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? Using novel data on all workplace sexual harassment appellate precedent from 1982-2002 and randomly assigned judges, we find that pro-plaintiff sexual harassment precedent reduced gender inequality and spurred the adoption of sexual harassment human resources policies. The effects were comparable to the Equal Employment Opportunity Act's impact on black employment share, greatest in the heavily-litigated construction industry and where male sexism was highest, and explains one-third of the adoption of sexual harassment human resources policies. Our results are consistent with an insider-outsider theory of involuntary unemployment.

Keywords: Sexual harassment, Gender discrimination, Gender Inequality, Insider-Outsider Model

JEL codes: J71, J16, K38

*Daniel L. Chen, daniel.chen@iast.fr, Toulouse School of Economics, Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France; Jasmin Sethi, Sethi Clarity Advisers, jasminsethi1@gmail.com. First draft: July 2007. Current draft: October 2018. Latest version available at: http://nber.org/∼dlchen/papers/Insiders_Outsiders_and_Involuntary_Unemployment.pdf. We thank research assistants and numerous colleagues at several universities and conferences. Work on this project was conducted while Chen received financial support from the European Research Council (Grant No. 614708), Swiss National Science Foundation (Grant Nos. 100018-152678 and 106014-150820), Agence Nationale de la Recherche, Ewing Marion Kauffman Foundation, Petrie Flom Center, Harvard Law School Summer Academic Fellowship, and Templeton Foundation (Grant No. 22420). We also acknowledge joint financial support from the John M. Olin Center for Law, Economics, and Business at Harvard Law School.

1 Introduction

Does sexual harassment affect labor market choices? Despite an explosion of policy interest in association with recent events, the consequences of sexual harassment are understudied. Survey evidence suggests that lower job satisfaction results from sexual harassment (Newman et al. 2003) and that many women leave their jobs to escape the harassing environment (McLaughlin et al. 2017). Sexual harassment has received the attention of one economic model–through the lens of compensating differentials (Basu 2003; Hersch 2011)–yet it receives no mention in two recent summaries of the empirical literature on discrimination (Bertrand and Duflo 2016; Neumark 2016). On the one hand, an insider-outsider theory of involuntary unemployment suggests that forbidding harassment ameliorates gender inequality by opening job opportunities (Lindbeck and Snower 1988). Empirical support for the insider-outsider theory has thus far been cross-sectional (Lindbeck and Snower 2001) and lab-based (Fehr and Fischbacher 2002). On the other hand, past experience suggests that sexual harassment law could exacerbate gender inequality if the primary function of the 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, act as a tax on labor demand, and make women more costly to hire (Epstein 1995). We find causal, field evidence in support for the insider-outsider theory.

Much of U.S. policy surrounding labor market discrimination, and sexual harassment in particular, is carried out through its common law court system. Legal precedents include (i) *defining* what constitutes sexual harassment (e.g., retaliation against a female employee for rejecting sexual advances of her boss), (ii) *shifting standards* (e.g., from a reasonable person standard to a reasonable woman standard for what constitutes sexual harassment), (iii) *moving thresholds* (e.g., waiving the need for a plaintiff to prove emotional harm in court), or (iv) *imposing liabilities* (e.g., declaring that the firm can be subject to liability for hostile environment created by a supervisor). U.S. legal precedents thus have great scope for large impacts. To identify causal effects, the ideal experiment would randomize court decisions. Since doing so violates justice, our quasi-experiment leverages random assignment of judges, as their biographies predict sexual harassment rulings. We use the federal appellate courts whose rulings establish precedent for jurisdictions of 4-9 states. Federal appellate precedents comprise almost the totality of U.S. court-made law, since the Supreme Court hears less than 2% of appellate cases. We collect and analyze all sexual harassment appellate and lower court rulings from 1982-2002. We chose 1982 as the starting point, since the EEOC (Equal

Employment and Opportunity Commission) first recognized sexual harassment in 1980, and there were very few cases before 1982.

Judges are repeatedly randomly assigned to panels of three, and the composition of these panels varies by case. Prior research has documented that politics and gender predict sexual harassment decisions (Farhang and Wawro 2004; Boyd et al. 2010; Peresie 2005), but politics is more predictive (Asmussen 2011) as female conservatives may vote against female interests (Eisenman 1991). In our data, Republicans were 13 percentage points less likely than Democrats to vote in favor of sexual harassment plaintiffs,[1] but female Republicans were 18 percentage points less supportive of sexual harassment plaintiffs. We leverage biographical characteristics since each judge is assigned to only a handful of federal appellate sexual harassment cases. There are 251 cases and roughly 180 life-tenured judges. We also collect data on judicial biographies to leverage random variation in judicial decisions that may arise from random combinations of biographical characteristics.

Using quasi-experimental variation in legal precedent, we find that pro-plaintiff sexual harassment precedents increased female employment shares by 1.6 percent. This effect is similar to what the Equal Employment Opportunity Act achieved for black employment shares (Chay 1998). The effects were greatest in the U.S. South, which scores highest in male sexism,[2] and in the construction industry, which had the highest rate of sexual harassment claims. This suggests, on net, employers did not experience sexual harassment law as an unfunded mandated benefit. Moreover, consistent with an insider-outsider model of involuntary unemployment, forbidding harassment *increased* earnings per hour of female labor force participants. Pro-plaintiff sexual harassment also led to the adoption of sexual harassment human resources policies (Dobbin and Kelly 2007); the impacts explain one-third of its adoption since 1982.

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 harassment. In India and Mexico, female-only trains and buses were introduced so women would face less harassment on their way to work (women are the usual victims of harassment). 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. This paper sheds light on the impact of laws making it more difficult to sexually harass. Future research can investigate whether these findings extend to the impacts of safe spaces, trigger

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

[2] Charles et al. (2010) measure sexism in the General Social Survey as an aggregate of social attitudes towards women.

warnings, trauma, and micro-aggression, and the broader economic consequences of recognition and dignity.

