at a level that transcends Circuit boundaries, estimated effects would be biased towards 0.

[21]Democrats were also 13 percentage points more likely to vote in favor of sexual harassment plaintiffs. Because female Republicans were more conservative than female Democrats were liberal, female judges on net were 3 percentage points less likely to vote in favor of sexual harassment plaintiffs. We use male Democrat as our instrumental variable to mirror female Republican as instrumental variable.

[22]The R-square increases due to the indicator for whether there were cases, $1\,[M_{ct} > 0]$.

[23]The thirty characteristics are: Democrat, male, male Democrat, female Republican, non-White, Black, Jewish,

TABLE III

First Stage: Relationship Between Judicial Diversity and Pro-Plaintiff Sexual Harassment Precedents, 1982-2002

	(1)	(2)	(3)	(4)	(5)
Female Republican	-0.122+	-0.407	-0.810**	-0.839**	-0.593
	(0.0616)	(0.234)	(0.217)	(0.167)	(0.428)
Male Democrat	0.110**	0.298*	0.459*	0.467**	0.509**
	(0.0353)	(0.110)	(0.150)	(0.147)	(0.0990)
N	752	251	252	5418564	5584
R-sq	0.112	0.128	0.669	0.691	0.733
F-statistic			63.92	50.31	13.95
Pro-Plaintiff measure	Judge Vote	Panel Vote	Law_{ct}	Law_{ct}	Law_{ct}
Controls	θ_c, θ_t	θ_c, θ_t	θ_c, θ_t	θ_c, θ_t	θ_c, θ_t
Analysis level	Judge	Panel	Circuit-Year	MORG CPS	HR Policies

Notes: Significant at +10%, *5%, **1%. Heteroskedasticity-robust standard errors clustered at the Circuit level are in parentheses. Column 1 reports a regression of judicial votes on biographical characteristics (male Democrat and female Republican). Column 2 presents a regression of the panel votes on fraction of the panel comprising of male Democrats (respectively, female Republicans). Column 3 regresses the percent of pro-plaintiff decisions on the percent of male Democrat (respectively, female Republican) panelists in sexual harassment cases. Columns 4 and 5 present the same relationship in the analytical dataset with CPS, respsectively, human resources policies.

increases to 33 for the Circuit-year level and up to 130 for the analysis-level.[24] The use of the LASSO-selected instruments provides a check of over-identification. The results are also unaffected by the inclusion of instruments for $\mathbf{1}[M_{ct} > 0]$, a discussion of which is presented in the appendix.

5 Estimating the Impact of Sexual Harassment Law on Gender Inequality

5.1 Employment Outcomes Table V presents the results on labor market outcomes. The odd-numbered columns report OLS estimates while the even-numbered columns report 2SLS. Roughly speaking, one pro-plaintiff sexual harassment precedent increased female employment relative to males by 1.6 percentage points in the likelihood of working part-time or full-time (Column 2), which translates to roughly an additional 30 minutes worked per year (Column 4). One pro-plaintiff sexual harassment decision reduced the likelihood that males were working by 1.3 percentage points,

---

Catholic, No religion, Mainline Protestant, Evangelical, BA received from same state of appointment, BA from a public institution, JD from a public institution, having an LLM or SJD, elevated from District Court, born in the 1910s, 1920s, 1930s, 1940s, 1950s, appointed when president and congress majority were from the same party, ABA score, above median wealth, appointed by president from an opposing party, prior federal judiciary experience, prior law professor, prior government experience, previous assistant U.S. attorney, and previous U.S. attorney. Adding panel-level interactions (e.g., fraction of judge seats assigned to Democrats multiplied by fraction of judge seats assigned to Blacks) yielded a total of 450 possible instruments.

[24]At the Circuit-year level, the LASSO procedure selected the following three instruments: the interaction between the number of male Democrats per seat and the number of judges born in the 1920s per seat, the interaction between the number of female Republican per seat and the number of judges having an LLM or SJD per seat, and the interaction between the number of female Republican per seat and the number of judges with above median wealth per seat.

TABLE IV

FALSIFICATION TEST: RELATIONSHIP BETWEEN PRO-PLAINTIFF SEXUAL HARASSMENT COURTS OF APPEALS
PRECEDENTS AND COMPOSITION OF SEXUAL HARASSMENT PANELS IN OTHER YEARS, 1982-2002

| | Law_{ct} | | | |
	(1)	(2)	(3)	(4)
Female Republican$_t$	-0.771*	-0.888**	-0.828**	-0.714*
	(0.249)	(0.226)	(0.178)	(0.249)
Male Democrat$_t$	0.442*	0.375*	0.433*	0.460**
	(0.146)	(0.147)	(0.154)	(0.142)
Female Republican$_{t-1}$	-0.226	-0.292		
	(0.320)	(0.374)		
Male Democrat$_{t-1}$	0.0486	0.00643		
	(0.120)	(0.107)		
Female Republican$_{t-2}$		-0.0758		
		(0.263)		
Male Democrat$_{t-2}$		-0.0948		
		(0.181)		
Female Republican$_{t+1}$			-0.301	-0.174
			(0.173)	(0.203)
Male Democrat$_{t+1}$			-0.196	-0.0741
			(0.134)	(0.129)
Female Republican$_{t+2}$				-0.0118
				(0.198)
Male Democrat$_{t+2}$				0.127
				(0.113)
Controls	θ_c, θ_t	θ_c, θ_t	θ_c, θ_t	θ_c, θ_t
N	240	228	240	228
R-sq	0.660	0.665	0.702	0.713

Notes: Significant at +10%, *5%, **1%. Heteroskedasticity-robust standard errors clustered at the Circuit level are in parentheses. The dependent variable is the percent of pro-plaintiff decisions in a Circuit and year. In Column 1, the explanatory variables are the percentages of male Democrat (respectively, female Republican) panelists in sexual harassment cases in the current year and in the previous year. Column 2 includes explanatory variables from each year until two years previous. In Column 3, the explanatory variables are the panel compositions of the current year and the subsequent year. Column 4 includes explanatory variables from each year until two years into the future.

