





ARTICLE

“Let Me Just Interrupt You”: Estimating Gender Effects in Supreme Court Oral Arguments

Erica Cai¹ , Ankita Gupta¹ , Katherine A. Keith² , Brendan O'Connor³  and Douglas Rice³

¹Ph.D. Student, University of Massachusetts Amherst, Amherst, Massachusetts, USA; ²Assistant Professor, Williams College, Williamstown, Massachusetts, USA and ³Associate Professor, University of Massachusetts Amherst, Amherst, Massachusetts, USA

Corresponding authors: Katherine A. Keith and Douglas Rice; Emails: kak5@williams.edu; drrice@umass.edu

(Received 22 December 2023; Revised 07 June 2024; Accepted 05 August 2024)

Abstract

Oral argument is the most public and visible part of the U.S. Supreme Court’s decision-making process. Yet what if some advocates are treated differently before the Court solely because of aspects of their identity? In this work, we leverage a causal inference framework to quantify the effect of an advocate’s gender on interruptions of advocates at both the Court-level and the justice-level. Examining nearly four decades of U.S. Supreme Court oral argument transcript data, we identify a clear and consistent gender effect that dwarfs other influences on justice interruption behavior, with female advocates interrupted more frequently than male advocates.

Keywords: U.S. Supreme Court; oral argument; gender bias; causal inference; text-as-data

Introduction

Oral argument is the most visible part of the U.S. Supreme Court’s decision-making process. Advocates on conflicting sides of the nation’s most important controversies press their arguments, while justices interject questions and comments. After, the Court meets in a private conference where the justices offer their initial votes and begin the process of drafting opinions explaining those votes. A rich literature documents how these arguments inform the decisions of the Court, with oral argument influencing the information justices have (e.g., Johnson 2001, 2004), the issues discussed in decisions (e.g., Black, Johnson, and Wedeking 2012), and the justice votes (e.g., Johnson, Wahlbeck, and Spriggs 2006).

Yet prior research raises an important question: What if some advocates are treated differently by the justices solely because of aspects of their identity unrelated to the case? Consider the Supreme Court’s oral argument in *United States v. Texas* in

November of 2021. The case dealt with a Texas state law that effectively banned abortion after six weeks. Arguing on behalf of the United States, Solicitor General Elizabeth Prelogar – a woman – was interrupted regularly by the justices, with one particularly poignant example occurring in this exchange with Associate Justice Samuel Alito:

GENERAL PRELOGAR: While I certainly acknowledge, Justice Alito, that an injunction that would bind state court judges is extremely rare, it's not unheard of, and I think, in the unprecedented facts of this case, it's appropriate relief. And –
 JUSTICE ALITO: Well, judges have been enjoined –
 GENERAL PRELOGAR: – and the reason for that is –
 JUSTICE ALITO: – let me just interrupt you – judges have been enjoined from performing unlawful acts.

Interruptions of this sort are not unique, with recent work demonstrating that female advocates and justices are, in the aggregate, interrupted more frequently and given less time to speak than their male colleagues (see, for example, Jacobi and Schweers 2017). This aggregate pattern of gendered interruptions relates to an emerging consensus on the influence of gender on judicial behavior more generally that covers the gamut from oral argument (e.g., Jacobi and Schweers 2017; Patton and Smith 2017) to deliberation (e.g., Boyd, Epstein, and Martin 2010) and through to voting (e.g., Collins and Moyer 2008; Boyd, Epstein, and Martin 2010; Szmer, Sarver, and Kaheny 2010; Szmer *et al.* 2013; Gleason and Smart 2022). Clearly, women being interrupted more frequently has stark consequences for their participation in political and legal spaces. Interruptions are a way for someone to assert dominance over a conversation (Zimmermann and West 1996), and that the justices might more frequently assert their dominance over female advocates than male advocates detracts from the ability of women to fully participate at the Court.

But establishing that women are being interrupted more frequently *because* they are women is a difficult causal problem. In the case of Supreme Court oral argument, for instance, women may systematically represent different ideological perspectives where the probability of being interrupted is greater. Likewise, as Solicitor General Prelogar's affirmative action quote demonstrates, given systemic disadvantages, the average woman who pursues and gains the role of advocate before the Supreme Court is likely to be systematically different than the average man who does the same, and their relationship to the Court's underlying ideological divisions may reflect this. In both cases, differences in interruption rates are not *directly* attributable to gender, though they are *indirectly*. By way of example, consider gender differences in judicial appointments, the pinnacle achievement for those in the legal profession, by the party of recent presidents. During his four years in office, approximately 24% of federal judges appointed by then-President Donald Trump were women, compared to 42% of appointments during the tenure of then-President Barack Obama (Gramlich 2021); as of the end of July 2022, fully 77% of the appointments by then-President Joe Biden had been women. These appointment dynamics mean the ideology of the politician in power influences judicial career prospects differently by gender and may analogously influence the distribution of observed ideologies of justices and advocates, generating difficult causal entanglements that render standard modeling approaches dangerous for causal interpretations.

We untangle the difficult causal question of whether female advocates are interrupted more often at oral argument *because* they are women. To do so, this paper makes three important contributions to the existing literature on gender and oral argument. First, we provide a justice-level analysis, estimating individual justice interruption rates to isolate behaviors and disentangle a single justice's interruption propensity from others. Second, it introduces a research design tailored for the complex nature of conversations like Supreme Court oral arguments, employing a new Markov assumption for conversational chunks and normalizing interruption rates by the number of tokens spoken by advocates, thus allowing for direct gender-based comparisons. Third, the study compares multiple causal pathways of gender bias with other potential causal pathways such as ideological alignment and style, using precise definitions and non-parametric estimation to ensure clarity and interpretability of assumptions and strategies.

Our careful causal research design is employed with data from over 3,000 oral argument sessions featuring over 770,000 advocate statements (or "utterances") by more than 4,000 unique advocates. The analysis demonstrates the bias that female advocates face before the Court. Throughout, we find consistent evidence most justices interrupt female advocates more often than male advocates. Critically, we find that the effect of gender on justice interruption behavior far outweighs that of other hypothetical influences on interruption behavior. Even ideological factors pale in comparison to the gender effect, with the gender effect anywhere from three to five times greater than the effect of ideological disagreement, or arguments where justices are likely – based on their ideological predispositions and the argument being made by the advocate – to disagree with the advocate. The effect is identifiable among most justices on the Court, but particularly prominent among the moderate and conservative justices. A series of corroborative analyses further establish the primacy of gender as the primary factor driving interruption behavior among most justices. Moreover, in establishing this effect, our causal framework permits us to state all our assumptions explicitly rather than let them be left up to interpretation (Lundberg, Johnson, and Stewart 2021), and provides a more careful and robust approach for this critical area of legal and political inquiry.

Gender Bias and Oral Argument

Women face significant barriers to participation in deliberative settings and the legal profession. Women speak less (Karpowitz and Mendelberg 2014; Mendelberg and Karpowitz 2016) and must navigate different linguistic expectations compared to men (e.g., Hancock and Rubin 2015; Gleason and Smart 2022). In Congress, women get fewer leadership opportunities (Kanthak and Krause 2012; Barnes 2016) and are interrupted more often (Miller and Sutherland 2022). Despite increasing law school admissions, the vast majority of prestigious legal positions remain male-dominated (St. Eve and Luguri 2021). At the Supreme Court, the disparity is stark, with only 18% of attorneys arguing before the Court in the 2020-2021 term being women (St. Eve and Luguri 2021).

During the 2022 oral arguments in two cases addressing affirmative action in college admissions, Solicitor General Prelogar highlighted the extreme gender disparities among advocates arguing before the Court, emphasizing the broader implications for women's career prospects in law. She remarked that out of 27 advocates

the Court would hear from that term, only two were women, despite women comprising over half of law school graduates.