2 Background

2.1 Conceptual Framework In the insider-outsider model of involuntary unemployment (Lindbeck and Snower 1988), outsiders are unable to find jobs even though they are prepared to work for less than the prevailing wages of incumbent workers (insiders). In brief, while harassment is allowed, the outsiders cannot underbid insiders; if they did and were to become 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 sexual harassment law 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 who had more aversion to harassment. This model assumes that firms do not forbid harassment on their own, an assumption that seems reasonable in light of the #MeToo movement.

Economic theory also suggests several arguments that makes this assumption reasonable. There may be labor turnover costs preventing firms from simply replacing all the insiders with outsiders. Managers may not have proper estimates of the link between harassment and productivity. Agency and transaction costs may also prevent Coasian bargaining between insiders and the firm. Then the firm has no incentive to implement the new contract that transfers a share of the additional profit resulting from the insider's cooperation in a manner that exceeds the losses in market power as the insider's wage falls towards his reservation wage. Further, this only happens if the firm relinquishes a share of gross profit, something that may make it a net loser compared to other firms. Entry of firms that hire the outsiders may not occur due to setup costs and capital market imperfections. However, a legal regime equalizes the playing field across all firms when no firm by itself would have the incentive to forbid harassment.

There are a few alternative theories on the potential effects of sexual harassment law. A broad class of economic models of anti-discrimination law posit that these laws act as an unfunded mandated benefit, making it more costly to hire the intended beneficiaries, lowering their wages or employment (Acemoglu and Angrist 2001; Autor et al. 2006; Epstein 1995; Gruber 1994). 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). However, sexual harassment law is different from an unfunded mandated benefit for two reasons. First, it may be difficult to know in advance who is going to be a sexual harassment plaintiff.[3] Second, men could bear some of the cost of the mandated benefit as potential harassers.[4] This broad class of mandated benefits models can be distinguished from the insider-outsider theory by the overall impact of sexual harassment law on gender inequality.

Other theories, such as a redistribution from females to males, or tax on the hiring of men, can be distinguished from the insider-outsider theory by examining the impact on insider women who had less aversion to harassment (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 theory 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.

2.2 Institution Our quasi-experiment leverages the U.S. common law system where decisions become precedent for future cases in the same jurisdiction. In the federal appellate courts (also known as Circuit Courts[5]), judges are randomly assigned. Moreover, decisions are binding precedent for future cases within the Circuit (there are 12 Circuits, each in charge of a geographic region comprising 4 to 9 U.S. states as seen in Figure 1; cases originate from one of the 94 District courts, numbering 1 to 4 per state). Less than 2% of Circuit cases reach the U.S. Supreme Court, so the Circuit decisions comprise the majority of precedents.

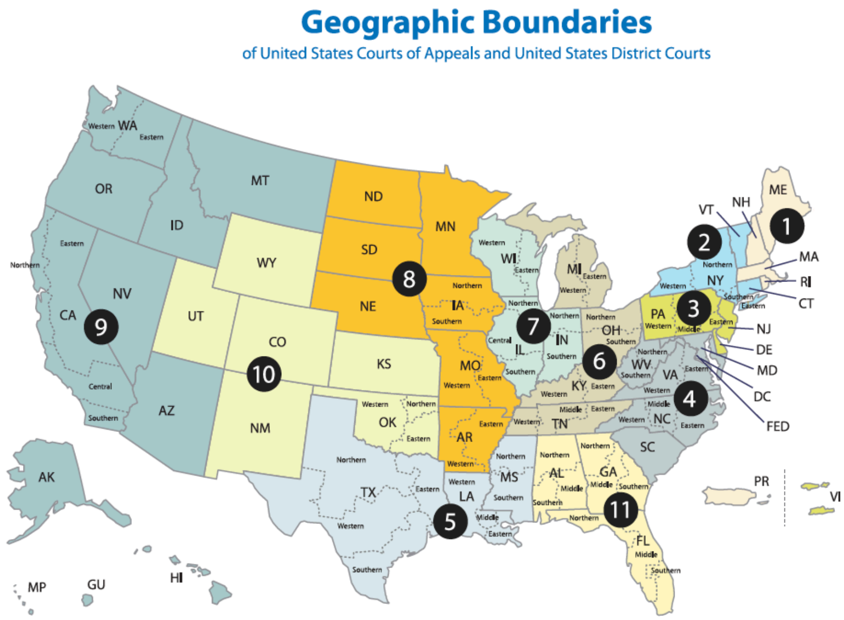
We use three datasets on cases. The first is from Boyd et al. (2010), which codes case char-

[3]For example, in Acemoglu and Angrist (2001)'s study of the 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.

[4]Unlike the cost of complying with the ADA or the federal requirement of providing maternity mandates. studied in Gruber (1994), 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.

[5]The name, "Circuit", refers to the fact that judges used to ride a horse in a circuit to reach the entire jurisdiction.

