

The Impact of Oral Argument Attendance

Damon Cann
Professor
Department of Political Science
Utah State University
damon.cann@usu.edu

Greg Goelzhauser
Professor
Department of Political Science
Utah State University
greg.goelzhauser@usu.edu*

Abstract

How does oral argument attendance impact public perceptions of the judiciary? Judicial independence is partly contingent on public support, but the conditions that generate institutional good will are not well understood. We examine how judicial outreach and court exposure inform public attitudes. Leveraging a field-experiment randomizing in-person attendance at oral arguments conducted by a federal circuit court of appeals on a university campus, we find that exposure increases perceptions of institutional legitimacy and the extent to which judicial decisions are motivated by law versus politics. The results have important implications for judicial politics and policy debates concerning reform initiatives involving circuit riding, courtroom cameras, and public outreach.

*Thanks to participants at the 2020 QuantLaw Conference hosted by the University of Arizona College of Law, and colloquium participants at Brigham Young University, for helpful feedback. Replication files are available on Dataverse.

Introduction

How does oral argument attendance impact public perceptions of the judiciary? Judicial independence is partly contingent on public support (Caldeira 1986; Stephenson 2004; Vanberg 2004), but the conditions that generate institutional good will are not well understood. The public can provoke political attacks on courts that threaten judicial independence (Clark 2009; Helmke 2010; Staton 2004). To forge support, courts can, for example, decide cases consistent with public opinion (Casillas, Enns, and Wohlfarth 2011; Epstein and Martin 2010; McGuire and Stimson 2004) or engage in various outreach efforts (Curry and Fix 2019; Glennon and Strother 2019; Savchak and Edwards 2016). To the extent public understanding of the judicial process is limited, outreach may be appealing to judges, though the extent to which it impacts public attitudes is unclear.

We evaluate the causal effect of oral argument attendance on public perceptions of the judiciary using a field experiment. Specifically, we randomized in-person attendance at oral argument proceedings conducted by a traveling federal circuit court of appeals on a university campus. This type of outreach initiative is common among state and lower federal courts. Consistent with positivity theory (Gibson and Caldeira 2009; Gibson, Lodge, and Woodson 2014; Gibson and Nelson 2017), we find that oral argument attendance increases perceptions of institutional legitimacy. We also find that exposure increases the extent to which people think judicial decisions are motivated by law as opposed to politics. The results are consistent across two oral argument sessions, with nonoverlapping control and treatment groups in each session.

This project has important implications for our understanding of judicial politics and the effect of exposure to government proceedings. With respect to the former, the results enhance our understanding of why oral argument matters (Black, Johnson, and Wedeking 2012; Johnson 2004; Ringsmuth and Johnson 2013) and the impact of court exposure on public attitudes (Black et al. n.d.; Benesh 2006; Krewson 2019). The results also contribute to policy debates concerning reform initiatives involving circuit riding, courtroom cameras, and outreach initiatives. With respect to the latter, the results contribute to our broader understanding of how exposure to government proceedings (Blair, Karim, and Morse 2019; Malesky and Taussig 2019; Peyton, Sierra-Arevalo, and Rand 2019) as well as deliberation and transparency initiatives (Gottlieb 2016; Hollyer, Rosendorff, and Vreeland 2019; Simon and Sulkin 2002) impact the public.

Oral Argument Exposure

Oral argument can serve a variety of purposes. Internally, judges may use oral argument to acquire information for purposes of informing decisions on the merits, coalition formation, opinion writing, and separation of powers considerations (Black, Johnson, and Wedeking 2012; Johnson 2004; Ringsmuth and Johnson 2013). Externally, oral argument can enhance institutional legitimacy, increase public awareness, guard against noncompliance, and provide judges with a platform for political position taking (Black et al. n.d.; Jacobi and Sag 2019; Krehbiel 2016).

The external impact of oral argument remains relatively understudied notwithstanding its potential importance. As discussions of judicial legitimacy proliferate, oral argument exposure has been at the forefront. With respect to the Supreme Court, for example, it has been said that “the public spectacle of oral argument assures the parties in the case at hand that their arguments have been heard and considered,” while also “allow[ing] the public to see the Court as an impartial tribunal exploring issues of national importance through a balanced adjudicative process” (Jacobi and Sag 2019, 1168).

Institutional maintenance is complicated at the appellate level by a public information deficit. It may be that “to know courts is to love them, because to know them is to be exposed to a series of legitimizing messages focused on the symbols of justice, judicial objectivity, and impartiality” (Gibson, Caldeira, and Baird 1998, 345). As one federal circuit court judge put it, however, “People simply do not understand what appellate courts are supposed to do” (Medina 1961, 155). As the only part of the appellate decision-making process open to the public, oral argument may inform the public’s perception of how these courts operate.

Judges engage in a variety of outreach activities for institutional maintenance purposes (Curry and Fix 2019; Glennon and Strother 2019; Savchak and Edwards 2016), but traveling to conduct oral arguments is a particularly intriguing example because it exposes people to official court proceedings. State and lower federal courts regularly conduct oral argument at law schools.¹ Moreover, several state courts have expanded their oral argument outreach initiatives to the public more broadly. Example institutionalized programs include Appeals on Wheels by the Indiana Court of Appeals, Courts in the Community by the Hawaii Supreme Court, and Justice on Wheels by the

¹See, e.g., *Ninth Circuit Holds Oral Arguments at Law School for First Time Since COVID-19*, <https://www.uscourts.gov/news/2023/02/27/ninth-circuit-holds-oral-arguments-law-school-first-time-covid-19>.

Wisconsin Supreme Court. Officials hope these initiatives enhance judicial legitimacy and improve public perception (Deits and Keenan 2004, 243; Nelson I. 2005, 170), but the extent to which they do so is unclear.

Bringing courts to the people is an old idea. Supreme Court circuit riding may have primarily been a cost saving measure, but it was also thought to be of “great value in keeping the federal judiciary in touch with the local communities” (Warren 1924, 58). When justices lobbied to abolish the practice, legislators resisted in part because of the public-interfacing component. In 1848, for example, one senator argued that abolishing circuit riding would result in the Court “losing... that responsive confidence of the people, which adds so essentially to the sanction of all the acts of the officers of government.”² Although circuit riding has long been abolished, calls for renewal regularly emphasize the value of public exposure (Calabresi and Presser 2006; Stras 2006).

Positivity theory provides a theoretical lens through which one can assess the potential impact of oral argument exposure on public attitudes. The core idea is simple: “[I]ncreased exposure to the judicial process, whatever the circumstances and even when citizens are displeased, results in collateral exposure to the symbols of judicial legitimacy, thereby tending to reinforce rather than undermine institutional support” (Gibson and Caldeira 2009, 4). Symbol exposure “triggers learned associated thoughts, which for most people in the United States have become connected with these symbols largely through socialization processes and experience, and which are typically ones of legitimacy and positivity” (Gibson, Lodge, Woodson 2014, 842). As a result, comprehending court proceedings is not a necessary condition for attitude change.

