

# Covering: Mutable Characteristics and Perceptions of Voice in the U.S. Supreme Court

Daniel Chen<sup>1</sup>   Yosh Halberstam<sup>2</sup>   Alan Yu<sup>3</sup>

<sup>1</sup>Toulouse Institute for Advanced Study

<sup>2</sup>University of Toronto

<sup>3</sup>University of Chicago

# Elite Law Hiring Preferences (Peppers 2006)

Byron White's preference for clerks who could perform on the basketball court

Rehnquist's penchant for hiring 3 clerks instead of 4 so they could play doubles tennis

Those preferences don't **cross** any **lines** unlike

Kavanaugh's alleged gravitation toward conventionally attractive female clerks (Lazarus 2018)

HOW DO WE MEASURE WHEN LINES ARE CROSSED?

# Civil Rights Law

Populated with plaintiffs who were punished for being different

- Courts distinguish between **immutable** and **mutable** (Yoshino 2006)
- Legal theorists suggest that discrimination, once aimed at entire groups, now aims at the subset of the group that refuses to **cover**
- **“to assimilate to dominant norms”** (Goffman 1963)
- Mutable characteristics gradually entering economic models (Fryer et al. 2010)
  - ▶ African-Americans wearing cornrows
  - ▶ Women becoming mothers
  - ▶ Jews wearing yarmulkes
  - ▶ Muslims wearing the hijab
  - ▶ Female Indonesian police applicants failing virginity tests
- One sentence (but no data) in two summaries of the empirical literature
  - ▶ **“Recently, the emphasis on ‘fit’ [of] a prospective employee .. has raised the spectre of a new form of subtle discrimination”** (Bertrand and Duflo 2016)

# Civil Rights

When courts make employment depend on covering, it is perceived to legitimize second-class citizenship for the subordinated group (Balkin 2011)

- Is this legally-sanctioned differential treatment a harm? (Yoshino 2000)
- This distinction between *being* and *doing* incentivizes assimilation
- Importance of assimilation in identity formation (Fryer & Torelli 2010)
  - ▶ Document the costs of assimilation in the NLSY (Mueller-Smith 2014)
  - ▶ Linguistic choices of minorities  $\sim$  labor market outcomes (Grogger 2011)
  - ▶ Female lawyers pay coaches to sound more masculine (Starecheski 2014)
  - ▶ Sears screened for masculine voices (EEOC vs. Sears)
- Homophily: speaking in a certain manner  $\rightarrow$  trust? (Alesina & Ferrara 2005)

Vocal similarity important for maintaining social relationships (Toft and Wright 2016)

  - ▶ Hard to know - content of speech may change

# Statistical discrimination versus prejudice

- Prejudice is differential treatment because of animus (Becker 1957)
  - ▶ Preferences
- Statistical discrimination is when people are treated differently because they are *believed* to be less productive (Phelps 1972, Arrow 1973)
  - ▶ Information
- Modern view blurs the line between information and preferences
  - ▶ Information acquisition endogenous to preferences (Brewer 1998)
- Correlations persist only if  $\exists$  animus at the margin (Becker 1957)
  - ▶ Identifying prejudice for immutable characteristics is challenging
    - ★ Need to identify “marginal” candidates (Ayres 2002; Arnold, et al. 2018)
  - ▶ But for mutable characteristics, everyone is “marginal”

# Empirical support

- Becker predicts no prejudice in profit-maximizing equilibrium.
  - ▶ **Audit experiments** and correlations **difficult to interpret** (Bertrand et al. 2004)
  - ▶ Hard to separate prejudice  $\perp$  statistical discrimination (Heckman et al. 1993)
    - ★ We typically **lack data on productivity** (Neumark 2016)
- Innovative experiments disentangle the two channels
  - ▶ Mazes, sport cards, school children (Mobius et al. 2006; List 2004; Rao 2013)
  - ▶ Hard to disentangle in real-world high-stakes setting
- U.S. Supreme Court oral arguments 1999-2013
  - ▶ High-skill setting
    - ★ “[Y]our oral argument ... the most effective weapon you have got.”  
(Chief Justice John Roberts Jr 2005) “Building a cathedral” (Justice Jackson 1951)  
Blackmun graded the attorneys (Johnson et al. 2006; McAtee et al. 2007)
  - ▶ Winning is productivity measure  $\rightarrow$  140B\$ stock prices (Katz et al. 2015)
    - ★ **Sample 1**
    - ★ **Sample 2**

# Do voice impressions matter?

- First impressions of voice are associated with behavior of perceiver
  - ▶ in mate selection, leader elections, housing and consumer choices, and stock market outcomes (in a vocal analysis of earnings conference calls) (Nass and Lee 2001, Klofstad et. al 2012, Purnell et. al 1999, Scherer 1979, Tigue et. al 2012)
  - ▶ Facial cues predict electoral outcomes (Berggren et. al 2010; Todorov et. al 2005)
  - ▶ Nonverbal behavior predict teacher evaluations (Ambady and Rosenthal 1993)
- Even when visual cues are present, potential employers rely more on **voice-based impressions** of a job applicant (Schroeder and Epley 2015)

# Model

- Advocate with voice  $V \in \{M, F\}$  wins the case or not
- Consider the following utility:

$$U_v = \alpha \pi_v + V$$

where  $\pi_v = \text{Prob}(v = \text{win})$  represents the individual's **beliefs** about whether the advocate chosen by him actually wins and  $V$  represents **taste** for advocate  $V$  apart from the economic consequences