Figure 1.—



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

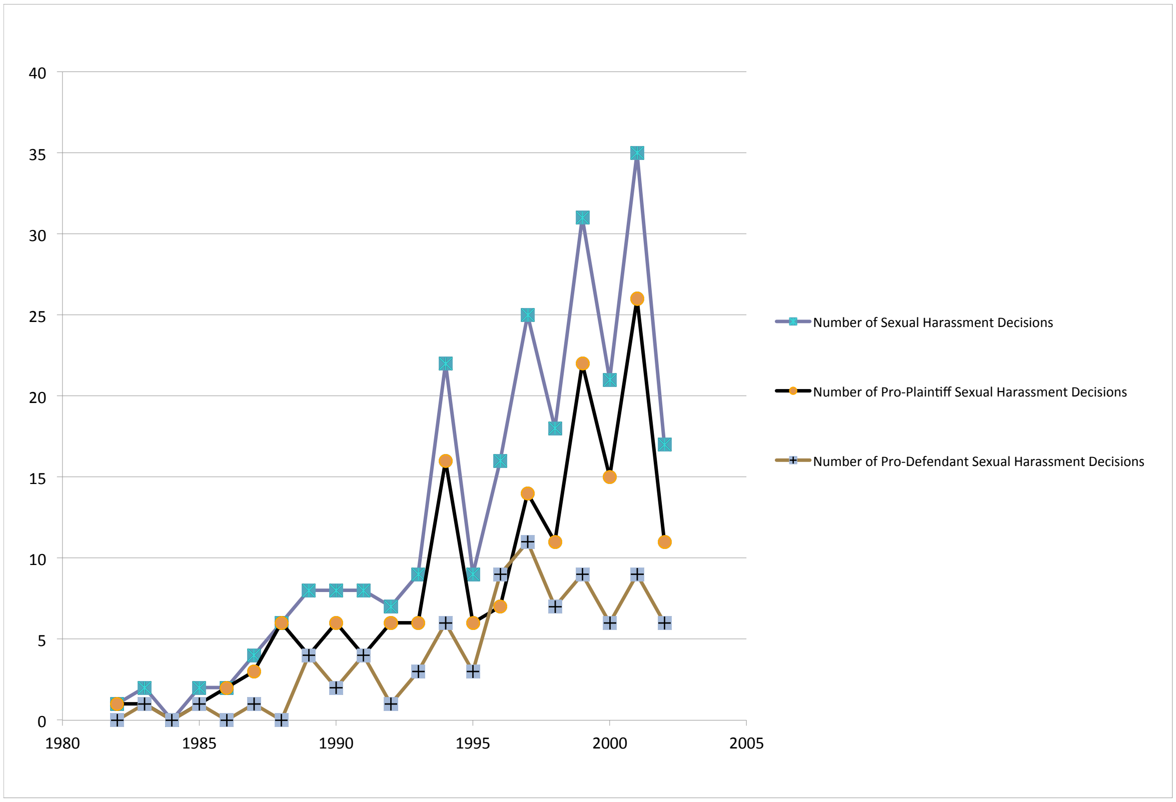
acteristics, such as the presence of certain fact patterns and legal issues for a subset of gender discrimination claims in the Chicago Judges Project data (Sunstein et al. 2006). We use this to perform a randomization check. Our second dataset is composed from our own collection of sexual harassment cases in Circuit and District Courts back to 1982. Prior to 1982, there were very few sexual harassment cases. Only in 1980, the EEOC (Equal Employment Opportunity Commission) first issued guidelines forbidding "sexual harassment" as a form of sex discrimination. We follow the data collection and coding method in the Chicago Judges Project; if the plaintiff was afforded any relief, then the vote was coded as pro-plaintiff.

Table I indicates that on average, there was roughly 1 sexual harassment precedent per Circuit-year for a total of 251. Appendix A presents an overview of key doctrinal developments. Roughly two-thirds of the decisions were pro-plaintiff. Our third dataset is all 3,754 sexual harassment cases between 1982 and 2002 in District Courts.[6] 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.[7] We use this data to perform a robustness check.

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

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

Figure 2.— Sexual Harassment Precedents, 1982-2002



Notes: X-axis is year and Y-axis is number of cases. This shows the plots the growth in sexual harassment cases, number of pro-plaintiff decisions, and number of pro-defendant decisions.

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	0.035	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.

Since the typical judge saw very few sexual harassment cases (there are roughly 180 life-tenured judges and 251 cases), we focus on shared biographical characteristics rather than use a judge-specific "leniency" measure that is used by other papers leveraging the random assignment of judges. For example, one-third of seats would be assigned a male Democrat in expectation, providing enough data to estimate attitudes towards sexual harassment. Data on judicial background characteristics comes from the Appeals Court Attribute Data, District Court Attribute Data,[8] Federal Judicial Center, and newspapers.[9] We filled in missing data by searching transcripts of Congressional confirmation hearings and other official or news publications on Lexis.

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, allowing us to isolate the causal effects.

3 Design of Study

3.1 Specification We use regressions of the form:

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

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 we vary the lag structure for robustness. We focus mainly on the average coefficient and joint significance of the lag coefficients. We also present the full set of individual coefficients for female labor share and human resources policies, as the temporal pattern is consistent with Dobbin and Kelly (2007)'s interviews with hundreds of firms indicating that labor lawyers and human resources consultants were quite effective in translating Circuit decisions into human resources policy changes. Millions of dollars were spent on training programs and establishing grievance procedures aimed to reduce the risk of lawsuit. The implementation of these programs may take a few years, which could then have downstream consequences on

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

[9]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.

female labor share.

- 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.
- $1[M_{ct-n} > 0]$ 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 $t-n$. We also consider state fixed effects, dropping 1 Circuit at a time, collapsing the data to the Circuit-year level, randomization inference, and clustering standard errors in different ways. Since random assignment is at the Circuit-year level (unlike differences-in-differences analyses of state law changes that turn "on" or "off" once), we expect to see similar results whether clustering standard errors at the Circuit or Circuit-year level.[10] Our final robustness check exploits variation in $1[M_{ct-n} > 0]$ (the decision to appeal) using the random assignment of District court judges. Since random assignment occurs at two levels of court hierarchy, we will be able to causally identify Law_{ct} in a distributed lag specification.

We describe our first stage more formally as follows. Let N_{ct} be a judicial characteristic. 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. Our moment condition for causal inference is: $\mathbf{E}[\frac{N_{ct}}{M_{ct}}\varepsilon_{ict}|\mathbf{E}(\frac{N_{ct}}{M_{ct}}), 1[M_{ct} > 0]] = 0$.[11] All lags and leads of Law_{ct} are instrumented. As standard, we interact the instruments with the female indicator as we interact Law_{ct}.

[10] Barrios et al. (2012) show that random assignment of treatment addresses serial and spatial correlation across 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. Our results are also unaffected if we cluster at the state level or use wild bootstrap. The coefficients on the leads serve as an omnibus falsification check for spurious significance.

[11] 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).

Our main identification strategy is illustrated in Figures 3A and 3B. 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.