TABLE V
The Effect of Sexual Harassment Law on Gender Inequality

	Employment Status		Hours Worked		Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
β_3	0.007	0.016	0.196	0.508	0.051	0.113
Joint F		8.53		5.42		7.60
β_1	-0.006	-0.013	-0.220	-0.430	-0.043	-0.086
Joint F		28.11		5.07		10.86
Controls	Y	Y	Y	Y	Y	Y
IV	N	Y	N	Y	N	Y
Mean Dep. Var. - Male	0.813	0.813	34.33	34.33	4.910	4.910
Mean Dep. Var. - Female	0.646	0.646	22.78	22.78	3.654	3.654
N	3736671	3736671	3608012	3608012	3410738	3410738
R-sq	0.095	0.095	0.131	0.131	0.133	0.133

Notes: Significant at +10%, *5%, **1%. All regressions use MORG CPS, restrict to individuals between the ages of 18 and 65, and use CPS survey weights. Heteroskedasticity-robust standard errors are clustered at the Circuit level. In Columns 1 and 2, the dependent variable is an indicator for part- or full-time employment. In Columns 3 and 4, the dependent variable is hours worked, and in Columns 5 and 6, log weekly earnings, both set to 0 for individuals who are not employed. Controls are demographic characteristics (age, race dummies, educational attainment dummies, and marital status dummy), Circuit fixed-effects, year fixed-effects, and a dummy for whether there were no cases in that Circuit-year. The regressions include interactions between the female indicator and the dummy for there being no cases and sexual harassment law. OLS estimates presented in odd-numbered columns and instrumental variables estimates presented in even-numbered columns. The instruments are the percentages of male Democrat (respectively, female Republican) panelists in sexual harassment cases and interactions of these percentages with the female indicator.

reducing hours worked by 26 minutes per week. We see a reduction in the wage gap by 0.11 in log real weekly earnings on average (Column 6).

5.2 Robustness Table VI presents a series of robustness checks on the labor force participation results.[25] The point estimate of 1.6 percentage points in Table V is essentially unchanged with the inclusion of Circuit-specific time trends in row A, removal of Circuit and year fixed effects in row B, removal of almost all controls in row C, addition of controls for the expected composition of judicial panels in row D, and addition of state fixed effects in row E. The joint F test increases when we do not use CPS weights in row F or vary the lag specification in row G. The estimates also change little when we drop one Circuit at a time in row H, though the effects may have been largest in Circuit 5, which includes Texas, Louisiana, and Mississippi. Results are also similar whether we cluster standard errors at the state level in row I, collapse the data to the Circuit-year level in row J, use LASSO IV in row J, or use District IV in row L. We also employ Monte Carlo placebo simulations

---

[25]To streamline the presentation, we focus on the impact on female employment shares.

that randomly assign the laws and panel assignments to different Circuits. The most conservative simulation assigns the complete time series of legal variation for one Circuit to another Circuit. The point estimate for the collapsed Circuit-year data is at the 90th percentile in these simulations. Notably, the lead effects are insignificant in rows M and N and the point estimates are very small relative to the lag effects.

6   Mechanisms

6.1   Disaggregating the Effects of Sexual Harassment Law by Region   To explore the mechanisms, we take advantage of Charles et al. (2010), which constructs a measure of Male Sexism by region using the General Social Survey. According to the theory, the empirical patterns should be attenuated in regions where males exhibit less sexism. The two census regions scoring highest in male sexism are East South Central and West South Central. In Table VI, we see that dropping Circuit 5 (a subset of East and West South Central) renders the smallest estimates, meaning the biggest effects are found in Circuit 5. In contrast, dropping Circuit 9 renders the largest estimates, meaning the smallest effects are found in Circuit 9, which is a subset of Pacific and Mountain and where male sexism scores are low.

6.2   Disaggregating the Effects of Sexual Harassment Law by Industry   Next, we disaggregate the effects of sexual harassment law and find that impacts were larger in the construction industry. Sexual harassment rates per 100,000 women are highest in the construction industry (Hersch 2011) and the bulk of sexual harassment plaintiffs (38%) are blue-collar workers (Juliano and Schwab 2000). Table VII documents the ameliorative effects on inequality across all outcomes employment status, hours worked, and earnings especially in the construction industry. This table is not comparable to Table V because we need data on industry, and labor force participants are more likely to report their industry.

6.3   Insiders and Outsiders   Table VIII reports the estimated effects of sexual harassment law for labor force participants. We find that pro-plaintiff sexual harassment precedents *increased* gender inequality by 0.16 hours worked last week, 0.004 in log real weekly earnings, and 0.7 percentage points in the likelihood to be a manager. Next, we assess how much outsider females gained relative to insider females, insider males, and outsider males. Column 1 in Table VIII indicates that insider men gained by 0.008 log real weekly earnings while insider women gained by 0.004 log real weekly earnings. Column 6 in Table V indicates that insider and outsider men lost 0.086 log real weekly

TABLE VI

Robustness of Effect of Sexual Harassment Law on Employment

	β_3 (1)	Joint F (2)
A. Add Circuit-Specific Trends	0.016	8.35
B. Drop θ_c, θ_t	0.016	8.17
C. Only $1\left[M_{ct-n} > 0\right], F_{ict}$	0.017	8.08
D. Add $\mathbf{E}(\frac{N_{ct}}{M_{ct}})$	0.016	8.31
E. Add State Fixed Effects	0.016	8.00
F. No CPS Weights	0.013	16.49
G. Add 2-year Lead	0.021	19.25
H. Drop 1 Circuit		
Circuit 1	0.015	6.57
Circuit 2	0.017	14.22
Circuit 3	0.016	13.81
Circuit 4	0.017	17.12
Circuit 5	0.007	37.15
Circuit 6	0.017	6.61
Circuit 7	0.017	8.72
Circuit 8	0.013	6.33
Circuit 9	0.019	5.13
Circuit 10	0.018	34.03
Circuit 11	0.014	17.23
Circuit 12	0.016	8.76
I. Cluster at State Level	0.016	11.88
J. Collapsed at Circuit-Year level	0.017	14.64
K. Collapsed with Lasso IV	0.011	25.47
L. Collapsed with Lasso and District IV	0.013	9.40
M. Falsification: Lead Effect on Inequality	0.00163	0.02
N. Falsification: Lead Effect on Males	0.00549	0.39

Notes: Significant at $+10\%$, *5%, **1%. All regressions use MORG CPS, restrict to individuals between the ages of 18 and 65, and use CPS survey weights. Heteroskedasticity-robust standard errors are clustered at the Circuit level. The baseline specification is Column 2 of Table V. The dependent variable is an indicator for part- or full-time employment. All specifications are instrumental variables estimates.