Legitimizing symbols can be specific or abstract. Example symbols, all of which tend to be present at oral argument, include “judges’ black robes” (Gibson and Nelson 2017, 593), judges occupying an elevated position in the courtroom (Gibson, Caldeira, and Spence 2003b, 553), and litigants using “honorific forms of address” (Gibson, Lodge, and Woodson 2014, 838). Experimental evidence indicates that exposure to such symbols can heighten the positive relationship between support and decision acquiescence for low-frequency consumers (Gibson, Lodge, and Woodson 2014), sever the link between decision disagreeableness and diffuse support for those expressing relatively low levels of disappointment with a court’s judgment (Gibson and Nelson 2016), and enhance perceptions of institutional legitimacy (Armaly 2018).

²Congressional Globe, 30th Congress, 1st Session 596 (1848).

Court proceedings also expose consumers to more abstract legitimizing symbols such as “decorum” (Nelson and Gibson 2017, 134); “judicial objectivity” (Gibson, Caldeira, and Baird 1998, 345); “emphasizing reliance on the Constitution, precedent, and legal norms” (Nicholson and Hansford 2014, 621); and “impartiality and insulation from ordinary political pressures” (Gibson and Caldeira 2009, 9). Consistent with courtroom settings, studies in other institutional contexts have found legitimizing effects associated with exposures emphasizing transparency (Gottlieb 2016; Hawkins et al. 2017; Hollyer, Rosendorff, and Vreeland 2019), deliberation (Simon and Sulkin 2002; Persson, Esaiasson, and Gilljam 2013; Tyler, Rasinski, and Spodick 1985), legal frames (Baird and Gangl 2006; Farganis 2012; Zink, Spriggs, Scott 2009), and procedural justice (Baird 2001; Gibson 1989; Tyler and Rasinski 1991).

There is little existing empirical evidence on the impact of court exposure on public attitudes. Most relevant for our purpose, Black et al. (n.d.) conducted survey experiments exposing observers to short video clips of oral argument exchanges in two state supreme courts. Leveraging variation in audio versus video presentations, contentious versus neutral interpersonal exchanges, and dynamic versus static camera angles, the authors find mixed evidence concerning the impact of exposure on perceptions of legitimacy. With respect to court exposure more broadly, Benesh (2006) finds that respondents who report having been a juror express more confidence in local courts, while those who report having been a direct party express less confidence. And Grimmelikhuisen and Klijn (2015) find that a randomized instruction to watch a television show featuring real trial court proceedings, commentary, and interviews with participating judges, yielded increased trust in judges as the self-reported number of shows watched increased.

Building on this literature, we hypothesize that oral argument attendance will increase perceptions of institutional legitimacy and the extent to which judges are thought to base decisions on law versus politics. While there have been relatively few examinations of how court exposure impacts perceptions of decision making compared to legitimacy, the underlying theory is similar. Legitimizing symbols presumably activate the “myth of legality,” which “holds that cases are decided by the application of legal rules formulated and applied through a politically and philosophically neutral process of legal reasoning” (Scheb and Lyons 2000, 929). While so-called “legalist” and “realist” perspectives on judging are theoretically distinct from the concept of legitimacy (Gibson and Caldeira 2011; Cann and Yates 2016), the same psychological processes that lead symbols to

enhance legitimacy can reasonably be construed to reinforce perceptions that judges make decisions on the basis of law rather than politics. Evidence indicates, for example, that exposure to a lunch or speech with a Supreme Court justice increases the extent to which decisions are thought to be driven by law relative to ideology (Krewson 2019).

The Field Experiment

To examine the causal effect of oral argument exposure on perceptions of the judiciary, we conducted a field experiment randomizing attendance at proceedings conducted by the U.S. Court of Appeals for the Tenth Circuit at Utah State University. The Tenth Circuit is one of thirteen federal circuit courts of appeals. It is an intermediate appellate court, sitting between state-based district courts and the Supreme Court in the federal judicial hierarchy. The Tenth Circuit hears appeals arising from district courts in six states.³ Oral arguments are heard at the court’s discretion, generally in three-judge panels. According to Tenth Circuit officials attending this event, the court travels for oral argument as an outreach initiative for students and the public.

At this event, a three-judge panel held two oral argument sessions, with two cases in each session and a short break between sessions. Case details are provided in the appendix. As is common with circuit court proceedings, the cases were not particularly salient. Advance session-specific tickets were required for admission to the event, which filled to capacity. Student access was primarily allocated through this study.⁴ Prior to the visit, we recruited undergraduate and graduate students to participate in a study concerning courts on event day in exchange for a \$10.00 Amazon gift card.⁵ In the recruitment email, students were told they would be randomly assigned to attend an oral argument session or participate in a different court-related activity. Participating students selected one of two time slots corresponding to the oral argument sessions. Once students were assigned to a session slot, they were randomized within sessions at a 50-50 split between

³These states are Colorado, Kansas, New Mexico, Oklahoma, Utah, and Wyoming.

⁴In addition to the tickets we allocated, a select group of student leaders obtained tickets from the event organizer. Other tickets were dispersed to university officials, community groups, and secondary schools.

⁵Student recruitment occurred using the primary email address on file with the university. Of the resulting sample, about 33 percent were freshmen, 20 percent sophomores, 15 percent juniors, and 15 percent seniors, with the remainder being graduate students or selecting the “other” category (e.g., non-degree-seeking students). Over fifty majors were represented in the sample. The university has a Law and Constitutional Studies major, but only seven students were pursuing that course of study. Sixteen students were political science majors. As a result, the vast majority of participants were not studying law or politics as majors. The university does not have a law school.

treatment and control groups.⁶

The treatment and control groups were instructed to appear on event day at separate but nearby campus locations. On event day, control group participants completed a survey immediately upon arrival at their scheduled time and were subsequently dismissed. Treatment group participants were admitted to their respective oral argument sessions, after which they were provided the same survey with a day’s end completion deadline. All surveys were completed electronically on event day.⁷ Overall, 248 students participated in the experiment. The first session respectively included 48 and 58 students in the control and treatment groups; the second session respectively included 70 and 72 students in the control and treatment groups.⁸ As described in more detail in the appendix, control and treatment groups were balanced within and across sessions by gender, race, ideology, and prior court exposure. Although student samples present obvious generalizability issues, ours is “relevant for theory” (Coppock and McClellan 2019, 3) since students are a target audience for traveling courts.

We follow standard practice in measuring our outcomes of interest. To capture perceptions of legitimacy, we use an established battery of statements (see, e.g., Gibson and Caldeira 2009; Gibson, Caldeira, and Spence 2003a; Gibson and Nelson 2015) that capture whether institutions enjoy “a widely accepted mandate to render judgments for a political community” (Gibson, Caldeira, and Spence 2003a). The items include:

- Judges on the federal courts who consistently make decisions at odds with what a majority of the people want should be removed from their position as a judge.
- If the federal courts started making a lot of decisions that most people disagree with, it might be better to do away with the federal courts altogether.

⁶We used a random number generator to assign participants to treatment and control groups within sessions.