To check for randomization, we regress case characteristics (from Boyd et al. (2010)) on male Democratic (female Republican) judges per seat and find that most characteristics are not correlated with the judicial panel composition. Table II shows that of 19 case characteristics, one is correlated with male Democrats per seat and one is correlated with female Republicans per seat, significant at the 10% level. Appendix B presents 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). We weight our analysis with CPS-provided weights.

- Employment status: We create an indicator to distinguish between no-employment (including non-labor force participants) vs. part- or full-time employment.[12] Non-labor force participants include discouraged workers. These individuals are 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.[13]
- 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.

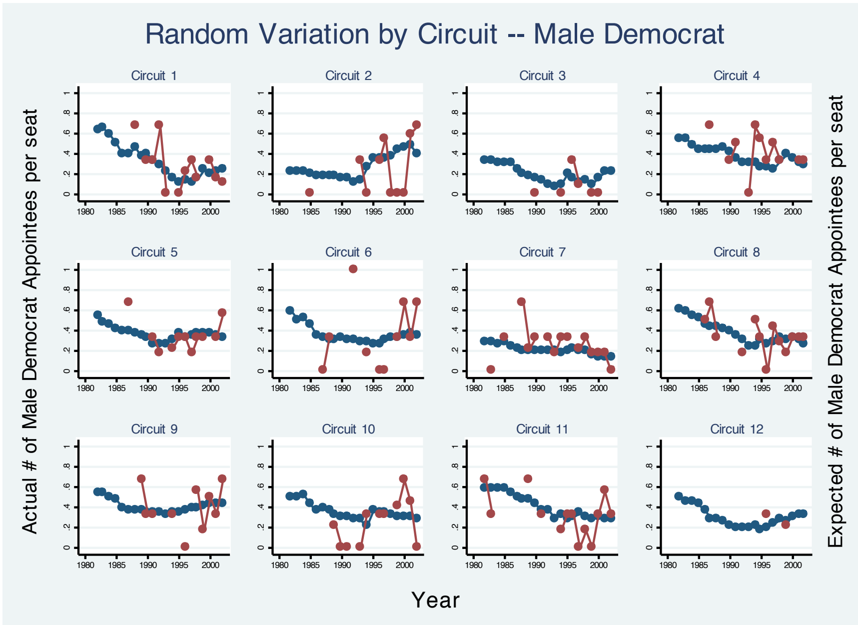
Some studies drop non-labor force participants and find smaller differences in labor market outcomes between men and women. When we analyze insiders, we follow the same restriction.

[12] 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."

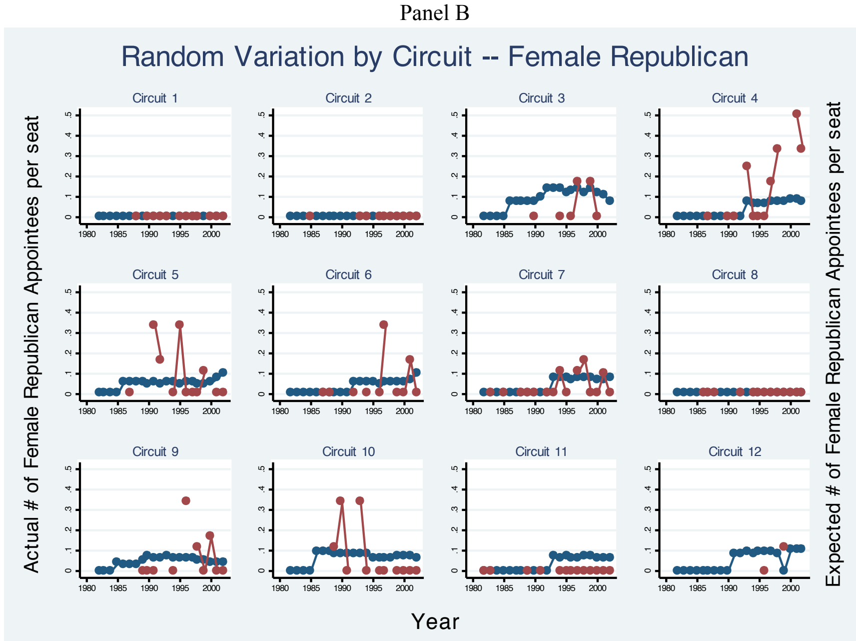
[13] Occupation is available for about 90% of the unemployed and 33% of those not in the labor force, about 10% of which are managerial. Respondents may interpret this question as being about their previous job.

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.

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).

To study the impact on human resources policies, we use a dataset compiled by Dobbin and Kelly (2007). We measure the presence of firm-level sexual harassment policies (adoption of sexual harassment grievance procedures and policies) in a national sample of 389 workplaces interviewed in 1997 on the history of human resources practices dating back to 1965. 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.[14] 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

What determines judicial decisions? We would like to believe it is "the law". To be sure, the law may be hard to determine or even indeterminate. The identity of the judge can thus make a large difference. This simple fact was statistically established at least a century ago (Everson 1919). Inter-judge differences show that the meaning of "the law" is not unique in practice, but they are consistent with each judge consistently applying his or her version of the law.

When it comes to gender issues, Dixon (2010) reports that judges vote more along party lines than along gender lines. This may be due to the fact that presidents who appoint women candidates take the opportunity to appoint more politically extreme individuals (Asmussen 2011) or due to female conservatives exhibiting prejudice against females (Eisenman 1991). In our dataset of sexual harassment precedents, Republicans were 13 percentage points less likely to vote in favor of sexual harassment plaintiffs, but female Republicans were 18 percentage points less likely to vote pro-plaintiff. Male Democrats were 13 percentage points more likely to vote pro-plaintiff. 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 in our dataset. We use "male Democrat" as an instrumental variable to mirror "female Republican" as instrumental variable. Table III Column 1 includes both characteristics together.