TABLE VII

The Effect of Sexual Harassment Law on the Construction Industry

Construction Industry	Employment Status		Hours Worked		Earnings	
	Yes	No	Yes	No	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
β_3	0.010	-0.007	0.378	-0.389	0.044	-0.033
Joint F	13.66	2.99	25.83	26.27	10.18	3.14
β_1	-0.018	0.000	-1.045	0.087	-0.132	-0.002
Joint F	19.27	3.51	49.25	1.45	20.09	2.86
Controls	Y	Y	Y	Y	Y	Y
IV	Y	Y	Y	Y	Y	Y
Mean Dep. Var. - Male	0.836	0.892	33.91	38.31	4.977	5.517
Mean Dep. Var. - Female	0.793	0.826	27.45	29.47	4.385	4.775
N	210153	2949731	201678	2825198	163297	2666305
R-sq	0.048	0.044	0.079	0.094	0.071	0.108

Notes: Significant at +10%, *5%, **1%. All regressions use MORG CPS, restrict to individuals between the ages of 18 and 65, use CPS survey weights, and restrict to those who report an industry category. Heteroskedasticity-robust standard errors are clustered at the Circuit level. In Columns 1 and 2, the dependent variable is an indicator for part- or full-time employment. In Columns 3 and 4, the dependent variable is hours worked, and in Columns 5 and 6, log weekly earnings, both set to 0 for individuals who are not employed. Controls are demographic characteristics (age, race dummies, educational attainment dummies, and marital status dummy), Circuit fixed-effects, year fixed-effects, and a dummy for whether there were no cases in that Circuit-year. The regressions include interactions between the female indicator and the dummy for there being no cases and sexual harassment law. OLS estimates presented in odd-numbered columns and instrumental variables estimates presented in even-numbered columns. The instruments are the percentages of male Democrat (respectively, female Republican) panelists in sexual harassment cases and interactions of these percentages with the female indicator.

TABLE VIII

The Effect of Sexual Harassment Law on Insiders

| | Labor Force Participants Reporting Non-Zero Wages | | |
| | Earnings | Hours Worked | Management |
	(1)	(2)	(3)
β_3	-0.004	-0.160	-0.007
Joint F	35.16	58.21	30.38
β_1	0.008	0.103	0.003
Joint F	24.71	13.10	6.25
Controls	Y	Y	Y
IV	Y	Y	Y
Mean Dep. Var. - Male	6.298	42.90	0.144
Mean Dep. Var. - Female	5.854	36.04	0.120
N	2424997	2622664	2755279
R-sq	0.296	0.081	0.057

Notes: Significant at +10%, *5%, **1%. All regressions use MORG CPS, restrict to individuals between the ages of 18 and 65, use CPS survey weights, and restrict to those who report an industry category. Heteroskedasticity-robust standard errors are clustered at the Circuit level. In Columns 1 and 2, the dependent variable is hours worked, and in Columns 3 and 4, log weekly earnings. In Columns 5 and 6, the dependent variable is an indicator for being in management. Controls are demographic characteristics (age, race dummies, educational attainment dummies, and marital status dummy), Circuit fixed-effects, year fixed-effects, and a dummy for whether there were no cases in that Circuit-year. The regressions include interactions between the female indicator and the dummy for there being no cases and sexual harassment law. OLS estimates presented in odd-numbered columns and instrumental variables estimates presented in even-numbered columns. The instruments are the percentages of male Democrat (respectively, female Republican) panelists in sexual harassment cases and interactions of these percentages with the female indicator.

earnings and insider and outsider women gained 0.027 log real weekly earnings. Assuming the 65% of women with part- or full-time employment to be insiders (similarly for 81% of men), then accounting for net movements into and out of labor force participation from Column 2 of Table V yields the following summary for the effects for females and males:[26]

- $(0.35 - 0.003) * 0 + (0.003) * 5.9 + (0.65) * 0.004 = 0.020$, which is near 0.027

- $(0.19) * 0 + (0.013) * -6.3 + (0.81 - 0.013) * 0.008 = -0.075$, which is near -0.086

These attributions account only for net flows. If there were significant inflows *and* outflows in the labor force, insider women would be worse off and outsider women better off than our calculations indicate.

---

[26]Describing in words, female outsiders who remained outsiders constitute nearly 35%, female outsiders who enter the labor force constitute 0.3%, and insider females are 65%. Male outsiders constitute 19%, insider males who became outsiders constitute 1.3% and the remaining insiders constitute nearly 80%. The numbers 0.004 and 0.008 come from Table VIII Column 1. The numbers 5.9 and 6.3 are the group-mean dependent variables in Table VIII. The numbers +0.3% and -1.3% come from Table V Column 2.

In sum, forbidding sexual harassment caused average wages of male *and* female labor force participants to increase. This result is inconsistent with theories of sexual harassment law such as a redistribution from females to males, tax on the hiring of men, or simply an (economically irrelevant) elimination of masculine culture. The resulting gains for females were tilted towards outsider females who gained six times more than insider females; outsider females also gained three times more than remaining insider males. An indication of the increase in productivity, particularly for females, is that while insider females gained 0.004 log real weekly earnings, they lost in hours worked (a decline of 0.06 hours per week) and also in managerial status (0.4 percentage point decline in likelihood to be a manager). Taken together, these results suggest that forbidding sexual harassment in large part encouraged the entry of outsider women, who then obtained part- or full-time employment and received wages, and increased the productivity of women in the labor force. Previously insider men lost the most in terms of employment status.

6.4   Human Resources Policies   Table IX reports the effect of pro-plaintiff sexual harassment precedent on the adoption of sexual harassment human resources policies. Firms were 5.7 percentage points more likely to have a sexual harassment policy after a pro-plaintiff precedent (Column 2). From 1982 to 1997, an average of 54% of establishments reported having sexual harassment policies, growing from 15% in 1982 to 96% in 1997. The impact of sexual harassment law is equivalent to 37% of the yearly increase.[27]

7   Discussion and Extensions

To compare with the Equal Employment Opportunity Act (EEOA), which increased black employment shares by 0.5 to 1.1 points per year (Chay 1998), we take into account how often there is a case and in which direction a case is typically resolved (half of the Circuit-years had a case and two-thirds of the cases were pro-plaintiff while one-third were pro-defendant). Multiplying 1.3 by 0.67 and 0.5 indicates that the effects of pro-plaintiff sexual harassment precedents are in the range of the effects of EEOA found by Chay (1998). For another benchmark, a standard deviation increase in median male sexism in a state is associated with a 0.031 log point decrease in female

---

[27]We take into account how often there is a case and in which direction a case is typically resolved (half of the Circuit-years had a case and two-thirds of the cases were pro-plaintiff while one-third were pro-defendant). Multiplying 0.67, 0.5, and 5.7 percentage points suggests that, since the origination of sexual harassment law, its average impact in a typical Circuit-year caused firms to be 1.9 percentage points more likely to have a sexual harassment policy. Assuming a linear 81 percentage point increase in sexual harassment policy during the 16 years, firms were 5.1 percentage points more likely to have a sexual harassment policy in any given Circuit-year. Under these assumptions, pro-plaintiff sexual harassment law appears to have played an important role in the change of human resources policies to address sexual harassment, equivalent to 37% of the yearly change.