⁷With survey completion time varying across treatment and control groups, there is some risk that treatment group participants were exposed to stimuli other than oral argument that could have impacted their perceptions of the judiciary, though the risk seems insubstantial for several reasons. First, there were no salient court-related events that afternoon or evening that would have provided broad exposure to judicial symbols. Second, not much time passed between oral argument and survey completion by the treatment groups—less than two hours for most participants. Third, if an intervening event impacted perceptions of the judiciary, we should observe weaker effects in the treatment-control comparison when omitting treatment group participants who completed surveys later in the day. However, dropping treatment group participants who completed their survey after 5:00 PM increases the gap between control and treatment groups on both outcome variables.

⁸A post-hoc power analysis using GPower shows that using a two-tailed test and the conventional $\alpha = .05$ criterion for statistical significance, our sample sizes (118 and 130 respectively) have “power” of .975, meaning that 97.5% of sample pairs would be able to detect a statistically significant difference if there was, in fact, an effect of at least .5 standard deviations in the population. Respectively, power for the first and second sessions were .719 and .841. Attrition details are presented in the appendix.

- The right of the federal courts to decide certain types of controversial issues should be reduced.
- The federal courts get too mixed up in politics.

Given the institutional context of our experiment, we modified these questions as developed for a study concerning the Supreme Court by replacing “Supreme Court” with “federal courts” (cf. Scherer, Benesh, and Steigerwalt 2010).⁹ Response choices were five-point Likert scales ranging from strongly disagree to strongly agree. A principal components analysis on the scores shows that items load on a single dimension.¹⁰ We used factor scores from the principal components analysis to create a legitimacy index and normalized scores to the [0,1] interval.¹¹

To operationalize our decision-making hypothesis, we use a battery of questions developed by Gibson and Caldeira (2011) to measure what they call “legal realism.” In essence, this battery captures the extent to which people think judicial decision making is driven by law versus politics. The items include:

- Judges always say that their decisions are based on the law and the Constitution, but in many cases, judges are really basing their decisions on their own personal beliefs.
- Judges’ values and political views have little to do with how they decide cases before the federal courts.
- Judges’ party affiliations have little to do with how they decide cases before the federal courts.

Again, given the institutional context of our experiment we modified the questions as designed by replacing “Supreme Court” with “federal courts” (cf. Scherer, Benesh, and Steigerwalt 2010). Response choices were five-point Likert scales ranging from strongly disagree to strongly agree. We used principal components analysis to create an index concerning decision-making attitudes and normalized scores to the [0,1] interval.¹² As detailed in the appendix, the result is similar using an alternative measure of beliefs about judicial decision making developed by Cann and Yates (2016).

⁹Although question wording and inclusion varies across studies, the shared theme is that battery items capture diffuse support, which is defined as “institutional loyalty” (Gibson, Caldeira, and Spence 2003a, 356) or the “reservoir of favorable attitudes or good will that helps [people] accept or tolerate outputs to which they are opposed or the effects of which they see as damaging to their wants” (Easton 1965, 273).

¹⁰This holds using the Scree test and Kaiser criterion.

¹¹As a robustness check, we re-ran the principal components analysis on a polychoric correlation matrix to determine whether the ordinal nature of the Likert items posed a threat to this approach. The scores from the principal components analysis on the polychoric correlation matrix correlate with the index from the original principal components analysis at $r > .99$.

¹²The Scree test and Kaiser criterion indicated a one-dimensional solution for this index. As with the legitimacy index, we re-ran the principal components analysis on a polychoric correlation matrix and found that the scores based on the polychoric matrix correlate with scores from a regular Pearson’s r matrix at $r > .99$.

Analysis and Results

Legitimacy

Beginning with the legitimacy hypothesis, Figure 1 plots means, 95 percent confidence intervals, and underlying data distributions (jittered for clarity) for the control and treatment groups. Between control and treatment, the perceived legitimacy mean increases from 0.48 to 0.60 ($p < .001$). Substantively, this change indicates that oral argument attendance increased perceived legitimacy by 57 percent of a standard deviation. As detailed in the appendix, these results hold across sessions. There is evidence of heterogeneous session effects, with perceived legitimacy higher in the first session, but exposure to the second session nonetheless increased perceived legitimacy by 39 percent of a standard deviation.

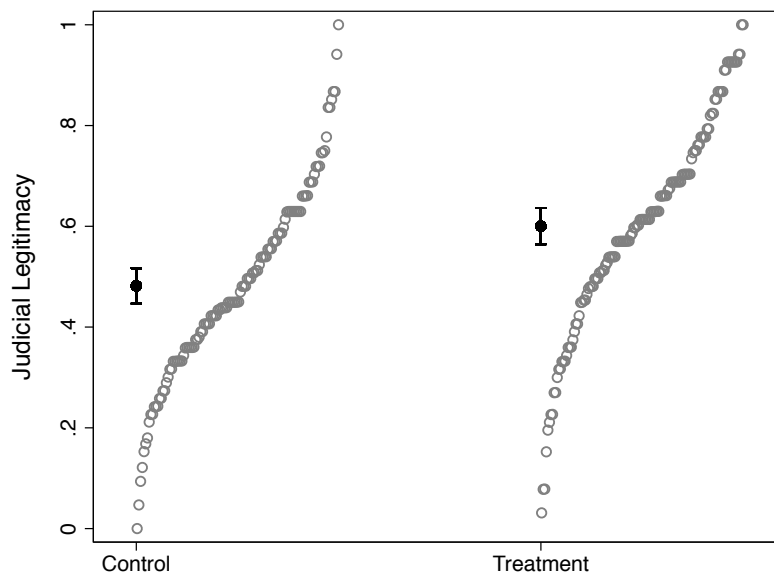


Figure 1: The Effect of Oral Argument Attendance on Legitimacy

Judicial Decision Making

Turning to the decision-making hypothesis, Figure 2 plots means, 95 percent confidence intervals, and underlying data distributions for the control and treatment groups. Between control and treatment, the mean index score increases from 0.34 to 0.46 ($p < .001$). Substantively, this indicates that oral argument attendance increased the extent to which individuals perceive judicial decision

making to be driven by law versus politics by 61 percent of a standard deviation. As detailed in the appendix, the result holds across sessions. Here, however, there is no evidence of heterogeneous session effects.

