Insiders and Outsiders: Does Forbidding Sexual Harassment Exacerbate Gender Inequality?

Daniel L. Chen and Jasmin Sethi[1]

November 2011

ABSTRACT

Sexual harassment is perceived to be a major impediment to female labor force participation. Using novel data on all workplace sexual harassment precedent in US Circuit Courts from 1982-2002, we exploit the random assignment of US appellate judges and the fact that a judge's gender and party of appointment predict decisions in sexual harassment cases to demonstrate the causal impact of pro-plaintiff sexual harassment precedent on the adoption of sexual harassment human resources policies and reduction in gender inequality. Consistent with an insider-outsider model of involuntary unemployment, forbidding sexual harassment encouraged entry of outsiders and reduced gender inequality along the dimensions of quantity and price, particularly in the construction industry, which was heavily affected by sexual harassment litigation, but these ameliorative effects are reduced for insider women. Pro-plaintiff decisions spurred the adoption of sexual harassment human resources policies by 4.8 percentage points and increased female employment shares by 0.4 percentage points.

I. Introduction

Sexual harassment is perceived to be a major impediment to female labor force participation, and policymakers in both developed and developing countries have taken steps to address this problem. For example, in India and Mexico, policy-makers have introduced female-only trains and buses so that women would face less harassment on their way to work.[2] In the US, making the work environment friendlier to women has been one of the most dramatic labor market changes in the past half-century. Yet we still do not know the consequences of these developments for female labor force outcomes. Forbidding sexual harassment could have ameliorated gender inequality by opening up job opportunities in previously harassing work environments, as suggested by an insider-outsider model of involuntary unemployment (Lindbeck and Snower 2001), or it could have exacerbated gender inequality if its primary function is to mandate a benefit, imposing costs through lower wages or lower employment of the targeted group (Gruber 1994).

Since forbidding sexual harassment may have been a consequence of the increase in female labor force participation, establishing a causal relationship between sexual harassment law and labor market outcomes is challenging. Social scientists have long speculated on the relationship between court-made law and socio-economic conditions and scholarship to date has been unsatisfactory in exploring questions of causality vis-à-vis court rulings. Robust empirical methods to evaluate the impact of judicial decisions may help judges who are interested in the broader empirical consequences of their decisions. Several notable judges have recognized the importance of considering empirical consequences in their adjudications. Judge Richard Posner has lamented that, "[judicial] opinions lack the empirical support that is crucial to sound constitutional adjudication,"[3] and similarly Justice Breyer remarked, "I believe that a[n] interpretive approach that undervalues consequences, by undervaluing related constitutional

---

[2] http://www.nytimes.com/2009/09/16/world/asia/16ladies.html, and
http://www.nytimes.com/2008/02/11/world/americas/11mexico.html
[3] "I would like to see an entirely different kind of constitutional theorizing. … Above all, what are the actual and likely effects of particular decisions and doctrines?" (Posner, Richard. 1998. "Against Constitutional Theory." *NYU Law Review* 73 (April): 1-22.).

objectives, exacts a constitutional price that is too high."[4]

We use a natural experiment provided by the random assignment of appellate judges to three-judge panels and by the fact that a judge's gender and party of appointment predict decisions in sexual harassment cases to estimate the causal impact of forbidding sexual harassment on gender inequality. Sexual harassment law is primarily court-made. Between 1982 and 2002, over 250 appellate cases addressing sexual harassment were decided in the United States. Our empirical strategy exploits the fact that judges are randomly assigned to their appellate cases within each circuit and the fact that Democratic appointees, particularly male Democratic appointees, are more likely to decide in favor of sexual harassment plaintiffs as compared to other judges and Republican appointees, particularly female Republican appointees, are more likely to decide in favor of sexual harassment defendants as compared to other judges. Because judicial composition of sexual harassment appellate panels is unlikely to be correlated with subsequent labor market outcomes other than through sexual harassment decisions, the random assignment of female Republican appointees and male Democratic appointees to three-judge panels creates exogenous variation in appellate precedent forbidding sexual harassment that can be used to estimate the causal impact of court-made sexual harassment law on gender inequality.

In the following sections, we present an analysis of data from the Current Population Survey and data on sexual harassment appellate decisions collected by the authors as well as by other authors. Section 2 discusses practical and theoretical aspects of sexual harassment law, emphasizing how labor lawyers and human resources consultants greatly exaggerated the risk of sexual harassment lawsuits after major appellate decisions to motivate employers to make human resources policy changes that forbid sexual harassment. Insiders would no longer be able to harass outsiders in order to capture economic rents. Outsider women would enter the labor force since they can now

---

[4] "Why should courts try to answer difficult … questions on the basis of logical deduction from text or precedent alone? Why not ask about the consequences of decision-making … I think a focus on consequences will itself constrain subjectivity." (Breyer, Stephen. 2004. Active Liberty: Interpreting Our Democratic Constitution. The Tanner Lecture on Human Values delivered at Harvard University).

compete for jobs previously dominated by insider men and women who tolerated sexual harassment.

Section 3 establishes that composition of the judicial panel is indeed related to sexual harassment appellate decisions. We first assess randomization by examining whether judicial panel characteristics are correlated with case characteristics as determined by lower courts and determining how similar the string of judicial panel assignments is to a random string. We then document that Republican appointees, particularly female Republican appointees, are more likely to vote in favor of sexual harassment *defendants* while Democratic appointees, particularly male Democratic appointees, are more likely to vote in favor of sexual harassment *plaintiffs*. Thus, the less frequently that female Republican appointees are assigned and the more frequently that male Democratic appointees are assigned, the more pro-plaintiff will be the aggregate sexual harassment precedent, making it easier for subsequent sexual harassment plaintiffs to bring and win suit.

A variety of evidence presented in Section 3 establishes that the random assignment of judges differentially impacts appellate decisions creating precedent in sexual harassment law. In Section 4, we consider whether firms subsequently adopt sexual harassment human resources policies and whether females see improvements in labor market outcomes relative to males in circuits and years with more precedent favorable to sexual harassment plaintiffs. Two-stage least squares estimates using this variation suggest that forbidding sexual harassment spurs adoption of human resources policies to address sexual harassment in the first three years after the pro-plaintiff precedent and increase employment status, hours worked, and earnings for females by the fifth year. Employment status, hours worked, and earnings decrease, to a lesser extent, for males. Our baseline estimates indicate that pro-plaintiff decisions spurred the adoption of sexual harassment human resources policies by 4.8 percentage points and increased female employment shares by 0.4 percentage points. We explore the robustness of our empirical design by varying our controls and data in a variety of specifications. Social outcomes are also not related to appellate decisions before they are

made. We take our results as evidence that forbidding sexual harassment has real economic consequences, which is inconsistent with theories of legal precedent that considers court-made law only as a product as opposed to producer of societal trends.

In Section 5, we present the evidence that is consistent with an insider-outsider model of harassment and involuntary unemployment: forbidding sexual harassment encouraged the entry of outsiders and had less ameliorative effects for insider women in the labor force. When the population sample is restricted to those in the labor force who report non-zero wages, the effects of sexual harassment law on gender inequality are negative on hours worked, earnings, and management status. In Section 6, we further explore heterogeneous treatment effects. The ameliorative effect of sexual harassment law is much stronger in the construction industry, which had the highest rates of sexual harassment claims.

Observing the insider-outsider model of involuntary unemployment empirically is difficult. Previous tests of the insider-outsider model use cross-sectional analyses or focus on the correlation between firm productivity, labor turnover costs, wages and lay-offs rather than on the causal consequences of requiring insiders to stop harassing outsiders.[5] Our findings support the insider-outsider model of harassment and involuntary unemployment and when combined with the existing survey evidence indicating higher job satisfaction resulting from the presence of sexual harassment law (Baker, Jackson, and Newman 2003), suggest that other prohibitions on identity-based (Akerlof and Kranton 2000) harassment may be welfare-improving for targeted groups with respect to their economic consequences. While other studies have employed the random assignment of judges to cases to identify the impact of judicial decisions on outcomes of individuals or firms in litigated cases (Kling 2006; Chang and Schoar 2007), a methodological innovation of this paper is using the random assignment of appellate judges to identify the causal impact of law on economy-wide outcomes. We further show how the random assignment of district court judges to instrument for the presence of appellate cases and the application of LASSO techniques (Belloni, Chen, Chernozhukov,

---

[5] For a summary of the empirical research, see Lindbeck and Snower (2001).

Hansen 2011) to select among a combinatorial possible number of instrumental variables aid our endeavor. This paper also contributes to debate on how much of gender inequality in labor markets is due to unobserved physiological differences, labor market choices, or discrimination (Summers 2005) by suggesting that some labor market choices, e.g. labor market participation, can be due to discrimination.[6]

II. Background

At a practical level, sexual harassment law, which is primarily court-made, and the imposition of direct and indirect litigation costs has impacted firm behavior in at least four ways. First, for those cases that were actually litigated, the law resulted in direct litigation costs and potentially large damage awards. Second, the potential for litigation likely led many firms to settle with potential plaintiffs, particularly after such plaintiffs filed complaints with the Equal Employment Opportunity Commission (EEOC) and received a right to sue letter from the agency. Third, many firms invested in formal grievance procedures, in part because of the advice of personnel experts who believed that instituting formal grievance procedures analogous to those for civil rights violations would help defend against damage awards (Dobbin and Kelly 2007). Fourth, some firms mandated training akin to the diversity training implemented in response to the Civil Rights Act of 1964 to educate their workers about sexual harassment. This last approach was controversial since some firms and attorneys feared that such training might make potential plaintiffs more aware of harassment and, therefore, more likely to sue (Dobbin and Kelly 2007).

Sociologists and legal ethnographers (e.g., Dobbin and Kelly 2007, Edelman 1992, 2002, Bisom-Rapp 2001a, 2001b, Edelman, Abraham, and Erlanger 1992, Edelman, Uggen, and Erlanger 1999) have documented how labor lawyers and human resources consultants greatly exaggerated the risk of sexual harassment suit subsequent to

---

[6] This paper also contributes to the literature on the gender wage gap (Blau and Kahn 2006; Card and DiNardo 2002; Weinberger and Kuhn 2006; O'Neill 2003; Black and Strahan 2001), anti-discrimination law (Basu 2003; Hellerstein, Neumark, and Troske 1997; Neumark and Stock 2006; Beller 1979; Eberts and Stone 1985; Acemoglu and Angrist 2001; Jolls and Prescott 2004; Chay 1998), and prejudice (Charles and Guryan 2007; Charles, Guryan, and Pan, 2010).

major appellate decisions, in human resources management publications and in law review articles.[7] Rapid change in case law added to the uncertainty.[8] Millions of dollars were spent on training programs and establishing grievance procedures aimed to reduce the risk of lawsuit more than to reduce the incidence of harassment.[9] Interviews with hundreds of firms from a representative sample of US businesses with 50 or more employees suggest that these labor lawyers and human resources consultants were quite effective in translating appellate decisions into human resources policy changes (Dobbin and Kelly 2007).[10] We employ this dataset to investigate the timing of adoption of sexual harassment grievance procedures and policies in response to sexual harassment law in Section 4.[11]

Forbidding sexual harassment would have significant labor market consequences in an insider-outsider model of harassment and involuntary unemployment (Lindbeck and

---

[7] For example, initial surveys reported that 90% of women experienced sexual harassment, when later surveys found that only a small fraction actually did.

[8] Our own discussions with labor lawyers indicate how fearful employers are of Title VII suits. Furthermore, sexual harassment cases tend to be personal allegations, which could cause risk-averse employees to change behavior, in a way exceeding that of their response to gender discrimination, ADA, or maternity mandate lawsuits, whose resolution can turn on evidence that is likely to be more statistical in nature.

[9] For example, more grievance procedures were established than maternity leaves, even though the law mandated maternity leaves in no uncertain terms and grievance procedures were not part of a bright-line rule until 1998, when the Supreme Court said grievance procedures greatly reduced the liability faced by firms.

[10] The imposition of direct and indirect litigation costs has impacted firm behavior in multiple ways. For those cases that were actually litigated, the law resulted in direct litigation costs and potentially large damage awards. The potential for litigation then likely led many firms to settle with potential plaintiffs, particularly after such plaintiffs filed complaints with the Equal Employment Opportunity Commission (EEOC) and received a right to sue letter from the agency. In addition, many firms invested in formal grievance procedures, in part because of the advice of personnel experts who believed that instituting formal grievance procedures analogous to those for civil rights violations would help defend against damage awards (Dobbin and Kelly 2007). Some firms even mandated training akin to the diversity training implemented in response to the Civil Rights Act of 1964 to educate their workers about sexual harassment. This last approach was controversial since some firms and attorneys feared that such training might make potential plaintiffs more aware of harassment and, therefore, more likely to sue (Dobbin and Kelly 2007). In the end, it is an empirical question whether firms responded to sexual harassment precedent.

[11] Changes in human resources policies are but 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 they were more likely to win in the event of a sexual harassment suit.

Snower 1988, Schultz 1990). In this model, while harassment is allowed, outsiders are unable to find jobs even though they are prepared to work for less than the prevailing wages of incumbent workers (insiders). The outsiders cannot underbid insiders; if they did and were successfully to become new entrants, insiders would withdraw cooperation and make the work experience of these entrants unpleasant. In other words, 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. Outsiders are unable to find work even though they would be just as profitable to the firm as the insiders, provided they faced identical conditions of employment. The insiders' harassment activities, however, ensure that conditions are not the same for insiders and outsiders.

Applying this insider-outsider model in its simplest form assumes that females are outsiders and men are insiders. Under these assumptions, forbidding harassment can increase the employment and wages of outsiders. Insider males would no longer be allowed to engage in harassment activities, thereby raising the productivity of females. Firms would be willing to hire females, so their employment and wages would increase. Of course, not all females are outsiders; hence, some females—those who previously obtained the insider rents—may see less benefit in their employment outcomes, especially as they faced increased competition from previously outsider females.

An alternative view suggests that sexual harassment law may have been a tax on the hiring of women, making it more costly to hire women.[12] The law, like an unfunded, mandated benefit, may act like a tax on labor demand (Summers 1989).[13] Similar to the

---

[12] This would be the conventional view among many law and economics scholars (Epstein 1995).
[13] These theoretical effects are described and analogized to those of accommodation mandates (Jolls 2000 and 2001). "Disadvantaged employees will be more willing to supply labor at any given wage once a particular benefit must be provided to them. Effects parallel to those of accommodation requirements also occur for labor demand. Employers are subject to a potential lawsuit over every adverse incident on the job suffered by a disadvantaged employee. These costs shift down the marginal revenue product of labor for disadvantaged employees, just as the costs associated with accommodation requirements shift down this marginal revenue product of labor." (Jolls 2001 p. 690)

unfunded mandate in the Americans with Disabilities Act (ADA), all of the direct and indirect costs of sexual harassment law on firms—from establishing internal infrastructures conducive to complaint to marginal costs associated with each female worker who has some probability of filing a complaint or becoming a litigant—may have been passed onto women and lowered their wages or employment relative to men.[14] Assuming that the effects of sexual harassment law were predominantly experienced by women, forbidding harassment may further lower female wages by increasing the supply of female labor by making it more pleasant for women who would be willing to work for lower wages in work environments that previously allowed harassment. If employees value the benefit at cost, the resulting equilibrium will result in the same level of employment but with the full cost reflected in lower wages (Basu 2004). If there are wage rigidities, then the cost of the benefit cannot be reflected in wages, and thus unemployment may result (Summers 1989).[15] Whether forbidding sexual harassment exacerbates gender inequality is a priori ambiguous and this motivates our empirical investigation.[16]

III. Design of Study

A. Data

Our empirical analysis draws on three sources of data on sexual harassment cases—an established dataset as well as our own data collection. The first dataset is from Boyd, Epstein, and Martin (2010), which codes case characteristics, such as the presence

---

[14] While the Equal Pay Act of 1963 (see U.S.C. § 206(d)) makes lower wages for women for the same work illegal, women's wages could be constrained through a failure to promote given the same qualifications or through the offering of fringe benefits packages designed to appeal to some workers but not others.

[15] There are, however, reasons to think that some of the mandated benefit mechanisms are weakened on the employer side in the case of sexual harassment law since it may be difficult to know in advance who is going to be a sexual harassment plaintiff and men could bear some of the cost of the mandated benefit as potential harassers.

[16] Some recent empirical work has found that similar social policies regulating labor markets, such as the ADA and maternity mandates, had detrimental effects on the groups they were intended to protect and undoing, in part, the redistributive goals of these policies (ADA: Acemoglu and Angrist 2001 (for a more nuanced perspective, see Jolls and Prescott 2004); Employment protection: Autor, Donohue, and Schwab 2006; Maternity mandates: Gruber 1994).

of certain fact patterns and legal issues for a subset of Title VII discrimination claims in the Chicago Judges Project data (Sunstein, Schkade, and Ellman 2006). We use this to perform a randomization check. Our second dataset is composed of our own collection of cases from 1982 to 2002 of all sexual harassment cases brought in an employment context; if the plaintiff was afforded any relief, then the vote was coded as a pro-plaintiff vote.[17] Figure 1 plots the growth in sexual harassment cases, number of pro-plaintiff decisions, and number of pro-defendant decisions during this time period. Table 1 indicates that on average, there are 0.996 sexual harassment panels per circuit-year for a total of 251 cases. A sizeable portion of circuit-years, 51%, had no sexual harassment panel. Two-thirds of the decisions are pro-plaintiff. Our third dataset comprises of all district court cases involving sexual harassment.[18] This resulted in 3,754 cases between 1982 and 2002.

We collect additional information on judge characteristics from the following sources: the Federal Judicial Center website;[19] the Appeals Court Attribute Data compiled by Zuk, Barrow, and Gryski;[20] data on judges' religious attributes compiled by Goldman[21] and by Sisk;[22] and additional religious and personal attributes that we assembled ourselves.[23] The average circuit-year has 18.50 judges available for assignment to panels. The expected number of Democrats per seat is 0.41 (there can be 0 to 3 male Democratic appointee judges on a 3-judge panel, which translates to 0 to 1 male Democratic appointee judges per seat).[24] Additional summary statistics are displayed in Table 1.

---

[17] This method follows the one outlined in Sunstein, Schkade, and Ellman (2006), which searches Lexis using the keywords "sex! harassment" but only collects cases from 1995-2002.

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

[19] http://www.fjc.gov/history/home.nsf.

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

[21] See Goldman (1997). Raw data were obtained directly from the author.

[22] http://courseweb.stthomas.edu/gcsisk/religion.study.data/cover.html

[23] See Chen and Yeh (2011). We calculate the expectations based on the frequency the typical senior judge sits on cases and weight senior judges accordingly.

[24] There can be 0 to 3 male Democratic appointee judges on a 3-judge panel, which translates to 0 to 1 male Democratic appointee judges per seat. "Per capita" is an alternative nomenclature.