TABLE IX
THE EFFECT OF SEXUAL HARASSMENT LAW ON HUMAN RESOURCES SEXUAL HARASSMENT POLICY

	Presence of Sexual Harassment Policy	
	(1)	(2)
β_1	0.029	0.057
Joint F	9.34	24.09
Controls	Y	Y
IV	N	Y
Mean dep. var.	0.543	0.543
N	4014	4014
R-sq	0.260	0.259

Notes: Significant at +10%, *5%, **1%. Heteroskedasticity-robust standard errors are clustered at the Circuit level. The dependent variable is the presence of a sexual harassment human resources policy. Controls are firm characteristics (number of employees, percent of women at location, sex of top executive named, percent women among named executives, age of establishment, dummies for manufacturing, service, and trucking, regional unemployment rate, and percent female in industry), Circuit fixed-effects, year fixed-effects, Circuit-specific time trends, and a dummy for whether there were no cases in that Circuit-year. OLS estimates presented in odd-numbered columns and instrumental variables estimates presented in even-numbered columns. The instruments are the percentages of male Democrat (respectively, female Republican) panelists in sexual harassment cases.

wages relative to males (Charles et al. 2010), so sexual harassment law is comparable to the effects of a standard deviation in median male sexism.[28]

Another counter-factual of potential interest is one that accounts for both β_3 and β_4, which is the effect of $\mathbf{1}[M_{ct} > 0]$. That is, instead of estimating the effect of pro-plaintiff precedents relative to pro-defendant precedents, another counterfactual could be the effect of pro-plaintiff precedents relative to the absence of a precedent. To have a causal interpretation of β_4, we leverage the random assignment of District Court judges (the appendix provides a randomization check). Let $w_{ct} = \frac{\sum_{d=1}^{J} K_{cdt} * \left(\frac{L_{cdt}}{K_{cdt}} \right)}{\sum_{d=1}^{J} K_{cdt}}$, where K_{cdt} denotes the number of cases filed in District *court* d within Circuit c at time t.[29] L_{cdt} denotes the number of judges with a particular characteristic assigned to cases. Variation from assigning District judges, who are disproportionately appealed, leads to the presence of a case in the Circuit, $\mathbf{1}[M_{ct} > 0]$.[30] Sen (2015) reports that ethnicity of the District judge, in particular, African-American judges' opinions were treated differently, which we also find and document in the appendix. Instrumenting with randomly assigned judges in both Circuit and District courts, we

[28]We take coefficients β_3 of 0.11 and β_4 of -0.05. Multiplying by 0.67 and 0.5 indicates that during the development of pro-plaintiff sexual harassment precedents, a typical year saw a 0.02 log points increase in female wages relative to males, roughly equivalent to two-thirds of a standard deviation in median male sexism.

[29]J goes from 5 to 13 depending on the District.

[30]Note that the terms in the numerator need $K_{cdt} > 0$. An approximation is to define $K_{cdt} * \left(\frac{L_{cdt}}{K_{cdt}} \right)$ as 0 if $K_{cdt} = 0$. Then, the instrument can be constructed if $\sum_{d=1}^{J} K_{cdt} > 0$, i.e., the denominator is non-zero.

found that the coefficients for Law_{ct} are very similar (Table VI). To add the effect of a presence of a Circuit case with the effect of a pro-plaintiff decision, we use the District IV specification from row L of Table VI where β_3 is 0.013 and β_4 is -0.009. Accounting for the typical representation of cases yields a smaller net effect. What this suggests is that pro-defendant precedents have effects that are opposite relative to the status quo of no precedent.

8   Conclusion

Making the workplace friendlier to women has been one of the most dramatic labor market changes in the past half-century, yet, the consequences of forbidding sexual harassment on female labor force outcomes remain unknown. Interpreting anti-discrimination law to forbid sexual harassment has been an important contribution of feminist legal theory (MacKinnon 1979) and has recently received renewed policy attention with the #MeToo movement. Unlike other employment laws, sexual harassment law is generally considered "good" social policy and has not come under fire for its potential negative consequences in the way that other employment protections, such as ADA and maternity mandates, have. Yet, economic theory, at first glance, suggests that the potential effects of forbidding sexual harassment may be similar to those of other employment mandates. It may exacerbate gender inequality overall because it could be viewed as a tax on the hiring of women. This paper puts forward an alternative theory of insiders, outsiders, and involuntary unemployment—where insiders harass outsiders to capture economic rents.

We identify the impact of pro-plaintiff sexual harassment precedent on gender inequality by using the fact that federal judges are randomly assigned to Circuit cases along with the fact that gender and party of appointment of judges affect sexual harassment decisions. We find that pro-plaintiff sexual harassment precedent increases female wages and employment relative to that of men. These findings are more consistent with an insider-outsider theory of involuntary unemployment outweighing standard models of compensating wage differentials or mandated benefits. In addition, we find that the gains experienced by females exceeds the losses of insider males, which suggests that firms were not profit-maximizing and forbidding harassment on their own in the absence of sexual harassment law and that there was some degree of rent capture by insiders. Our findings are consistent with the persistence of beliefs that women should play only certain roles, or should know their "place", and that this persistence accounts for a substantial portion of inequality.

Methodologically, for jurists interested in evaluating the impacts of their decisions, the empirical framework developed here provides causal estimates of court precedent holding all else equal

including unobserved factors and the general equilibrium effects are those which we would want to include. Furthermore, it has the advantages that the LATE interpretation of the IV estimates are policy relevant. We hope it proves fruitful for policy-makers and judges interested in assessing the impact of court-made law as well as for scholars and theorists interested in evaluating theories of societal responses to law.

References

Acemoglu, Daron, and Joshua D. Angrist, 2001, Consequences of Employment Protection? The Case of the Americans with Disabilities Act, *The Journal of Political Economy* 109, 915–957.

Ashenfelter, Orley, Theodore Eisenberg, and Stewart J. Schwab, 1995, Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes, *Journal of Legal Studies* 24, 257–281.

Asmussen, Nicole, 2011, Female and Minority Judicial Nominees: President's Delight and Senators' Dismay?, *Legislative Studies Quarterly* 36, 591–619.

Autor, David H., John J. Donohue, and Stewart J. Schwab, 2006, The Costs of Wrongful-Discharge Laws, *The Review of Economics and Statistics* 88, 211–231.

Barrios, Thomas, Rebecca Diamond, Guido W. Imbens, and Michal Kolesár, 2012, Clustering, Spatial Correlations and Randomization Inference, *Journal of the American Statistical Association* 107, 578–591.