Beyond this disparity in representation, what of disparities in treatment by gender once advocates are participating at the Court? We focus on one area, oral argument, where even the female *justices* on the Court – equals by design of their male colleagues – are interrupted more frequently than the male justices (e.g., Jacobi and Schweers 2017). The advocate, with limited time to present, can be freely interrupted during that time by justices who may ask questions, respond to particular arguments, or simply speak. Importantly, these arguments are opportunities for the justices to gain information (e.g., Johnson 2004), elucidate issues, and possibly change another justice's vote (e.g., Johnson 2004; Black, Johnson, and Wedeking 2012; Ringsmuth, Bryan, and Johnson 2013). The arguments are also an increasingly public presentation of the Court, with live audio broadcast as well as recorded audio and transcripts released shortly afterwards. In all, oral argument is a critical, public-facing component of the work of a justice on the Court.

What, therefore, influences justice behavior at oral argument? As elsewhere (e.g., Zorn and Bowie 2010), scholars have demonstrated the influence of a justices' ideological preferences on their behavior during oral arguments. Indeed, recent work suggests a growing influence of ideology on behavior at oral argument (Jacobi and Sag 2019). Liberal justices are more likely to interrupt conservative arguments, and conservative justices are more likely to interrupt liberal arguments (Johnson, Wahlbeck, and Spriggs 2006; Jacobi and Schweers 2017). Notably, justices are particularly likely to pose questions to legal representatives aligned with ideologies opposite their own; for example, the number of questions asked to each side has emerged as a consistent predictor of the likely winner of a case in the past, as those sides asked more questions are viewed as the likely loser (Roberts 2005). Taken together, the evidence demonstrates that a justice is more likely to interrupt, more likely to speak to, and more likely to ask questions of an attorney arguing against the justice's perceived ideological inclinations.

Perhaps unsurprisingly then, the justice's behavior at oral argument is predictive of how the justice votes in the case (see, for example, Jacobi and Rozema 2018; Dietrich, Enos, and Sen 2019a). Both the number of questions (Roberts 2005) and the tone of the questions (Black *et al.* 2011; Dietrich, Enos, and Sen 2019a) are predictive of how justices ultimately vote. For advocates, the time is therefore invaluable in structuring the discussion of the case, setting the table for a possible successful outcome, and making their pitch to the broader public (Sullivan and Canty 2015; Jacobi and Schweers 2017).

Interruptions at Oral Argument

The growing importance of oral argument has been accompanied by attention to the conversational dynamics at play, and particularly the extent to which the conversations are marked by interruptions. The earlier exchange between Justice Alito and Solicitor General Prelogar is one example. Interruptions can be broken down into two different types: justices interrupting the advocates and justices interrupting other justices. In the latter case, the interruptions are often deemed inappropriate; the justices are equals, and thus none should interrupt another. Responding to concerns that the female *justices* were not getting the opportunity to speak, the Court recently

instituted new rules around oral argument, with each justice permitted to take a turn to ask questions uninterrupted after an attorney's allotted time has expired.¹

In the former case, interruptions of advocates are in the design of oral argument. Indeed, the Court's *Guide for Counsel* states, "If you are speaking and a Justice interrupts you, cease talking immediately and listen" (9). Interruptions stop the development of an argument so that a justice may move the argument in a different direction. Importantly, justice interruptions occur more frequently today than in the past (Sullivan and Canty 2015). Yet while the Court sought to actively deal with interruptions of *justices*, for interruptions of *attorneys*, the Court's response has been more limited, with the Court only instituting a new norm around trying to allow attorneys to speak uninterrupted for the first two minutes.²

The new norm of brief uninterrupted time recognizes the challenge that interruptions introduce for advocates attempting to develop their argument. Accordingly, scholars have likewise turned their attention to conversational dynamics at oral argument. In Table 1, we summarize prior research on this, and specifically work examining gender-related bias and U.S. Supreme Court oral arguments. As the table makes clear, scholars have identified many different manifestations of bias, generally falling into three categories: speaking time, interruptions, and case outcomes. Overwhelmingly, scholars have measured each at the level of the case (or attorney), and have then estimated some form of parametric model. Importantly, only Patton and Smith (2020), in their analysis of justice speaking time, examine differences at the level of individual justices. Our analysis therefore diverges from this prior work by shifting the focus to the individual justice and their interactions with litigants, better specifying the causal pathways and relationships, improving the modeling and measurement of conversational dynamics through modeling of a justice-level token-normalized interruption rate, and limiting parametric assumptions. Of course, doing so requires several important empirical advances, and we outline these at length below.

Before doing so, though, we note that we root our analysis in this rich prior work and build our expectation that two factors – beyond gender – might influence interruption behavior: whether a justice agrees or disagrees with the argument being made (or *ideological alignment* with the advocate) and the stylistic *quality* of the argument. Unfortunately, both factors are potentially intertwined with advocate gender. In terms of style, the Court publicly lays out guidance. The *Guide for Counsel* notes advocates should "speak in a clear, distinct manner" and "[b]e careful to use precise language" (*Guide for Counsel*, 10). Further, justices have beliefs about the quality of oral argument; former Justice Blackmun even graded the quality of an advocate's oral argument (Johnson, Wahlbeck, and Spriggs 2006). Thus, justices may be more likely to interrupt when the stylistic quality of the argument is low. On this front, there are two ways we might expect female counsel to be different than male. First, given the barriers we note above, women may, in general, have higher stylistic quality than men arguing before the Court. By way of comparison, prior work on Congress has noted that women work harder and better represent their constituents (Lazarus and Steigerwalt 2018). Considered in the context of a legal profession where

¹Adam Liptak, "Supreme Court Tries to Tame Unruly Oral Arguments", *New York Times*, November 1, 2021.

²Dan Berman and Ariane de Vogue, "Supreme Court gives lawyers 2 minutes with no interruptions," *CNN*, October 3, 2019.

Table 1. Summary of Prior Work Examining Interruption Behavior During U.S. Supreme Court Oral Argument. We compare this to the contribution of this work in the final line in which Justices are interrupting Advocates (Justice→Advocate).

| Authors | Dependent variable(s) | Model(s) | Justice-level? |
|-------------------------------------|---|-------------------------|----------------|
| Jacobi and Schweers (2017) | Justice→Justice Interruptions | Many | None |
| Patton and J. Smith (2017) | Advocate Length of First Speech; Average Length of Speech | Truncated Poisson | None |
| Lindom, Gregory, and Johnson (2017) | Advocate Speaking Time; Justice Sentiment; Justice→Advocate Interruptions | OLS and Tobit | None |
| Feldman and Gill (2019) | Justice→Justice Interruptions | Logit | None |
| Jacobi and Sag (2019) | Justice→Advocate Interruption Count Per Case | OLS | None |
| Patton and J. L. Smith (2020) | Proportion of Words by Justices | Justice match for Tobit | Partial |
| Hack and Jenkins (2022) | Petitioner Win | Logit | None |
| Gleason and Smart (2022) | Justice Vote | Logit | None |
| This work | Justice→Advocate Interruption Rate, Token-Normalized | Non-parametric | Yes |

they are consistently excluded from high-prestige positions, women participating in oral argument are likely to make arguments of exceptionally high quality. That being the case, quality differences may mask underlying gender differences.

Turning to the entwining of ideology and gender, the issue is nicely illustrated by Jacobi and Schweers (2017), who note in their analysis of justice interruptions of other justices that “ideology and gender seem to work together in an interactive effect” (1487). Disentangling these is therefore critical to identifying the gender effect, as “results confirm that there is a clear ideological direction to interruptions on the Supreme Court” (Jacobi and Schweers 2017, 1379). The issue is further complicated because there is a substantial gender effect on the type of cases that women lawyers represent at the appellate level (St. Eve and Luguri 2021). As an analysis of the Seventh Circuit (St. Eve and Luguri 2021, 21) revealed,

Women argued more often in criminal cases and other cases that involved the government – like immigration and habeas cases – and at lower rates in civil cases. And among civil cases, the gender gap was particularly pronounced in cases that may be perceived as complex due to the structure of the case (class actions), its content (e.g., antitrust), or money at stake (as indicated through the involvement of a major law firm). Women were more likely to argue cases that involved the government, especially federal and local governments. An attorney representing the government was twice as likely to be a woman (compared to attorneys who represented non-governmental entities).

The gender disparity in representation means that women are more likely to represent particular types of arguments at the Supreme Court, arguments that might make them systematically more likely to be ideological allies (or opponents) of the justices.