Our main outcome variables are from the Merged Outgoing Rotation Groups (MORG) Current Population Survey (CPS), which contains individual employment outcomes, including weekly earnings, amount of time worked, employment status, and management status. Our main outcome variable is the distinction between no-employment (including non-labor force participants) vs. part- or full-time employment. Non-labor force participants include discouraged workers.[25] We use hours last week instead of usual weekly hours because usual weekly hours are not consistently available. As a result of the CPS redesign in 1994, workers who report that their weekly hours vary are not asked to report usual weekly hours, yielding a non-report rate of 7.0 to 8.5 percent of workers in 1994 to 2003 (Autor, Katz, and Kearney 2005). We recode the number of hours worked for individuals who are either not in the labor force or unemployed as zero.[26] We also recode earnings as zero for individuals who are not in the labor force or unemployed.[27] Earnings are adjusted to be in 2000 real terms. We do not recode management status, which is constructed from the occupation variable. 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, however, so we do not analyze this outcome when our sample includes those unemployed or not in the labor force. The CPS also contains demographic controls, including 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. We restrict to individuals between the ages of 18 and 65. We obtain data on the presence of firm-level sexual harassment policies from Dobbin

---

[25] According to the BLS, "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. Information is collected on their desire for and availability for work, job search activity in the prior year, and reasons for not currently searching."

[26] In the March CPS dataset, the number of hours worked last week is coded as zero for individuals who are not in the labor force and for individuals who are unemployed. The number of hours worked last week for the same demographic group is coded as missing in the MORG dataset.

[27] When we take logs, we add 1 to the zero, so in essence we are setting log real weekly earnings to 0.

and Kelly (2007), a national sample of 389 workplaces interviewed in 1997 on the history of human resources practices dating back to 1965.

B. Identification Strategy

Understanding five aspects of the US judicial system is important for the development and understanding of the identification strategy. First, the United States has a common law system where American judges not only apply the law but also, to some extent, make the law. Judge's decisions in current cases become precedent for use in decisions in future cases in the same court and in lower courts of the same jurisdiction. Second, there are three layers of courts in the US federal judicial system. District courts are the courts of general jurisdiction and hold trials. When trial cases are appealed, they go to appellate courts, referred to as circuit courts, which typically decide issues of new law or determine whether the district court was in error. A small portion of appellate cases is appealed again to the Supreme Court.[28] Therefore, appellate courts are quite active in shaping law. They handle the vast majority of cases deciding issues of new law and provide new interpretations or distinctions of pre-existing precedents or statutes.[29]

Third, there are twelve appellate (circuit) courts, each in charge of a geographic region of the United States, known as a circuit.[30] Appellate decisions are binding precedent only in the circuit of the court delivering the opinion; that is, the district courts within a circuit and the circuit court itself must follow the precedent set by the circuit court's prior decisions. As such, appellate decisions in one circuit do not establish precedent that other circuits must necessarily follow.[31] Fourth, judges are randomly

---

[28] In a random sample of about 20,000 appellate cases from 1925 to the present, only 2% are heard by the Supreme Court (Berdejo and Chen 2010).

[29] According to one view, appellate courts are continually finding new distinctions with which to expand or contract the space under which an actor would be found liable (Gennaioli and Shleifer 2007).

[30] See Figure 1.

[31] When circuits choose to adopt the precedent of another circuit, it is typically with some delay. For example, a new case bringing the same issue of law must be filed in a district court, appealed

assigned to cases in appellate courts. Some judges take a reduced caseload, but all are randomly assigned by a computer algorithm and are typically not revealed to the litigating parties until after they file their briefs, sometimes only a few days before the hearing, if there is a hearing. Fifth, appellate courts assign three judges to a case; because a circuit can have twenty to forty judges in the pool of judges available to be assigned, the number of possible combinations of judges or combinations of judicial demographic characteristics on a panel is very large.

It has been documented that judges' personal attributes, such as gender, religion, and political persuasion, can predict how an individual judge may vote on a certain type of case.[32] These facts, and in particular, the fact that assigning judges with a particular gender or party of affiliation increases the likelihood of a pro-plaintiff sexual harassment decision, allow us to construct a setting akin to a randomized experiment in the establishment of precedent across different regions of the United States.[33]

In estimating the impact of sexual harassment law, the ideal research design would be a randomized controlled experiment, where pro-plaintiff precedent is randomly assigned to a treatment group that is initially identical to the control group. After administering the treatment, the researcher would measure the outcomes for each group, and conclude that any differences in the outcomes between the treatment and control group constitute the effect of treatment. Randomization is necessary to ensure that any observed differences between the two groups arise solely because of the precedent.[34]

---

to the circuit court, decided upon, and have an opinion issued before the doctrine becomes binding precedent in the new circuit.

[32] Boyd, Epstein, and Martin 2007; Chang and Schoar 2006; Ellman, Sunstein, Schkade 2003; Peresie 2005.

[33] The appendix provides a timeline of the major developments in sexual harassment doctrine. For examples of how doctrinal shifts could make it easier for subsequent sexual harassment plaintiffs to bring and win suit, consider the replacement of a "reasonable person" standard with a "reasonable woman" standard for determining whether sexual harassment occurred and the elimination of the requirement to prove psychological harm.

[34] The correlation between court-made law and economic outcomes is generally difficult to interpret since the causality may run in both directions and the relationship may reflect omitted variables. Momentous judicial decisions may be caused by rather than be causes of political or socioeconomic changes. Judges have relied on personal values, heavily influenced by larger historical forces, in decisions such as *Brown v. Board of Education* (Klarman 2004).

Since this randomization is impractical to implement, an alternative is to identify an exogenous source of variation that allows us to approach the conditions of the ideal experiment. In our empirical strategy, we exploit a natural experiment where pro-plaintiff sexual harassment precedent varies randomly by circuit and over time due to the random assignment of judges to appellate panels. We use this variation to identify the effects of sexual harassment law on human resources policies and gender inequality in the labor market.

We should expect to see an effect of appellate judicial decisions if judges follow precedent and appellate decisions on the margin make it easier for subsequent plaintiffs to bring and win suit. For example, an appellate precedent shifting from a reasonable person standard to a reasonable woman standard in sexual harassment claims (*Ellison v. Brady*) would make it easier for subsequent plaintiffs to bring and win suit. We might then expect firms and individuals to respond to appellate decisions, whether through newspaper publicity (Kritzer and Drechsel 2011), advocates, lawyers, or information consultants greatly exaggerating the risk of suit after major appellate decisions.

Our research design can be further clarified by the following illustration. Consider the Ninth Circuit, a circuit with a high proportion of judges that are Democratic appointees. The empirical strategy does not rely on cases getting more Democratic appointees in the Ninth Circuit as opposed to the Fourth Circuit, which could be different for a variety of reasons. Rather, the strategy relies on the fact that, from year to year, the proportion of cases for a particular case category in the Ninth Circuit that are assigned Democratic appointees varies in a random manner. The idiosyncratic variation is not expected ahead of time since judicial assignment is not revealed to parties until very late in the appellate process and after each litigant's briefs are filed. In the years when an unexpectedly high number of Democratic appointees are assigned to panels for a particular case category, the proportion of cases for that case category that will result in pro-plaintiff precedent is also high. Even though we cannot ask any particular Circuit to randomize their decisions, a randomized control trial is in effect created through the random assignment of judges who interpret the facts and the law differently.

Random variation in the assignment of appellate judges is an attractive instrument for a number of reasons. The random assignment of judges is exogenous and unexpected. It varies in both the cross-section and the time-series, so it does not rely on strong assumptions about the comparability of different regions (e.g. circuits) and years. The enormous variation in legal decisions due to the judicial panel composition also makes the empirical design an ideal setup to study the consequences of law. Additional advantages are discussed after presenting the basic specification.

C. Specification

Our basic specification models the changes in sexual harassment precedent at the circuit-year level and its relationship to individual outcomes of persons or firms in those circuits over time:[35]

$$(1a) \quad Y_{ict} = \beta_0 + \beta_1 Law_{ct} + \beta_2 Law_{ct} *Female_{ict} + \beta_3 Female_{ict} + \varepsilon_{ict}$$

The dependent variable, Y_{ict}, is a measure of outcomes of individual i in circuit c and year t. Outcomes are employment status,[36] hours worked last week,[37] log weekly real earnings,[38] and management status[39] of individual i in circuit c and year t.[40] The key coefficient of interest is β_2 on the interaction of Law_{ct} and $Female_{ict}$, where Law_{ct} is the measure of sexual harassment precedent issued in circuit c and year t. Let Law_{ct} be the proportion of cases with a pro-plaintiff outcome.[41]

If sexual harassment law and employment outcomes are systematically correlated with omitted variables, then β_2 is biased. A critical concern with judge-made law is that there is so much cross-fertilization across different areas of legal doctrine. If different,

---

[35] For our firm-level analysis we do not interact the legal precedent with the gender indicator.
[36] Employment status is a binary indicator for whether the individual has any part or full time employment.
[37] Hours worked is set to 0 if an individual is not employed or not in the labor force.
[38] Earnings are normalized to account for inflation. In addition, logs are taken of 1+earnings and earnings are set to 0 if an individual is not employed or not in the labor force.
[39] Management status is a binary indicator for whether an individual has an administrator, official, public administration, executive, or other management-related occupation.
[40] We use CPS weights when examining labor market outcomes.
[41] With about one sexual harassment decision per circuit per year, we did not consider quadratic or non-monotonic functions of the number of pro-plaintiff decisions.

but related, doctrinal areas have independent effects on employment outcomes, social changes may be misattributed to one legal rule when many legal rules are changing simultaneously. Distinguishing correlation from causation is particularly challenging in observational studies, which are frequently confounded by other factors moving in tandem with the treatment of interest. If, for example, social trends drive judicial decisions, how do we ascertain a causal effect from judicial decisions to social trends?

The common approach of controlling for potential confounders:

(1b) $Y_{ict} = \beta_0 + \beta_1 Law_{ct} + \beta_2 Law_{ct} *Female_{ict} + \beta_3 C_c + \beta_4 T_t + \beta_5 C_c *Year + \beta_6 W_{ct}$
$+ \beta_7 W_{ct} *Female_{ict} + \beta_8 X_{ict} + \varepsilon_{ict}$

in a multivariate regression can accentuate the problem of omitted variable bias (Clarke 2005). With a research design involving random treatment assignment, however, adding controls can add precision to the estimates if the controls are strong predictors of the outcomes (Duflo, Glennerster, Kremer 2006). In robustness checks we show that our main estimates are invariant to the inclusion or exclusion of: circuit fixed effects, C_c, and year fixed effects, T_t, to address whether fixed unobservable differences within circuits and within years are correlated with pro-plaintiff sexual harassment precedent and employment outcomes; circuit-specific time trends, $C_c *Year$, to allow different circuits to be on different trajectories with respect to outcomes;[42] state fixed effects to address the possible influence of state-specific sexual harassment statutes or state interpretation of federal laws; a vector of observable individual characteristics, X_{ict}, such as age, gender, educational attainment, and race, which each enter as dummies with the exception of age; and time-varying circuit-level controls, W_{ct}, such as the characteristics of the pool of judges available to be assigned. Since employment outcomes are serially correlated, ε_{ict} is not *i.i.d.* Hence, all specifications cluster standard errors at the circuit level. Bester, Conley, and Hansen (2011) suggest scaling up the cluster t-statistic with relatively few clusters (e.g. 12 circuits), but the scaling is very close to one with large N.[43] Barrios, Diamond, Imbens and Kolesar (2010) indicates that the use of clustered standard errors,

---

[42] For example, the historically liberal Ninth Circuit may be increasingly pro-plaintiff on sexual harassment decisions as women enter the labor force in increasing numbers.
[43] See Footnote 6 in their paper.

along with the random assignment of treatment, address possible spatial correlation in the errors as well. We also execute a wild bootstrap suggested by Cameron, Miller, and Gelbach (2008) for small number of clusters and a Monte Carlo simulation, where we randomly assign the legal variation to another circuit.

Since Law_{ct} and ε_{ict} may be correlated due to uncontrolled-for social trends or other legal developments that correlate both with Law_{ct} and outcomes Y_{ict}, β_1 and β_2 may be biased. We need an instrumental variable for Law_{ct} that is uncorrelated with ε_{ict}. Figure 3 roughly depicts the intuition for our 2SLS identification strategy, in which we exploit the random variation that arises from using the random deviation in the actual number of male Democratic appointees per seat and the actual number of female Republican appointees per seat in sexual harassment cases. These numbers are plotted in red in Figure 3 for each of the 12 Circuits. The blue lines in indicate the expected number of male Democratic appointees per seat and the expected number of female Republican appointees per seat for each Circuit. Circuit-years receiving an unexpectedly high proportion of male Democratic appointees (low proportion of female Republican appointees) on their sexual harassment panels receive an unexpectedly higher proportion of pro-plaintiff sexual harassment decisions. Each spike in the actual number of male Democratic appointees per seat above the expected number of male Democratic appointees per seat corresponds to the circuit-year randomly receiving a "treatment" of more pro-plaintiff sexual harassment precedent. Thus, changes in outcomes can be attributed to the "treatment" of pro-plaintiff sexual harassment and not to other unobserved social trends or legal developments.

Figure 3 suggests the first stage equation:

(2a) $Law_{ct} = \gamma_0 + \gamma_1 Treatment_{ct} + \gamma_2 C_c + \gamma_3 T_t + \gamma_4 C_c * Year + \gamma_5 X_{ict} + \gamma_6 W_{ct} + \varepsilon_{ict}$

where Law_{ct} is defined as the percentage of decisions that are pro-plaintiff, conditional on there being any decision in that circuit and year. The "Treatment" group ($Treatment_{ct} = 1$) comprises people who experience an unexpectedly higher percentage of pro-plaintiff decisions due to an unexpectedly higher actual number of male Democratic appointees

(lower number of Female Republican appointees[44]) being assigned to the panels. The "Control" group (*Treatment$_{ct}$* = 0) comprises people who experience an unexpectedly lower percentage of pro-plaintiff decisions. Formally, *Treatment$_{ct}$* = *1*[(N_{ct}/M_{ct} > **E**(N_{ct}/M_{ct})], where N_{ct} is the number of male Democratic appointees assigned to all sexual harassment cases in that circuit-year divided by three and M_{ct} is the number of sexual harassment cases in that circuit year. N_{ct}/M_{ct} is the actual number of male Democratic appointees per seat and **E**(N_{ct}/M_{ct}) is the expected number of male Democratic appointees per seat. The moment condition for causal inference is **E**[*Treatment$_{ct}$* ε_{ict}] = **E**[*1*[(N_{ct}/M_{ct} - **E**(N_{ct}/M_{ct}))] ε_{ict}] = 0, which holds since being above or below the threshold is uncorrelated with social trends or legal developments that might otherwise be correlated with outcomes through ε_{ict}. The effect of law on outcomes is the difference in *Outcome$_{ict}$* for *Treatment$_{ct}$* = 1 or 0, divided by the difference in *Law$_{ct}$* for *Treatment$_{ct}$* = 1 or 0. That is, we could simply look at the Wald estimator for outcomes in any given year during or after the treatment to estimate the treatment effect.[45]

Before moving on to extensions of the basic model, we make two remarks. First, the exclusion restriction is likely to hold, and we will thus be able to interpret the 2SLS estimates as the causal impact of sexual harassment precedent rather than simply the causal impact of assigning male Democratic appointees to sexual harassment cases. Here, the identity of judges sitting on sexual harassment panels is not likely directly to affect economy-wide outcomes that are of interest except through the appellate precedent alone.[46] Second, the LATE interpretation of the instrumental variables estimate is

---

[44] For expositional simplicity, we suppress discussion of female Republican appointees for now.

[45] Note that labor market outcomes are unlikely to zig-zag in the manner suggested by the Figure 3. Rather, we look for deviations from an underlying trend. Estimates of the treatment effect would be obtained from differencing the average of these deviations for treated circuit-years with the average of these deviations for non-treated circuit-years.

[46] This aspect of the research design also highlights that causation from demographic characteristics to decision-making is not necessary for the methodology, rather, only correlation is needed between demographic characteristics and decision-making. The exact mechanism for why demographic characteristics affect decision-making is also irrelevant. For example, decisions could be different because litigants tailor their oral arguments to the judge.

Why use demographic characteristics instead of judicial attitudes? Demographic characteristics provide a multivariate characterization of judges that can be used across judges.

restricted in terms of external validity. Here, only cases where there is enough controversy to allow judicial biographical characteristics to matter are going to be the subject of the study. These cases may very well be the difficult decisions that set new precedent, and the sorts of cases in which judges interested in empirical consequences of decisions, like Judge Richard Posner or Justice Stephen Breyer, seek guidance.[47]

For more statistical power, we can employ the entire excess proportion of male Democratic appointees per seat as a continuous instrumental variable. That is, we can write:

(2b) $Law_{ct} = \pi_0 + \pi_1 Z_{ct} + \pi_2 C_c + \pi_3 T_t + \pi_4 C_c*Year + \pi_5 X_{ict} + \pi_6 W_{ct} + \omega_{ict}$

where Z_{ct} is the difference between the actual number of male Democratic appointees per seat and the expected number of male Democratic appointees per seat. The moment condition for causal inference is $\mathbf{E}[(N_{ct}/M_{ct} - \mathbf{E}(N_{ct}/M_{ct}))\ \varepsilon_{ict}] = 0$. In words, the greater the excess proportion of male Democratic appointees per case, the more pro-plaintiff is the sexual harassment precedent in that circuit-year. We attribute the degree to which outcomes change to this excess proportion.

Laws are not likely to have an immediate impact. Firms may need time to adjust to a new legal regime; alternatively, the effects of a law change may fade as expectations adjust.[48] We build on our basic model with a distributed lag specification that includes five years of lags of the law and one lead. The use of leads helps assess whether trends in labor market inequality precipitate sexual harassment precedent. We therefore estimate a distributed lag specification to study the dynamic effects of law over time:

(3) $Outcome_{ict} = \beta_0 + \beta_{1y}\sum Law_{c(t-y)} + \beta_{2y}\sum Law_{c(t-y)}*Female_{ict} + \beta_3 Circuit_c +$

$\beta_4 Year_t + \beta_5 Year*Circuit_c + \beta_{6y}\sum W_{c(t-y)} + \beta_{7y}\sum W_{c(t-y)}*Female_{ict} + \beta_8 X_{ict} +$

ε_{ict}

These characteristics provide additional power relative to attitudinal scores, which are unidimensional, or to individual judge fixed effects, which would be imprecisely estimated if each judge hears only a handful of sexual harassment cases.

[47] See the introduction for the initial mention of Judge Posner and Justice Breyer.

[48] Judicial precedent can also be modified by the Supreme Court or legislature.

A problem with this dynamic treatment effect specification is that circuit-years with no cases greatly reduce sample size.[49] If no appellate cases appear in any of the last, say, 5 years, the observation would drop from the estimation.