Basu, Kaushik, 2003, The Economics and Law of Sexual Harassment in the Workplace, *The Journal of Economic Perspectives* 17, 141–157.

Belloni, Alexandre, Daniel L. Chen, Victor Chernozhukov, and Christian Hansen, 2012, Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain, *Econometrica* 80, 2369–2429.

Bertrand, Marianne, and Esther Duflo, 2016, Field Experiments on Discrimination, Technical report, National Bureau of Economic Research.

Bird, Susan Willett, 1975, The Assignment of Cases to Federal District Court Judges, *Stanford Law Review* 27, 475–487.

Black, Sandra E., and Philip E. Strahan, 2001, The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry, *The American Economic Review* 91, 814–831.

Blau, Francine D., and Lawrence M. Kahn, 2006, The U.S. Gender Pay Gap in the 1990s: Slowing Convergence, *Industrial and Labor Relations Review* 60, 45–66.

Boyd, Christina, Lee Epstein, and Andrew D. Martin, 2010, Untangling the Causal Effects of Sex on Judging, *American Journal of Political Science* 54, 389–411.

Breyer, Stephen, 2006, *Active Liberty: Interpreting Our Democratic Constitution* (Vintage Books).

Brown Jr., Robert J., and Allison Herren Lee, 2000, Neutral Assignment of Judges at the Court of Appeals, *Texas Law Review* 78, 1037–1116.

Card, David E., and John Enrico DiNardo, 2002, Technology and U.S. Wage Enequality: A Brief Look, *Economic Review* 87, 45–62.

Charles, Kerwin K., Jonathan Guryan, and Jessica Pan, 2010, Sexism and Women's Labor Market Outcomes, Working paper, University of Chicago.

Charles, Kerwin Kofi, and Jonathan Guryan, 2008, Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination, *The Journal of Political Economy* 116, 773–809.

Chay, Kenneth Y., 1998, The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employment Opportunity Act of 1972, *Industrial and Labor Relations Review* 51, 608–632.

Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad, 2014, Family Welfare Cultures, *Quarterly Journal of Economics* 129, 1711–1752.

Dixon, Rosalind, 2010, Female Justices, Feminism, and the Politics of Judicial Appointment: A re-examination, *Yale Journal of Law and Feminism* 21, 297–338.

Dobbin, Frank, and Erin L. Kelly, 2007, How to Stop Harassment: Professional Construction of Legal Compliance in Organizations1, *American Journal of Sociology* 112, 1203–1243.

Eberts, Randall W., and Joe A. Stone, 1985, Male-Female Differences in Promotions: EEO in Public Education, *The Journal of Human Resources* 20, 504–521.

Edelman, Lauren B., 1992, Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law, *American Journal of Sociology* 97, 1531–1576.

Eisenman, Russell, 1991, Gender and Racial Prejudice of Conservative College Women, *Psychological Reports* 68, 450–450.

Epstein, Richard A., 1995, *Forbidden Grounds: The Case Against Employment Discrimination Laws* (Harvard University Press).

Farhang, Sean, and Gregory Wawro, 2004, Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making, *Journal of Law, Economics, and Organization* 20, 299–330.

Fehr, Ernst, and Urs Fischbacher, 2002, Why Social Preferences Matter - The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives, *The Economic Journal* 112, C1–C33.

Fitzpatrick, Brian T, 2010, An Empirical Study of Class Action Settlements and Their Fee Awards, *Journal of Empirical Legal Studies* 7, 811–846.

Goldman, Sheldon, 1999, *Picking Federal Judges: Lower Court Selection from Roosevelt Through Reagan* (Yale University Press).

Gruber, Jonathan, 1994, The Incidence of Mandated Maternity Benefits, *The American Economic Review* 84, 622–641.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske, 2002, Market Forces and Sex Discrimination, *The Journal of Human Resources* 37, 353–380.

Hersch, Joni, 2011, Compensating Differentials for Sexual Harassment, *The American Economic Review Papers and Proccedings* 101, 630–634.

Jordan, Samuel P., 2007, Early Panel Announcement, Settlement, and Adjudication, *Brigham Young University Law Review* 2007, 55–107.

Juliano, Ann, and Stewart J. Schwab, 2000, The Sweep of Sexual Harassment Cases, *Cornell Law Review* 86, 548–592.

Keele, Denise M., Robert W. Malmsheimer, Donald W. Floyd, and Lianjun Zhang, 2009, An Analysis of Ideological Effects in Published Versus Unpublished Judicial Opinions, *Journal of Empirical Legal Studies* 6, 213–239.

Kling, Jeffrey R., 2006, Incarceration Length, Employment, and Earnings, *The American Economic Review* 96, 863–876.

Lindbeck, Assar, and Dennis J. Snower, 1988, Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach, *The American Economic Review* 78, 167–188.

Lindbeck, Assar, and Dennis J. Snower, 2001, Insiders versus Outsiders, *The Journal of Economic Perspectives* 15, 165–188.

MacKinnon, Catharine A., 1979, *Sexual Harassment of Working Women: A Case of Sex Discrimi-*

nation, Yale Fastback Series (Yale University Press).

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand, 2013, Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt, *American Economic Review* 103, 1797–1829.

McLaughlin, Heather, Christopher Uggen, and Amy Blackstone, 2017, The economic and career effects of sexual harassment on working women, *Gender & Society* 31, 333–358.

Merritt, Deborah Jones, and James J. Brudney, 2001, Stalking Secret Law: What Predicts Publication in the United States Courts of Appeals, *Vanderbilt Law Review* 54, 69–121.

Nash, Jonathan R., 2015, Examining Federal District Judges' Referrals to Magistrate Judges, in *2015 Annual Meeting of the International Society of New Institutional Economics (ISNIE)*, number 1-44.

Neumark, David, 2016, Experimental Research on Labor Market Discrimination, Technical report, National Bureau of Economic Research.

Newman, Meredith A., Robert A. Jackson, and Douglas D. Baker, 2003, Sexual Harassment in the Federal Workplace, *Public Administration Review* 63, 472–483.

Nielsen, Laiura B., Robert L. Nelson, and Ryon Lancaster, 2010, Individual Justice or Collective Legal Mobilization? Employment Discrimination Litigation in the Post Civil Rights United States, *Journal of Empirical Legal Studies* 7, 175–201.

O'Neill, June, 2003, The Gender Gap in Wages, circa 2000, *The American Economic Review* 93, 309–314.

Peresie, Jennifer L., 2005, Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts, *The Yale Law Journal* 114, 1759–1790.

Posner, Richard A., 1998, Against Constitutional Theory, *New York University Law Review* 73, 1–22.