Column 2 examines the case-level, where there are three judges to a panel. We use a per-seat measure ($\frac{N_{ct}}{M_{ct}}$), so coefficients need to be divided by three to interpret the effect of an additional

[14] 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.

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, respspectively, human resources policies.

male Democrat on a three-judge panel, which is roughly the same as the effect reported in Column 1. Column 3 reports estimates at the Circuit-year level.[15] Column 4 presents the first-stage for the individual-level CPS analysis. Column 5 presents the first-stage for the firm analysis.[16] In each of the last three columns, the F-statistics are strong.

An unusual feature of U.S. federal appellate courts is that, technically, cases should only appear in the appellate courts if they present new legal issues (a matter of doctrinal interpretation). Cases with identical fact patterns should not be appealed. Therefore, we can present another check of our identification strategy: We should not expect the assignment of judges in a previous year to predict the decisions in a subsequent year. Table IV 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 result assures us that our instrument is not picking up general societal trends correlated with the composition of judicial panels and the outcomes of cases. Furthermore, since each instrument is affecting the corresponding contemporaneous endogenous variable, we will be isolating the causal effects of Law_{ct} in a distributed lag specification.

[15] Coefficients at the Circuit-year level can differ from the case level due to uneven bunching across Circuit-years. The R-square increases due to the indicator for whether there were cases, $1\,[M_{ct} > 0]$.

[16] Coefficients differ since the number of firms per Circuit-year is not constant.

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.

To assuage concerns that the previous literature has idiosyncratically focused on gender and political party of appointment of judges, we also employed LASSO to select biographical features as instruments for Law_{ct} (Belloni et al. 2012) and the results are similar. The joint F statistic increases to 33 for the Circuit-year level and up to 130 for the analysis-level.[17] The use of the LASSO-selected instruments provides a check of over-identification that causal effects of Law_{ct} remain similar regardless of whose tendencies to vote in favor of sexual harassment plaintiffs is affecting the decision.

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). This 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, 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. We focus on β_3, the impact on female employment shares. We find the results are essentially unchanged (the point estimate of 1.6 percentage points in Table V) 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 on the effects of the previous five years of sexual harassment laws increases when we do not use CPS weights in row F or vary the lag specification in row G. The estimates change little when we drop one Circuit at a time in row H.

[17] The thirty biographical characteristics we collected are: Democrat, male, male Democrat, female Republican, non-White, Black, Jewish, 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. 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 V

The Effect of Sexual Harassment Law on Gender Inequality

	Employment Status		Hours Worked		Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Pro-Plaintiff * Female	0.007	0.016	0.196	0.508	0.051	0.113
Joint F		8.53		5.42		7.60
Proportion Pro-Plaintiff	-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.

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 to select instruments in row K, or use District court judge assignment to instrument for $1\left[M_{ct-n}>0\right]$ in row L. We also employ Monte Carlo placebo simulations 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 90^{th} 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

Sexual harassment is very difficult to measure. In one study, Banerjee et al. (2012) find that only 50% of sexual harassment cases are registered by the police. Changing social mores can also make it difficult to measure whether a change in registered sexual harassment complaints is due to an increasing willingness to report, increasing sensitivity to harassment, or increasing harassment. Therefore, it is challenging to directly measure the impact of sexual harassment law on incidence of sexual harassment.[18]

6.1 Disaggregating the Effects of Sexual Harassment Law by Region We can, however, disaggregate the effects of sexual harassment law and evaluate its effects by region. In Table VI, we see that dropping Circuit 5, renders the smallest estimates, meaning the biggest effects are found in Circuit 5, which includes Texas, Louisiana, and Mississippi. This last finding is interesting because it parallels Charles et al. (2010)'s measure of male sexism that varies by region. Using the General Social Survey, they found that the two census regions scoring highest in male sexism are East South Central and West South Central, which includes Circuit 5.[19] 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 census regions and where they found male sexism scores to be low. According to the theory, the empirical patterns should be attenuated in regions where males exhibit less sexism.

6.2 Disaggregating the Effects of Sexual Harassment Law by Industry Next, we evaluate its effects specifically in an industry with a reputedly high amount of sexual harassment–the construction industry. Sexual harassment rates per 100,000 women were highest in the construction industry according to complaints filed with the Equal Employment Opportunity Commission (Her-

[18]For example, the number of sexual harassment claims filed increased after the 1991 Clarence Thomas hearings.
[19]Unfortunately, the census region boundaries do not coincide with the Circuit boundaries.

TABLE VI

Robustness of Effect of Sexual Harassment Law on Female Employment Share

	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

	Employment Status		Hours Worked		Earnings	
Construction Industry	Yes	No	Yes	No	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Pro-Plaintiff * Female	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
Proportion Pro-Plaintiff	-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.

sch 2011). The bulk of sexual harassment plaintiffs (38%) are in the blue-collar industry (Juliano and Schwab 2000).

Table VII documents the ameliorative effects on inequality are positive and statistically significant for employment status, hours worked, and earnings in the construction industry but negative for the other industries. Note that we restrict the analysis to workers who report an industry, so the analysis necessarily relies on a less complete picture of the entire population, since labor force participants were more likely to report their industry. This finding is also noteworthy as it suggests that the estimated impacts of sexual harassment law are not simply capturing a general time trend in female labor market outcomes.

6.3 Insiders and Outsiders Thus far, we have shown positive effects of sexual harassment law on female employment outcomes on average. We now turn to some evidence for the insider-outsider theory of harassment and involuntary unemployment. If some women are insiders, then forbidding sexual harassment should result in less ameliorative impacts for insider women. We do not have a panel of individuals followed before and after legal changes, so in this analysis, we must make the albeit rough approximation that labor force participants are insiders and non-participants 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 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:[20]

- $(0.35 - 0.003) * 0 + (0.003) * 5.9 + (0.65) * 0.004 = 0.020$, 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.

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

[20]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.