To address this missing data problem, consider the moment condition for causal inference:

(4) $\mathbf{E}[(N_{ct}/M_{ct} - \mathbf{E}(N_{ct}/M_{ct})) \, \varepsilon_{ict}] = 0.$

We seek to construct an instrumental variable whose moment conditions will imply the aforementioned moment condition. Consider an instrument, $p_{ct} - \mathbf{E}(p_{ct})$. The moment condition for this instrumental variable is:

(5) $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct})) \, \varepsilon_{ict}],$

where p_{ct} is the number of male Democratic appointees per seat in sexual harassment cases in circuit c and time t and p_{ct} is defined as 0 when there are no cases. Specifically, let:

(6) $p_{ct} = N_{ct}/M_{ct}$ if $\mathbf{1}[M_{ct} > 0] = 1,$

$\quad\quad\ p_{ct} = 0$ if $\mathbf{1}[M_{ct} > 0] = 0.$

When $\mathbf{1}[M_{ct} > 0] = 1$, $p_{ct} = N_{ct}/M_{ct}$ returns the original moment condition $\mathbf{E}[(N_{ct}/M_{ct} - \mathbf{E}(N_{ct}/M_{ct})) \, \varepsilon_{ict}] = 0$. When $\mathbf{1}[M_{ct} > 0] = 0$, then $p_{ct} = 0$ and $\mathbf{E}(p_{ct}) = 0$, so $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct})) \, \varepsilon_{ict}] = 0$. However, these two conditional moment conditions do not imply $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct})) \, \varepsilon_{ict}] = 0$ unconditionally. The presence of appellate cases, $\mathbf{1}[M_{ct} > 0]$, may be a function of ε_{ict}, so it needs to be controlled. After controlling for it, then $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct})) \, \varepsilon_{ict}] = 0$ unconditionally.

In addition, since $\mathbf{E}[(p_{ct} - \mathbf{E}(p_{ct})) \, \varepsilon_{ict}] = \mathbf{E}[p_{ct} \, \varepsilon_{ict}] - \mathbf{E}[\mathbf{E}(p_{ct})) \, \varepsilon_{ict}] = \mathbf{E}[p_{ct} \, \varepsilon_{ict}] - \mathbf{E}(p_{ct}) \, \mathbf{E}[\varepsilon_{ict}] = \mathbf{E}[p_{ct} \, \varepsilon_{ict}]$, we can ignore $\mathbf{E}(p_{ct})$, the expected number of male Democratic appointees per seat. We have now constructed our instrumental variable, p_{ct}. This allows a distributed lag specification since there is no dividing by 0, but one needs to include a binary indicator $\mathbf{1}[M_{ct} > 0]$ representing whether there are cases. This binary indicator

---

[49] Cellini, Ferreira, and Rothstein (2010) address this kind of issue in the context of a repeat regression discontinuity design.

ensures that circuit-years with no male Democratic appointees assigned to cases are treated differently if there are no appealed cases in that circuit-year.

It is important to note the difference between the conditional and unconditional effect of Law_{ct}, which is also set to 0 when there are no cases. This distinction is important since the conditional effect is the one of policy-interest to a judge making a decision on a case already in front of him or her. The unconditional effect is the one of policy-interest to an advocate or historian interested in the social change that is due to court-made law. For example, to calculate the effect of 1 pro-plaintiff decision when there is only 1 decision, we would need to add the effect of $\mathbf{1}[M_{ct} > 0]$ with the effect of Law_{ct} to obtain the unconditional estimates of going from 0 to 1 case.

The inclusion of $\mathbf{1}[M_{ct} > 0]$, however, threatens the moment condition in a distributed lag specification. Whether there are any cases in a given year may respond to previous years' realization of the instrument. That is, having many male Democratic appointees being assigned to sexual harassment cases in prior years may affect litigants' willingness to appeal in the current year. If this is the case, then treatment affects both the left and right-hand side and leads to downward bias in the estimates of interest with the greatest downward bias in the most lagged treatment and the least downward bias in the least lagged treatment (such as lead coefficients). This form of downward bias makes the distributed lag specification difficult to interpret.

To address this potential downward bias, we instrument for $\mathbf{1}[M_{ct} > 0]$ with the random assignment of district court judges to their cases. One district court judge is randomly assigned per case (Bird 1975).[50] Figure 1 displays the boundaries of each district court with dashed lines. Whether the district court cases got disproportionately assigned to certain types of judges will be uncorrelated with treatment (the random assignment of appellate judges) but may affect the likelihood of subsequent appeal. Theoretically, this could occur, for example, if some district judges are less likely to be

---

[50] We search on District Court database in Westlaw using (SEX! +2 HARASSMENT). We restrict our attention to cases authored by district court judges and exclude recommendations by magistrate judges because litigants cannot directly appeal a magistrate judge's recommendation (28 U.S.C. § 636(c)(1)).

reversed and this discourages litigating parties from pursuing an appeal. A handful of papers examine the correlation between district judge demographic characteristics and their reversal rates (Steinbuch 2009, Barondes 2011, Haire, Songer, and Lindquist 2003).[51]

D. Randomization

Our empirical strategy is to use the number of male Democratic and female Republican appointees assigned to sexual harassment cases to approximate a true experiment. This requires that appellate judges be randomly assigned. At the circuit-year level, the number of male Democratic (female Republican) appointees per seat needs to be as good as randomly assigned, conditional on having a case. A few scholars argue that certain circuits have not used random assignment (Hall 2010), so we consider three tests of this assumption.

First, 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" with regard to the 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.

---

[51] The effect of $\mathbf{1}[M_{ct} > 0]$ has multiple interpretations. The absence of an appellate decision may be due to a widely publicized pro-plaintiff district court decision. Alternatively, the presence of appellate sexual harassment cases could signal workplace norms regarding sexual harassment (see e.g. formal models along the lines of Benabou and Tirole 2011). The conditional causal effects are more easily interpretable than the unconditional causal effects and, in any event, are the ones of likely policy interest to appellate judges.

In another court, random assignment of panels occurs before the random assignment of cases. Panels of judges are set up 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 in to those spots by the chief judge. This same program ensures that the same judges are not seated together repeatedly on the same panel. Thus, the judges know at the beginning of the year which days they will be hearing cases and the composition of the panel.

Once all of the briefing is completed, a case goes to 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, etc.). 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 of 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.[52]

As a second randomization check, we use data from Boyd, Epstein, and Martin (2010), which codes some case characteristics for a subset of 415 gender discrimination cases in the Chicago Judges Projects data (Sunstein, Schkade, and Ellman 2006). We regress case characteristics on male Democratic (female Republican) appointees per seat and find that most characteristics are not correlated with the judicial panel composition. Table 2 shows that of 19 case characteristics, one is correlated with male Democratic appointees per seat and one is correlated with female Republican appointees per seat. Both correlations are statistically significant at the 10% level, which are no longer significant with Bonferroni corrections for multiple hypotheses testing. For the 79 cases that we coded from 1982 to 1995, we also noted whether the plaintiff was the victim of

---

[52] 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 appellate level, see Brown, Jr. and Lee (2000) and http://law.du.edu/images/uploads/neutral-assignment/Neutral_assignment_links.pdf

sexual harassment and which party appealed. The last two rows of Table 2 show that these case characteristics are also uncorrelated with judicial panel composition. Judicial characteristics do appear orthogonal to case characteristics as determined before the assignment of appellate judges.

As a third randomization check, we examine whether the sequence of actual numbers of male Democrats (female Republicans) per seat is like a random process. These checks are important because the decision to publish may introduce non-randomness.[53] For example, the decision not to publish may be a compromise among judges who disagree about the correct outcome (Law 2005, Wald 1999). If Democratic appointees publish and Republican appointees choose not to publish sexual harassment decisions, then a correlation may arise between the egregiousness of the sexual harassment case and the judicial panel composition. This correlation could reintroduce possible endogeneity between social trends related to the egregiousness of sexual harassment cases and the panel composition of published cases. Examining whether characteristics of the case determined by lower courts are unrelated to the composition of the appellate panel provides an assessment of this concern. Moreover, if panel composition significantly affects the decision to publish, panel composition of published cases would be serially correlated.

Figure 3 suggests visually that panel composition is not serially correlated. Formally, the general approach to assessing randomness is analogous to a Fisher exact test, except that we use simulations. The methodology we follow is:

1. Propose a statistic that can be computed from the sequence of numbers of male Democrats (female Republicans) 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., $s_1, s_2, s_3 \ldots s_n$. Since there were changes in the expected number

---

[53] More specifically, the editorial decision to be included in Westlaw or Lexis is left to the discretion of the individual companies for unpublished cases. Many "unpublished" cases are actually published. Even if editors desire to be as inclusive as possible, some decisions may go truly undetected. If judges grant motions for summary judgment, the opinion would not include the words "sexual harassment". Therefore, it is important to check for randomization.

of male Democrats (female Republicans) per seat over time, we treat our bootstrap samples as a vector a 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 $s_1, s_2, s_3 \ldots s_n$.

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 jth case depends on the outcome in the j-1th case. This statistic can detect whether judicial assignments are "clustered," meaning a higher than expected number of back-to-back high number of seat assignments to a male Democrats (female Republicans). 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 nth case is correlated with the assignment in previous $n - 1$ cases. This test tells us whether judges or their assignors were attempting to equilibrate their presence, considering whether a judge was "due" for a sexual harassment case.

Longest-Run: We test whether there are abnormally long "runs" of number of male Republicans (female Democrats) 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.[54]

While this process generates a collection of p-values, it is not intuitively obvious what should be the rejection criteria. Since p-values from a truly random process with a sufficient number of possible states is uniformly distributed, even with just 10 units and 3 statistics, the probability of not having even one p-value less than .025 or greater than .975 is only about 21%. With a truly random process, we would expect that collection of all unit p-values to be uniformly distributed. (Imagine that you generate summary

---

[54] Some sources suggest that courts do batch cases dealing with similar issues to one panel in order to more quickly depose of cases without duplication of effort (Wallace 2005).

statistics for 1000 random strings. The 1001th random string should have a summary statistic that is equally likely to be anywhere from 1 to 1000.) Of course, since there are only 12 units, we would not expect a kernel density estimate to "look" uniform. We use Kolmogorov-Smirnov Test to test whether the empirical distribution of p-values approaches the CDF of a uniform distribution using the one-sided critical value with n = 12.[55] Figure 4 plots the empirical distribution for our 3 test statistics and both sets of instruments. Table 3 confirms the visual intuition[56] that our p-values are uniformly distributed for all 6 tests except autocorrelation for female Democrats, which is largely due to Circuit 6. In our robustness checks, we drop 1 circuit at a time.

E. First Stage

We begin our analysis by examining whether sexual harassment cases assigned to judges with different background characteristics do in fact have different outcomes. A number of papers have documented the effect of judges' demographic background on sexual harassment cases (Farhang and Wawro 2004, Epstein 2007, Peresie 2005[57]).[58] They find, for example, that having a female judge in the three-judge panel increases the probability of a pro-plaintiff decision from 22% to 41% and, at the individual-level, Republican appointees vote for plaintiffs at a rate of 37% while Democratic appointees vote for plaintiffs at a rate of 52%.

In our data sample, a regression of the judge's decision on gender and party of appointment shows that Democratic appointees are 13% more likely to vote in favor of sexual harassment plaintiffs while female judges are 3% less likely to vote in favor of sexual harassment plaintiffs and are 8% less likely to vote pro-plaintiff when also

---

[55] http://www.ciphersbyritter.com/JAVASCRP/NORMCHIK.HTM#KolSmir.
[56] One adds up the space between the 45-degree line and the p-values to ascertain how similar the distribution of p-values is to a uniform distribution.
[57] See Schultz and Petterson (1992) on judicial decisions in employment discrimination.
[58] We are focusing on appellate precedent; some papers find no effect of judicial background on civil rights and employment discrimination case outcomes in lower courts (Nielsen, Nelson, and Lancaster 2010, Ashenfelter, Eisenberg, and Schwab 1995), where judges may have less discretion in applying appellate precedent.

controlling for party of appointment (Table 4 Panel A).[59] This result suggests that judges may be voting more along party lines than by gender.[60]   Indeed, female Republican appointees are 18% less likely to vote pro-plaintiff while male Democratic appointees are 13% more likely to vote pro-plaintiff. The point estimates and statistical significance change little when including circuit and year fixed effects, the expected number of male Democratic appointees per seat, and the expected number of female Republican appointees per seat (comparing Columns 6-9).

At the case level, an additional actual female Republican appointee per seat on a three-judge panel reduces the chances of a pro-plaintiff decision by 55%, and an additional actual male Democratic appointee per seat increases the chances of a pro-plaintiff decision by 32% (Table 4 Panel B). In other words, a panel that is 100% female Republican would be 55% less likely to have a pro-plaintiff decision.   The expected number of female Republicans per seat is 0.035 (displayed in Table 1 and Figure 3). Again, the point estimates and statistical significance change little with the inclusion of fixed effects and the expected judge type per seat.

At the circuit-year level, an additional actual female Republican appointee per seat reduces by just over 100% the proportion of pro-plaintiff decisions (there are, of course, no cases with 3 female Republicans during this time period).[61]   An additional actual[62] male Democratic appointee per seat increases by 51% the proportion of pro-plaintiff decisions (Table 4 Panel C). These estimates are slightly different from the case level since cases are not evenly distributed across circuit-years. The F-statistic is 21 for

---

[59] All analyses in this section cluster standard errors at the circuit level.

[60] This possibility has been noted by previous legal scholars (Dixon 2010).

[61] The coefficient at the circuit-year level can be different from the coefficient at the case level due to the fact that not every circuit-year has a case and cases can bunch up unevenly across circuit-years. For an arithmetic example, suppose there are 4 cases, one case each with 0, 1, 2, or 3 judges that are male Democrat, and suppose that the panel makes a pro-plaintiff decision when there are 2 or 3 male Democrat judges. If 1 circuit-year has the case with 0 judges that are male Democrat and the other circuit-year has the remaining 3 cases, the coefficient at the circuit-year level is 0.33 but when the 1 circuit-year with the case has the case with 1 male Democrat judge, the coefficient at the circuit-level is 1. Note this example also illustrates how an increase in the probability of a pro-plaintiff decision due to an additional judge of a particular type can be greater than 1.

[62] In what follows, we always mean actual as opposed to expected when the term is omitted.

the female Republican instrument and 9.8 for the male Democrat instrument. When both are included, the joint F is 22.5. The estimates and statistical significance change not at all regardless of whether the circuit-years with no cases are dropped or are dummied and the proportion of pro-plaintiff decisions and judge type per seat are set to 0.[63] The addition of fixed effects, expectations, and circuit-specific time trends also do not affect the point estimates but do increase the F-statistic on the instruments up to 63.9. The R-square does not change much with the inclusion of these controls and changes not at all when expectations are added on top of the fixed effects.

Finally, we examine the first stage relationships at the level of our analysis in Table 4 Panel D. The estimates are slightly different because of the differing numbers of firms and individuals per circuit-year. The joint F statistic on the two instruments is well past the conventional threshold for weak instruments (Stock and Yogo 2005) at 10 and 17, respectively at the firm-level and at the individual-level, and these F statistics again increase with the inclusion of firm/individual-level controls, fixed effects, and additional circuit-year controls up to 16 and 50 respectively. Table 5 presents a falsification test of the instrument and shows that the proportion of pro-plaintiff decisions is not related to the number of male Democrat appointees per seat or the number of female Republican appointees per seat in the one or two years before and after the true instrument.[64]

To check whether our (linear) specifications miss important aspects of the data, we re-estimate the effect of judicial panel composition on proportion of pro-plaintiff decisions using nonparametric local polynomial estimators. Figure 5A presents nonparametric local polynomial estimates of the effect of actual number of male Democrat (female Republican) appointees per seat on the proportion of pro-plaintiff decisions.[65] We use an Epanechnikov kernel.[66] It turns out that the relationship is monotonically increasing (decreasing) for male Democrats (female Republicans).

---

[63] The R-square increases significantly.

[64] Typically, appellate courts handle cases that present new legal issues, so we should not necessarily expect future proportions of pro-plaintiff decisions to respond to past assignment of treatment.

[65] Estimation proceeds in two steps. In the first step, we regress proportion pro-plaintiff and number of male Democrat appointees per seat on circuit and year fixed effects. Then we take the

Some econometricians recommend larger first stage F-statistics[67] in order to ensure the first stage is sufficiently strong. With a large number of possible judicial background characteristics and a combinatorial possible number of panel compositions, our approach has the benefit of a surfeit of experimental variation. The statistician must trade-off, however, between the power of the first stage regression with the addition of instruments[68] and avoiding the weak instruments problem stemming from too many instruments.[69] Choosing among a large number of instruments quickly becomes a challenging statistical issue. We use a LASSO technique to address the issue of instrument selection.[70] LASSO addresses a problem with OLS, which has low bias but large variance: change the data a bit, and you get different subsets of covariates deemed important. LASSO is a sparse model that automatically sets small estimated coefficients to 0 to reduce model complexity. As such, LASSO is an effective tool in selecting instrumental variables from available judicial biographical characteristics. Since using too many instruments effectively renders the instruments weak, such a selection device is necessary. LASSO minimizes the sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. Because of the nature of this constraint, it tends to produce some coefficients that are exactly 0 and hence gives interpretable models. Intuitively, it is a data penalty for having too many covariates. Additional covariates, whose coefficients are not 0, are more likely to violate that constraint. We have a very large number of valid instruments, but not every demographic characteristic

---

residuals from these two regressions and use the nonparametric local polynomial estimator to examine the relationship between male Democrat appointees and pro-plaintiff decisions.

[66] The bandwidths are the default bandwidths selected by Stata.

[67] Such as 25 or 50 to allow for heteroskedasticity and serial autocorrelation (Olea and Pflueger 2011)

[68] Joshua D. Angrist and Guido W. Imbens, *Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity*, 90 JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION 431 (1995).

[69] James H. Stock & Motohiro Yogo, *Testing for Weak Instruments in IV Regression*, in Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg, p. 80-108 (Donald W. K. Andrews & James H. Stock eds., Cambridge University Press 2005).

[70] Alex Belloni, Daniel L. Chen, Victor Chernozhukov, & Chris Hansen, *Instrument Selection using LASSO with an Application to Eminent Domain*, revise and resubmit, *Econometrica*.

matters. LASSO enhances statistical precision when using the random assignment of appellate and district court judges.[71] LASSO is theoretically optimal under certain conditions including sparsity described by Belloni, Chen, Chernozhukov, and Hansen (2011).[72] The use of the LASSO instruments provides a check of over-identification.

To construct our potential LASSO instruments, we use the following biographical characteristics and their interactions at the panel level:[73] 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 US attorney, and previous US attorney, for a total of 450 possible instruments. At the circuit-year level, the LASSO procedure selected the following three instruments: the interaction between number of male Democrats per seat and number of judges born in the 1920s per seat, the interaction between number of female Republican per seat and number of judges having an LLM or SJD per seat, and the interaction between number of female Republican per seat and

---

[71] As a conservative check of our standard errors, we employ Monte Carlo placebo simulations that randomly assign the laws and instrumental variables to different circuits. We do not need to adjust the standard errors further since the theory shows that the increased uncertainty due selecting among many instruments does not show up to first order. When there is sparsity, you do a good enough job finding relevant variables so that this uncertainty is small relative to the usual sampling variability.

[72] In the research design described thus far, we identify a causal impact of sexual harassment appellate precedent but we cannot be sure that it is the causal impact of pro-plaintiff sexual harassment precedent rather than some other dimension of sexual harassment precedent that is correlated with the assignment of male Democratic appointees and female Republican appointees. While the pro-plaintiff dimension appears to be the salient aspect of sexual harassment doctrine to some practitioners and legal scholars, the selection of relevant instrumental variables allows us to code different dimensions of appellate precedent and use different judicial panel characteristics to identify the causal impact of these different dimensions. We leave this application of LASSO for future work.