Rosen, Sherwin, 1974, Hedonic prices and implicit markets: product differentiation in pure competition, *Journal of political economy* 82, 34–55.

Sen, Maya, 2015, Is Justice Really Blind? Race and Appellate Review in U.S. Courts, *Journal of Legal Studies* 44, (In Press).

Summers, Lawrence H., 1989, Some Simple Economics of Mandated Benefits, *The American Economic Review* 79, 177–183.

Summers, Lawrence H., 2005, Remarks at NBER Conference on Diversifying the Science and Engineering Workforce, in *National Bureau of Economic Research Conference on Diversifying the Science and Engineering Workforce. Cambridge, MA. Retrieved March*, volume 9.

Sunstein, Cass R., David Schkade, Lisa M. Ellman, and Andres Sawicki, 2006, *Are Judges Political?: An Empirical Analysis of the Federal Judiciary* (Brookings Institution Press).

Taha, Ahmed E., 2004, Publish or Paris? Evidence of How Judges Allocate Their Time, *American Law and Economics Review* 6, 1–27.

Taha, Ahmed E., 2009, Judge Shopping: Testing Whether Judges' Political Orientations Affect Case Filings, *University of Cincinnati Law Review* 20, 101–135.

Waldfogel, Joel, 1995, The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory, *The Journal of Political Economy* 103, 229–260.

Web Appendix:

A *Major Doctrinal Developments in Sexual Harassment Law*

1964 – **Title VII** – prohibits sex discrimination in employment.

1976 – **Williams v. Saxbe** – Court recognized sexual harassment as a form of sex discrimination when sexual advances by male supervisor towards female employee, if proven, would be deemed an artificial barrier to employment placed before one gender and not another.

1977 – **Barnes v. Costle** – U.S. Court of Appeals for the Second District ruled that retaliation against a female employee for rejecting sexual advances of her boss is a violation of Title VII's prohibition against sex discrimination.

1980 – **EEOC** issues guidelines forbidding "sexual harassment" as a form of sex discrimination.

1985 – **McKinney v. Dole** - U.S. Court of Appeals for the DC Circuit ruled that physical violence, even if it is not overtly sexual, can be sexual harassment if the unwelcome conduct is based on the victim's gender.

1986 – **Meritor Savings Bank, FSB v. Vinson** – The Supreme Court first recognized "sexual harassment" as a violation of Title VII and established the standards for analyzing whether the conduct was welcome and levels of employer liability.

1988 – **Hall v. Gus Construction** - U.S. Court of Appeals for the Eighth District finds that when male construction workers "hazed" three female colleagues, even if the conduct was not specifically sexual in nature, was gender based harassment.

1991 – **Ellison v. Brady** – Changed analysis of conduct from reasonable person to reasonable women test when determining whether actionable sexual harassment occurred.

1991 – **Civil Rights Act of 1991** provides for jury trials and for increased damages in Title VII sexual harassment suits.

1993 – **Harris v. Forklift Systems, Inc** – plaintiff may bring sexual harassment claim without necessarily showing psychological harm. In addition to Meritor, the factors when analyzing whether sexual harassment occurred include: (i) Frequency of conduct; (ii) Severity; (iii) Whether the conduct is physically threatening or humiliating; (iv) Or is a mere offensive utterance; (v) And whether the conduct unreasonably interferes with employees work performance; (vi) No Single Factor is Required but Totality of the Circumstances Test.

1998 – **Faragher v. City of Boca Raton** - Supreme Court decision that establishes that an employer is subject to vicarious liability for hostile environment created by a supervisor unless the employer can demonstrate that it exercised reasonable care to prevent and correct promptly any sexually harassing behavior and that the employee unreasonably failed to take advantage of any preventative or corrective opportunities provided by the employer.

1998 – **Burlington Industries, Inc v. Ellerth** - Companion Supreme Court decision to Faragher that further elaborates that the employer's "Faragher" defense to vicarious liability is not available if the employee suffers a tangible job consequence as result of supervisor's actions.

First Stage Relationship Between Absence of Circuit Cases and Judicial Composition of District Court Cases, 1986-2002

	$1\,[M_{ct} = 0]$			
	(1)	(2)	(3)	(4)
Black judges$_t$	0.470*	0.485**	0.419*	0.459**
	(0.176)	(0.138)	(0.145)	(0.144)
Black judges$_{t+1}$		-0.0492	-0.0248	
		(0.369)	(0.357)	
Black judges$_{t+2}$			-0.0192	
			(0.374)	
Circuit-year controls	N	N	N	θ_c, θ_t
N	203	190	177	203
R-sq	0.019	0.020	0.014	0.372

Notes: Significant at +10%, *5%, **1%. Heteroskedasticity-robust standard errors clustered at the Circuit level are in parentheses. The dependent variable is the absence of any decisions in a Circuit and year. In Column 1, the explanatory variables are the percentages of black panelists in District Court sexual harassment cases in the current year. Column 2 adds an explanatory variable from the subsequent year. Column 3 includes explanatory variables from each year until two years into the future. Column 4 includes dummy indicators for Circuit and for year.

B   District IV

We report a first-stage correlation between the presence of a Circuit case and the proportion of District cases in a Circuit that received a Black judge, with an F-statistic of 7. With the inclusion of Circuit and year fixed effects, the point estimates are identical and the F-statistic goes up to 10 as can be seen from comparing Columns 1 and 4 in Appendix Table I. The average Circuit-year had 8% of District cases ruled by a Black judge, which decreased the probability the Circuit-year had an appeal by 46% (Appendix Table I and Appendix Figure 1).

Columns 2 and 3 of Appendix Table I show that the proportion of District cases heard by a Black judge is not related to the presence of a Circuit case in the previous one or two years.[31] Since we may expect a lag between District and Circuit Court rulings, we also considered the relationship between the presence of a Circuit case and the previous year's assignment of District judges to sexual harassment cases. Regardless of the source of variation, the two-stage least squares estimates are very similar, which is consistent with the 2SLS results being unaffected by the inclusion of District instruments.

---

[31] For 1982-1985, most Circuits did not have any District-level sexual harassment cases so when the District IV is employed, those years are dropped from the sample.

C   Additional Remarks

Our empirical strategy needs the assignment of judges to sexual harassment cases to approximate a true experiment. This section reports an additional assessment of this assumption.

C.1   Omnibus Test for Circuit Courts: Random Strings   We examine deviations from random assignment by seeing whether the sequence of proportions of judges is similar to a random process. Figures 3A and 3B suggest visually that panel composition is not serially correlated. We formally investigate this by:

1. Propose a statistic that can be computed from the sequence of numbers of female Republicans (male Democrats) per seat within a Circuit.