TABLE VIII

The Effect of Sexual Harassment Law on Insiders

	Labor Force Participants Reporting Non-Zero Wages		
	Earnings	Hours Worked	Management
	(1)	(2)	(3)
Proportion Pro-Plaintiff * Female	-0.004	-0.160	-0.007
Joint F	35.16	58.21	30.38
Proportion Pro-Plaintiff	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.

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

	Presence of Sexual Harassment Policy	
	(1)	(2)
Proportion Pro-Plaintiff	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.

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.[21]

Temporal patterns in the effects of sexual harassment law are consistent with a role for human resources policies. We present the set of coefficients on the effects on female labor share in Table X and visualize the coefficients and 95% confidence intervals in Figure 4. The strongest effects on

[21] 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.

human resource policies are observed three years after the decision, while the strongest effects on female labor share occur five years later.

To be sure, the adoption of human resources policies is only one channel through which sexual harassment law can have labor market consequences. Women could be encouraged to enter the work force even in the absence of firm policy changes if they felt safer.

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 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.[22]

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 (Appendix C 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.[23] 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, $\mathbf{1}[M_{ct} > 0]$.[24] 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 Appendix C.

Instrumenting with randomly assigned judges in both Circuit and District courts, we found that

[22] 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.

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

[24] 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.

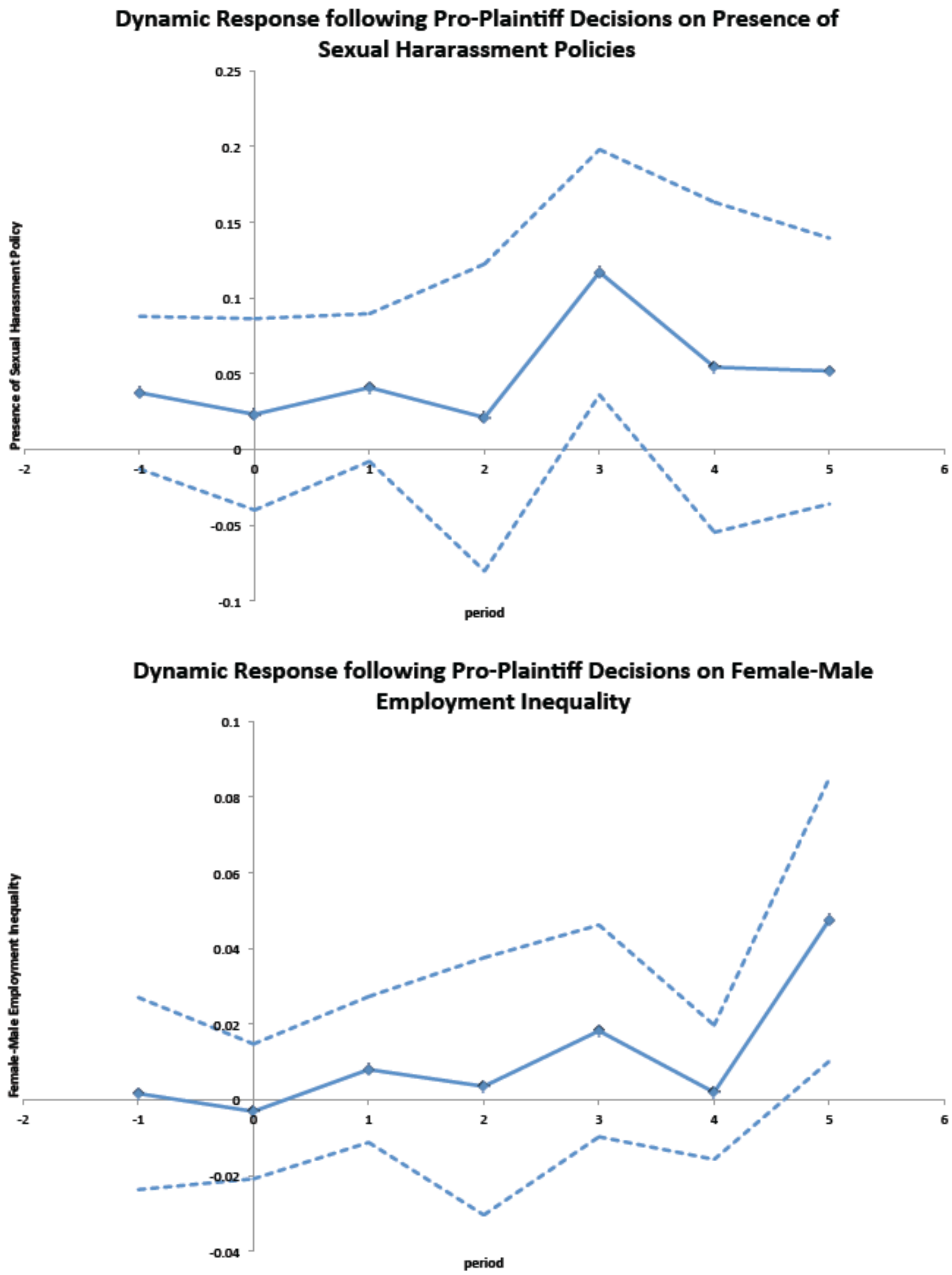
TABLE X

The Effect of Sexual Harassment Law on Gender Inequality

	Employment Status (1)	(2)	Hours Worked (3)	(4)	Earnings (5)	(6)
Proportion Pro-Plaintiff	-0.000709	0.00549	0.0690	0.527	0.00981	0.0717
Circuit Decisions$_{t+1}$	(0.00308)	(0.00876)	(0.147)	(0.466)	(0.0251)	(0.0549)
Proportion Pro-Plaintiff	0.00208	0.00163	-0.152	-0.360	-0.00837	-0.0248
Circuit Decisions$_{t+1}$ * Female	(0.00454)	(0.0127)	(0.136)	(0.595)	(0.0285)	(0.0667)
Proportion Pro-Plaintiff	-0.00119	0.00336	0.0655	0.147	0.000603	0.0366
Circuit Decisions$_t$	(0.00300)	(0.00718)	(0.149)	(0.334)	(0.0217)	(0.0424)
Proportion Pro-Plaintiff	0.00207	-0.00309	-0.188	-0.552	-0.0000162	-0.0398
Circuit Decisions$_t$ * Female	(0.00527)	(0.00891)	(0.196)	(0.438)	(0.0284)	(0.0532)
Proportion Pro-Plaintiff	-0.00483	-0.00609	-0.170	-0.300	-0.0234	-0.0166
Circuit Decisions$_{t-1}$	(0.00361)	(0.00694)	(0.152)	(0.405)	(0.0248)	(0.0424)
Proportion Pro-Plaintiff	0.00701	0.00798	0.135	0.105	0.0378	0.0176
Circuit Decisions$_{t-1}$ * Female	(0.00685)	(0.00964)	(0.245)	(0.373)	(0.0382)	(0.0501)
Proportion Pro-Plaintiff	-0.00649	-0.0117	-0.303	-0.739	-0.0456	-0.0912
Circuit Decisions$_{t-2}$	(0.00367)	(0.0135)	(0.176)	(0.681)	(0.0265)	(0.0967)
Proportion Pro-Plaintiff	0.00631	0.00353	0.258	0.539	0.0490	0.0653
Circuit Decisions$_{t-2}$ * Female	(0.00509)	(0.0170)	(0.216)	(0.868)	(0.0293)	(0.113)
Proportion Pro-Plaintiff	-0.00657	-0.0147	-0.293	-0.273	-0.0520	-0.0992
Circuit Decisions$_{t-3}$	(0.00440)	(0.0106)	(0.209)	(0.604)	(0.0322)	(0.0902)
Proportion Pro-Plaintiff	0.00476	0.0182	0.214	0.522	0.0506	0.142
Circuit Decisions$_{t-3}$ * Female	(0.00613)	(0.0140)	(0.278)	(0.646)	(0.0397)	(0.109)
Proportion Pro-Plaintiff	-0.00726	-0.00640	-0.202	-0.117	-0.0463	-0.0552
Circuit Decisions$_{t-4}$	(0.00487)	(0.00513)	(0.216)	(0.259)	(0.0344)	(0.0500)
Proportion Pro-Plaintiff	0.00771	0.00195	0.124	-0.156	0.0559	0.0466
Circuit Decisions$_{t-4}$ * Female	(0.00721)	(0.00887)	(0.229)	(0.354)	(0.0423)	(0.0632)