[73] That is, we interact the number of Democrat appointees per seat with the number of Black judges per seat, for example.

number of judges with above median wealth per seat. The joint F statistics are 33 at the circuit-year level and similarly larger than F-statistic for the instruments displayed at the analysis level in Panel D of Table 1 (doubling from the low of 10 without using LASSO to 20 with LASSO, ranging up to 130 with additional controls).

We consider a similar set of biographical characteristics and instrumental variables for the district judges to identify an exogenous component of the existence of an appeal. The first stage F-statistic for the relationship between presence of an appellate case and the proportion of district cases in a circuit that received a Black judge, is 7.[74] The average circuit-year had 8% of district cases ruled by a Black judge, which decreased the probability the circuit-year had an appeal by 46% (Appendix Table 1 and Appendix Figure 1).[75] The point estimates are identical and the F-statistic goes up to 10 with the inclusion of circuit and year fixed effects as can be seen from comparing Columns 1 and 4 in Appendix Table 1. Columns 2 and 3 show that the proportion of district cases receiving a Black judge is not related to the presence of an appellate case in the previous one or two years. The proportion of district cases in a circuit that received a Black judge is uncorrelated with the number of male Democrat appointees (female Republican appointees) per seat in appellate cases as indicated in the second and third figures of Appendix Figure 1. 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.

IV. Estimating the Impact of Sexual Harassment Law on Gender Inequality
A. Human Resources Policies

---

[74] In unreported analyses, one of the authors finds a similar relationship for eminent domain (Chen and Yeh 2010).

[75] Since we may expect a lag between district and appellate court rulings, we also consider the relationship between the presence of an appellate case and the previous year's assignment of district judges to sexual harassment cases. The average circuit-year had 13% of district cases ruled by a female judge, which increased the probability that the circuit had an appeal the following year by 45%. The first stage F-statistic is 6.7. The remaining analyses obtain similar results.

We should expect to see an effect if the following three assumptions are met: judges follow precedent; on the margin, pro-plaintiff decisions in appellate courts make it easier for subsequent sexual harassment plaintiffs to bring and win suits[76]; and firms respond to appellate decisions. While the first two assumptions are less contested, the third bears further examination. Using a national sample of workplaces and their human resources policies (Dobbin and Kelly 2007), we find that conditional on an appellate court's having rendered a decision in a sexual harassment case, firms are 5.7 percentage points more likely to have a sexual harassment policy on average in each year in the 5 years after a pro-plaintiff precedent than after a pro-defendant precedent[77] (Table 6 displays the OLS and IV estimates in Columns 1 and 2), with the most statistically significant effects found in the first three years after the precedent (the 95% confidence intervals are graphically displayed in Figure 6A). The joint F test for statistical significance is 24.09. The lead effect is insignificant with an F of 2.2. Similar findings are found with the LASSO instruments (Column 3) and when both the LASSO and district-level instruments are used (Column 4). These results are robust to collapsing the data to circuit-year means and using LASSO and district-level instruments.

From 1982-1997, an average of 54% of establishments reported having sexual harassment policies, growing from 15% in 1982 to 96% in 1997. Three-eighths of this annual increase may be attributable to the conditional causal effects of pro-plaintiff precedent of a typical circuit-year. Since the typical circuit-year had sexual harassment panels half of the time and 0.67 proportion of decisions were pro-plaintiff (Table 1), multiplying 0.67, 0.5, and 5.7 percentage points suggests that during the development of sexual harassment law, firms in a typical circuit-year were 1.9 percentage points more likely to have a sexual harassment policy, a small but economically significant effect. 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

---

[76] For examples of pro-plaintiff precedent, see the Appendix for a list of major doctrinal developments.

[77] Since the average number of sexual harassment panels per circuit-year is 0.996 (Table 1), we interpret the effect of going from 0 to 100% proportion pro-plaintiff as being the effect of a decision being pro-plaintiff instead of pro-defendant.

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 38% of the yearly change. This back-of-the-envelope calculation is similar to assuming that a pro-plaintiff decision has a permanent effect on human resources policies of a firm and multiplying the average 1-year effect by the number of pro-plaintiff decisions within a circuit during the 16-year time period. Since our specification comes from a 5-year lag distribution, this suggests that judicial sexual harassment decisions in the last 5 years caused 9.5% more establishments to have a sexual harassment human resources policy.[78]

It is important to note that the estimates in Table 6 are conditional effects: conditional on the presence of an appellate case, we identify the causal effect of a pro-plaintiff decision instead of a pro-defendant decision. To examine the unconditional effect, consider the average coefficient (not displayed) on the dummy indicator for presence of an appellate case is -0.043 for Column 2, -0.038 for Column 3, and -0.027 for Column 4. Focusing on Column 4, which uses district IV to identify the effect of having an appellate case, a pro-plaintiff precedent increases by 0.075 - 0.027 or 4.8 percentage points the likelihood of having sexual harassment human resources policy while a pro-defendant precedent decreases by 2.7 percentage points the likelihood of having sexual harassment human resources policy. The joint F statistic on the effect of presence of appeals is 9.2. Since 67% of decisions are pro-plaintiff, but only half the circuit-years see an appellate sexual harassment case, we multiply 0.67, 0.50 and 4.8 to obtain 1.6 percentage point increase in sexual harassment human resources policy due to pro-plaintiff decisions in a typical circuit-year and multiply 0.33, 0.50, and 2.7 to obtain 0.4 percentage point decrease in sexual harassment human resources policy due to pro-defendant decisions in a typical circuit-year. The net effect is 1.6 - 0.4, or 1.2 percentage points out of 5.1 percentage point annual increase in sexual harassment human resources policies. Assuming a linear increase in sexual harassment policy or permanent adoption of human resources policies by firms, the impact of sexual harassment appellate decisions

---

[78] That is, we multiply 1.9 by 5.

from a typical circuit-year is equivalent to 24% of the annual change in human resources policies addressing sexual harassment. The unconditional effect of court-made sexual harassment law in the last 5 years is that 6% more establishments adopted a sexual harassment human resources policy.

B. Labor Force Participation

Forbidding sexual harassment increased female employment status relative to males. Table 7 Columns 1 and 2 display the OLS and IV estimates and indicates that a pro-plaintiff sexual harassment appellate decision increases female employment relative to males by 1.6 percentage points in the likelihood of working part-time or full-time on average in each year during the five years after the decision (Column 2). A pro-plaintiff sexual harassment decision reduces the likelihood that males are working part-time or full-time by 1.3 percentage points per year during the five years after the decision. The strongest effects are found in the fifth year after the decision, a few years after the strongest effects are found for sexual harassment human resource policies (the 95% confidence intervals are graphically displayed in Figure 6B). The joint F of statistical significance is 8.53 on the lag interaction effects and 28.11 on the lag level effects. None of the lead coefficients are statistically significant.

In Table 8 we present a number of robustness checks. The average of the interaction lags are displayed in Column 1 and the joint F test of the lags is displayed in Column 2. As would be predicted by Table 4 (where we observed that the first stage estimates are invariant to the inclusion or exclusion of controls), when we add circuit specific time trends (row A), remove circuit and year fixed effects (row B), remove all control variables except gender and presence of appellate cases[79] (row C), control for the expected number of male Democrat appointees per seat and the expected number of female Republican appointees per seat (row D), add state fixed effects (row E), the point estimates and statistical significance hardly change. When we do not use CPS weights, the joint F test increases significantly to 16.49 (row F). When we add a 2-year lead, the

---

[79] The presence of an appellate case is also interacted with gender.

average interaction lag increases from 1.6 percentage points to 2.1 percentage points and has a joint F test of 19.25 (row G); the joint F statistics on the leads are still insignificant. When we drop 1 circuit at a time, the estimates change little (row H), though the joint significance varies. When we cluster standard errors at the state level (row I), the results are still similar, as is the case when we collapse the data to the circuit-year level (row J), use LASSO IV (row J), and use district IV (row L). The point estimates are very stable, which is consistent with the use of randomization to estimate causal effects. Finally, we 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 (Shoag 2011). The point estimate for the collapsed circuit-year data is at the 90th percentile in these simulations.

To understand what the conditional effect of 1.6 percentage points means in terms of its economic significance, multiplying by 0.67, the typical proportion of pro-plaintiff decisions, and 0.5, the proportion of circuit-years with an appellate sexual harassment case, suggests that women are 0.55 percentage points more likely to have any employment as compared to men in the typical circuit-year due to the development of sexual harassment law during this time period. Across all circuits and years, 81% of men are employed and 65% of women are employed on average. Between 1982 and 2002, the percentage of men with full or any employment has stayed relatively constant from 78.8% to 80.5%, but for women, it has increased dramatically from 56.7% to 68.6%, or about 0.57 percentage points per year. Since our specification comes from a 5-year lag distribution, this suggests that judicial sexual harassment decisions in the last 5 years caused 2.75 percentage points impact on gender inequality in employment. Examining the impact on female employment alone, we would add the average level effect lag with the average interaction lag, which results in 0.3 percentage points. The joint F significance of statistical significance is 12.70. Multiplying by 0.67, 0.5, and 5 suggests that appellate decisions regarding sexual harassment cases in a typical 5-year time period caused 0.5 percentage points increase in female employment status, a small but economically significant amount.

We can compare these estimates with the estimates of the impact of other anti-discrimination laws. For example, the Equal Employment Opportunity Act increased black employment shares by 0.5-1.1 points per year (Chay 1998). In order to do the comparison, we would need to add the effect of a presence of an appellate case with the effect of a pro-plaintiff decision. To calculate the unconditional effect of sexual harassment precedent taking into account the effect of the presence of an appellate case, we use the district IV estimates from row L. The average effect of a pro-plaintiff decision on female employment shares is 0.013. The average coefficient on the dummy indicator for the interaction of gender and presence of an appellate case is -0.009. The unconditional effect of pro-plaintiff decision is to decrease female employment shares by 0.4 percentage points. Multiplying 1.3 by 0.67 and 0.5, the proportion of pro-plaintiff decisions and the frequency with which an appellate case is present, indicates that the conditional effects of pro-plaintiff sexual harassment decisions are on the lower range (0.44) of the effects of EEOA found by Chay (1998) but multiplying 0.4 by those factors indicates that the unconditional effects are one-third of those effects (0.13). Multiplying by 5 to considering the effect of sexual harassment decisions in a typical 5-year time period indicates that sexual harassment law increased female employment shares by 0.67-2.1 points per year, suggesting an effect similar to those of EEOA.

C. Hours Worked

Forbidding sexual harassment also increased female hours worked relative to males. Table 7 Columns 3 and 4 display the OLS and IV estimates and indicate that a pro-plaintiff sexual harassment appellate decision increases female hours worked relative to males by 0.51 hours worked per week on average in each year during the five years after the decision (Column 4). A pro-plaintiff sexual harassment decision reduces hours worked by males by 0.43 hours (per week) on average in each year during the five years after the decision. The strongest effects are again found in the fifth year after the decision. The joint F of statistical significance is 5.42 on the lag interaction effects and

5.07 on the lag level effects.[80] None of the lead coefficients are statistically significant.

To understand what 0.51 hours worked means in terms of its economic significance, multiplying by 0.67, the typical proportion of pro-plaintiff decisions, suggests that women are working 0.34 more hours per week compared to men in the typical circuit-year due to the development of sexual harassment law during this time period. Across all circuits and years, men work 34.33 hours per week and women work 22.78 hours per week. (We define hours worked to be 0 if an individual is not employed or in the labor force. In Section V we restrict our attention to labor force participants and do not make this imputation.) Between 1982 and 2002, the male hours worked per week rose from 32.8 to 33.7 while female hours worked increased from 19.0 to 24.6, or about 0.26 hours worked per week per year.[81]

D. Earnings

Pro-plaintiff sexual harassment precedent increased female wages earned relative to males by 0.11 in log real weekly earnings on average in each year during the five years after the decision (Table 6 Column 6). It reduced male wages by 0.09 log real weekly earnings on average per year during the five years after the decision. The strongest effects are again found in the fifth year after the decision. The joint F of statistical significance is 7.60 on the lag interaction effects and 10.86 on the lag level effects. None of the lead coefficients are statistically significant.

Multiplying 0.11 by 0.67, the typical proportion of pro-plaintiff decisions, suggests that the development of sexual harassment law helped women close the wage gap by 0.07 log real weekly earnings in the typical circuit-year. Across all circuits and years, men receive 4.91 log real weekly earnings and women receive 3.65 log real weekly earnings. (We define log real weekly earnings to be 0 if an individual is not employed or in the labor force. In Section V we restrict our attention to labor force participants and do

---

[80] This does not pass conventional significance tests.

[81] The same caveats in Sub-section B apply here as well in trying to explain variation in social change that is due to sexual harassment law, since we measure only the conditional causal effect of pro-plaintiff decisions.

not make this imputation.) Between 1982 and 2002, the male log real weekly earnings rose from 4.75 to 4.95 while female log real weekly earnings increased from 3.11 to 4.02, or about 0.033 decrease in gender wage gap per year. These changes represent a decrease of approximately 0.033.[82]

It is important to note that our analysis of wages includes both insiders and outsiders. We do this since the insider-outsider theory of involuntary employment is primarily about labor force participation and lower actual wages of outsiders. Analyses of wage gaps typically focus on insiders, the labor force participants who usually report non-zero wages. Focusing on insider wages, the Equal Employment Opportunity Act narrowed the black-white earnings gap by 0.11-0.18 log points per year (Chay 1998) and state laws barring race discrimination increased black men's relative earnings by 0.28% per year (Neumark and Stock 2006). When analyses include outsiders, studies make different wage imputations about non-labor force participants. Some studies of the gender wage gap correct for selection by imputing wages above or below the median based on educational attainment for women not working full-time and then estimating median regressions (see e.g. Neal (2004) and Olivetti and Petrongolo (2008)). Using this method to include outsiders, 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, Guryan, and Pan 2010).

To compare our estimates with the estimates of the impact of male sexism, we calculate the unconditional effect of sexual harassment law using the district IV. A pro-plaintiff decision increases female wages relative to males by 0.11 log points but the presence of an appeal decreases female wages relative to males by 0.05 log points. In other words, a pro-defendant decision has an unconditional effect of decreasing female wages relative to males by 0.05 log points and a pro-plaintiff decision has an unconditional effect of increasing female wages relative to males by 0.06. Multiplying

---

[82] The same caveats in Sub-section B apply here as well in trying to explain variation in social change that is due to sexual harassment law, since we measure only the conditional causal effect of pro-plaintiff decisions, which otherwise is roughly equivalent to double the yearly shift in gender wage gap.

by 0.67 and 0.5, the proportion of pro-plaintiff decisions and the frequency with which an appellate case is present, indicates that during the development of stricter sexual harassment law, a typical year saw 0.02 log points increase in female wages relative to males, roughly equivalent to two-thirds of a standard deviation in median male sexism. Part of this may be due to the increase in the adoption of human resources policies to address sexual harassment.

V. Evidence for the Insider-Outsider Model of Involuntary Unemployment

Thus far, we have shown positive effects of sexual harassment law on female employment outcomes overall. We now turn to some evidence for the insider-outsider model of harassment and involuntary unemployment. We show that forbidding sexual harassment results in less ameliorative impacts for insider women.

Table 9 displays the estimated effects of sexual harassment law for labor force participants. When the analysis is restricted to the labor force we find that pro-plaintiff sexual harassment decisions exacerbated 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. The joint F tests of statistical significance are very large, 35, 58, and 30, respectively, for the lag interaction effects and 25, 13, and 6 for the lag level effects. One of the lead coefficients is statistically significant for earnings in Column 1. Given the number of F tests for lead coefficients, it may be expected that at least one would be statistically significant.

The mean dependent variables for males and females are similar to those found in other studies of labor force participants (Blau and Kahn 2006). It is helpful to assess exactly how much outsider females are gaining relative to insider females, insider males, and outsider males. Focusing on the average yearly lag effect over 5 years, Column 1 in Table 9 indicates that insider men gain by 0.008 log real weekly earnings while insider women gain by 0.004 log real weekly earnings. Column 6 in Table 6 indicates that insider and outsider men lose 0.086 log real weekly earnings and insider and outsider women gain 0.027 log real weekly earnings. Roughly approximating the 65% of women

with part or full-time employment to be insiders and similarly for 81% of men, and accounting for net movements into and out of labor force participation[83] (Table 6 Column 2), the effects for females and males can be summarized as follows:

(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

with some rounding error in the above calculation.[84]

Forbidding sexual harassment causes wages of all labor force participants to increase, and a portion of insider males to exit in response to or because of the entrance of outsider females. The resulting gains for females are tilted towards outsider females who gain 6 times more than insider females; outsider females also gain 3 times more than remaining insider males. Another 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 managerial status (0.4 percentage point decline in likelihood to be a manager).[85] Taken together, these results suggest that forbidding sexual harassment in large part encouraged the entry of outsider women, who then obtained part- or full-time employment and received wages, and increased productivity of women in the labor force.[86] Previously insider men lost the most in terms of employment status. We only analyze labor market outcomes, however, and a proper welfare calculation would need to take into account the effect along all margins including, for example, firm profits and child outcomes, which we leave for future work.

VI. Disaggregating the Effects of Sexual Harassment Law by Industry

Sexual harassment is very difficult to measure objectively. In a study using decoy

---

[83] This analysis assumes there are not large amounts of inflows and outflows in labor force participation.

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

[85] It is not obvious why management would shift to being a little more male except perhaps the senior colleagues stepping into the role of managing newcomer entrants were more likely to be male.

[86] An alternative interpretation is that women asked for and received higher wages.

crime victims, Banerjee et al (2009) find that only 50% of sexual harassment cases are registered by the police, in contrast to 92% of break-ins and more than 64% of motorcycle thefts. 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 increase in harassment. We therefore cannot directly measure the impact of sexual harassment law on incidence of sexual harassment. We can, however, disaggregate the effects of sexual harassment law and 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 the highest in the construction industry according to complaints filed with the Equal Employment Opportunity Commission (Hersch 2011). Table 10 documents in the odd-numbered Columns that the ameliorative effects on gender inequality are positive and statistically significant for employment status, hours worked, and earnings but negative for the other industries.[87] Note that we restrict the analysis to workers who report an industry, so the analysis necessarily relies more on insiders since labor force participants are more likely to report their industry. The effect on managerial status is negative for both construction and non-construction industries. Because workers, particularly male workers, are much more likely to be managers in the construction industry, the negative impact on this measure of gender inequality is larger for the construction industry. Dividing by the overall inequality in management status results in a similar negative impact in terms of fraction of overall inequality so sexual harassment laws hurt the promotion of women to management in construction and non-construction industries to a similar degree.

VII. Potential Concerns and Additional Remarks

There are a number of reasons to not find an effect.

---

[87] A priori it is not obvious that the effects would be more ameliorative for gender inequality in the construction industry. The bulk of plaintiffs (38%) are blue collar and blue collar plaintiffs win more often in district courts, but high-status (white collar) victims may have more resources to bring suit (Juliano and Schwab 2001), so the effects may offset each other.

Miscoding: Even though this empirical strategy identifies a causal effect of law, the estimated effect may still be smaller than the true effect. Suppose our appellate database miscodes pro-plaintiff and pro-defendant decisions. In the extreme case, judicial characteristics will not predict decisions and a weak first stage will result. If the miscoding errors are not as extreme, there will be a first-stage relationship, and classical measurement error is solved by instrumental variables (Ashenfelter and Krueger 1994).