2. Compute the statistic for the actual sequence, s^*.

3. Compute the statistic for each of 1,000 bootstrap samples from the actual sequence, i.e., s_1, s_2, s_3 . . .s_n. Since there were changes in the expected number of female Republicans (male Democrats) per seat over time, we treat our bootstrap samples as a vector of realized random variables, with the probability based on the expectation during the Circuit-year.

4. Compute the empirical p-value, p_i by determining where s^* fits into s_1, s_2, s_3 . . .s_n.

5. Repeat steps 1-4 and calculate p_i for each unit.

We use the following statistics:

Autocorrelation: We see if the value in the j$^{\text{th}}$ case depends on the outcome in the j-1$^{\text{th}}$case. This statistic can detect whether judicial assignments are "clustered," meaning a higher than expected number of back-to-back seat assignments to a particular type of judge. This test tells us whether certain judges sought out sexual harassment cases, perhaps in sequence.

Mean-Reversion: We test whether there is any form of mean reversion in the sequence, meaning that the assignment in the n$^{\text{th}}$ case is correlated with the assignment in previous $n-1$ cases. This test tells us whether judges or their assignors were attempting to equilibrate their presence, considering whether a judge was "due" for a sexual harassment case.

Longest-Run: We test whether there are abnormally long "runs" of certain types of judges per seat. This test tells us whether certain Circuits may have assigned certain judges with sexual harassment cases during certain time periods, for example, to achieve specialization.

Number of Runs: Instead of simulating 1000 random strings, we compute the exact statistic for number of runs. This test captures violations of randomization at the case level rather than Circuit-year.

	Female Republican		Male Democrat	
Auto-correlation	0.34*		0.24	
Mean reversion	0.32		0.16	
Longest run	0.22		0.23	
	Kolmogorov-Smirnov Test			
	Values of D_n for Various P			
	$n = 9$		$n = 12$	
	$\mathrm{Prob}(\sqrt{n}D_n < b)$	$b/\sqrt{n}$	$\mathrm{Prob}(\sqrt{n}D_n < b)$	$b/\sqrt{n}$
	0.01	0	0.01	0.009
	0.05	0.0373	0.05	0.0345
	0.1	0.0619	0.1	0.0553
	0.25	0.1091	0.25	0.09
	0.5	0.1804	0.5	0.1574
	0.75	0.2608	0.75	0.2275
	0.9	0.3392	0.9	0.2958
	0.95	0.3874	0.95	0.3381
	0.99	0.4795	0.99	0.4288

Notes: Significant at +10%, *5%, **1%. This table reports Kolmogorov-Smirnov tests for differences in the distributions of P-values relative to a uniform distribution for a sample size of 9 for female Republican (3 Circuits never had a female Republican) and 12 for male Democrat.

With a truly random process, the collection of all unit p-values should be uniformly distributed. (Imagine that you generate summary statistics for 1000 random strings. The 1001th random string should have a summary statistic that is equally likely to be anywhere from 1 to 1000.) A visual examination suggests that the empirical distributions for our p-values approach the CDF of a uniform distribution (Figure 4), which we formally test using a Kolmogorov-Smirnov test statistic from an empirical null (Table II). In our robustness checks, we drop one Circuit at a time. We also checked one or two years before the true instrument, that judicial decision-making is not correlated with future judicial assignment.

These checks address potential issues that arise because our data comprise published opinions. First, settlement is not an issue: judges are revealed after litigants file their briefs in Circuit Courts, sometimes only a few days before the hearing, which gives little opportunity and incentive for settlement upon learning the identity of the panel.[32] Second, unpublished cases do not have precedential authority.[33] To see the random strings test as an omnibus test: Suppose female Republicans are more likely to publish cases or use the keywords that generate our sample, then we should expect

---

[32]Notably, settlement rates were unaffected by the D.C. Circuit announcing their judges earlier (Jordan 2007).

[33]Unpublished cases are routine cases. Judicial ideology has no impact on unpublished cases (Keele et al. 2009) nor the decision to publish (Merritt and Brudney 2001).

FIGURE 4.— Randomization Check
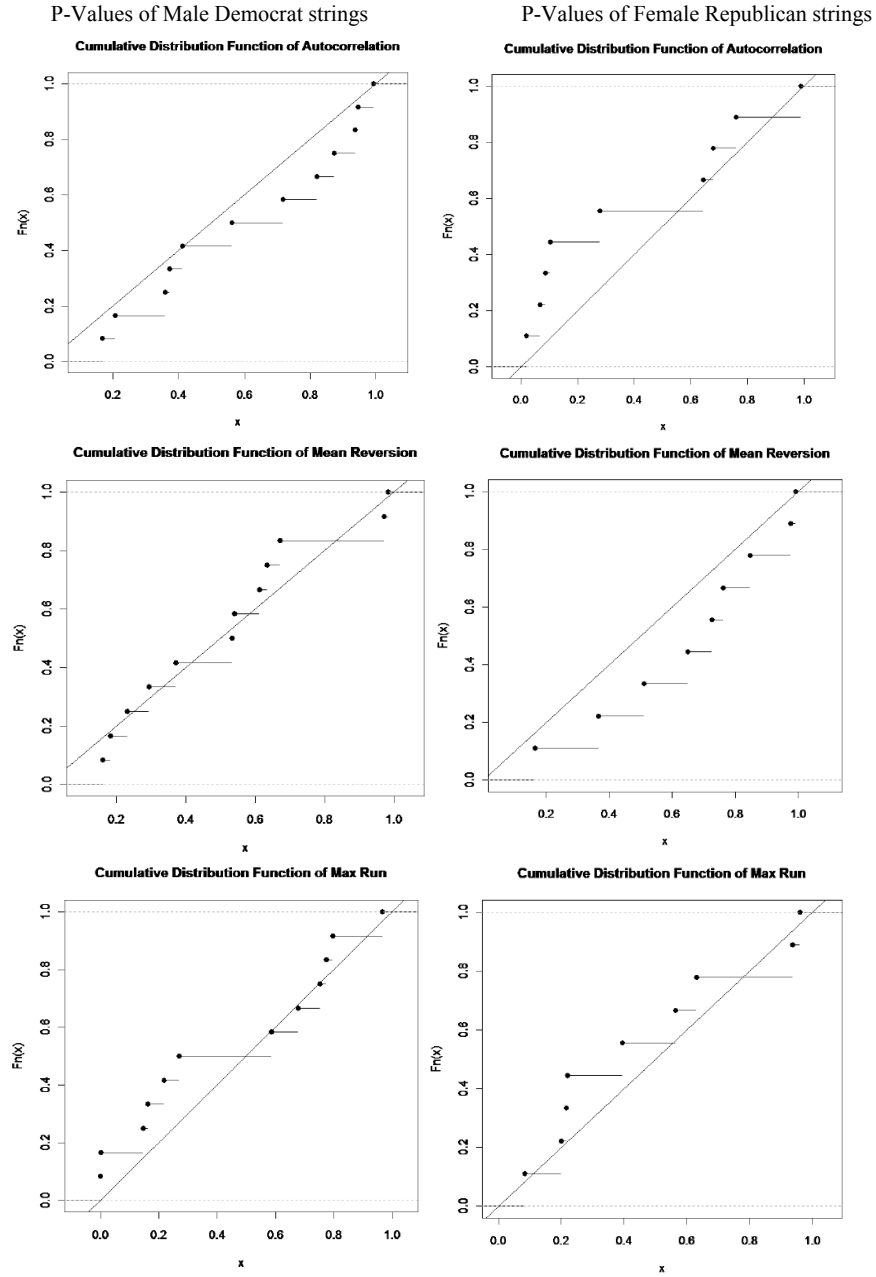
Notes: Each plot presents the cumulative distribution function of P-values for an autocorrelation test, mean reversion test, and longest run test of the sequence of judge assignments.