In all, prior work is constrained by the struggle to disentangle the influence of gender on interruptions from other factors. The challenge we confront echoes many others encountered by scholars studying hypothetical influences on judicial behavior, as justice's ideological preferences are entwined with other justice- and case-level characteristics. As a result, we have preliminary evidence from prior work that gender matters, but little understanding of *how* and *how much* it matters. To address this gap, we introduce a causal framework that lays out how gender influences interruption behavior at the Court. Our approach demonstrates that the magnitude of the effect of gender on interruption behavior is well beyond the magnitude of the effect of ideology (and style) on interruption. We turn now to describing our approach.

Research Design

Our goal of quantifying the effect of advocate gender in oral argument is inherently a causal goal. As Lundberg, Johnson, and Stewart 2021 explain, in the social sciences, the “dominant mode of quantitative inquiry” has been “hypotheses about regression coefficients.” However, the focus on regression coefficients limits a researcher to a particular class of models and parametric assumptions. In contrast, stating an explicit theoretical causal estimand allows researchers to thoroughly investigate *what* the aim of the statistical analysis is and make explicit the assumptions behind the analysis.

In our setting, we ideally want to estimate a counterfactual: if everything else in an oral argument remained the same, but we swapped a female advocate for a male advocate, would judges behave differently? In causal terminology, we are interested in the average treatment effect of advocate gender on a justice interrupting that advocate. Following causal frameworks for quantitative social science (Morgan and Winship 2015; Lundberg, Murray, and Stewart 2021), we discuss the decisions behind our research design – from the idealized causal experiment to causal estimands that can be estimated from data.

Unit of Analysis. In order to precisely state a theoretical causal estimand – for example, the average treatment effect of gender on interruption – we first need to define our unit of analysis, the quantity that is our idealized counterfactual and over which we aggregate.

Definition of a valid chunk. Oral arguments are composed of *utterances* – a continuous piece of speech from a single speaker – made by both advocates and justices. We are interested in aggregating them to short multi-utterance *chunks* of oral arguments, each representing an exchange between justice and advocate, to investigate our causal question about justices interrupting advocates. We define a *valid chunk* as four or more contiguous utterances in which there are exactly two speakers – a single justice and a single advocate – where the advocate makes the first utterance, and each speaker has two or more utterances.³ Valid chunks are extracted with a greedy algorithm, iterating through utterances in the transcript, adding each utterance to the current chunk as long as they are from the same advocate–justice dyad, otherwise starting a new candidate chunk. We require a chunk to begin with the advocate so there is an opportunity to be interrupted by the justice and to have a

³Our approach deviates from that of Miller and Sutherland (2022) as we define *chunks* as exchanges between justice and advocate lasting at least four turns, whereas they define a chunk as an individual turn, or what we are calling an utterance.

minimum of 20 tokens total across all advocate utterances (for definition of Y below). The earlier exchange between General Prelogar and Justice Alito is a valid chunk.

Formally, we define our unit of analysis for a chunk with Chunk ID i , having Justice ID j_i , and Advocate ID a_i . For convenience, we often refer to chunk i with notation ij to also indicate its justice j , since our primary analysis is of justice's interruption behavior within chunks. Although we report aggregate measures, most of our analysis focuses on individual justices since we know each justice has their own background, experience, and implicit or explicit biases. Because our unit of analysis consists of a single advocate, we can condition on specific attributes of the advocate, including their gender and ideology.

Operationalization of variables. One contribution of this paper is precise, interpretable, and careful operationalization of the following variables: gender signal of an advocate (T), ideological alignment of an advocate and justice (A), and token-normalized interruption rates (Y).

Gender signal of an advocate (T). Thus far, we have been relatively loose in our description of the "gender" of an advocate and what it means to define a causal quantity of gender. We formally conceptualize "gender" as the binary⁴ gender signal given by an advocate's outward appearance. As previous researchers have discussed at length, we should be wary of estimating the causal effects of "immutable characteristics" such as gender or race (Berk, Li, and Hickman 2005; Holland 2008; Sen and Wasow 2016). However, we follow previous work that argues that *gender signal* is a variable we could hypothetically manipulate. For example, we could hire an actor to say the exact same script but flip the gender presentation of the actor. We operationalize the *gender signal* of advocates through a two-stage process, detailed in the [Supplementary material](#).

Ideological alignment of advocate and justice (A). To capture the ideological alignment of the justice and the advocate, we require measures of advocates' and justices' ideological preferences. For the justice, we measure ideology using the average (arithmetic mean) of the justice's time-varying Martin-Quinn score (Martin and Quinn 2002), and treat any values less than zero as a liberal justice and values greater than zero as a conservative justice. For the advocate, we rely on the Supreme Court Database's (Spaeth *et al.* 2021) coding of the ideological direction of the decision of the Court and whether or not the advocate won; if the advocate won and the direction of the decision is conservative, we indicate the advocate was making a conservative argument to the Court and vice versa for liberal arguments. With a measure of justice ideological proclivities and a measure of the ideological direction of the attorney's argument, we create our ideological alignment variable for each chunk i :

$$A_i = \begin{cases} 1 & \text{if advocate } a_i \text{ and justice } j_i \text{'s ideological preferences match} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

We define ideological alignment for an advocate and justice in the context of a specific case; an advantage of this approach is that attorneys – many of whom appear

⁴While we understand there are more than two genders, the U.S. Supreme Court never introduces an advocate with an explicitly non-binary gender title like "Mx.", and thus we restrict our analysis to binary gender.

multiple times before the Court – can be ideologically aligned with a justice in one case but not in another case, as the measure is argument-based.

Token-normalized interruption rates (Y). In creating the outcome variable, Y , we aim to account for variation in the speaking rates of various advocates. Thus, for a single valid chunk i , we define the token-normalized interruption rate as

$$Y_{ij} = \frac{\text{number of advocate utterances interrupted by justice } j \text{ in chunk } i}{(\text{number of advocate tokens in chunk } i)/1000} \quad (2)$$

Intuitively, this represents the following: if an advocate is trying to say 1,000 words during arguments, how many interruptions from a justice would the advocate endure (on average) by the time they got to 1,000 words?

In our data, interruptions are identified deterministically by two different markers in the transcript: double dashes and two dots. Note, these transcripts are manually created and follow standardized protocols for recording interruptions, a fact we leverage to automatically identify all interruptions. Later, we look at the robustness of the transcription of interruptions over the transcript years in [Figure 1](#).

Unit specific quantity. Following Lundberg, Murray, and Stewart (2021), we define a counterfactual quantity of interest for a single unit: for a specific justice and chunk, the difference in interruption rates if the advocate was female or male. We use *potential outcomes notation* (Rubin 1974) in which $Y(T=0)$ and $Y(T=1)$ represent the potential (counterfactual) outcomes given that the unit is treated, $T=1$, and not treated, $T=0$, respectively. Formally, for a given justice, j , and chunk, i , we are interested in the the following *unit-specific quantity* of interruption rate given counterfactual genders:

$$Y_{ij}(T_i = F) - Y_{ij}(T_i = M) \quad (3)$$

Target population. Lundberg, Murray, and Stewart (2021) emphasize that a careful research design must also state a target population, “over whom or what do we aggregate [the] unit-specific quantity?” For the remainder of our analysis, we only include justices who participate in more than 1,000 valid chunks in order to ensure we have enough data to reliably support any research conclusions.⁵ Given this set of valid justices, our target population is all advocates a justice has or would encounter.

Theoretical causal estimand. For a fixed justice, how much would intervening to change the gender signal of the advocate change the rate the justice interrupted the advocate? The target estimand, the average treatment effect for justice j , is the average of this unit-specific quantity over all advocates and chunks.

$$\tau_{\text{Gender}}^j := E(Y_{ij}(T_i = F) - Y_{ij}(T_i = M)) \quad (4)$$

⁵Based on our filtering criteria, we considered fifteen justices: Anthony Kennedy, Antonin Scalia, Byron White, David Souter, Elena Kagan, John Roberts Jr., John Paul Stevens, Ruth Bader Ginsburg, Samuel Alito Jr., Sandra Day O’Connor, Sonia Sotomayor, Stephen Breyer, Thurgood Marshall, Warren Burger, and William Rehnquist.