Cross-Circuit Transmission: Now suppose circuits follow each other. In this case, the treatment is more similar to the control, and in the extreme case, no effect would be identified. If it is less extreme, we would underestimate the true effect.[88] This can be addressed with a spatial lag specification:[89] rather than only estimating the effect of lagged laws in a distributed lag specification, we would estimate the effect of other circuits' precedent on outcomes. We leave this model for future work.

Trivial Decisions: Now consider that this measure of law conflates momentous and trivial decisions. Very strong pro-plaintiff precedent gets treated the same as weak pro-plaintiff precedent, which leads to underestimates. In the extreme case where every decision is trivial, no effect of the law would be found. Our measure of law, however, is more appropriately interpreted as being about the average decision. This average effect perspective applies to other sources of potential heterogeneities, such as the baseline legal or economic environment at the time of decision (or assignment of the judicial panel) or subsequent treatment by state laws and higher courts.

Jurisdiction: Other issues arise depending on the outcome being observed. For example, for person-level outcomes, labor mobility across circuits could cause outcomes in employment to converge. In addition, multi-jurisdictional entities, such as firms, could

---

[88] To see this, consider the following numerical example. We measure 3 pro-plaintiff decisions in treatment and 3 pro-defendant decisions in control, but in reality, the precedent including peer effects has the strength of 2 pro-plaintiff decisions and 2 pro-defendant decisions. We underestimate the true effect since we measure 10 units of outcome change in response to 6 units of law change rather than the actual 4 units of law change. Alternatively, we could have an overestimate if circuits choose to do the opposite of what neighboring circuits do.

[89] Spatial lag specifications are typically difficult to interpret because of the Manski reflection and common shocks problem; however, the instrumental variables strategy proposed here provides identification of these peer effects.

enact similar policies across the circuits. Finally, the ability of plaintiffs to choose jurisdictions in which to file suit weakens the precedential effect in any given circuit. These considerations would all tend to result in underestimates. However, allowing factor mobility to reduce the size of the estimates is appropriate since we are interested in the general equilibrium effects of law.

Settlement: We have now discussed a number of possible violations of SUTVA (Stable Unit Treatment Value Assumption), but even with the violations discussed thus far, the estimates described still have causal interpretation. Certain other considerations, however, could weaken causal inference. First, suppose parties settle after observing the judicial panel. Parties typically do not see the panel composition before they file an appeal.[90] Appellate judges are revealed to parties very late in the process, usually after briefs are filed. Parties are unlikely to settle after filing briefs because the relatively short interval between learning panel members' identities and announcement of the judges' decision imposes small additional costs relative to the cost of litigation prior to learning the judges' identities. Many decisions are based solely on the briefs since in many cases there is no oral argument. However, to the extent plaintiffs settle when they discover the judges' identities are unfavorable to their winning the case, the first stage would be biased towards zero. Even with a significant first stage, the pro-plaintiff precedent would no longer be otherwise identical to the pro-defendant precedent in unobserved characteristics. To address this, we tested for orthogonality between judicial characteristics and case characteristics as determined by the lower court. If only Democratic appointees get cases that have fact patterns that favor a pro-plaintiff finding, then we should have observed a correlation between lower court case characteristics and the appellate panel.

Randomization: For a second potential SUTVA violation that affects causal interpretation, suppose case category is endogenous to the panel composition. That is, if

---

[90] *See* Richard L. Revesz, *Litigation and Settlement in the Federal Appellate Courts: Impact of Panel Selection Procedures on Ideologically Divided Courts*, 29 JOURNAL OF LEGAL STUDIES 685 (2000); *see also* Samuel P. Jordan, *Early Panel Announcement, Settlement and Adjudication,* 2007 BYU L. REV. 55 (2007).

judges selectively use the keywords "sexual harassment," the placement of a case into a particular category will reflect the judicial panel, undermining random assignment within a case category. If this was the case, we should have observed serial correlation in panel composition. We examined how similar the assignment of judges to 3-judge panels is to a random string of assignments. We considered characteristics of these strings, such as auto-correlation in appellate panel characteristics, as a test for possible endogeneity of case category.

Information: A third consideration, more like a potential violation of the exclusion restriction, also affects causal interpretation. Even if political and socioeconomic trends respond to appellate precedent, people still may simply be responding to the information provided by the decision on the political preferences of the judges. In other words, appellate decisions might not affect law per se, but only reveal how judges may decide future cases. To address this, we considered whether our analyses are robust to different specifications that address circuit-specific differences. When there are many judges in the circuit pool, any given decision reveals less information about how future panels might decide.

Case Selection: Finally, two concerns affect interpretation of the estimates but do not affect their validity. Case selection at the appellate level is ignorable for estimating internally valid treatment effects since the case selection happens before the judges are randomly assigned. Societal and litigant preferences are held constant across treatment and control.

Rational Expectations: Rational expectations by firms that the pro-plaintiff precedent is due to idiosyncratic variation in judicial assignment is not relevant as long as lower court judges follow appellate precedent. Note that firms are unlikely to respond to the judicial pool composition[91] since there are many kinds of cases that could affect firm behavior.[92]

---

[91] E.g., expected number of male Democrats (female Republicans) per seat.
[92] We also do not find a relationship between pro-plaintiff decisions and expected number of male Democratic appointees per seat or the expected number of female Republican appointees per seat.

VIII. Alternative Theories

In this section, we consider several alternative theories to explain our findings.

Tax on Hiring of Men: An alternative view of sexual harassment law is simply that it is a tax on the hiring of men, who are potential harassers.[93] A less sympathetic view is that sexual harassment law mandates a transfer from females to males. If sexual harassment law is a tax on the hiring of men, why didn't insider male outcomes fall relative to insider female outcomes?

Machismo: The insider-outsider model suggests intentional harassment, but the productivity of females could have been lower simply due to their unfamiliarity with the machismo culture prevalent before females were hired in substantial numbers. However, if this were the case, wages should not be related to the elimination of machismo, but we find that both male and female insider wages increase.

Compensating Differentials: Compensating differentials do not appear sufficient to explain the findings. If insider women were compensated for having to face sexual harassment, their wages would decline with as sexual harassment law developed; instead, their wages increase. Moreover, a pure compensating differentials story would not explain why outsider women join the work force. Wages inside the labor force would have adjusted downwards for the decrease in sexual harassment and outsider women on the margin would have been indifferent to entry.

Change in the Composition of the Female Labor Force: Perhaps in the absence of sexual harassment law, firms chose to hire less productive women and after forbidding sexual harassment, firms hired more productive women. This may be the case, but is insufficient by itself to explain the exit of insider males.

Mandated Benefit: While the mandated benefits view is not an alternative theory for the results, explaining why maternity mandates and the ADA had different effects from sexual harassment law is worth mentioning.[94] First, unlike maternity mandates, sexual harassment directly improves the productivity of women by making a better work

---

[93] The vast majority of sexual harassment cases are filed by women regarding sexual harassment by men.
[94] See footnote 17.

environment. Second, unlike ADA, it is far more difficult to determine precisely which women are likely to impose the costs of sexual harassment on a firm. 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. Third, unlike the cost of complying with the ADA or the federal requirement of providing of maternity mandates, the cost of compliance with sexual harassment law could be reduced by not hiring either the group being harassed or the group doing the harassing. Men were hired less as well as women hired more, unlike what happened with ADA. Fourth, the costs of sexual harassment law are quite high, although exact figures for all the direct and indirect costs are difficult to obtain. Some labor lawyers observe that the ADA further required large fixed costs upfront in physical infrastructure, whereas the fear of a Title VII suit is always looming.

IX. Conclusion

Interpreting anti-discrimination law to forbid sexual harassment has been a key contribution of feminist legal theory. 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 appellate 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. When, however, 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.

We remark on the application of the insider-outsider model of involuntary unemployment to anti-discrimination law and the regulation of labor markets more generally. Existing models of anti-discrimination law suggest that profit-maximizing firms have their own incentives not to discriminate (Becker 1971, Epstein 1995). Similar arguments could be applied in the context of sexual harassment law. Profit-maximizing firms should have their own incentive to forbid harassment and retain outsider workers at higher productivity, so forbidding sexual harassment with its accompanying litigation and human resources cost should have no effect and could be inefficient. However, there are at least three reasons for why firms might not forbid harassment on their own (Lindbeck and Snower 1988). First, insider employees may be risk averse; forbidding harassment could change the insider profit sharing scheme and thereby impose additional risk on insider employees, who then suffer a utility loss. The firm may be unable to compensate them for this loss. Second, an insider cooperates with entrants if his gains, a share of the additional profit resulting from his cooperation, exceed his losses in market power as his wage falls towards his reservation wage. However, this only happens if the firm relinquishes a share of gross profit, something that may make it a net loser compared to other firms. Agency and transaction costs may prevent Coasian bargaining between insiders and the firm. Then the firm has no incentive to implement the new contract. Third, there may be additional sources of labor turnover costs preventing firms from simply replacing all the insiders with outsiders. Entry of firms that hire the outsiders may not occur due to setup costs, capital market imperfections, scarcity of entrepreneurial skills, and reduction of product prices.

A legal regime equalizes the playing field across all firms when no firm by itself would have the incentive to forbid harassment. The fact that female labor market

46

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 that firms were not profit-maximizing and forbidding harassment on their own in the absence of sexual harassment law. As the workplace evolves, courts continue to struggle with exactly how to define sexual harassment and whether to include, for example, cyber-harassment (Franks 2011), as a form of sexual harassment. A recent nationally representative study, moreover, finds that 30% of students between 7th and 12th grade in the US experienced sexual harassment in the previous school year and 44% experienced in-person harassment. Negative effects such as absenteeism, poor sleep, and stomachaches were reported by 87% of the harassed students (American Association of University Women 2011). Our empirical analysis sheds light on these contemporary debates.[95]

In conclusion, 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, the general equilibrium effects are those which we would want to include, and the impulse response function 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.

---

[95] Men are also increasingly likely to bring sexual harassment claims, so the reduction in gender inequality may be more muted in the future.

REFERENCES

Acemoglu, Daron, and Joshua D. Angrist. 2001. "Consequences of Employment Protection? The Case of the Americans with Disabilities Act." *Journal of Political Economy*, 109(October): 915-57.

Akerlof, George and Rachel Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(August): 715-53.

American Association of University Women. 2011. "Crossing the Line: Sexual Harassment At School."

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90(430): 431-442.

Ashenfelter, Orley, and Alan Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review*, 84(Dec): 1157-1173.

Attributes of U.S. Federal Judges Database. 2011. The Judicial Research Initiative (JuRI) at the University of South Carolina. www.cas.sc.edu/poli/juri/attributes.htm (accessed May 16, 2011).

Autor, David H., John J. Donohue, and Stewart J, Schwab. 2006. "The Costs of Wrongful-Discharge Laws." *The Review of Economics and Statistics*, 88(May): 211-31.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2005. "Rising Wage Inequality: The Role of Composition and Prices." Working Paper no. 11628 (September), NBER, Cambridge, MA.

Baker, Douglas D., Robert A. Jackson, and Meredith A. Newman. 2003. "Sexual Harassment in the Federal Workplace." *Public Administration Review*, 63(July): 472-83.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo and Daniel Keniston. 2009. "Rajasthan Police Performance and Public Perceptions." Poverty Action Lab Report. http://www.povertyactionlab.org/sites/default/files/publications/118_Duflo _Rajasthan_Police_Performance.pdf.

Barondes, Royce. 2011. "Federal District Judge Gender and Reversals." http://ssrn.com/abstract=1640876.

Basu, Kaushik. 2003. "The Economics and Law of Sexual Harassment in the Workplace." *Journal of Economic Perspectives,* 17(Summer): 141-57.

Basu, Kaushik. 2004. *Prelude to Political Economy: A Study of the Social and Political Foundations of Economics.* U.K.: Oxford: Oxford University Press.

Becker, Gary. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(April): 169-217.

Beller, Andrea H. 1979. "The Impact of Equal Opportunity Laws on the Male-Female Earnings Differential." In *Women in the Labor Market*, ed. Cynthia Lloyd, Emily Andrews, and Curtis Gilroy, 203-30. New York: Columbia University Press.

Belloni, Alex., Victor Chernozhukov, Daniel L. Chen, and Chris Hansen. 2010. "Instrument Selection using LASSO" MIT mimeo.

Bester, C. Alan, Timothy G. Conley, and Christian B. Hansen. Forthcoming 2011. "Inference with Dependent Data Using Cluster Covariance Estimators." *Journal of Econometrics*.

Bird, Susan. 1975. "The Assignment of Cases to Federal District Court Judges." *Stanford Law Review*. 27(2): 475-487.

Bisom-Rapp, Susan. 2001a. "Fixing Watches with Sledgehammers: The Questionable Embrace of Employee Sexual Harassment Training by the Legal Profession." *University of Arkansas at Little Rock Law Review*, 24(April): 147–68.

Bisom-Rapp, Susan. 2001b. "An Ounce of Prevention Is a Poor Substitute for a Pound of Cure: Confronting the Developing Jurisprudence of Education and Prevention in Employment Discrimination Law." *Berkeley Journal of Employment and Labor Law*, 22(February): 1–46.

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

Blau, Francine D., and Lawrence M. Kahn. 2006. "The US Gender Pay Gap in the 1990s: Slowing Convergence." *Industrial & Labor Relations Review*, 60(October): 45-66.

Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science*, 54(2): 389-411.

Breyer, Stephen. 2004. Active Liberty: Interpreting Our Democratic Constitution. The Tanner Lecture on Human Values delivered at Harvard University November 17-19.

Brown, J. Robert and Allison Herren Lee, Esq. 2000. "The Neutral Assignment of Judges at the Court of Appeals." *Texas Law Review*, 78(April): 1037.

Cameron, A. Colin, Douglas Miller & Jonah B. Gelbach, 2008. "Bootstrap-Based Improvements for Inferences with Clustered Errors." Review of Economics and Statistics.

Card, David, and John E. DiNardo. 2002. "Technology and U.S. Wage Inequality: A Brief Look." *Federal Reserve Bank of Atlanta Economics Review*, 87(Third Quarter): 45-62.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics*, 215-224.

Chang, Tom, and Antoinette Schoar. 2007. "Judge Specific Differences in Chapter 11 and Firm Outcomes." Working Paper (June), MIT Sloan School of Management, NBER and CEPR, Cambridge, MA.

Charles, Kerwin K. and Jonathan Guryan. 2008. "Prejudice and Wages: An Empirical Assessment of Becker's *The Economics of Discrimination*." *Journal of Political Economy*, 116 (October): 773-809.

Chay, Kenneth Y. 1998. "The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employment Opportunity Act of 1972." *Industrial and Labor Relations Review*, 51 (July): 608-32.

Dixon, Rosalind. 2010. "Female Justices, Feminism & the Politics of Judicial Appointment: A Re-Examination." University of Chicago Law School mimeo.

Dobbin, Frank and Erin L. Kelly. 2007. "How to Stop Harassment: Professional Construction of Legal Compliance in Organizations." *American Journal of Sociology*, 112(January): 1203-43.

Eberts, Randall W., and Joe A. Stone. 1985. "Male-Female Differences in Promotions: EEO in Public Education." *Journal of Human Resources*, 20(Fall): 504-21.

Edelman, Lauren B. 1992. "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law." *American Journal of Sociology*, 97(May): 1531–76.

Edelman, Lauren B. 2002. "Legality and the Endogeneity of Law." Pp. 187–203 in *Legality and Community: On the Intellectual Legacy of Philip Selznick*, edited by Robert A. Kagan, Martin Krygier, and Kenneth I. Winston. Lanham, MD: Rowman & Littlefield.

Edelman, Lauren B., Steven E. Abraham, and Howard S. Erlanger. 1992. "Professional Construction of the Law: The Inflated Threat of Wrongful Discharge." *Law and Society Review*, 26(March): 47–84.

Edelman, Lauren B., Christopher Uggen, and Howard S. Erlanger. 1999. "The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth." *American Journal of Sociology*, 105(September): 406–54.

Ellman, Lisa Michelle, Cass. R. Sunstein, and David Schkade. 2003. "Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation. AEI-Brookings Joint Center for Regulatory Studies." Working Paper no. 03-9 (September), University of Chicago, Chicago, IL.

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

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

Franks, Mary Anne. 2011. "Unwilling Avatars: Idealism and Discrimination in Cyberspace." *Columbia Journal of Gender and Law*. Forthcoming.

Gennaioli, Nicola and Andrei Shleifer. 2007. "The Evolution of Common Law." *Journal of Political Economy*, 115(February): 43-68.

Goldman, Sheldon. 1997. Picking Federal Judges. New Haven, CT: Yale University Press.

Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *The American Economics Review*, 84(June): 622-42.

Haire, Susan, Donald Songer, and Stefanie Lindquist. 2003. "Appellate Court Supervision in the Federal Judiciary: A Hierarchical Perspective," *Law and Society Review*, 37(1): 143-168.

Hall, Matthew. 2010. "Randomness Reconsidered: Modeling Random Judicial Assignment in the U.S. Courts of Appeals." *Journal of Empirical Legal Studies* 7(3): 574-589.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1997. "Market Forces and Sex Discrimination." Working Paper no. 6321 (December), NBER, Cambridge, MA.

Hersch, Joni. 2011. "Compensating Differentials for Sexual Harasment." Research Paper no. 11-06, Vanderbilt Law & Economics Research Paper.

Jolls, Christine. 2000. "Accommodation Mandates." *Stanford Law Review*, 53(November): 223-306.

Jolls, Christine. 2001. "Anti-discrimination and Accommodation." *Harvard Law Review*, 115(December): 642-99.

Jolls, Christine, and J.J. Prescott. 2004. "Disaggregating Employment Protection: The Case of Disability Discrimination." Working Paper no. 10740 (September), NBER, Cambridge, MA.

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

Juliano, Ann and Stewart Schwab. 2001. "The Sweep of Sexual Harassment Cases." *Cornell Law Review*, 87(March): 548-602.

Klarman, Michael. 2004. *From Jim Crow to Civil Rights: The Supreme Court and the Struggle for Racial Equality.* New York: Oxford University Press.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *American Economic Review*, 96(June): 863-76.

Kritzer, Herbert M., and Robert E. Drechsel. 2011. "A Portrait of Local News Reporting of Civil Litigation." University of Minnesota Working Paper.

Law, David S. 2005. "Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit." *University of Cincinnati Law Review*, 73: 817.

Neumark, David, and Wendy A, Stock. 2006. "The Labor Market Effects of Race and Sex Discrimination Laws." *Economic Inquiry*, 44(July): 385-419.

Nielsen, Laura Beth, Robert L. Nelson, and Ryon Lancaster. 2010. "Individual Justice or Collective Legal Mobilization? Employment Discrimination Litigation in the Post-Civil Rights United States." *The Journal of Empirical Legal Studies*, (June): 24.

O'Neill, June. 2003. "The Gender Gap in Wages, circa 2000." *The American Economics Review*, 93(May): 309-14.

Olea, P. M., and C. Pflueger (2010): "Is F > 10 enough? TSLS Weak Instrument Bias with Heteroskedasticity and Autocorrelation," Harvard University mimeo.

Peresie, Jennifer L. 2005. "Female Judges Matter: Gender and Collegial Decision making in the Federal Appellate Courts." *Yale Law Journal*, 114(May): 1759-91.