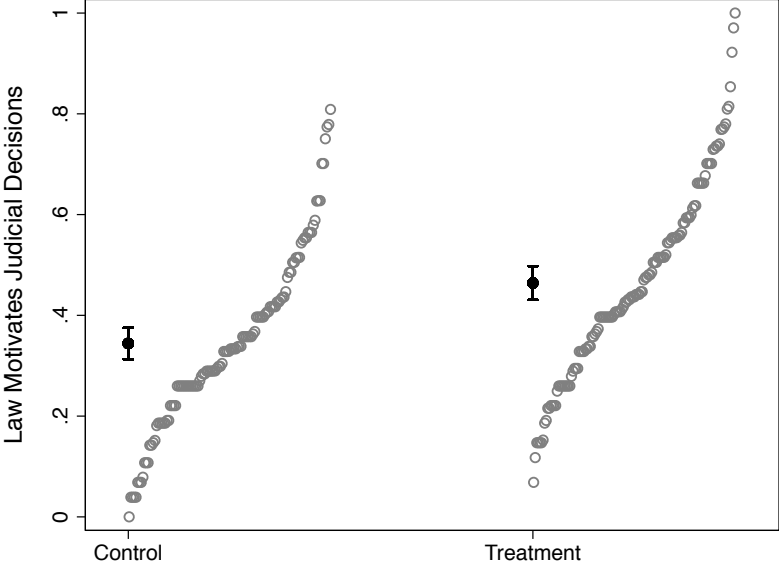


Figure 2: The Effect of Oral Argument Attendance on Beliefs About Judicial Decision Making

Conclusion

Judicial independence is widely thought to be contingent on public support. Given increasing political attacks on the judiciary, courts are engaging in outreach to enhance and maintain institutional legitimacy. What remains unclear, however, is the extent to which courts can unilaterally manage public relations—particularly given that people are often unfamiliar with how they operate. We consider whether court exposure can change public attitudes with a field experiment randomizing in-person attendance at oral argument proceedings conducted by a traveling federal circuit court of appeals. We find that oral argument attendance increases perceived legitimacy and the extent to which judges are thought to make decisions based on law versus politics.

The normative implications are mixed. From a public policy perspective, commentators have long debated reform to increase the public’s exposure to court proceedings. The results presented here suggest that proposals to institute circuit riding and install cameras in courtrooms may in-

crease perceived legitimacy and change the way people think about how judges make decisions, which may be arguments in their favor. But these policy questions are complex, and a comprehensive assessment will require more empirical evidence concerning a variety of potential costs and benefits. With respect to cameras in courtrooms, for example, impact may depend on a variety of circumstances (Black et al. n.d.), which cautions against interpreting the evidence presented here as indicating that oral argument exposure is necessarily positive.

A potentially concerning normative implication is that oral argument exposure may generate misleading public perceptions of the judiciary. Although we find that exposure increases the extent to which judges are thought to base decisions on law versus politics, for example, it is not clear that such updating is warranted based on what transpires at oral argument. Interestingly, some Supreme Court justices have argued against courtroom cameras out of concern that excerpts may mislead viewers (Segall 2015, 793–794), but any misdirection may work the other way as well. To the extent oral arguments disproportionately emphasize neutral principles relative to our broader empirical understanding of the determinants of judicial decision making, exposure may distort public perception.

The results have important implications for judicial politics. While much of the empirical research on oral arguments emphasize internal dynamics, we show that these proceedings can impact the public as well. We also contribute to the literature on positivity theory with application to a new setting using a real-world exposure, both of which enhance generalizability. Moreover, notwithstanding the literature’s emphasis on legitimacy, it is important to understand how exposure impacts other outcomes of interest (cf. Black et al. n.d., 17). We advance this goal by examining how oral argument attendance impacts public perceptions of judicial decision making. Last, our research design is transportable to other areas where exposure to court business may impact public attitudes, such as briefs, conference deliberations, and internal correspondence bargaining over opinion language.

Notwithstanding this project’s contributions, there are several important limitations. Generalizability is limited by reliance on college students who self-selected into exposure. Although students are “relevant for theory” (Coppock and McClellan 2019, 3) here insofar as they are a target demographic for judicial outreach efforts, it may be, for example, that older participants would be less impressionable. Theoretically, there is reason to expect heterogeneous treatment effects based

on factors such as case salience and topic familiarity, but the proceedings analyzed here did not involve constitutional or otherwise politically contested issues. Moreover, we do not explore treatment effect persistence or isolate the mechanisms underlying observed attitude changes. Future research will help develop a more comprehensive evaluation of how exposure to court proceedings impact the public.

References

- Armaly, Miles T. 2018. "Politicized Nominations and Public Attitudes toward the Supreme Court in the Polarization Era." *Justice System Journal* 39(3): 193–209.
- Baird, Vanessa A. 2001. "Building Institutional Legitimacy: The Role of Procedural Justice." *Political Research Quarterly* 52(2): 333–354.
- Baird, Vanessa A., and Amy Gangl. 2006. "Shattering the Myth of Legality: The Impact of the Media's Framing of Supreme Court Procedures on Perceptions of Fairness." *Political Psychology* 27(4): 597–614.
- Benesh, Sara C. 2006. "Understanding Public Confidence in American Courts." *Journal of Politics* 68(3): 697–707.
- Black, Ryan C., Timothy R. Johnson, Ryan J. Owens, and Justin Wedeking. n.d. "Televised Oral Arguments and Judicial Legitimacy: An Initial Assessment." Forthcoming, *Political Behavior*.
- Black, Ryan C., Timothy R. Johnson, and Justin Wedeking. 2012. *Oral Arguments and Coalition Formation on the U.S. Supreme Court*. Ann Arbor: University of Michigan Press.
- Blair, Robert A., Sabrina M. Karim, and Benjamin S. Morse. 2019. "Establishing the Rule of Law in Weak and War-torn States: Evidence from a Field Experiment with the Liberian National Police." *American Political Science Review* 113(3): 641–657.
- Calabresi, Steven G., and David C. Presser. 2006. "Reintroducing Circuit Riding: A Timely Proposal." *Minnesota Law Review* 90(5): 1386–1416.
- Caldeira, Gregory A. 1986. "Neither the Purse nor the Sword: Dynamics of Public Confidence in the Supreme Court." *American Political Science Review* 80(4): 1209–1226.
- Cann, Damon M., and Jeff Yates. 2016. *These Estimable Courts: Understanding Public Perceptions of State Judicial Institutions and Legal Policy-Making*. New York: Oxford University Press.
- Casillas, Christopher J., Peter K. Enns, and Patrick C. Wohlfarth. 2011. "How Public Opinion Constrains the U.S. Supreme Court." *American Journal of Political Science* 55(1): 74–88.
- Clark, Tom S. 2009. "The Separation of Powers, Court Curbing, and Judicial Legitimacy." *American Journal of Political Science* 53(4): 971–999.
- Coppock, Alexander, and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research and Politics* 6(1): 1–14.
- Curry, Todd A., and Michael P. Fix. "May it Please the Twittersverse: The Use of Twitter by State High Court Judges." *Journal of Information Technology and Politics* 16(4): 379–393.