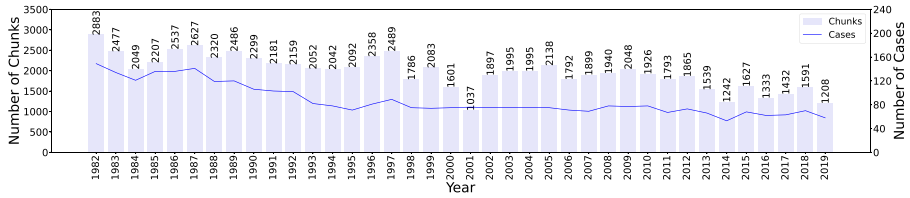


Figure 1. Number of Chunks and Cases per Supreme Court Term. Plots indicate the number of valid chunks (bars) and the number of cases (line) per October Term.

Causal identification assumptions. In order to estimate causal effects from data, one needs to state the causal *identification* assumptions that will link estimates from data to the theoretical causal estimand. We state these assumptions explicitly, as many of these assumptions cannot be measured empirically.

Markov assumption for conversational chunks. We make the assumption that the conversational dynamics between a justice and an advocate in one chunk do not influence the conversational dynamics in a subsequent chunk. This is a Markov independence assumption over conversational chunks. We acknowledge this is a strong assumption; however, if we do not make this assumption and instead assume previous interactions influence the current one, we need to account for all the text and interruptions of the previous chunks as potential confounders. In this scenario, we very quickly run into causal positivity violations, $0 < Pr(T = 1 | X = x) < 1 \forall x$, where X is a variable representing previous conversational dynamics. This would make the causal effects not *identified*, and we would be unable to reach any conclusions. Thus, we opt for our Markov assumption.

No unmeasured confounding. We assume there are no unmeasured confounding variables that cause both treatment – gender signal – and outcome – interruption rates. See Table A1 of the [Supplementary material](#) for one measured confounder, the case topic. See our section on corroborative analyses for evidence results remain the same despite possible threats to this assumption such as speech disfluencies, advocate experience, and ideological alignment as causal mediators between gender signal and interruption.

Empirical estimands. Given the aforementioned causal identification assumptions, we translate the theoretical estimand – Equation 4 – to an empirical estimand that can be estimated from data. We use τ to refer to theoretical estimands and θ for their corresponding empirical estimands. In these empirical formulations, T_i refers to the observed advocate gender in chunk i , heuristically inferred from the text as described previously. The empirical estimate of average treatment effect for a given justice is simply the difference in mean interruption rates between female and male advocates with whom the justice interacts during valid chunks,

$$\theta^j_{\text{Gender}} := \left(\frac{1}{n_{j,T_i=F}} \sum_{i:j_j=T_i=F} Y_{ij} \right) - \left(\frac{1}{n_{j,T_i=M}} \sum_{i:j_j=T_i=M} Y_{ij} \right), \quad (5)$$

where, for example, $n_{j,T_i=F}$ indicates the number of chunks for justice j where advocates’ gender is female.

We can compare this to an analogous empirical quantity about difference due to ideological alignment:

$$\theta_{\text{Ideological Alignment}}^j := \left(\frac{1}{n_{j, A_i=1}} \sum_{i:j_i=j, A_i=1} Y_{ij} \right) - \left(\frac{1}{n_{j, A_i=0}} \sum_{i:j_i=j, A_i=0} Y_{ij} \right) \quad (6)$$

Then for each justice, we can compare the magnitudes of θ_{Gender}^j and $\theta_{\text{Ideological Alignment}}^j$.

Data

We acquired the Supreme Court Oral Arguments Corpus (Danescu-Niculescu-Mizil et al. 2012) from the Cornell Conversational Analysis Toolkit (ConvoKit) (Chang et al. 2020). The corpus covers more than 8,000 oral argument transcripts from more than 7,700 cases decided between 1955 and 2019. Due to limitations related to the reliability and consistency of the transcription, particularly with respect to the identification of the speaking justices, we limit our analysis to 1982 through 2019. The corpus is organized at the level of an utterance. In all, we have 776,193 utterances across 3,424 cases. Of the 4,025 unique advocates, we identify 551 as female, or approximately 14% of all unique advocates. Overall, approximately 25% of all advocate utterances are interrupted, reflecting the extent to which the dialogue is driven by justices.

We segment the oral argument transcript utterances into valid chunks as described in the Unit of Analysis section above. In the bottom half of Table 2, we provide summary statistics on these chunks. We identify 65,768 valid chunks, accounting for more than 23 million tokens spoken across more than 520,000 utterances.

As evidence that the chunking approach offers a valid representation of oral argument, in Figure 1 we plot the number of valid chunks we identify each Court term (the bars) alongside the number of cases decided by the Court during that term (the line).⁶ The number of chunks tracks closely with the number of cases over time. Our chunking approach reflects the drop in the number of cases heard by the Court during an average term over time and provides broad and consistent coverage of oral argument during the entire time period under study.

Notably, the rate at which women participate in these exchanges *has* changed significantly over time. In Figure 2, we plot the number of valid chunks featuring a woman either as an advocate (left panel) or as a justice (right panel). The participation of women as advocates increased marginally from the 1980s until about the turn of the century, but has generally leveled out to an average of around 12% of arguments more recently.

On the other hand, the proportion of chunks featuring a female justice has shifted consistently, and unsurprisingly, with changes in the membership of the Court. The right panel of Figure 2 makes plain the shifts associated with the addition or subtraction of female justices. Indeed, the shifts are approximately proportional to the gender composition of the Court itself during each October Term, reflecting the

⁶The consistency of the chunking approach over time provides evidence that transcription issues related to identifying justices and advocates are not frequent during the period of analysis, as our chunking approach relies on identifying justice and attorney frequent in specifying valid chunks.

Table 2. Descriptive Statistics of the Corpus and Dataset

| Corpus details | | |
|---------------------------------------|------------|--------------|
| Years | | 1982–2019 |
| Cases | | 3,424 |
| Unique advocates | | 4,025 |
| Unique female advocates | | 551 |
| Chunk details | All data | Valid chunks |
| Count | – | 64,164 |
| Number of tokens | 37,880,545 | 23,065,962 |
| Number of utterances | 776,193 | 508,189 |
| Prop. advocate utterances interrupted | 0.25 | 0.21 |
| Median num. tokens per chunk | – | 296 |
| Median num. utterances per chunk | – | 6 |

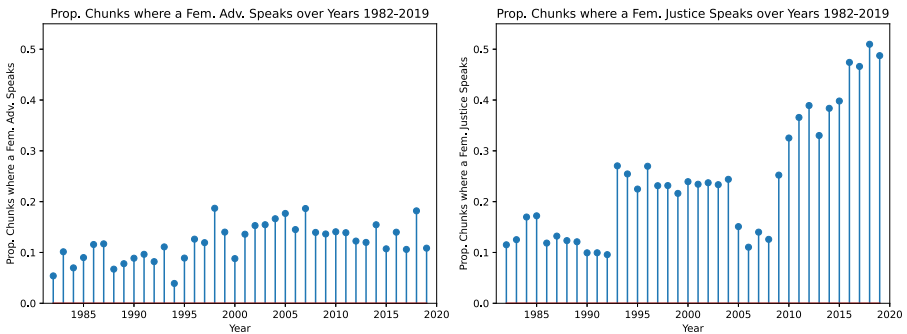


Figure 2. Participation of Women in Oral Argument Exchanges on the U.S. Supreme Court. These plots provide the proportion of exchange chunks featuring a female advocate (left panel) or female justice (right panel) by October Term.

extent to which the chunking approach validly captures the dynamics of oral argument on the Court.

Measuring Interruption Behavior

With the data in hand, we turn to measuring interruptions. We analyze the token-normalized interruption rate, where interruption is defined as the justice interrupting an advocate. By way of interpretation, a rate of 4.0 would indicate that for every 1,000 tokens of speech for the advocate, we would expect them to be interrupted four times.