Perongolo, Barbara, and Claudia Olivetti. 2008. "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps." CEPR Discussion Paper no. 5506.

Posner, Richard. 1998. "Against Constitutional Theory." *NYU Law Review*, 73(April): 1-22.

Revesz, Richard L. 2000. "Litigation and Settlement in the Federal Appellate Courts: Impact of Panel Selection Procedures on Ideologically Divided Courts." *Journal of Legal Studies*, 29:685.

Schultz, Vicki. 1990. "Telling Stories about Women and Work: Judicial Interpretations of Sex Segregation in the Workplace in Title VII Cases Raising the Lack of Interest Argument." *Harvard Law Review*, 103(June): 1749-1843.

Schultz, Vicki and Stephen Petterson. 1992. "Race, Gender, Work, and Choice: An Empirical Study of the Lack of Interest Defense in Title VII Cases Challenging Job Segregation." *University of Chicago Law Review*, 59(Summer): 1167-81.

Shoag, Daniel. 2011. "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns," Harvard University mimeo.

Steinbuch, Robert. 2009. "An Empirical Analysis of Reversal Rates in the Eighth Circuit during 2008," *Loyola of Los Angeles Law Review*, 43: 51-191.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in IV Regression." In *Identification and Inference for Economic Models: A Festschrift in Honor of Thomas Rothenberg*, eds. Donald W.K. Andrews, and James H. Stocks, 80-108. Cambridge, MA: Cambridge University Press.

Summers, Lawrence, 1989. "Some Simple Economics of Mandated Benefits." *American Economic Review*, 79(May): 177-83.

Summers, Lawrence, 2005. Speech at Conference on Diversifying the Science & Engineering Workforce at the National Bureau of Economic Research, January.

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

Wald, Patricia M. 1999. "A Response to Tiller and Cross," *Columbia Law Review*, 99:235.

Wallace, Cliffort. 2005. "Improving the Appellate Process Worldwide Through Maximizing Judicial Resources," *Virginia Journal of Transantional Law*, 38: 187.

Weinberger, Catherine, and Peter Kuhn. 2006. "The Narrowing of the U.S. Gender Earnings Gap, 1959-1999: A Cohort-Based Analysis." Working Paper no. 12115 (March), NBER, Cambridge, MA.

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** – US 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** - US 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** - US 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:
- Frequency of conduct
- Its Severity
- Whether the conduct is physically threatening or humiliating
- Or is a mere offensive utterance
- And whether the conduct unreasonably interferes with employees work performance
- 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.

Table 1 -- Summary Statistics

Circuit-Year Level	Mean [Standard Deviation]
Number of Judges	18.504 [7.356]
Number of Panels	0.996 [1.471]
Proportion of Circuit-Years with No Panels	51%
Proportion of Pro-Plaintiff Decisions when Circuit-Year has Panels	67%
Expected # of Females per Seat	0.117 [0.081]
Expected # of Democratic Appointees per per Seat	0.407 [0.121]
Expected # of Female Republican Appointees per Seat	0.035 [0.040]
Expected # of Male Democrat Appointees per Seat	0.326 [0.119]
N (circuit-years)	252

Table 2 -- Randomization Check: Orthogonality with Case Characteristics as Determined by Lower Court

Case Characteristics as Determined by Lower Court	Male Democratic appointees per seat (1)	Female Republican appointees per seat (2)
Direction of Lower Court Decision	0.0115	-0.171
	(0.0856)	(0.187)
P 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)
P suing under state law	0.0677	-0.283
	(0.0830)	(0.181)
P claims illegally denied promotion	-0.0591	-0.0465
	(0.0755)	(0.165)
P claims illegally not being hired	-0.0909+	0.105
	(0.0529)	(0.116)
P claims illegally fired	0.0460	-0.159
	(0.0961)	(0.210)
P claims unequal pay	-0.0235	-0.0868
	(0.0675)	(0.148)
P sued under 14th Amendment	0.0606	-0.167+
	(0.0429)	(0.0938)
P 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)
P not victim of harassment	-0.0312	-0.338
	(0.131)	(0.340)
P is appellant	-0.109	-0.349
	(0.208)	(0.541)

Notes: Heteroskedasticity-robust standard errors are in parentheses. Each coefficient represents a separate regression of a distinct case characteristic on male Democratic (female Republican) appointees per seat. "P" refers to plaintiff. + Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 3 -- 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 Dn for Various P

n = 9		n = 12	
Prob(sqr(n)*Dn<b)	b/sqr(n)	Prob(sqr(n)*Dn<b)	b/sqr(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

Panel A: Judge Level	Outcome: Pro-plaintiff Decision								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.0332		-0.0785+						
	(0.0394)		(0.0416)						
Democratic Appointee		0.131**	0.143**						
		(0.0358)	(0.0371)						
Female Republican Appointee				-0.183*		-0.145+	-0.122+	-0.112+	-0.117
				(0.0757)		(0.0739)	(0.0616)	(0.0612)	(0.0659)
Male Democratic Appointee					0.130**	0.121**	0.110**	0.122**	0.111*
					(0.0361)	(0.0362)	(0.0353)	(0.0377)	(0.0373)
Circuit-year controls	N	N	N	N	N	N	Fixed Effects	Expectations	Both
F-statistic	0.71	13.48	7.81	5.87	12.89	7.88	6.06	6.45	5.02
N	752	752	752	752	752	752	752	752	752
R-sq	0.001	0.018	0.021	0.006	0.016	0.019	0.112	0.026	0.112

Panel B: Case Level	Outcome: Pro-Plaintiff Decision					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Republican Appointees	-0.553*		-0.434+	-0.418+	-0.379+	-0.407
per Seat	(0.200)		(0.217)	(0.211)	(0.208)	(0.234)
Male Democratic Appointees		0.322**	0.286*	0.285*	0.314**	0.298*
per Seat		(0.0876)	(0.0933)	(0.103)	(0.0967)	(0.110)
Circuit-year controls	N	N	N	Fixed Effects	Expectations	Both
F-statistic of instruments	7.66	13.48	7.91	6.17	7.33	4.61
N	251	251	251	251	251	251
R-sq	0.019	0.032	0.043	0.126	0.046	0.128

Panel C: Circuit-Year Level	Outcome: Proportion of Pro-Plaintiff Decisions								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female Republican Appointees	-1.015**		-0.848**	-1.015**		-0.848**	-0.810**	-0.804**	-0.845**
per Seat	(0.224)		(0.256)	(0.224)		(0.256)	(0.217)	(0.221)	(0.199)
Male Democratic Appointees		0.513**	0.445*		0.513**	0.445*	0.459*	0.455*	0.494**
per Seat		(0.164)	(0.165)		(0.164)	(0.165)	(0.150)	(0.155)	(0.145)
Circuit-years with no cases	Dropped	Dropped	Dropped	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied
Circuit-year controls	N	N	N	N	N	N	Fixed Effects	FE, Expect	FE, Trends
F-statistic of instruments	20.53	9.84	22.46	20.53	9.84	22.56	63.92	41.23	32.95
N	124	124	124	252	252	252	252	252	252
R-sq	0.067	0.080	0.125	0.605	0.611	0.630	0.669	0.669	0.706

Panel D: Analysis Level Firm HR Policies	Outcome: Proportion of Pro-Plaintiff Decisions						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female Republican Appointees	-1.284*		-0.874	-0.871	-0.593	-0.600	-0.528
per Seat	(0.568)		(0.508)	(0.510)	(0.428)	(0.409)	(0.436)
Male Democratic Appointees		0.663**	0.535**	0.534**	0.509**	0.523**	0.479**
per Seat		(0.174)	(0.130)	(0.131)	(0.0990)	(0.0943)	(0.102)
Circuit-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied
Firm-level controls	N	N	N	Y	Y	Y	Y
Circuit-year controls	N	N	N	N	Fixed Effects	FE, Expect	FE, Trends
F-statistic of instruments	5.12	14.52	10.30	10.15	13.95	16.01	11.09
N	5584	5584	5584	5584	5584	5584	5584
R-sq	0.647	0.663	0.679	0.680	0.733	0.735	0.751

Panel D: Analysis Level Individual MORG CPS	Outcome: Proportion of Pro-Plaintiff Decisions						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Female Republican Appointees	-1.163**		-0.973**	-0.971**	-0.839**	-0.849**	-0.767**
per Seat	(0.284)		(0.283)	(0.281)	(0.167)	(0.174)	(0.182)
Male Democratic Appointees		0.524**	0.441*	0.442*	0.467**	0.466**	0.517**
per Seat		(0.157)	(0.151)	(0.151)	(0.147)	(0.152)	(0.155)
Circuit-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied
Individual-level controls	N	N	N	Y	Y	Y	Y
Circuit-year controls	N	N	N	N	Fixed Effects	FE, Expect	FE, Trends
F-statistic of instruments	16.77	11.22	16.96	17.11	50.31	32.32	38.12
N	5429470	5429470	5429470	5418564	5418564	5418564	5418564
R-sq	0.627	0.628	0.652	0.652	0.691	0.691	0.711

Notes: Heteroskedasticity-robust standard errors are in parentheses. Observations are clustered at the circuit level. Fixed effects are dummy indicators for circuit and for year. Expectations are the expected number of Female Republican Appointees and Male Democratic Appointees per seat. Trends are circuit-specific time trends. In Panels C and D, proportions of pro-plaintiff decisions and judicial type per seat during circuit-years with no cases are defind to be 0.
+ Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 5 -- Falsification Test of Instrument: Relationship Between Pro-Plaintiff Sexual Harassment Appellate Decisions and Composition of Sexual Harassment Panels in Other Years, 1982-2002

Circuit-Year Level	Outcome: Proportion of Pro-Plaintiff Decisions, t			
	(1)	(2)	(3)	(4)
Female Republican Appointees	-0.771*	-0.888**	-0.828**	-0.714*
per Seat, t	(0.249)	(0.226)	(0.178)	(0.249)
Male Democratic Appointees	0.442*	0.375*	0.433*	0.460**
per Seat, t	(0.146)	(0.147)	(0.154)	(0.142)
Female Republican Appointees	-0.226	-0.292		
per Seat, $t-1$	(0.320)	(0.374)		
Male Democratic Appointees	0.0486	0.00643		
per Seat, $t-1$	(0.120)	(0.107)		
Female Republican Appointees		-0.0758		
per Seat, $t-2$		(0.263)		
Male Democratic Appointees		-0.0948		
per Seat, $t-2$		(0.181)		
Female Republican Appointees			-0.301	-0.174
per Seat, $t+1$			(0.173)	(0.203)
Male Democratic Appointees			-0.196	-0.0741
per Seat, $t+1$			(0.134)	(0.129)
Female Republican Appointees				-0.0118
per Seat, $t+2$				(0.198)
Male Democratic Appointees				0.127
per Seat, $t+2$				(0.113)
Circuit-years with no cases	Dummied	Dummied	Dummied	Dummied
Circuit-year controls	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
N	240	228	240	228
R-sq	0.660	0.665	0.702	0.713

Notes: Heteroskedasticity-robust standard errors are in parentheses. Observations are clustered at the circuit level. Fixed effects are dummy indicators for circuit and for year. Proportions of pro-plaintiff decisions and judicial type per seat during circuit-years with no cases are defind to be 0. + Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 6 -- The Effect of Sexual Harassment Law on Human Resources Sexual Harassment Policy

	Presence of Sexual Harassment Policy			
	(1)	(2)	(3)	(4)
Proportion Pro-Plaintiff Appellate Decisions$_{t+1}$	-0.00108 (0.0142)	0.0374 (0.0252)	0.0279 (0.0233)	0.0353 (0.0354)
Proportion Pro-Plaintiff Appellate Decisions$_t$	0.00984 (0.0135)	0.0231 (0.0316)	-0.0167 (0.0443)	0.0282 (0.0291)
Proportion Pro-Plaintiff Appellate Decisions$_{t-1}$	0.0157 (0.0215)	0.0408+ (0.0244)	0.0455* (0.0224)	0.0223 (0.0687)
Proportion Pro-Plaintiff Appellate Decisions$_{t-2}$	0.0291 (0.0188)	0.0210 (0.0507)	0.0978+ (0.0503)	-0.0102 (0.0542)
Proportion Pro-Plaintiff Appellate Decisions$_{t-3}$	0.0616** (0.0153)	0.117** (0.0405)	0.0612 (0.0413)	0.112* (0.0510)
Proportion Pro-Plaintiff Appellate Decisions$_{t-4}$	0.0494* (0.0162)	0.0542 (0.0545)	-0.00609 (0.0497)	0.146** (0.0502)
Proportion Pro-Plaintiff Appellate Decisions$_{t-5}$	-0.00874 (0.0154)	0.0517 (0.0439)	0.0533+ (0.0290)	0.103 (0.0677)
Controls	Y	Y	Y	Y
IV	N	Y	Lasso IV	Lasso IV, District IV
Mean dependant variable	0.543	0.543	0.543	0.646
Average lag	0.029	0.057	0.050	0.075
Joint F lags	9.34	24.09	30.20	13.70
Joint F leads	0.01	2.20	1.44	1.44
N	4014	4014	4014	2617
R-sq	0.260	0.259	0.257	0.121

Notes: Heteroskedasticity-robust standard errors are in parentheses. Observations are clustered at the circuit level. These regressions include firm-related controls (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 (this is lagged as are the measure of sexual harassment law and instruments).   + Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 7 -- The Effect of Sexual Harassment Law on Gender Inequality

	Employment Status		Hours Worked		Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Pro-Plaintiff	-0.000709	0.00549	0.0690	0.527	0.00981	0.0717
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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
Appellate 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*
Appellate 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**
Appellate 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: 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 (this is interacted with the female dummy as are the measure of sexual harassment law and instruments). + Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 8 -- Robustness of IV Estimates

| | The Effect of Sexual Harassment Law on Female Employment Share | |
| | Average of interaction lags | Joint F of interaction lags |
	(1)	(2)
A. Add Circuit-Specific Trends	0.016	8.35
B. No Fixed Effects	0.016	8.17
C. No Control Variables (except Presence of Appellate Case, Presence Interacted with Gender, and Gender)	0.017	8.08
D. Control for Expectation	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 to 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

Table 9 -- The Effect of Sexual Harassment Law on Insiders

	Earnings (1)	Hours Worked (2)	Management (3)
		Labor Force Partipants Reporting Non-Zero Wages	
Proportion Pro-Plaintiff Appellate Decisions$_{t+1}$	0.0256* (0.0101)	0.232 (0.244)	0.000853 (0.00182)
Proportion Pro-Plaintiff Appellate Decisions$_{t+1}$ * Female	-0.0208 (0.0135)	-0.381 (0.441)	-0.00212 (0.00464)
Proportion Pro-Plaintiff Appellate Decisions$_t$	0.0257 (0.0206)	0.0221 (0.212)	0.00545+ (0.00306)
Proportion Pro-Plaintiff Appellate Decisions$_t$ * Female	-0.0243 (0.0283)	-0.463 (0.391)	-0.0115* (0.00480)
Proportion Pro-Plaintiff Appellate Decisions$_{t-1}$	0.0299* (0.0118)	-0.0781 (0.185)	0.00486* (0.00233)
Proportion Pro-Plaintiff Appellate Decisions$_{t-1}$ * Female	-0.0343* (0.0140)	-0.191 (0.257)	-0.00873* (0.00403)
Proportion Pro-Plaintiff Appellate Decisions$_{t-2}$	-0.00465 (0.0138)	-0.306 (0.196)	0.00240 (0.00219)
Proportion Pro-Plaintiff Appellate Decisions$_{t-2}$ * Female	0.0126 (0.0175)	0.485 (0.419)	-0.00326 (0.00568)
Proportion Pro-Plaintiff Appellate Decisions$_{t-3}$	-0.00226 (0.0222)	0.343 (0.245)	-0.0000238 (0.00283)
Proportion Pro-Plaintiff Appellate Decisions$_{t-3}$ * Female	0.00444 (0.0184)	-0.328 (0.287)	-0.00636 (0.00569)
Proportion Pro-Plaintiff Appellate Decisions$_{t-4}$	0.00107 (0.0169)	0.186 (0.164)	0.00276 (0.00221)
Proportion Pro-Plaintiff Appellate Decisions$_{t-4}$ * Female	0.00688 (0.00926)	-0.389* (0.177)	-0.00665* (0.00285)
Proportion Pro-Plaintiff Appellate Decisions$_{t-5}$	0.0150 (0.0291)	0.370 (0.363)	0.00358+ (0.00205)
Proportion Pro-Plaintiff Appellate Decisions$_{t-5}$ * Female	-0.00821 (0.0330)	-0.379 (0.486)	-0.00876** (0.00260)
Controls	Y	Y	Y
IV	Y	Y	Y
Mean dependant variable - Male	6.298	42.90	0.144
Mean dependant variable - Female	5.854	36.04	0.120
Average interaction lag	-0.004	-0.160	-0.007
Average level effect lag	0.008	0.103	0.003
Joint F of interaction lags	35.16	58.21	30.38
Joint F of interaction leads	2.38	0.75	0.21
Joint F of level effect lags	24.71	13.10	6.25
Joint F of level effect leads	6.47	0.90	0.22
N	2424997	2622664	2755279
R-sq	0.296	0.081	0.057

Notes: 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. 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 (this is interacted with the female dummy as are the measure of sexual harassment law and instruments). + Significant at 10%; * Significant at 5%; ** Significant at 1%