Proportion Pro-Plaintiff	-0.00694+	-0.0284**	-0.133	-0.723+	-0.0460	-0.166*
Circuit Decisions$_{t-5}$	(0.00373)	(0.0105)	(0.164)	(0.427)	(0.0306)	(0.0689)
Proportion Pro-Plaintiff	0.00803	0.0476*	0.249	1.531+	0.0615	0.293**
Circuit Decisions$_{t-5}$ * Female	(0.00626)	(0.0187)	(0.210)	(0.799)	(0.0369)	(0.113)
Controls	Y	Y	Y	Y	Y	Y
IV	N	Y	N	Y	N	Y
Mean dependant variable - Male	0.813	0.813	34.33	34.33	4.910	4.910
Mean dependant variable - Female	0.646	0.646	22.78	22.78	3.654	3.654
Average interaction lag	0.007	0.016	0.196	0.508	0.051	0.113
Average level effect lag	-0.006	-0.013	-0.220	-0.430	-0.043	-0.086
Joint F of interaction lags		8.53		5.42		7.60
Joint F of interaction leads		0.02		0.37		0.14
Joint F of level effect lags		28.11		5.07		10.86
Joint F of level effect leads		0.39		1.28		1.70
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 and restrict to individuals between the ages of 18 and 65. Heteroskedasticity-robust standard errors are in parentheses. Observations are clustered at the Circuit level. Hours and Log earnings set to 0 for individuals not employed. All estimates are weighted using CPS survey weights. These regressions include individual controls (age, race dummies, educational attainment dummies, and marital status dummy), Circuit fixed-effects, year fixed-effects, a dummy for whether there were no cases in that Circuit-year. The regressions also include interactions between the female indicator and the dummy for there being no cases, sexual harassment law, and the instruments.

Figure 4.— Distributed Lag



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. This result suggests 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. The fact that female labor market outcomes improve more than insider male labor market outcomes decline is consistent with the idea that there was some degree of rent capture by insiders and suggests that firms were not profit-maximizing and forbidding harassment on their own in the absence of sexual harassment law. It is also consistent with the persistence of beliefs that women should play only certain roles, or should know their "place" (Charles et al. 2010), and that this persistence accounts for a substantial portion of inequality.

Interpreting anti-discrimination law to forbid sexual harassment has been a key contribution of academic scholarship in employment law (MacKinnon 1979). 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.

We identify the impact of court-made 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 sexual harassment law does not appear to exacerbate gender inequality. Pro-plaintiff precedent increases female wages and employment relative to that of men. However, when restricted to people previously in the work force, sexual harassment law has less ameliorative effects for females. These findings are more consistent with an insider-outsider theory of involuntary unemployment–where insiders harass outsiders in order to capture economic rents and forbidding harassment increases entry of outsiders, raising their employment and then wages–than with compensating wage differentials or mandated benefits models of sexual harassment.

Methodologically, the empirical framework developed here provides causal estimates of court precedent holding all else equal including unobserved factors. It overcomes the basic issues of omitted variables and reverse causality. Furthermore, it has the advantages that the exclusion restriction is likely to hold, the LATE interpretation of the IV estimates are policy relevant (difficult cases without strong legal precedent are the ones where judges seek guidance (Posner 1998; Breyer 2006)), the general equilibrium effects are those which we would want to include (allowing for factor migration across Circuit boundaries), and the impulse response function in distributed lag is well-identified. 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 behavioral responses to the law, exploring heterogeneity in terms of early or late decisions, or exploiting variation in the sequence of decisions.