In Figure 3, we plot over time the Court-level average of the token-normalized interruption rate (blue) and also, for comparison purposes, plot the token-normalized rate of justices being interrupted by advocates (red). Two dynamics bear mentioning and again speak to the validity of our measurement approach. First, there is a marked increase in the token-normalized interruption rate of advocates in the mid-1990s. This dynamic is consistent with Jacobi and Sag (2019) who identified this as the point justice behavior began to reflect the polarization of the political system. Second, the increase persists from that point forward. Indeed, the *lowest* rate during the 25-year span after the 1994 term is 9.9 in OT2001, an interruption rate greater than any of the preceding years. As the Court’s bar expanded to include marginally more female advocates, the rate of interruptions likewise increased. Of course, this

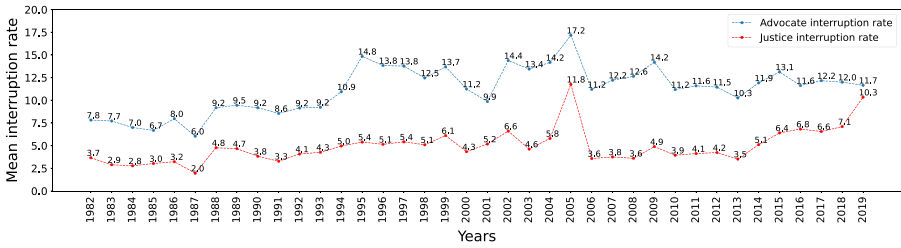


Figure 3. Interruption Rate Over Time. This plot presents the Court-level average of token-normalized interruption rate (y-axis) by October Term (x-axis). The blue line is the advocate being interrupted by justices and the red line is justices being interrupted by advocates.

also coincided with changes in the ideological composition of the Court as well as the gender composition of the Court, both factors that may contribute in different ways to the gendered differences in observed interruption behavior. In all, this further underscores the necessity of a careful causal research design to understand and disentangle the complex causal story.

The Causal Effect of Gender on Interruptions

We begin with an analysis of the rates that justices interrupt male and female advocates differently.⁷ In Table 3, we present the interruption effect of gender (column 2) and, for comparison, the interruption effect of ideological alignment (column 3). The rows calculate the interruption effects across different levels of justice aggregates: tabulated⁸ across all justices, across only male justices, and across only female justices. Recall, θ_{Gender} is equal to $E[Y|\text{Gender} = \text{Female}] - E[Y|\text{Gender} = \text{Male}]$, such that positive values indicate higher interruption rates for female advocates, and negative values indicate higher interruption rates for male advocates.

Across all justices, female advocates are interrupted more frequently, with a point estimate of 0.89 indicating they are interrupted on average nearly one additional time per 1,000 words in comparison to male advocates. The dynamic is more striking when the results are disaggregated by the gender of the justice; male justices interrupt female advocates more than one additional time per 1,000 tokens. In comparison, though female justices are also more likely to interrupt female advocates than male advocates, the effect size (0.43) is far smaller, well less than half the magnitude of the effect among male justices.⁹

⁷In the supplemental material, we explore differences in interruption behavior between gendered and non-gendered issues, demonstrating that in women’s issue cases—which cover approximately 2.4% of chunks—women are interrupted less frequently. Because they represent a unique dynamic and comprise such a small number of cases and chunks, we exclude these from the analyses below.

⁸This is a simple mean over the per-justice effects, $\theta_{\text{Gender}} = \frac{1}{N} \sum_j \theta_{\text{Gender}}^j$ and $\theta_{\text{Ideological Alignment}} = \frac{1}{N} \sum_j \theta_{\text{Ideological Alignment}}^j$

⁹In the supplemental material, we address and reject the possibility that differences in the rates of advocates interrupting justices may lead to different conversational dynamics and influence the differences in rates of justices interrupting advocates. We also address concerns that the results are driven by short, back-channel communications (“OK”, “Right”). The results are consistent when excluding these back-channel communications.

Table 3. Effects on advocate interruption rate, aggregated by justice gender. Positive θ_{Gender} indicates justices interrupt female advocates more often than they interrupt male advocates. Negative θ_{Gender} indicates justices interrupt ideologically opposed advocates more often than they interrupt ideologically aligned advocates. Here \pm indicates the 95% confidence interval via bootstrap sampling of valid chunks. The last column shows the absolute ratio of point estimates to aid interpretation.

| Justices | θ_{Gender} | $\theta_{\text{Ideological Alignment}}$ | $\frac{\theta_{\text{Gender}}}{\theta_{\text{Ideological Alignment}}}$ |
|----------|--------------------------|---|--|
| | | | |
| All | 0.89±0.36 | -0.25±0.23 | 3.59 |
| Male | 1.06±0.43 | -0.20±0.26 | 5.34 |
| Female | 0.43±0.71 | -0.39±0.45 | 1.12 |

The substantive magnitude of the effect is made clear by a comparison to the effect of ideological alignment on interruption behavior. Across all justices, and particularly for male justices, the gender effect on interruption behavior dwarfs the effect of ideological alignment. Among all justices, the effect of gender is approximately 3.5 times greater than the effect of ideological alignment. Among just the male justices, though, the effect is well more than five times greater than the effect of ideological alignment. In contrast, the effect among female justices is nearly identical to the effect of ideological alignment. In all, we have clear evidence that – in the aggregate – advocate gender influences the interruption behavior of Supreme Court justices at oral argument, and the effect is largely concentrated among male justices.

Justice-Level Analysis

As opposed to analyzing aggregate rates of interruptions, justice-level analysis not only offers a more nuanced understanding of the intricate interpersonal dynamics at play during oral arguments but also offers a far superior approach to causal estimation. By shifting our focus to the individual justices, we gain a sharper lens through which to identify differences in individual behavior that collectively contribute to the observed aggregate results discussed above. In doing so, our individual-level approach acknowledges that aggregate rates are not homogenous entities but rather composite outcomes shaped by the diverse styles, preferences, and biases of each justice. An individual-level approach disentangles these multifaceted influences, empowering us to better understand all of the factors that drive a justice's interruption behavior. Such an approach is therefore pivotal for unraveling the causal impact of gender on interruption behavior on the Court.

Our analytical framework entails calculating the effects of advocate gender and ideological alignment on the interruption rate at the level of the justice based on the same rigorous estimation approach outlined above, analyzing effects by focusing on justice-specific chunk subsets. This methodological approach is distinctive in its ability to dissect justice-specific interactions, allowing us to disentangle the unique behavioral patterns of each justice.

We present the results of the justice-level analyses in Figure 4. We order the justices along the y-axis, in descending order by θ_{Gender} , and indicate for each justice the ideological classification on the basis of the average Martin & Quinn scores (where *L* is *Liberal* and *C* is *Conservative*) along with the total number of valid chunks our estimates are based on for that justice (in parentheses).¹⁰ The far left panel

¹⁰Confidence intervals are inferred using a non-parametric bootstrap, detailed in the [Supplementary material](#).

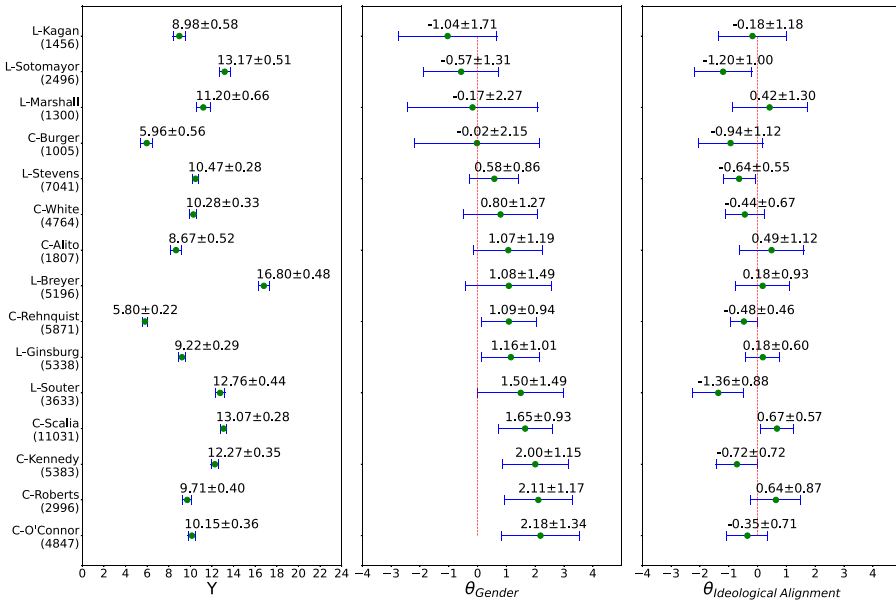


Figure 4. Plot of Justice Interruption Rates, Effect of Gender, and Effect of Ideological Alignment. Justice-specific estimates (points) and bootstrapped confidence intervals (line segments) for justice interruption rate (left panel), gender effect (center panel), and ideological alignment (right panel). Higher values indicate more interruptions. A table of these results appears in the [Supplementary material](#).

presents the rate justices are interrupting advocates in all valid chunks for which they participate, the center panel presents the gender effect on interruption rates (where higher values indicate the justice is more likely to interrupt a female advocate), and the right panel presents the effect of ideological alignment on the justice’s interruption behavior (where higher values indicate the justice is more likely to interrupt an advocate whose ideology aligns with the justice).