- Deits, Mary J., and Lora E. Keenan. 2004. "Getting to Know Us: Judicial Outreach in Oregon." *The Journal of Appellate Practice and Process* 6(2): 237–248.
- Easton, David. 1965. *A Systems Analysis of Political Life*. New York: John Wiley & Sons.
- Epstein, Lee, and Andrew D. Martin. 2010. "Does Public Opinion Influence the Supreme Court? Possibly Yes (But We're Not Sure Why)." *University of Pennsylvania Journal of Constitutional Law* 13(2): 263–282.
- Farganis, Dion. 2012. "Do Reasons Matter? The Impact of Opinion Content on Supreme Court Legitimacy." *Political Research Quarterly* 65(1): 206–216.
- Gibson, James L. 1989. "Understandings of Justice: Institutional Legitimacy, Procedural Justice, and Political Tolerance." *Law & Society Review* 23(3): 469–496.
- Gibson, James L., and Gregory A. Caldeira. 2011. "Has Legal Realism Damaged the Legitimacy of the U.S. Supreme Court?" *Law & Society Review* 45(1): 195–219.
- Gibson, James L., and Gregory A. Caldeira. 2009. *Citizens, Courts, and Confirmations: Positivity Theory and the Judgments of the American People*. Chicago: University of Chicago Press.
- Gibson, James L., and Gregory A. Caldeira, and Vanessa A. Baird. 1998. "On the Legitimacy of National High Courts." *American Political Science Review* 92(2): 343–358.
- Gibson, James L., and Gregory A. Caldeira, and Lester Kenyatta Spence. 2003a. "Measuring Attitudes toward the United States Supreme Court." *American Journal of Political Science* 47(2): 354–367.
- Gibson, James L., and Gregory A. Caldeira, and Lester Kenyatta Spence. 2003b. "The Supreme Court and the US Presidential Election of 2000: Wounds, Self-Inflicted or Otherwise?" *British Journal of Political Science* 33(4): 535–556.
- Gibson, James L., and Michael J. Nelson. 2017. "Reconsidering Positivity Theory: What Roles do Politicization, Ideological Disagreement, and Legal Realism Play in Shaping U.S. Supreme Court Legitimacy." *Journal of Empirical Legal Studies* 14(3): 592–617.
- Gibson, James L., and Michael J. Nelson. 2016. "Change in Institutional Support for the U.S. Supreme Court: Is the Court's Legitimacy Imperiled by the Decisions it Makes?" *Public Opinion Quarterly* 80(3): 622–641.
- Gibson, James L., and Michael J. Nelson. 2015. "Is the U.S. Supreme Court's Legitimacy Grounded in Performance Satisfaction and Ideology?" *American Journal of Political Science* 59(1): 162–174.
- Gibson, James L., Milton Lodge, and Benjamin Woodson. 2014. "Losing, but Accepting: Legitimacy, Positivity Theory, and the Symbols of Judicial Authority." *Law & Society Review* 48(4): 837–866.
- Glennon, Colin, and Logan Strother. 2019. "The Maintenance of Institutional Legitimacy in

- Supreme Court Justices' Public Rhetoric." *Journal of Law and Courts* 7(2): 241–261.
- Gottlieb, Jessica. 2016. "Greater Expectations: A Field Experiment to Improve Accountability in Mali." *American Journal of Political Science* 60(1): 143–157.
- Grimmelikhuijsen, Stephan, and Albert Klijn. 2015. "The Effects of Judicial Transparency on Public Trust: Evidence from a Field Experiment." *Public Administration* 93(4): 995–1011.
- Hawkins, Darren, Lucas C. Brook, Ian M. Hansen, Neal A. Hoopes and Taylor R. Tidwell. 2019. "Do Citizens See Through Transparency? Evidence from Survey Experiments in Peru." *British Journal of Political Science* 49(1): 205–228.
- Helmke, Gretchen. 2010. "The Origins of Institutional Crises in Latin America." *American Journal of Political Science* 54(3): 737–750.
- Hollyer, James R., B. Peter Rosendorff, and James Raymond Vreeland. 2019. "Transparency, Protest and Democratic Stability." *British Journal of Political Science* 49(4): 1251–1277.
- Jacobi, Tonja, and Matthew Sag. 2019. "The New Oral Argument: Justices as Advocates." *Notre Dame Law Review* 94(3): 1161–1254.
- Johnson, Timothy R. 2004. *Oral Arguments and Decision Making on the United States Supreme Court*. Albany: State University of New York Press.
- Krehbiel, Jay N. 2016. "The Politics of Judicial Procedures: The Role of Public Oral Hearings in the German Constitutional Court." *American Journal of Political Science* 60(4): 990–1005.
- Krewson, Christopher N. 2019. "Save this Honorable Court: Shaping Public Perceptions of the Supreme Court Off the Bench." *Political Research Quarterly* 72(3): 686–699.
- Malesky, Edmund, and Markus Taussig. 2019. "Participation, Government Legitimacy, and Regulatory Compliance in Emerging Economies: A Firm-Level Field Experiment in Vietnam." *American Political Science Review* 113(2): 530–551.
- McGuire, Kevin T., and James A. Stimson. 2004. "The Least Dangerous Branch Revisited: New Evidence on Supreme Court Responsiveness to Public Preferences." *Journal of Politics* 66(4): 1018–1035.
- Medina, Harold R. 1961. "Some Reflections on the Judicial Function at the Appellate Level." *Washington University Law Review* 1961(2): 148–156.
- Nelson, Ingrid A. 2005. "More Judicial Outreach: 'Justice on Wheels' from the Supreme Court of Wisconsin." *Journal of Appellate Practice and Process* 7(1): 167–170.
- Nelson, Michael J., and James L. Gibson. 2017. "U.S. Supreme Court Legitimacy: Unanswered Questions and an Agenda for Future Research." In *Routledge Handbook of Judicial Behavior*. Robert M. Howard and Kirk A. Randazzo, eds. New York: Routledge. Pp. 132–150.

- Nicholson, Stephen P., and Thomas G. Hansford. 2014. "Partisans in Robes: Party Cues and Public Acceptance of Supreme Court Decisions." *American Political Science Review* 58(3): 620–636.
- Persson, Mikael, Peter Esaiasson, and Mikael Gilljam. 2013. "The Effects of Direct Voting and Deliberation on Legitimacy Beliefs: An Experimental Study of Small Group Decision-Making." *European Political Science Review* 5(3): 381–399.
- Peyton, Kyle, Michael Sierra-Arevalo, and David G. Rand. 2019. "A Field Experiment on Community Policing and Police Legitimacy." *Proceedings of the National Academy of Sciences* 116(40): 19894–19898.
- Ringsmuth, Eve M., and Timothy R. Johnson. 2013. "Supreme Court Oral Arguments and Institutional Maintenance." *American Politics Research* 41(4): 651–673.
- Savchak, Elisha Carol, and Amanda Ross Edwards. 2016. "Why Are State Judges Among Us?" *Judicature* 100(2): 21–30.
- Scheb, John M. II, and William Lyons. 2000. "The Myth of Legality and Public Evaluation of the Supreme Court." *Social Science Quarterly* 81(4): 928–940.
- Scherer, Nancy, Sara C. Benesh, and Amy Steigerwalt. 2010. "How Do Lower Federal Courts Attain Institutional Legitimacy When They Are Largely Unknown to the Public." Paper presented at the annual meeting of the American Political Science Association.
- Segall, Eric J. 2016. "Invisible Justices: How Our Highest Court Hides from the American People." *Georgia State University Law Review* 32(4): 787–848.
- Simon, Adam F., and Tracy Sulkin. 2002. "Discussion's Impact on Political Allocations: An Experimental Approach." *Political Analysis* 10(4): 403–412.
- Staton, Jeffrey K. 2004. "Judicial Policy Implementation in Mexico City and Merida." *Comparative Politics* 37(1): 41–60.
- Stephenson, Matthew C. 2004. "Court of Public Opinion: Government Accountability and Judicial Independence." *Journal of Law, Economics, and Organization* 20(2): 379–399.
- Stras, David D. 2006. "Why Supreme Court Justices Should Ride Circuit Again." *Minnesota Law Review* 90(5): 1710–1751.
- Tyler, Tom R., and Kenneth Rasinski. 1991. "Procedural Justice, Institutional Legitimacy, and the Acceptance of Unpopular U.S. Supreme Court Decisions: A Reply to Gibson." *Law & Society Review* 25(3): 621–630.
- Tyler, Tom R., Kenneth A. Rasinski, and Nancy Spodick. 1985. "Influence of Voice on Satisfaction with Leaders: Exploring the Meaning of Process Control." *Journal of Personality and Social Psychology* 48(1): 72–81.
- Vanberg, Georg. 2004. *The Politics of Constitutional Review in Germany*. New York: Cambridge

University Press.