5

female Republicans to appear autocorrelated relative to a set of simulated strings.

C.2   Randomization for District Judges   District Courts assign one judge to a case randomly or rotationally (Taha 2009; Bird 1975).[34] For example, one District told us that random assignment occurs within 24 hours of a case filing, which is handled in the order of its arrival. Waldfogel (1995) reports that one District Court uses three separate randomization wheels and each wheel corresponds to the anticipated case length.[35] Related[36] cases, if filed within a few weeks, may be consolidated.[37] Consolidation only occurs for relatively high-frequency case types, which is not relevant to our study. For the handful of District cases that do overlap such that they are consolidated, we assume the decisions about case relatedness occur in a manner exogenous to judge assignment.

District Courts judges are revealed much earlier than Circuit Court judges. Ideally, we would use docket filings in the Administrative Office of the U.S. Courts pertaining to free speech cases, but judges are omitted for most cases prior to 2000, so we must use published District opinions to construct our District IV. Two facts support the assumption that settlement, publication, and

---

[34]Cases being returned on remand from the Circuit court are not randomly assigned. We do not use remanded cases in our dataset.

[35]The ideal construction of $\tilde{w}_{ct}$ takes a weighted sum across wheels of deviation from expectations, $E\left(\frac{L_{c\tilde{d}t}}{K_{c\tilde{d}t}}\right)$, separately for senior and non-senior judges. Senior judges can elect not to be assigned to certain wheels. Another District Court uses, instead of wheels, thirteen computer generated decks of cards–one deck for each case category and an identical number of cards (two or five) for each active judge (http://www.mnd.uscourts.gov/cmecf/Order-for-Assignment-of-Cases.pdf). The decks refill when the majority of the deck has been exhausted. Senior judges can request to be assigned to certain decks. Even within a deck, senior judges can ex ante request a "bye" for specialized case types. Within each District Court are several courthouses (also referred to as Divisions). The appropriate Division is determined by where the parties are located and where the cause of action arose. Some Divisions get their own deck of cards. Taha (2009) reports that in 29 Districts, a case may be assigned to any judge in that District, while in the others, the cases are assigned to a geographic Division within the District and randomly assigned to one of the judges in that Division.

However, since $E\left(\frac{L_{c\tilde{d}t}}{K_{c\tilde{d}t}}\right)$ is uncomputable for senior judges since we would need to know the senior "byes" in every District courthouse, we drop senior District judges for calculating $\tilde{w}_{ct}$; we also drop visiting (judges routinely visit other courts to assist with caseload) and magistrate judges (they assist District Court judges but do not have life tenure and we do not have their biographical data) for similar reasons, collectively resulting in less than 10% sample loss. Non-ideological cases are referred to magistrate judges (Nash 2015), so omitting them will not matter. Identification is unaffected by dropping judges even if they are in the same wheel. Some courts spin separate random wheels for District judges and for magistrate judges. In some Districts, parties can decline assignment to a magistrate judge within a certain time period and request another random draw. This will not affect identification because it happens before the random assignment that we use. In some Districts, when the federal government is a litigant on the case, the U.S. attorney can pick the wheel.

In sum, conditional on case type, there is random assignment at the court or courthouse level, and we must only calculate the yearly expected composition of judges in District courthouses, $E\left(\frac{L_{c\tilde{d}t}}{K_{c\tilde{d}t}}\right)$, and we drop judges whose $E\left(\frac{L_{c\tilde{d}t}}{K_{c\tilde{d}t}}\right)$ is unknowable.

[36]Related means that one decision will substantially resolve all cases.

[37]Waldfogel (1995) reports that plaintiffs can argue the case is related to another pending case and, if the judge agrees, the cases will be consolidated. A clerk reported 8% of filed cases were accepted as related in 1991 in SDNY. In another District Court, if a clerk identifies and two judges agree that a new civil case is related to another open civil case, they will be consolidated in the interests of justice or judicial economy. The clerk brings the possible connection to the attention of the judge of the new case, who then confers with the judge of the earlier case to determine whether they are in fact related cases.

strategic use of keywords or citations are exogenous. First, in District Courts, judges are much more constrained and ideology has been found to play hardly any role. Judicial ideology does not predict settlement rates (Ashenfelter et al. 1995; Nielsen et al. 2010), settlement fees (Fitzpatrick 2010), publication choice (Taha 2004), or decisions in published or unpublished cases (Keele et al. 2009).[38] Second, we examine these issues directly:[39] we test whether District Court judicial biographical characteristics in *filed* cases jointly predict publication.[40] We are able to conduct this test because we link PACER filing data, which has judge identity, to AOC data, which has information on publication.[41] We assume that remaining deviations from random assignment, like vacation days, are ignorable.

---

[38]This is consistent with the District judge identity only affecting outcomes through the presence of an appeal but not through the District Court decision.

[39]The random strings test is ineffective because some Districts use rotational assignment or random drawing of judges from card decks without replacement.

[40]We use LASSO to select biographical characteristics and no characteristic was chosen.

[41]We obtained all freely available PACER (Public Access to Court Electronic Records) data on District cases from 32 districts for 1980 to 2008 for a total of 359,595 non-duplicated cases. This data contains the name of the District where the case was filed, the filing and termination date (missing for 10% of cases), the assigned docket number, and the name of the District or magistrate judge presiding on the case. We merge the names of the judges into the Administrative Office of the U.S. Courts (AOC) database.