Four justices stand out in the analysis for the substantive and statistical significance of the gender effect in their exchanges with advocates at oral argument: Sandra Day O’Connor, John Roberts, Anthony Kennedy, and Antonin Scalia. The effects among this group range from 1.65 (for Justice Scalia) to 2.18 (for Justice O’Connor). For these two justices, the results indicate that for every 1,000 tokens spoken by a female advocate, relative to a male advocate she will be interrupted 1.65 times more frequently in exchanges with Justice Scalia, and 2.18 times more frequently by Justice O’Connor. For each, the gender effect is more than twice the magnitude of the ideological alignment effect.

However, only four justices out of the fifteen in our study interrupt male advocates more frequently than female advocates at *any* magnitude, and for each, the 95% confidence intervals cover zero. Therefore, we cannot confidently say, based on this data, that *any* justice interrupts male advocates more frequently in comparison to female advocates.

Another notable dynamic emerging is the extent to which the gender effect covers nearly the entirety of the ideological elements of the Court’s composition. Three of the five most liberal justices in our analysis interrupt male advocates more frequently than female advocates, though – to reiterate – the confidence intervals for each cross

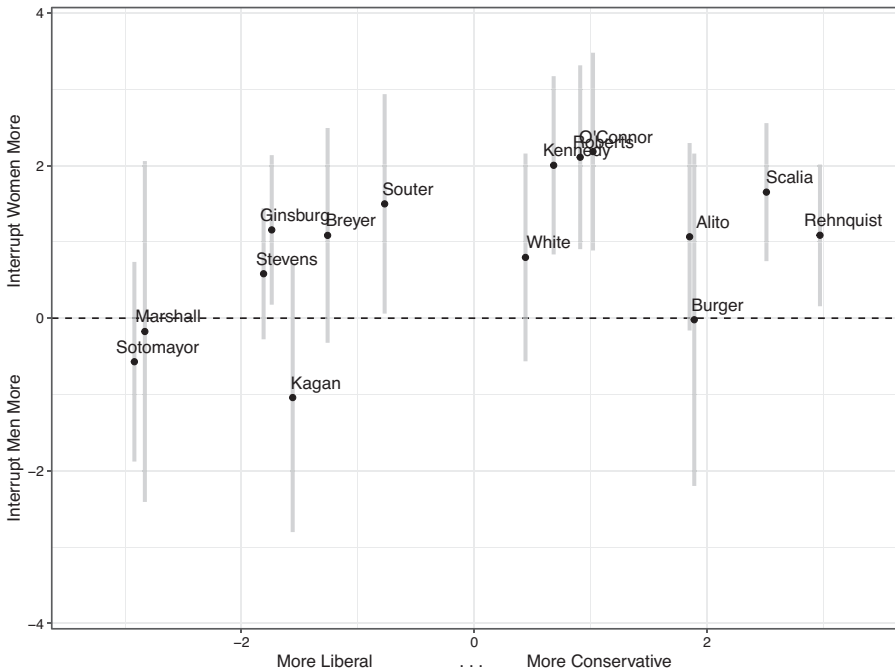


Figure 5. Justice Gender Effects θ_{Gender}^j (y-axis) by Martin & Quinn Ideology Scores (x-axis). Gender effects are calculated as in Figure 4.

zero. To illustrate justices' ideological preferences more clearly, in Figure 5, we plot the justice-level gender effects by the justice's Martin & Quinn ideology score. First, seven of the eight conservative justices interrupt female advocates more frequently than male advocates, with Burger being the only exception. Among these, the effect sizes we identify are statistically significant at the 95% level for five of the eight. The largest point estimates are concentrated among the more moderate conservatives (Kennedy, O'Connor, and Roberts); interestingly, each was also often portrayed as the median justice during at least part of their tenure on the Court.

Among the more liberal justices, two patterns bear mentioning. First, uncertainty around the point estimates in most of the cases cautions against firm conclusions of a gender effect. Of the seven justices rated as liberal, only two – Justices Ginsburg and Souter – demonstrate a statistically significant effect of advocate gender on their interruption behavior; in both cases, the justice interrupts female advocates more frequently than male advocates, and at rates generally commensurate with the group average among their more conservative colleagues.

Second, three liberal justices – Sonia Sotomayor, Thurgood Marshall, and Elena Kagan – interrupt men more frequently than they do women within our data. This group is particularly interesting in that it contains the two female justices (Sotomayor and Kagan) who have most recently joined the Court within our timeframe.¹¹ While

¹¹Justices Amy Coney Barrett and Ketanji Brown Jackson joined more recently and offer an avenue for future exploration.

O'Connor and Ginsburg were trailblazing women in the legal profession, within our data, both are more likely to interrupt female advocates than male advocates. In contrast, Sotomayor – who joined the Court in 2009 – and Kagan – who joined in 2010 – are more likely to interrupt male advocates. Taken together, we have clear evidence of a gender effect on interruption behavior that tracks across most justices on the Court.

Corroboration: Causal Mediation Analysis

The analysis above begs further corroboration to support the claim that the estimated effects relate to gender specifically. For instance, one might be concerned that the observed gender effect is not a function specifically of gender, but is instead a function of differences in the ideological orientations of male and female advocates and the members of the Court. Moreover, one might question whether the broader professionalization biases faced by women in the legal profession yields a qualitatively different pool of advocates, leading to differences in interruption behavior. In both cases, one might worry that the observed gender effect does not capture the *causal* effect of gender on interruption behavior, but rather that the observed effect of gender operates through ideological differences or differences in advocates' speech quality. To address this, we turn to a causal mediation analysis (Pearl 2001; Imai, Keele, and Tingley 2010; VanderWeele 2016).

Our analysis focuses on two pathways through which the gender effect may operate: differences in the ideological orientation of advocates and justices, and differences in the quality of arguments. In the former case, as discussed above, professionalization and prospects for career progression may be different for women representing particular types of arguments and cases before the Court, leading to different patterns of representation at the Court for particular arguments that may be more or less likely to yield interruption by justices. If female advocates are more likely to represent liberal arguments during the period under study while the Court is consistently composed of a majority of conservative justices, the effect of gender may operate through the ideological differences between advocate and justice. We maintain the same approach to measuring ideological alignment and incorporate it now through a causal mediation approach.

Further, we analyze two dimensions related to argument quality: speaking fluidity and advocate experience. Consider first speaking fluidity, or the extent to which an advocate can seamlessly present their argument. At a basic level, if an advocate is proceeding seamlessly through their argument, the justice must be more aggressive to interrupt; alternatively, if an advocate is stumbling over their words, there are more opportunities for an interruption. If we expect differences in the presentation styles of women and men at the Court, whether because of differences in the legal profession or because of the efforts of advocates to conform with gender schemas (e.g., Gleason and Smart 2022), then we would expect the effect of gender to operate through differences in oral argument quality.