Warren, Charles. 1924. *The Supreme Court in United States History*. Volume I. Boston: Little, Brown, and Company.

Zink, James R., James F. Spriggs II, and John T. Scott. 2009. "Courting the Public: The Influence of Decision Attributes on Individuals' Views of Court Opinions." *Journal of Politics* 71(3): 909–925.

Appendix

A. Attrition

Overall, 406 students expressed interest in participating and 248 students ultimately participated, yielding a 39 percent attrition rate. Attrition may have two consequences. First, there is some risk that those who expressed interest but did not participate are systematically different from those who did participate. If true, this may limit generalizability but does not in and of itself bias estimates in the observed population. Second, bias may result if factors systematically related to the outcome variables are correlated with a disproportionate study completion across control and treatment groups. Where this occurs, the problem is differential rather than overall attrition.

The respective attrition rates in our control and treatment groups were 42 and 36 percent, resulting in a 6 percent differential attrition rate. These overall and differential attrition rates are acceptable under U.S. Department of Education (DOE) standards for studies where the treatment itself is unlikely to induce attrition (U.S. Department of Education 2020, 13; U.S. Department of Education 2013; Broton et al. 2016; Deke et al. 2015), which we presume to be the case here. Deke et al. (2015, 12) suggest a maximum bias of .075 standard deviations for the use of these standards; the effect sizes presented here are substantially larger, which mitigates any concern about potential bias.

There were also attrition differences across sessions. For the first session, overall attrition was 43 percent and differential attrition was 11 percent. For the second session, overall attrition was 35 percent and differential attrition was 1 percent. The second session's low level of differential attrition satisfies the DOE's most cautious standard, leaving only minimal room for error. As discussed below, the second-session results standing alone are consistent with the combined results (as are the first-session results), which further alleviates any concern about bias.

B. Balance

Balance assessments help determine whether randomization yielded similar control and treatment groups. We test for balance on partisanship, gender, race, and prior court exposure using a χ^2 test. There are no statistically significant differences between the control and treatment groups on these dimensions.

Table 1: Party Identification Balance Test

	Strong Democrat	Weak Democrat	Lean Democrat	Independent	Lean Republican	Weak Republican	Strong Republican
Control	15 51.72%	13 50.00%	9 39.13%	8 34.78%	9 34.62%	34 54.84%	18 56.25%
Treatment	14 48.28%	13 50.00%	14 60.87%	15 65.22%	17 65.38%	28 45.16%	14 43.75%
Total	29	26	23	23	26	62	32

Note: $\chi^2 = 6.44$, $p = .376$. A similar test combining strong, weak, and leaning partisans (yielding a three-category measure of partisanship) also fails to reject the null hypothesis. Within-session balance tests also fail to reject the null hypothesis.

Table 2: Gender Balance Test

	Man	Woman	Other/Non-Binary
Control	49 43.75%	68 50.37%	1 100%
Treatment	63 56.25%	67 49.63%	0 0.00%
Total	112	135	1

Note: $\chi^2 = 2.18$, $p = .336$. Within-session balance tests also fail to reject the null hypothesis.

Table 3: Race Balance Test

	Person of Color	White
Control	12 38.71%	106 48.85%
Treatment	19 61.29%	111 51.15%
Total	31	217

Note: $\chi^2 = 1.12$, $p = .290$. Within-session balance tests also fail to reject the null hypothesis.

Table 4: Prior Court Exposure Balance Test

	No Jury Experience	Jury Experience
Control	104 47.71%	14 46.67%
Treatment	114 52.29%	16 53.33%
Total	218	30

Note: Because numbers of students with court experience other than jury-related experience are so small, we focus our balance test on jury service. $\chi^2 = .011$, $p = .915$. Within-session balance tests also fail to reject the null hypothesis.

C. Alternative Decision Making Outcome Measure

As an alternative measure of the extent to which people think judicial decision making is driven by law as opposed to politics, we employ the following question developed by Cann and Yates (2016):

- Do you think federal judges make decisions more based on their ideology and personal beliefs, or do you think they make decisions more on the basis of the law, regardless of their personal beliefs?

To answer, respondents located a position on a [0,100] sliding scale with the following marked thresholds:

- 0 = ideology and/or personal beliefs only
- 50 = an even mix of law and personal beliefs
- 100 = the law only, regardless of belief

Figure 3 plots means, 95 percent confidence intervals, and underlying data distributions for the control and treatment groups. Between control and treatment, the mean index score increases from 55 to 69 ($p < .001$). Substantively, this indicates that court exposure increases the extent to which individuals perceive decisions to be driven by law versus politics by 71 percent of a standard deviation.

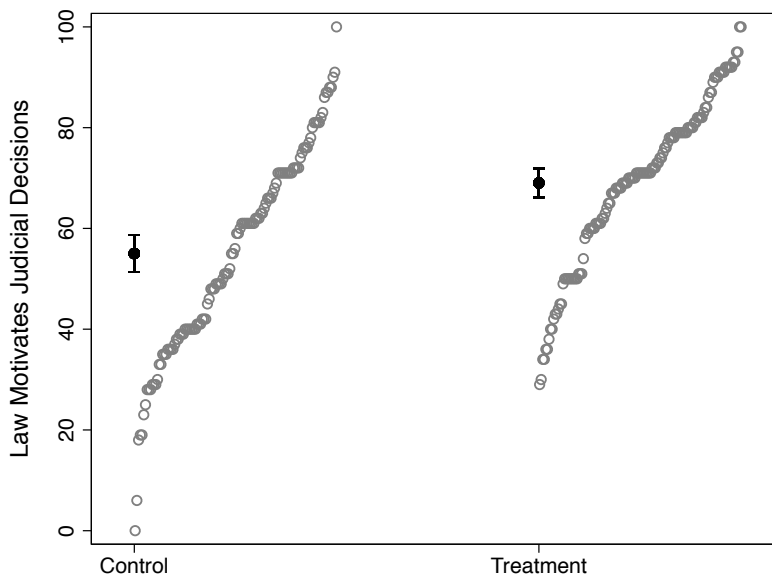


Figure 3: The Effect of Oral Argument Attendance on Beliefs About Judicial Decision Making

D. Effects by Session

Since we effectively have two separate field experiments—one for each oral argument session—we report disaggregated effects. With respect to legitimacy, Figure 4 plots means, 95 percent confidence intervals, and underlying data distributions for the control and treatment groups across sessions. For the first session (left panel), the mean legitimacy score increases from 0.47 to 0.64 ($p < .001$); this corresponds to an increase in legitimacy of 78 percent of a standard deviation. For the second session (right panel), the mean legitimacy score increases from 0.49 to 0.57 ($p < .001$); this corresponds to an increase in legitimacy of 39 percent of a standard deviation.