Table 10 -- The Effect of Sexual Harassment Law on the Construction Industry

	Construction	Non-Construction	Construction	Non-Construction	Construction	Non-Construction	Construction	Non-Construction
	Employment Status		Hours Worked		Earnings		Management	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion Pro-Plaintiff	0.00680	0.0118	0.408	0.617+	0.0830	0.0917*	0.00527	0.000504
Appellate Decisions$_{t+1}$	(0.0116)	(0.00777)	(0.595)	(0.348)	(0.0892)	(0.0466)	(0.00601)	(0.00211)
Proportion Pro-Plaintiff	-0.0172	-0.0137	-0.996	-0.725	-0.184	-0.0854	-0.0357+	-0.00120
Appellate Decisions$_{t+1}$ * Female	(0.0307)	(0.0130)	(0.775)	(0.573)	(0.170)	(0.0722)	(0.0197)	(0.00490)
Proportion Pro-Plaintiff	-0.000151	0.00362	-0.386	0.192	-0.00504	0.0405	-0.00638	0.00684*
Appellate Decisions$_t$	(0.00955)	(0.00643)	(0.562)	(0.310)	(0.0775)	(0.0402)	(0.00664)	(0.00344)
Proportion Pro-Plaintiff	-0.0524+	-0.00586	-2.086	-0.653	-0.358	-0.0461	-0.00621	-0.0128*
Appellate Decisions$_t$ * Female	(0.0295)	(0.0138)	(1.276)	(0.584)	(0.223)	(0.0807)	(0.0217)	(0.00500)
Proportion Pro-Plaintiff	-0.0126	0.00516	-1.578	0.129	-0.107	0.0441	0.00272	0.00620**
Appellate Decisions$_{t-1}$	(0.0156)	(0.00788)	(0.995)	(0.381)	(0.129)	(0.0457)	(0.00299)	(0.00219)
Proportion Pro-Plaintiff	-0.000377	-0.0178	0.942	-0.771	-0.0171	-0.114	-0.00183	-0.0104**
Appellate Decisions$_{t-1}$ * Female	(0.0243)	(0.0149)	(1.396)	(0.600)	(0.170)	(0.0823)	(0.0185)	(0.00354)
Proportion Pro-Plaintiff	-0.0186	0.00813	-1.254	-0.0238	-0.153	0.0302	0.000524	0.00288
Appellate Decisions$_{t-2}$	(0.0164)	(0.0114)	(0.831)	(0.533)	(0.122)	(0.0726)	(0.00682)	(0.00275)
Proportion Pro-Plaintiff	-0.0367+	-0.0255	-1.295	-0.444	-0.292	-0.116	-0.0620*	-0.00148
Appellate Decisions$_{t-2}$ * Female	(0.0211)	(0.0204)	(1.348)	(0.904)	(0.193)	(0.119)	(0.0266)	(0.00667)
Proportion Pro-Plaintiff	-0.0187	0.00279	-0.498	0.368	-0.124	0.00775	0.00713	-0.00159
Appellate Decisions$_{t-3}$	(0.0120)	(0.00832)	(0.788)	(0.459)	(0.0931)	(0.0602)	(0.0119)	(0.00260)
Proportion Pro-Plaintiff	0.000774	-0.00979	1.428	-0.598	0.0677	-0.0384	-0.0520*	-0.00337
Appellate Decisions$_{t-3}$ * Female	(0.0418)	(0.0109)	(1.546)	(0.442)	(0.263)	(0.0688)	(0.0239)	(0.00533)
Proportion Pro-Plaintiff	-0.0148	-0.00301	-0.699	0.0529	-0.147+	-0.0245	0.00652	-0.000364
Appellate Decisions$_{t-4}$	(0.00967)	(0.00691)	(0.536)	(0.346)	(0.0798)	(0.0555)	(0.00678)	(0.00253)
Proportion Pro-Plaintiff	-0.00615	0.000147	-1.469	-0.344	-0.0549	0.0142	-0.0215	-0.00236
Appellate Decisions$_{t-4}$ * Female	(0.0361)	(0.00986)	(1.978)	(0.329)	(0.304)	(0.0589)	(0.0352)	(0.00288)
Proportion Pro-Plaintiff	-0.0246	-0.0141	-1.194	-0.0913	-0.129	-0.0693	-0.0111+	0.00384
Appellate Decisions$_{t-5}$	(0.0217)	(0.0114)	(1.104)	(0.379)	(0.166)	(0.0728)	(0.00633)	(0.00266)
Proportion Pro-Plaintiff	0.0907*	0.0175	2.284**	0.210	0.515*	0.0884	0.0537+	-0.0117**
Appellate Decisions$_{t-5}$ * Female	(0.0359)	(0.0176)	(0.873)	(0.732)	(0.223)	(0.110)	(0.0326)	(0.00315)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
IV	Y	Y	Y	Y	Y	Y	Y	Y
Mean dependant variable - Male	0.836	0.892	33.91	38.31	4.977	5.517	0.174	0.129
Mean dependant variable - Female	0.793	0.826	27.45	29.47	4.385	4.775	0.106	0.098
Average interaction lag	0.010	-0.007	0.378	-0.389	0.044	-0.033	-0.017	-0.006
Average level effect lag	-0.018	0.000	-1.045	0.087	-0.132	-0.002	0.001	0.002
Joint F of interaction lags	13.66	2.99	25.83	26.27	10.18	3.14	11.13	32.54
Joint F of interaction leads	0.31	1.13	1.70	1.64	1.14	1.39	3.22	0.07
Joint F of level effect lags	19.27	3.51	49.25	1.45	20.09	2.86	5.94	15.43
Joint F of level effect leads	0.28	2.17	0.45	3.04	0.77	3.68	0.71	0.07
N	210153	2949731	201678	2825198	163297	2666305	210153	2949731
R-sq	0.048	0.044	0.079	0.094	0.071	0.108	0.107	0.059

Notes: All regressions use MORG CPS and restrict to individuals between the ages of 18 and 65 and who report an industry category. 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 (this is interacted with the female dummy as are the measure of sexual harassment law and instruments). + Significant at 10%; * Significant at 5%; ** Significant at 1%

Figure 1

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

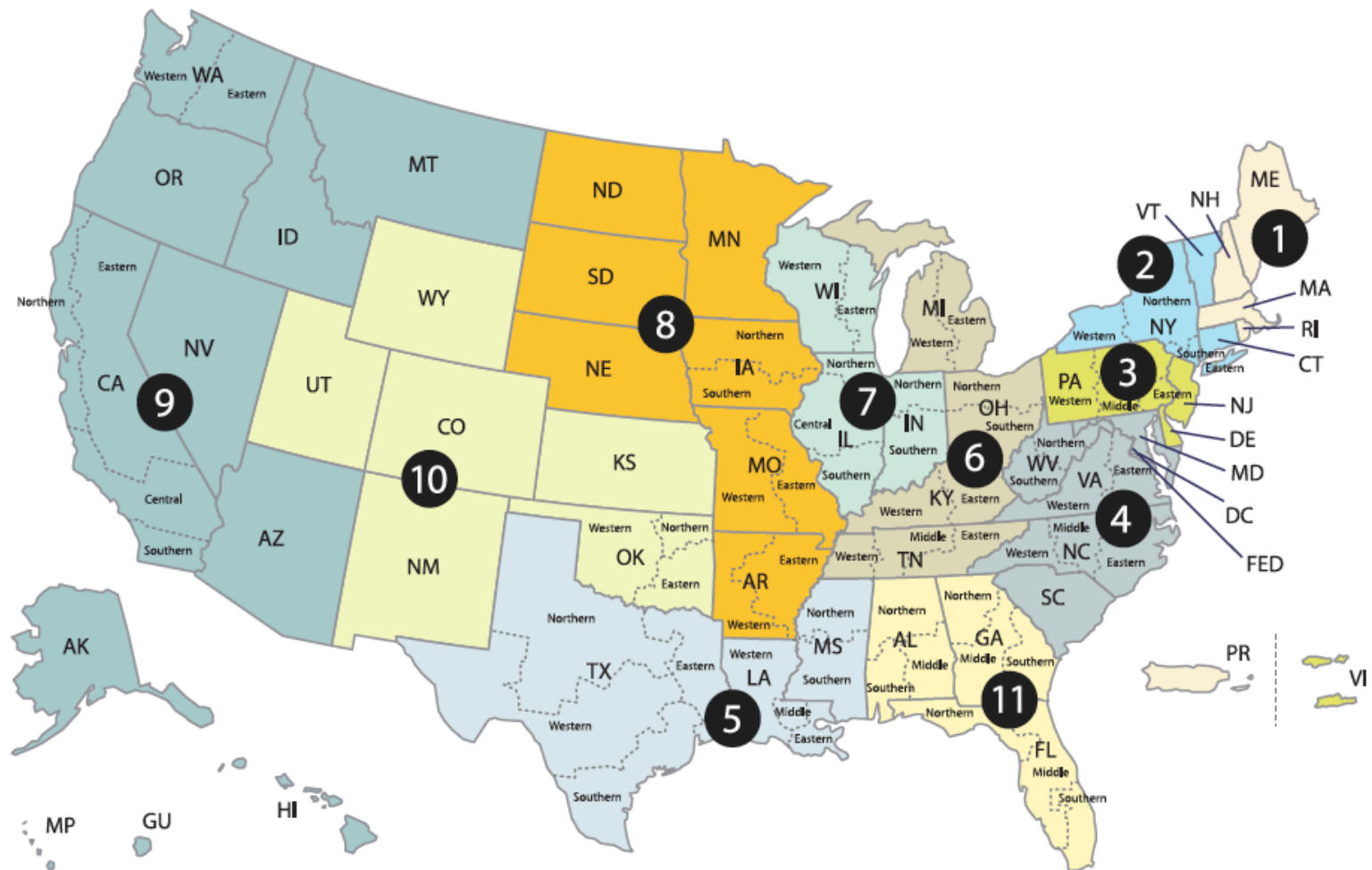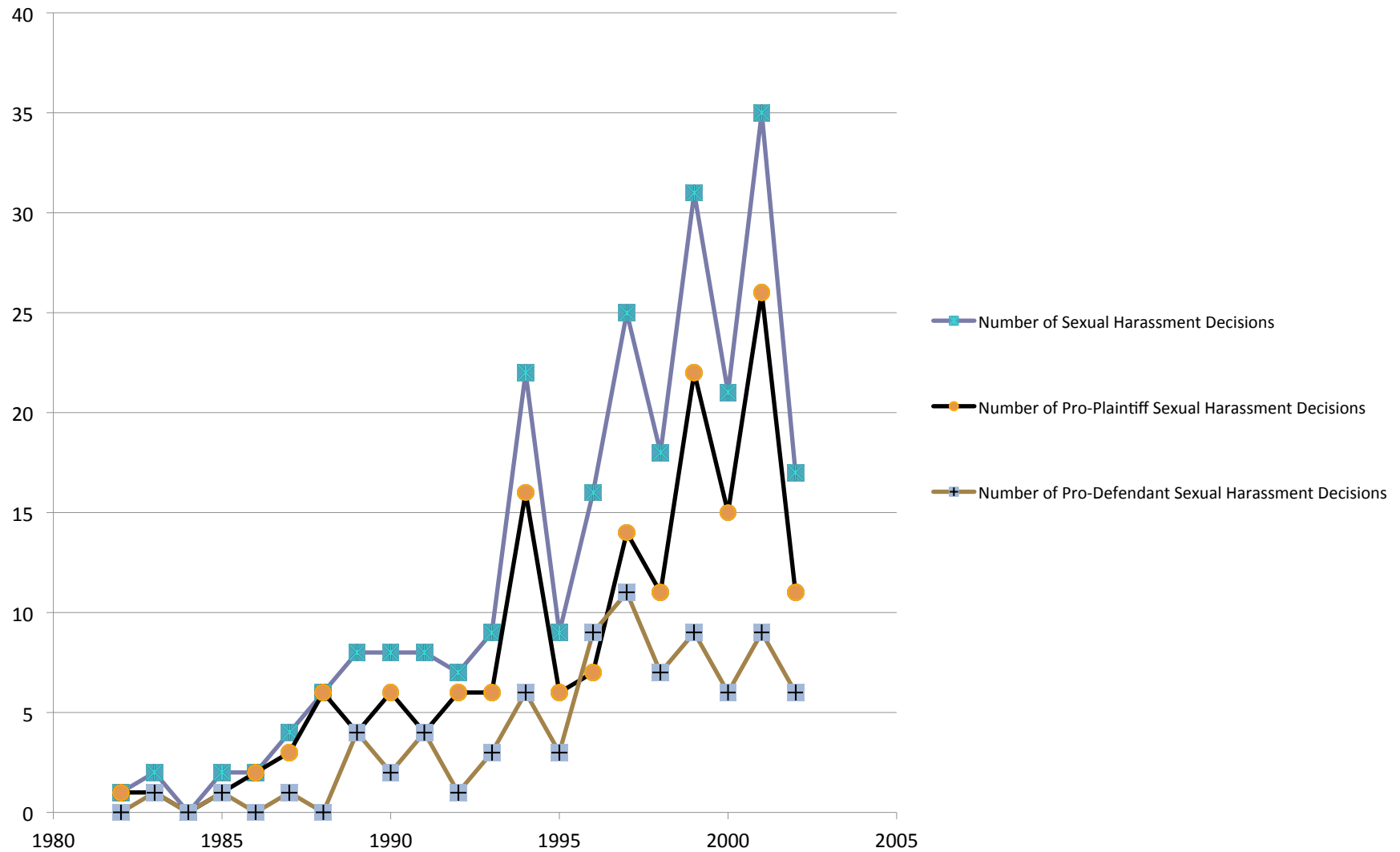Figure 2: Sexual Harassment Decisions, 1982-2002

- Number of Sexual Harassment Decisions
- Number of Pro-Plaintiff Sexual Harassment Decisions
- Number of Pro-Defendant Sexual Harassment Decisions