To capture speaking fluidity, we turn to counts of *speech disfluencies*, or breaks in the flow of a speech. Though there are many types of speech disfluencies, one of the most visible is stuttering or stopping and starting in order to repeat words or phrases. In oral argument transcripts, these types of speech disfluencies are indicated by the presence of a dash followed by a repeat of the word or phrase that had preceded the

dash. We automatically extracted these at the utterance level. In the process of doing so, we found that such disfluencies are unreliably indicated in some years of the digital versions of the transcripts, and therefore, we limit our analysis to the years 2007 through 2019.¹²

Beyond speaking fluidity, we also address *advocate experience* as a variable on the causal path between advocate gender and interruption. Again, given differences in the participation of women as advocates before the Court, women advocates may have less experience on average. If justices are more likely to interrupt advocates who are inexperienced, then the effect of gender would operate through an experience path. To capture experience, we employ a dichotomous indicator of prior experience arguing before the Court. The first time an advocate argues before the Court, the indicator takes on the value of 0. For any additional arguments, the indicator takes on the value of 1. Though our causal mediation analysis starts in 2007, we search for prior advocate experience back to 1980 to ensure full coverage.

With the updated data, we undertake a causal mediation analysis, including speech disfluencies, experience, and ideological alignment as mediators. We calculate the natural direct effects (NDE) and natural indirect effects (NIE) via the assumptions and formulas presented in Keith, Rice, and O'Connor (2021); these estimates are intentionally non-parametric (i.e., not based on the assumptions of linear regression) and simplify to the conditional means from the data. We report the results in Table 4. We find all the natural indirect effects are near zero, indicating that the effect of gender *directly* to interruption is the main driving force for the differences in interruption we observe.

Discussion

The U.S. Supreme Court, currently characterized by a conservative supermajority, is at the center of battles over power and politics in the United States. This article tackles a fundamental question regarding the Court's business: can men and women advocates participate equally? We address the complex causal challenges involved in this question by disentangling the effect of an advocate's signaled gender from other legal, political, and social factors affecting representation before the Court. Our approach elucidates the gender effects on justice interruption behavior and allows us to examine differentials in interruption behavior at the level of the individual justice, providing insights on the behavioral underpinnings of bias in interruption behavior at the Court.

We find a pronounced gender effect, with most justices interrupting female advocates more frequently than male advocates during oral arguments. This effect, which significantly outpaces other potential explanations of interruption behavior, scales with the length of oral arguments, revealing a substantial impact over cases and terms. That said, the dynamic reverses in areas traditionally thought of as women's issues, including the right to privacy. Moreover, we find that conservative justices exhibit the bias more strongly, while some liberal justices tend to interrupt male advocates more frequently. In all, the work conclusively demonstrates that conversational dynamics on the nation's highest court are significantly shaped by the genders of the lawyers appearing before the Court.

¹²This analysis includes ten judges who were involved in greater than 1,000 chunks: Anthony Kennedy, Antonin Scalia, David H. Souter, Elena Kagan, John Paul Stevens, John Roberts Jr., Ruth Bader Ginsburg, Samuel Alito Jr., Sonia Sotomayor, and Stephen Breyer.

Table 4. Corroborative analyses. Causal mediation estimates aggregated across justices of the natural direct effect (NDE) from gender to interruption and the natural indirect effect (NIE) from gender to interruption through one of the following mediators: speech disfluencies, ideological alignment, or advocate experience. The data is from 2007-2019 and n is the number of chunks.

| All justices (n=17,801) | NDE | NIE |
|-----------------------------------|--------------|--------------|
| Speech disfluencies as mediator | 0.73 ± 0.60 | -0.07 ± 0.29 |
| Ideological alignment as mediator | 0.83 ± 0.69 | 0.04 ± 0.05 |
| Advocate experience as mediator | 0.68 ± 0.70 | -0.03 ± 0.08 |
| Male justices (n=11,821) | NDE | NIE |
| Speech disfluencies as mediator | 1.20 ± 0.78 | 0.12 ± 0.41 |
| Ideological alignment as mediator | 1.36 ± 0.92 | 0.06 ± 0.08 |
| Advocate experience as mediator | 1.18 ± 0.91 | 0.00 ± 0.11 |
| Female justices (n=5,980) | NDE | NIE |
| Speech disfluencies as mediator | -0.38 ± 0.84 | -0.03 ± 0.26 |
| Ideological alignment as mediator | -0.38 ± 0.87 | -0.01 ± 0.04 |
| Advocate experience as mediator | -0.48 ± 0.86 | -0.10 ± 0.10 |

The results have stark implications going forward. The concentration of gendered interruption effects among conservative justices, coupled with the probability of a conservative super-majority going forward, indicates a major challenge for gender equity before the Court going forward. Given that more recent liberal additions to the Court have shown a tendency to operate in the opposite direction, interrupting male advocates more frequently than female advocates, future research should build from prior work (e.g., Haire and Moyer 2015) to examine how the Court's polarization affects advocates ability to participate equitably. Moreover, while we have considered interruptions generally, we have not explored whether the interruptions were to ask questions or to make statements, though the different types are associated with different probabilities of winning or losing the case (Sorenson 2023). Therefore, future work could build to better understand how the types of interruptions vary across the advocate gender by each justice.

On this, our analyses highlight significant differences in interruption behavior among female justices. Trailblazers like Justice O'Connor and, to a lesser extent, Justice Ginsburg frequently interrupted female attorneys, contrary to more recent appointees Justice Sotomayor and Justice Kagan. This indicates generational shifts and evolving norms within the Court (see, e.g., Haire and Moyer 2015). With two more recent female appointees, analysis of these generational differences going forward could provide deeper insights into the intersection of gender and judicial behavior at oral argument. Such longitudinal studies can also better capture how societal attitudes relate to changes in behavior at the Court.

Beyond the justices, future work might also consider how different advocacy styles and strategies influence the interruption behavior of the justices, considering for instance the extent to which advocates behaving in particular ways might discourage or encourage interruptions from particular justices (Gleason and Smart 2022). Likewise, our work has not considered other intersectional disadvantages that attorneys might face, including how race, ethnicity, or other social identities might intersect with gender to influence interruption behavior.

On these above opportunities, our approach is generalizable and, along with our findings, offers clear pathways for such future research, including how recent appointments have changed these dynamics and how the changing composition of the federal judiciary relates to the dynamics observed at the U.S. Supreme Court. The approach is also generalizable beyond the United States, and future comparative work, both in the context of the states and internationally, could contextualize whether the dynamics we observe are more general or are specific to particular places and courts. Likewise, the approach generalizes to other forms of marginalization in the courtroom beyond interruptions – for instance, changes in the tone (Rice and Zorn 2021) or vocal pitch (Dietrich, Enos, and Sen 2019b) of justices in their interactions, and other potential mediators (Goelzhauser, Kassow, and Rice 2021). Throughout these avenues for future work, our analysis suggests that to understand the U.S. Supreme Court, it is vitally important to understand the ways justice behavior is shaped by gender dynamics.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/jlc.2024.7>.

Acknowledgement. Authors are in alphabetical order by last name. Authors contributed as follows: implementation of data pipeline by EC, AG, KK, DR; research design by KK, BO, DR; first draft writing of manuscript by KK, DR; editing manuscript by EC, AG, KK, BO, DR. The authors thank Mohit Iyyer, David Jensen, Paul Musgrave, Jesse Rhodes, Kelsey Shoub, Donny Snyder, members of the American Politics Working Group at UMass Amherst, and conference participants at the 2021 Causal Inference + NLP Workshop and 2022 New Directions in Analyzing Text as Data conference for comments and suggestions on earlier versions of this work. All data and materials necessary to replicate the results reported in this article are available on the *Journal of Law & Courts* Dataverse and at <https://github.com/kakeith/interruptions-supreme-court>.