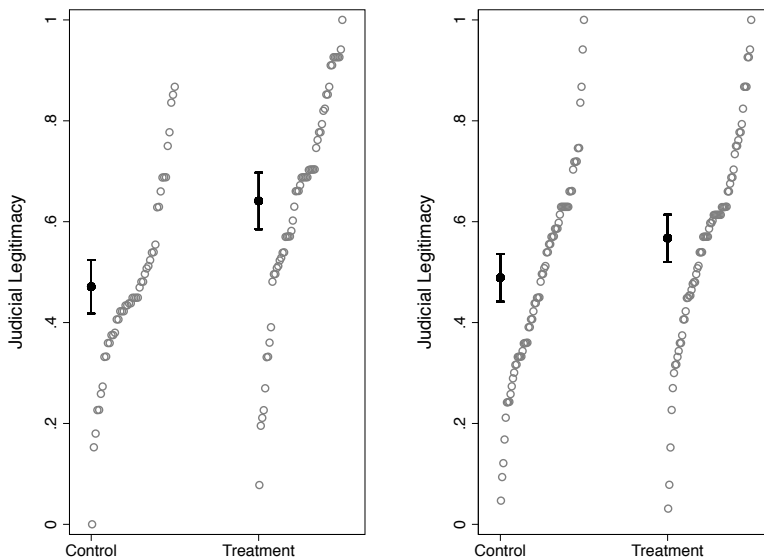


Figure 4: The Effect of Oral Argument Attendance on Perceived Legitimacy by Session

Figure 5 further explores heterogeneous session effects by displaying estimated coefficients and standard errors from a linear regression interacting indicator variables for treatment and session. Although the estimated coefficient on the interaction is not statistically distinguishable from zero at the conventional threshold ($p = 0.076$), the difference in mean predicted values for the treatment groups is statistically distinguishable from zero ($p < .05$).

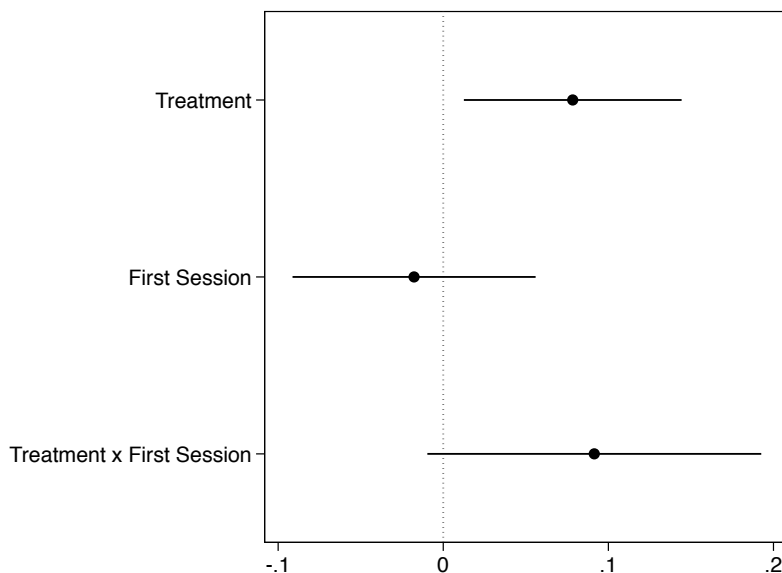


Figure 5: The Effect of Oral Argument Attendance on Perceived Legitimacy

Aside from chance, there are two general possibilities for why we observe heterogeneous session effects. First, since students self selected into time slots it is possible that some unmeasured characteristic explains sorting into sessions. The two sessions were, however, close in time: the first began at 8:45 AM, the second at 10:30 AM. Moreover, given the similarity in decision-making effects, it would have to be an unobserved characteristic that impacted legitimacy but not decision making. Second, the difference could be due to session-specific circumstances. Most prominently, each session heard arguments in cases concerning different topics. Moreover, perceived performance may have varied across sessions. Last, in addition to oral arguments the first session observed a generic event introduction while the second observed a brief question and answer session. The latter may have made the judges seem more like “regular people” relative to their specialized conduct during oral arguments, resulting in a legitimacy-moderating effect. In any event, notwithstanding this inter-session difference, both treatment groups experienced a substantial increase in perceived legitimacy. Moreover, as described in more detail below, there is no difference in treatment effects concerning beliefs about judicial decision making.

With respect to the extent to which judicial decision making is thought to be driven by law versus politics, Figure 6 plots means, 95 percent confidence intervals, and underlying data distributions for the control and treatment groups across sessions. For the first session (left panel), the mean legalism score increases from 0.35 to 0.46 ($p = .001$); this corresponds to an increase in legitimacy of 58 percent of a standard deviation. For the second session (right panel), the mean legitimacy score increases from 0.34 to 0.46 ($p < .001$); this corresponds to an increase in legitimacy of 65 percent of a standard deviation.

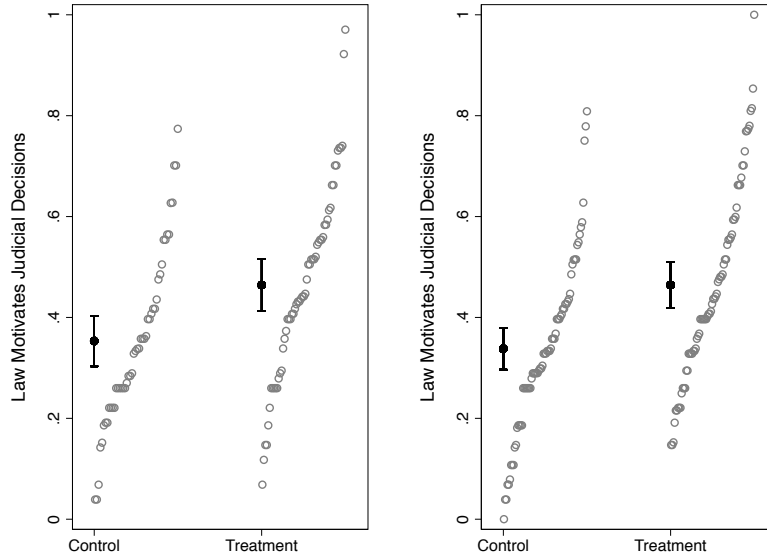


Figure 6: The Effect of Oral Argument Attendance on Beliefs About Judicial Decision Making by Session

Figure 7 further explores heterogeneous treatment effects by displaying estimated coefficients and standard errors from a linear regression interacting indicator variables for treatment and session. The difference in mean predicted values is not statistically distinguishable from zero ($p = 0.99$).