Figure 3A

Random Variation by Circuit -- Male Democrat

Actual # of Male Democrat Appointees per seat

Expected # of Male Democrat Appointees per seat

Year
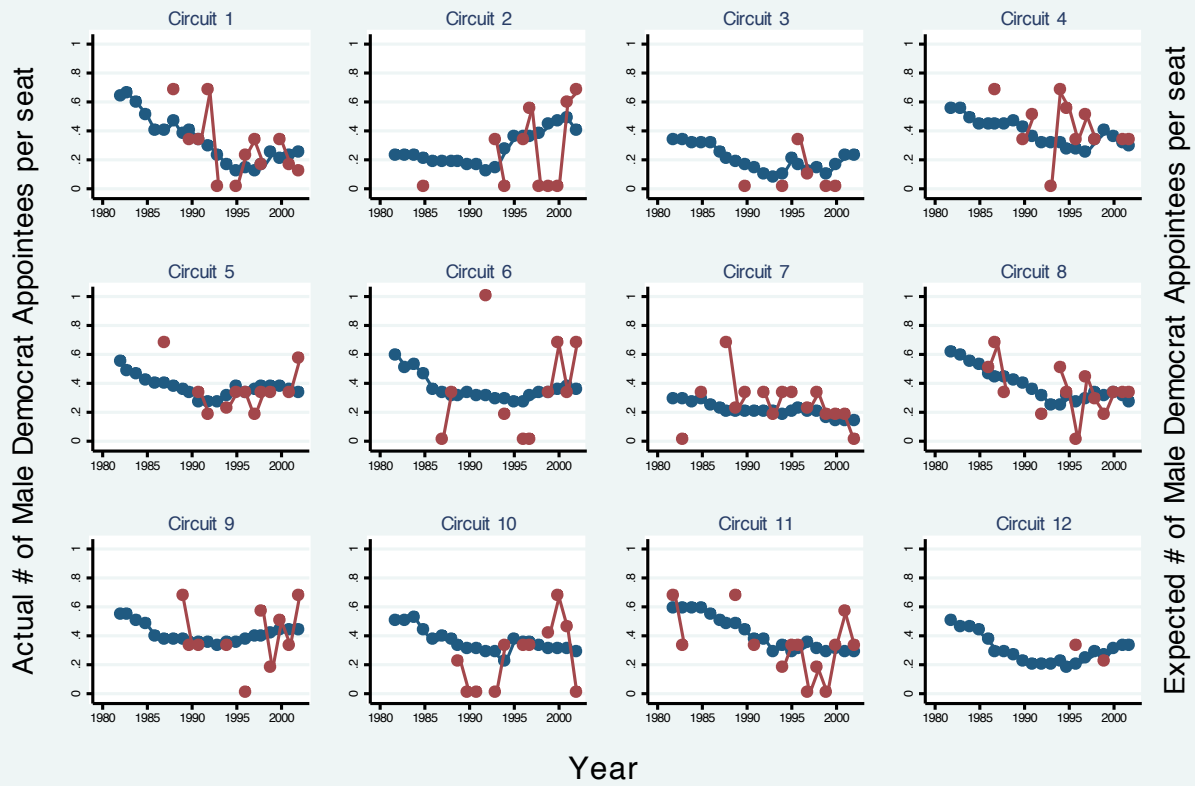
Figure 3B

Random Variation by Circuit -- Female Republican

Actual # of Female Republican Appointees per seat

Expected # of Female Republican Appointees per seat

Year
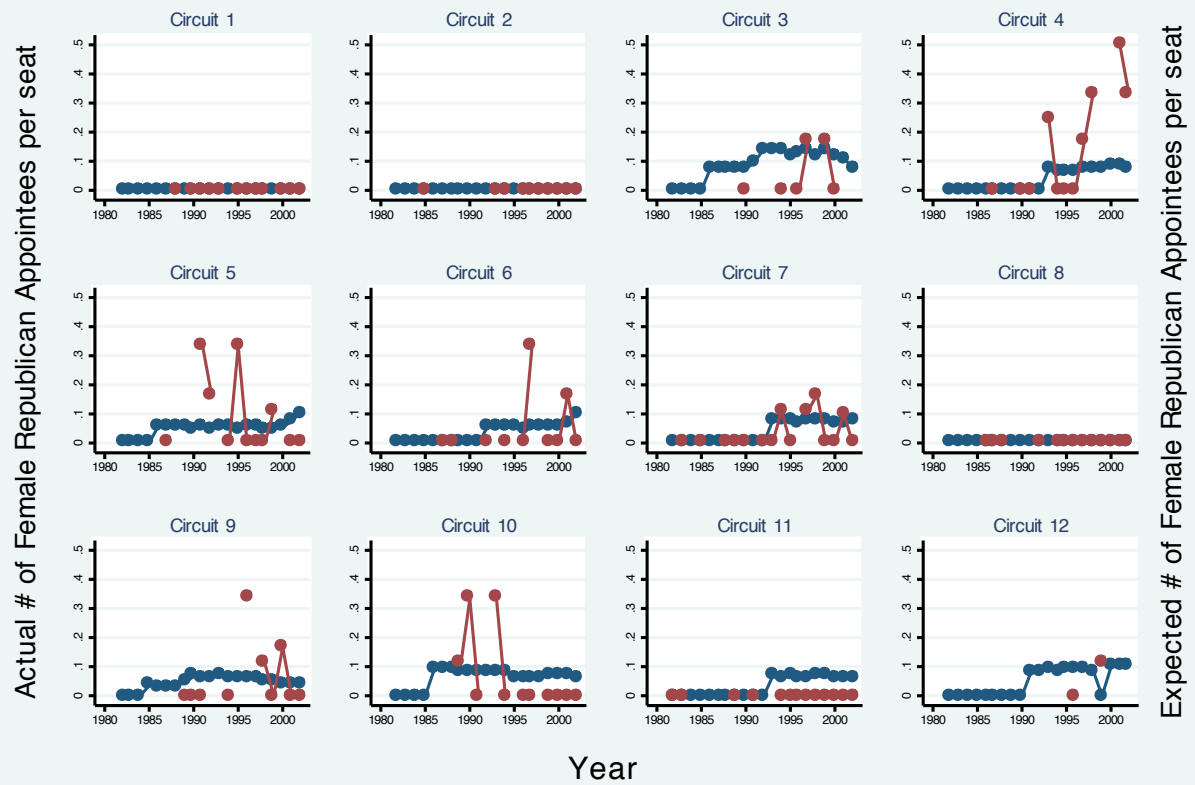
Figure 4: Randomization Check

P-Values of Male Democrat strings   P-Values of Female Republican strings
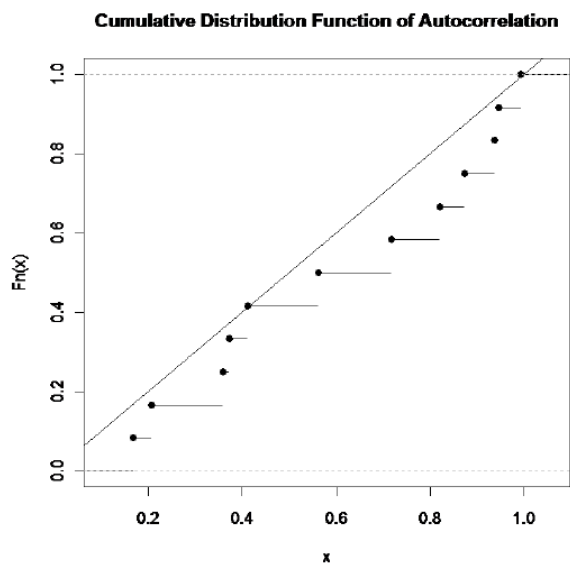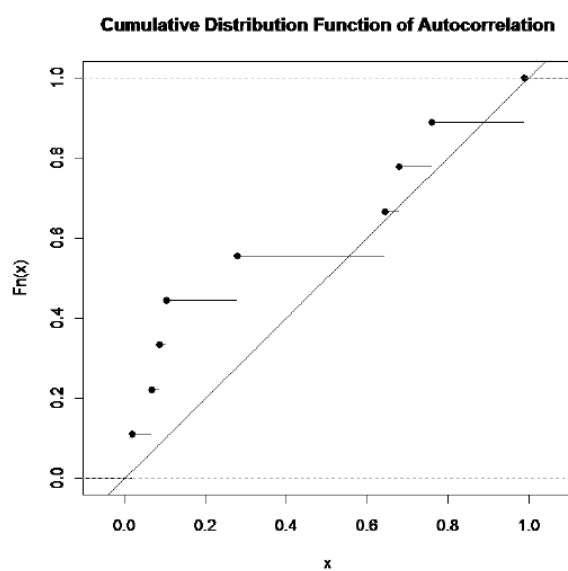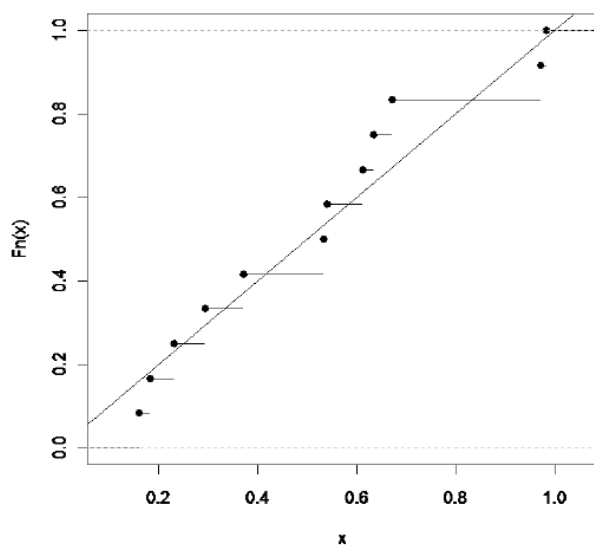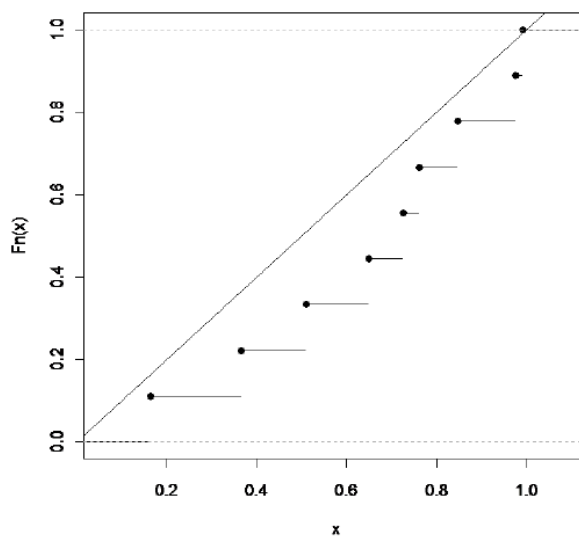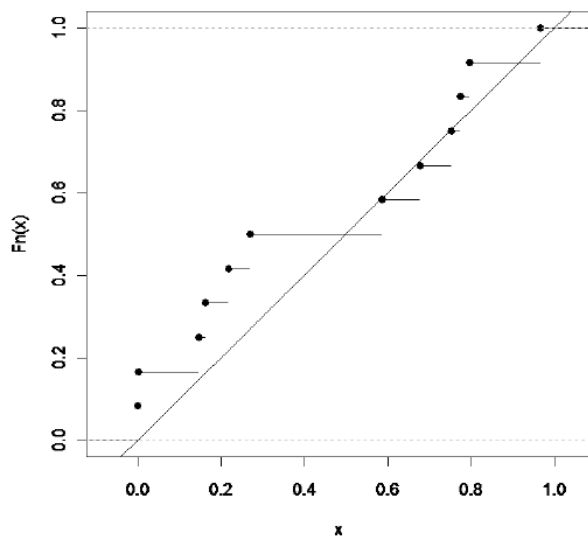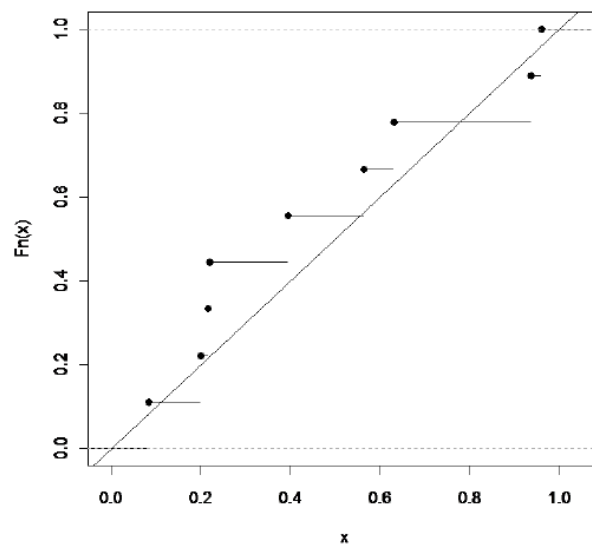
Figure 5A: Pro-Plaintiff Sexual Harassment Decisions and Composition of Judicial Panel
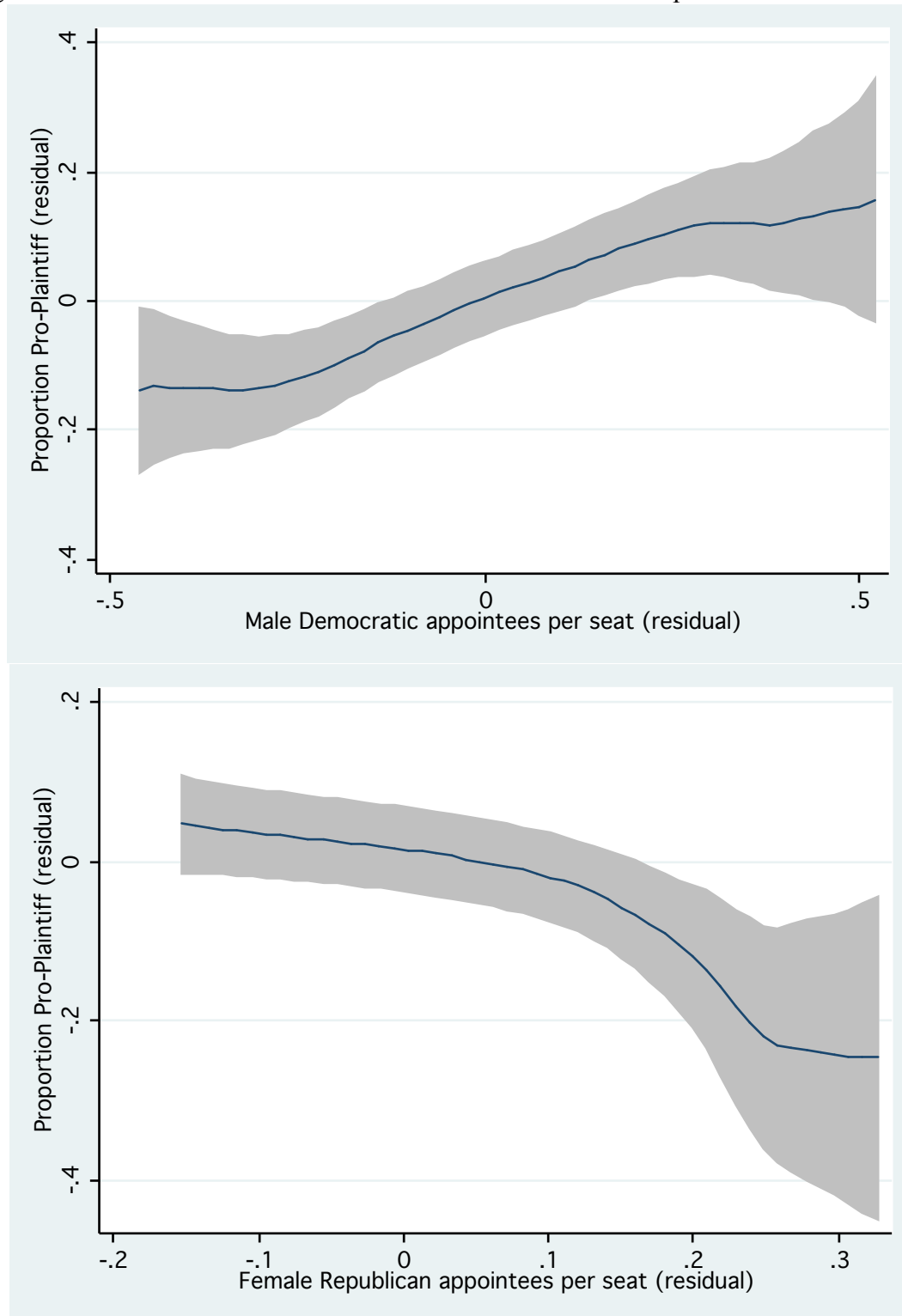

Nonparametric local polynomial estimates are computed using an Epanechnikov kernel. The bandwidth is 0.14 for the male Democrat graph and 0.09 for the female Republican graph. Shaded area indicates 90 percent confidence bands. The residuals are calculated removing circuit and year fixed effects.

Figure 6: Sexual Harassment Law and Labor Markets

Dynamic Response following Pro-Plaintiff Decisions on Presence of Sexual Hararassment Policies
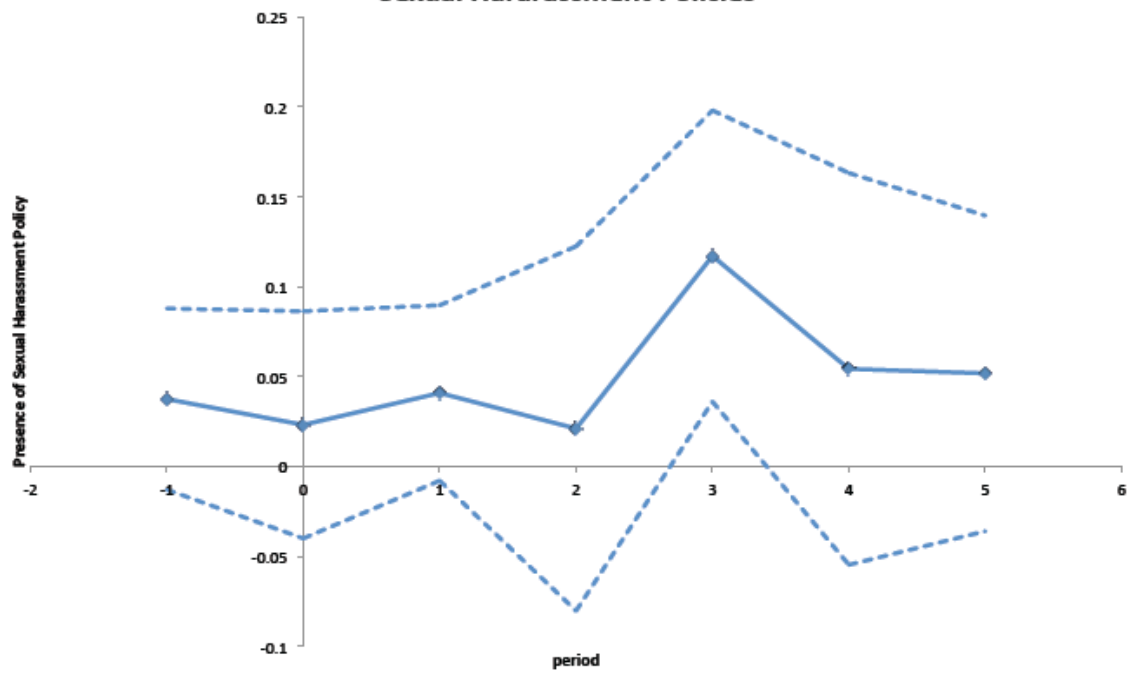
Dynamic Response following Pro-Plaintiff Decisions on Female-Male Employment Inequality
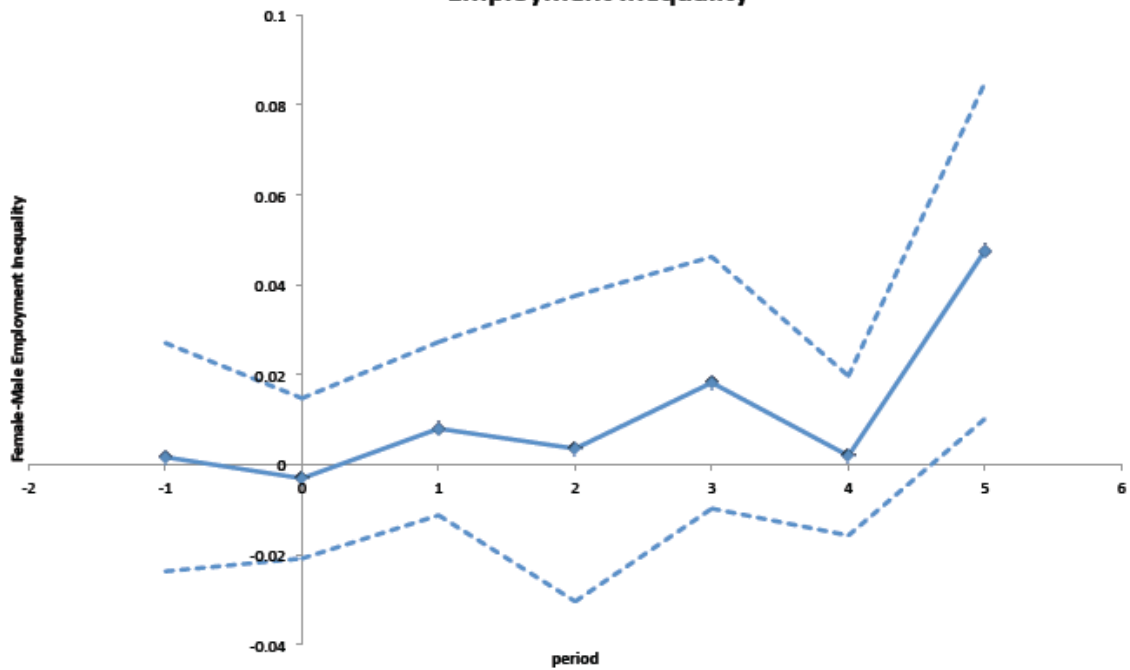
**Appendix Table 1 -- First Stage Relationship Between Absence of Appellate Cases
and Judicial Composition of District Court Cases, 1986-2002**

Circuit-Year Level	Outcome: Absence of Appellate Cases, t			
	(1)	(2)	(3)	(4)
Black judges per District case, t	0.470*	0.485**	0.419*	0.459**
	(0.176)	(0.138)	(0.145)	(0.144)
Black judges per District case, $t+1$		-0.0492	-0.0248	
		(0.369)	(0.357)	
Black judges per District case, $t+2$			-0.0192	
			(0.374)	
Circuit-year controls	N	N	N	Fixed Effects
N	203	190	177	203
R-sq	0.019	0.020	0.014	0.372

Notes: Heteroskedasticity-robust standard errors are in parentheses. Observations are clustered at the circuit level. Fixed effects are dummy indicators for circuit and for year. + Significant at 10%; * Significant at 5%; ** Significant at 1%

Appendix Figure 1: Composition of District Panels, Absence of Appellate Panels, and Composition of Appellate Panels
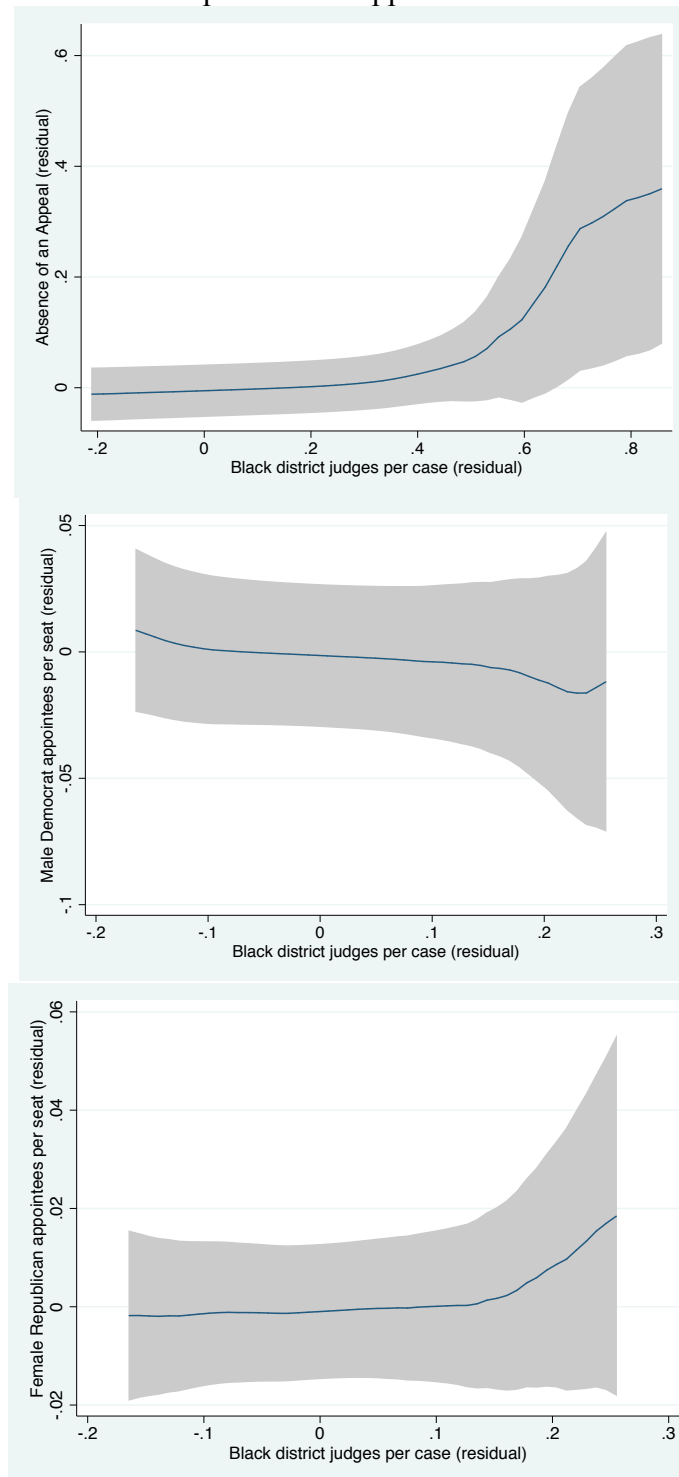
Nonparametric local polynomial estimates are computed using an Epanechnikov kernel. The bandwidth is 0.24 for the absence of appellate panels graph and 0.11 and 0.09 for the male Democrat and female Republican graph. Shaded area indicates 90 percent confidence bands. The residuals are calculated removing circuit and year fixed effects.