References

- Barnes, Tiffany. 2016. *Gendering legislative behavior: institutional constraints and collaboration*. Cambridge University Press.
- Berk, Richard, Azusa Li, and Laura J. Hickman. 2005. Statistical difficulties in determining the role of race in capital cases: a re-analysis of data from the state of maryland. *Journal of Quantitative Criminology* 21 (4): 365–390.
- Black, Ryan, Sarah Treul, Timothy Johnson, and Jerry Goldman. 2011. Emotions, oral arguments, and supreme court decision making. *The Journal of Politics* 73 (2): 572–581.
- Black, Ryan C., Timothy R. Johnson, and Justin Wedeking. 2012. *Oral arguments and coalition formation on the u.s. supreme court: a deliberate dialogue*. University of Michigan Press.
- Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. Untangling the causal effects of sex on judging. *American Journal of Political Science* 54:389–411.
- Chang, Jonathan P, Caleb Chiam, Liye Fu, Andrew Z Wang, Justine Zhang, and Cristian Danescu-Niculescu-Mizil. 2020. Convokit: a toolkit for the analysis of conversations. *Proceedings of SIGDIAL*.
- Collins, Todd, and Laura Moyer. 2008. Gender, race, and intersectionality on the federal appellate bench. *Political Research Quarterly* 62 (2): 219–227.
- Danescu-Niculescu-Mizil, Cristian, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of power: language effects and power differences in social interaction. In *WWW*.
- Dietrich, Bryce J., Ryan D. Enos, and Maya Sen. 2019a. Emotional arousal predicts voting on the u.s. supreme court. *Political Analysis* 27 (2): 237–243. <https://doi.org/10.1017/pan.2018.47>.
- Dietrich, Bryce J., Ryan D. Enos, and Maya Sen. 2019b. Emotional arousal predicts voting on the u.s. supreme court. *Political Analysis* 27 (2): 237–243.

- Feldman, Adam, and Rebecca D. Gill. 2019. Power dynamics in supreme court oral arguments. *The Justice System Journal* 40 (3): 173–195.
- Gleason, Shane A., and EmiLee Smart. 2022. You think; therefore i am: gender schemas and context in oral arguments at the supreme court, 1979–2016. *Political Research Quarterly*.
- Goelzhauser, Greg, Benjamin J. Kassow, and Douglas Rice. 2021. Measuring Supreme Court Case Complexity. *The Journal of Law, Economics, and Organization* 38, no. 1 (January): 92–118.
- Gramlich, John. 2021. How trump compares with other recent presidents in appointing federal judges. *Pew Research Center*.
- Hack, Jonathan S., and Clinton M. Jenkins. 2022. The attorneys' gender: exploring counsel success before the u.s. supreme court. *Political Research Quarterly* 75 (3): 632–645.
- Haire, Susan, and Laura Moyer. 2015. *Diversity matters: judicial policy making in the u.s. courts of appeals*. University of Virginia Press.
- Hancock, Adrienne B., and Benjamin A. Rubin. 2015. Influence of communication partner's gender on language. *Journal of Language and Social Psychology* 34 (1): 46–64.
- Holland, Paul W. 2008. Causation and race. ,*White logic, white methods: Racism and methodology* 93–109.
- Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. A general approach to causal mediation analysis. *Psychological methods* 15 (4): 309.
- Jacobi, Tonja, and Kyle Rozema. 2018. Judicial conflicts and voting agreement: evidence from interruptions at oral argument. *Boston College Law Review* 59.
- Jacobi, Tonja, and Matthew Sag. 2019. The new oral argument: justices as advocates. *Notre Dame Law Review* 94.
- Jacobi, Tonja, and Dylan Schweers. 2017. Justice, interrupted: the effect of gender, ideology, and seniority at supreme court oral arguments. *Virginia Law Review* 103 (7): 1379–1496.
- Johnson, Timothy R. 2001. Information, oral arguments, and supreme court decision making. *American Politics Research* 29:331–351.
- Johnson, Timothy R. 2004. *Oral arguments and decision making on the u.s. supreme court*. SUNY Press.
- Johnson, Timothy R., Paul J. Wahlbeck, and James F. Spriggs. 2006. The influence of oral arguments on the u.s. supreme court. *American Political Science Review* 100 (1): 99–113.
- Kanthak, Kristin, and George A. Krause. 2012. *The diversity paradox: political parties, legislatures, and the organizational foundations of representation in america*. Oxford University Press.
- Karpowitz, Christopher F., and Tali Mendelberg. 2014. *The silent sex: gender, deliberation, & institutions*. Princeton University Press.
- Keith, Katherine, Douglas Rice, and Brendan O'Connor. 2021. Text as causal mediators: research design for causal estimates of differential treatment of social groups via language aspects. In *Proceedings of the first workshop on causal inference and nlp*, 21–32.
- Lazarus, Jeffrey, and Amy Steigerwalt. 2018. *Gendered vulnerability: how women work harder to stay in office*. University of Michigan Press.
- Lindom, Tiffany, Charles Gregory, and Timothy R. Johnson. 2017. Gender dynamics and supreme court oral arguments. *Michigan State Law Review*, no. 5, 1033–1055.
- Lundberg, Ian, Rebecca Murray, and Brandon M Stewart. 2021. What is your estimand? defining the target quantity connects statistical evidence to theory. *American Sociological Review* 86 (3): 532–565.
- Martin, Andrew, and Kevin Quinn. 2002. Dynamic ideal point estimation via markov chain monte carlo for the u.s. supreme court, 1953–1999. *Political Analysis* 10:134–153.
- Mendelberg, Tali, and Christopher F. Karpowitz. 2016. Women's authority in political decision-making groups. *The Leadership Quarterly* 27 (3): 487–503.
- Miller, Michael G., and Joseph L. Sutherland. 2022. The effect of gender on interruptions at congressional hearings. *American Political Science Review*.
- Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and causal inference*. Cambridge University Press.
- Patton, Dana, and Joseph Smith. 2017. "lawyer, interrupted: gender bias in oral arguments at the u.s. supreme court". *Journal of Law and Courts*, 337–361.
- Patton, Dana, and Joseph L. Smith. 2020. Gender, ideology, and dominance in supreme court oral arguments. *Journal of Women, Politics, & Policy* 41 (4): 393–415.
- Pearl, Judea. 2001. Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence* 411–420.

- Rice, Douglas R., and Christopher Zorn. 2021. Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods* 9 (1): 20–35.
- Ringsmuth, Eve M., Amanda C. Bryan, and Timothy R. Johnson. 2013. Voting fluidity and oral argument on the u.s. supreme court. *Political Research Quarterly* 66 (2): 426–437.
- Roberts, John G. 2005. Oral advocacy and the re-emergence of a supreme court bar. *Journal of Supreme Court History* 30 (1): 68–81.
- Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66 (5): 688.
- Sen, Maya, and Omar Wasow. 2016. Race as a bundle of sticks: designs that estimate effects of seemingly immutable characteristics. *Annual Review of Political Science* 19:499–522.
- Sorenson, Mason W. 2023. Asking versus telling: the supreme court’s strategic use of questions and statements during oral arguments. *Political Research Quarterly* 76 (4): 1559–1572.
- Spaeth, Harold, Lee Epstein, Andrew Martin, Jeffrey Segal, Theodore Ruger, and Sara Benesh. 2021. *The supreme court database*. <http://Supremecourtdatabase.org>.
- St. Eve, Amy J., and Jamie B. Luguri. 2021. *How unappealing: an empirical analysis of the gender gap among appellate attorneys*. Technical report. American Bar Association.
- Sullivan, Barry, and Megan Canty. 2015. Interruptions in search of a purpose: oral argument in the supreme court, october terms 1958-60 and 2010-12. *Utah Law Review* 2015 (5).
- Szmer, John, Erin B. Kaheny, Tammy A. Sarver, and Mason DeCamillis. 2013. The impact of lawyer gender on decision making in the united states courts of appeals. *Journal of Women, Politics and Policy* 34 (1): 72–100.
- Szmer, John J., Tammy A. Sarver, and Erin B. Kaheny. 2010. Have we come a long way, baby? the influence of lawyer gender on supreme court decision making. *Politics and Gender* 6:1–36.
- VanderWeele, Tyler J. 2016. Mediation analysis: a practitioner’s guide. *Annual Review of Public Health* 37: 17–32.
- Zimmermann, Don H., and Candace West. 1996. Sex roles, interruptions, and silences in conversation. In *Amsterdam studies in the theory and history of linguistic science*, 211–36.
- Zorn, Christopher, and Jennifer Bowie. 2010. Ideological influences on decision making in the federal judicial hierarchy: an empirical assessment. *Journal of Politics* 72 (4): 1212–1221.

Cite this article: Cai, Erica, Ankita Gupta, Katherine A. Keith, Brendan O’Connor, and Douglas Rice. 2025. “Let Me Just Interrupt You”: Estimating Gender Effects in Supreme Court Oral Arguments.” *Journal of Law and Courts*, 1–22, doi:10.1017/jlc.2024.7