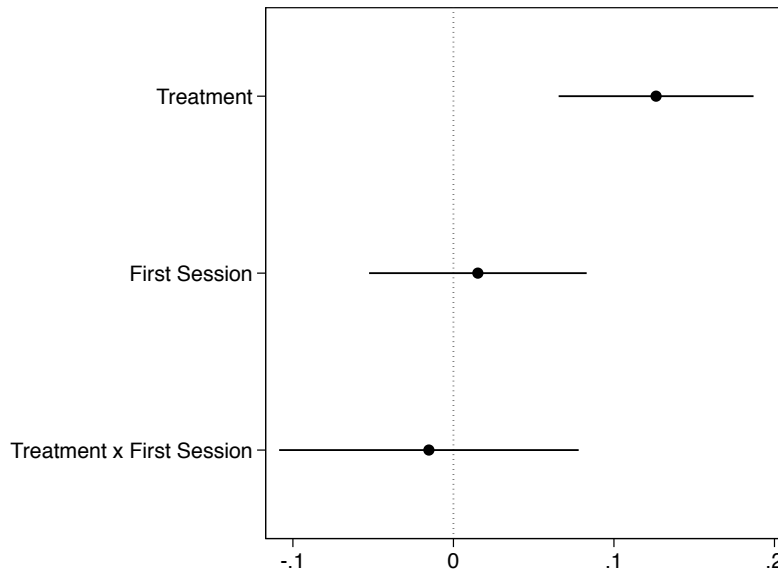


Figure 7: The Effect of Oral Argument Attendance on Beliefs About Judicial Decision Making

E. Question Wording

Participants answered the following questions:

- Judges on the federal courts who consistently make decisions at odds with what a majority of the people want should be removed from their position as a judge (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- The federal courts get too mixed up in politics (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- If the federal courts started making a lot of decisions that most people disagree with, it might be better to do away with the federal courts altogether (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- The right of the federal courts to decide certain types of controversial issues should be reduced (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- For the next question, use the slider to indicate where your opinion falls between two ends of the spectrum (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Do you think federal court judges make decisions more based on their ideology and personal beliefs, or do you think they make decisions more on the basis of the law, regardless of their personal beliefs? (respondent positions a 0-100 scale slider where 100 represents the law only, regardless of belief and 0 represents Ideology and/or personal beliefs only)
- A federal court's prior decisions play an important role in judicial decision making (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Judges always say that their decisions are based on the law and the Constitution, but in many cases, judges are really basing their decisions on their own personal beliefs. (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Judges' values and political views have little to do with how they decide cases before the federal courts (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Judges' party affiliations have little to do with how they decide cases before the federal courts. (Response set: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Do you happen to know whether judges on the U.S. federal courts serve a set number of years or whether they serve a life term? (Response set: set number of years, life term, don't know)
- Do you happen to know whether judges on the U.S. federal courts can invalidate laws they think are unconstitutional? (Response set: yes, no, don't know)
- Please tell us whether the U.S. federal courts have ruled on the following issues: Abortion, Rights of Black Americans, Maximum income tax rate (Response set: yes, no, don't know for each issue)

- Before participating in today's experience, have you had direct experience, contact, or involvement with a court case that brought you in to a courthouse, including being called for jury duty? Please check all that apply: (response set Called for Jury duty but not selected, Selected to sit on a jury, Being the person sued/accused, Being the person filing the lawsuit, Witness, Other (free response), None of these/I have no direct court experience).
- What is your racial/ethnic group (check all that apply): (Response set: White, Black, Hispanic, Asian, Native American, Other)
- Generally speaking, do you consider yourself a... (Response set: Democrat, Republican, Independent, Other)
- If Republican or Democrat: Do you consider yourself a strong (Democrat/Republican) or not a strong (Democrat/Republican)? (Response set: strong, not strong)
- If Independent: Do you lean more toward... (Response set: Republican Party, Democratic Party, Neither)
- What is your gender? (Response set: Man, Woman, Non-binary, other)
- What is your major? (Free response)
- What is your class standing (Response set: Freshman, Sophomore, Junior, Senior, Graduate Student, Other)

F. Oral Argument Cases

Each oral argument session included two cases. The cases and questions presented were as follows:

- *U.S. v. Rodriguez*, 18-1449, examining whether a lower court judge’s violation classification in a supervised release case was procedurally reasonable under the federal sentencing guidelines.
- *Hamilton v. Commissioner*, 16-8037, weighing whether a bank account owned by a taxpayer’s son but claimed by the government to be managed on behalf of the taxpayer counted as part of the taxpayer’s assets for purposes of determining insolvency under the Internal Revenue Code.
- *Payan v. UPS*, 18-4017, judging whether a plaintiff pled sufficient facts to defeat a summary judgment motion in a case alleging private workplace racial discrimination under federal law.
- *Scott v. Windgate*, 19-4052, assessing whether a wilderness therapy firm counted as a health care provider under state law for purposes of reducing the state statute of limitations period for bringing malpractice suits.

We could not find evidence that these cases were covered locally, regionally, or nationally outside of local coverage of the event itself. Although ex ante legal salience is more difficult to gauge, the cases were not covered by the prominent legal blog *How Appealing*, which provides daily updates about news concerning federal appellate litigation. Nor could we find evidence of coverage on searchable legal blog aggregators hosted by the American Bar Association and Justia, which individually amass content from thousands of domestic and global legal websites. Although the relative low salience and issue area limitations (e.g., lack of a constitutional case) have implications for generalizability, these also seem to be typical appellate cases.

As is common with traveling courts, neither cases nor panel members were randomly assigned to the event. According to the Tenth Circuit, conditioned on docket progression, cases are assigned to outreach events based on originating state and topic variety. The origination constraint is meant to take advantage of local attorney connections and reduce participating attorney travel costs. Whereas some traveling courts select case topics that “would be most interesting for the audience, choosing, for example, a search-and-seizure case over a case involving an issue of appellate procedure” (Nelson 2005, 165), the Tenth Circuit emphasizes diversifying that day’s calendar. After the event is scheduled, judges from the region are asked if they are interested in sitting, and available to sit, on the panel; if more than two judges are interested and available, there is random assignment from that group to the panel. In addition, the chief judge prefers to sit on outreach panels.

F. References

Broton, Katharine M., Sara Goldrick-Rab, and James Benson. 2016. “Working for College: The Causal Impacts of Financial Grants on Undergraduate Employment.” *Educational Evaluation and Policy Analysis* 38(3): 477–494.

Deke, John, Emily Sama-Miller, and Alan Hershey. 2015. *Addressing Attrition Bias in Randomized Controlled Trials: Considerations for Systematic Evidence Reviews*, https://homvee.acf.hhs.gov/sites/default/files/2019-06/HomVEE-Attrition-White_Paper-7-2015.pdf.

Nelson, Ingrid A. 2005. “More Judicial Outreach: ‘Justice on Wheels’ from the Supreme Court of Wisconsin.” *Journal of Appellate Practice and Process* 7(1): 167–170.

U.S. Department of Education. 2020. *What Works Clearinghouse Standards Handbook*, <https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf>.

U.S. Department of Education. 2013. *Assessing Attrition Bias*, https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v2.1.pdf.