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.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh, 2012, Can Institutions be Reformed from Within? Evidence from a Randomized Experiment with the Rajasthan Police, NBER Working Papers 17912, National Bureau of Economic Research.

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, Alex, Daniel L. Chen, Victor Chernozhukov, and Chris 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.

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.

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

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.

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.

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).

Everson, George, 1919, The Human Element in Judging, *Journal of the American Institute of Criminal Law and Criminology* 10, 90–99.

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.

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.

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 Discrimination*, Yale Fastback Series (Yale University Press).

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.

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).

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.

For Online Publication

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.

B Randomization Checks

B.1 Interviews 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 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).[25]

Other variations from random assignment include: en banc cases that are heard by the entire

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

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.

B.2 Random Strings Test We also examine deviations from random assignment by seeing whether the sequence of proportions of judges is similar to a random process. The figures in Section C.1 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

APPENDIX TABLE I

Randomization Check: P-values

	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

$n = 9$		$n = 12$	
Prob($\sqrt{n}D_n < b$)	$b/\sqrt{n}$	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.

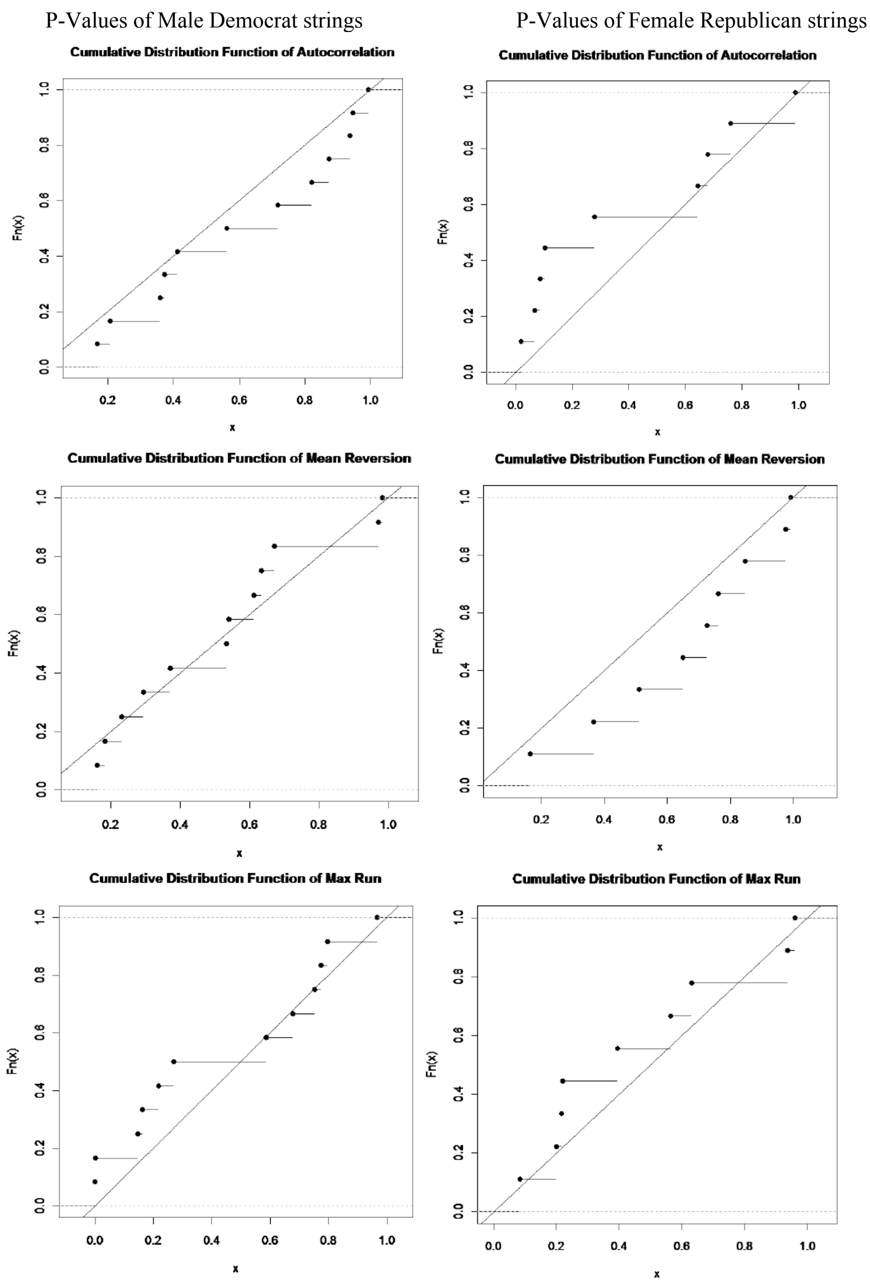
tells us whether judges or their assignors were attempting to equilibrate their presence, considering whether a judge was "due" for a sexual harassment case.

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.

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 1001$^{\text{th}}$ 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 (Appendix Figure 4), which we formally test using a Kolmogorov-Smirnov test statistic from an empirical null (Appendix Table I). 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.

Figure 5.— 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.

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.[26] Second, unpublished cases do not have precedential authority.[27] 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 female Republicans to appear autocorrelated relative to a set of simulated strings.

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

[27]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).

C District Courts

C.1 Randomization for District Judges District Courts assign one judge to a case randomly or rotationally (Taha 2009; Bird 1975).[28] 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.[29] Related[30] cases, if filed within a few weeks, may be consolidated.[31] 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 these 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 strategic use

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

[29] 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.

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

[31] 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.

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).[32] Second, we examine these issues directly:[33] we test whether District Court judicial biographical characteristics in *filed* cases jointly predict publication.[34] 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.[35] We assume that remaining deviations from random assignment, like vacation days, are ignorable.

C.2 District Judge Assignment as 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 II. 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%.

Columns 2 and 3 of Appendix Table II 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.[36] 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.

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

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

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

[35] 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.

[36] 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.

APPENDIX TABLE II

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

	$1\left[M_{ct}=0\right]$			
	(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.