- $\alpha$  represents **stakes**
  - ▶ “Billions of \$ at Stake in Supreme Court Power Market Fight” (Bloomberg 1/19/16)
- Individuals will choose advocate  $F$  over  $M$  if and only if

$$\alpha (\pi_F - \pi_M) \geq d$$

where  $d \equiv M - F$  is the relative taste for voice  $M$




# Experimental design

- Suppose individuals are more likely to choose  $M$  over  $F$ 
  - ▶ Deep-voiced men/women perceived competent, persuasive, confident, trustworthy (Apple, et al 1979; Klofstad et al., 2012; Tigue et al., 2011; Jones et al. 2010)
  - ▶ Margaret Thatcher & George H. W. Bush were coached to be less shrill (Kramer 1987); Via humming exercises, Thatcher reduced **half** the male-female difference (Atkinson 1984); her natural voice occasionally **slipped out** (Nallon 2014)
  - ▶ As women entered the work force and into positions of authority, their voices moved closer to a masculine standard (Pemberton 1993); rewarded for doing so (Case 1995)
  - ▶ Not limited to women: Employment discrimination case involving Sears, who asked applicants: "Do you have a low pitched voice?" EEOC vs. Sears, Roebuck & Co., 628 F. Supp. 1264, 1300
  - ▶ Employers prefer  $M$  over  $F$ , even if they do worse on the job (Case 1995)
- **Information** can be used to update one's beliefs about  $\pi_F - \pi_M$ 
  - ▶ Any changes in behavior are due to **mis-beliefs**
- **Incentives** to choose correctly erode the effect of taste on choices ( $\pi_F - \pi_M > \frac{d}{\alpha}$ )
  - ▶ Any changes in behavior are due to **preferences**

# Experimental design

- 748 MTurk raters evaluated 1,901 audio clips of 1,085 lawyers
  - Each rater evaluated 60 random clips, producing 20 ratings per clip

Recording 1 of 66



1. Please provide your impression of the voice recording in the matrix below:

Very Attractive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Unattractive
Very Masculine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not At All Masculine
Not Intelligent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Intelligent
Very Unaggressive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Aggressive
Not Trustworthy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Trustworthy
Very Confident	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Timid

2. Assuming that this is a lawyer arguing a case in front of a panel of Judges, how likely do you think this lawyer will win the case?

Will Definitely Lose ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Will Definitely Win

3. How good is the quality of the recording?

Very Bad ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very Good

Next

Figure: Screenshot of Experiment

# Measurement

- **Within-rater normalization**

$$\widehat{attribute}_{itw} = \frac{attribute_{itw} - \overline{attribute}_w}{\sigma(attribute)_w} \quad (1)$$

- ▶ Subject  $w$ 's perception of a given attribute of advocate  $i$  in case  $t$
- ▶ Adjusts for cross-subject variability in the cardinality and spread of ratings
- ▶ Mean 0, standard deviation 1

- **Raw scores**

- ▶ Weights raters providing more signal amid greater variance in their ratings
- ▶ If some raters occasionally change their numbers to look like they are doing work, normalization puts more weight on those providing less signal in their ratings, which would yield less precise associations with outcomes

- **Average scores**

$$\overline{\widehat{attribute}}_{it} = \frac{1}{|W_{it}|} \sum_w \widehat{attribute}_{itw} \quad (2)$$

- ▶  $W_{it}$  is the set of raters who rated the voice of advocate  $i$  in case  $t$

# Perceived Masculinity Predicts US Supreme Court Outcomes

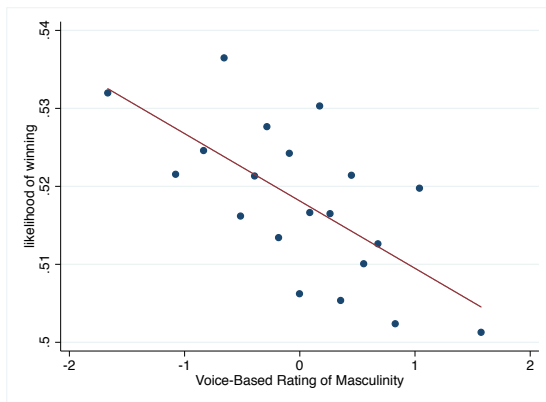


Figure: Voice-Based Masculinity Rating and Court Outcomes

- Does not appear driven by outliers
- No other personality dimension consistently predictive
- ◀ Results from Design and Model Perturbations

# Petitioner Masculinity Predicts US Supreme Court Outcomes

- The first person to argue in front of the Justices should exhibit a stronger vocal first impression effect, under a hypothesis of the primacy of first impressions on court decisions.
  - ▶ Respondent speaks 30 minutes later.

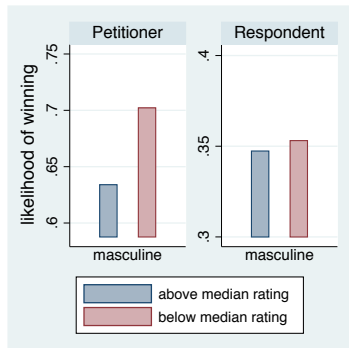


Figure: Perceived Masculinity and Win Rate by Petitioner-Respondent Status.

- Petitioner is main driver
- Below median masculinity rating  $\sim 7$  percentage points

# Robust to Other Voice Characteristics

Table: Baseline Results (raw scores)

Dependent Variable: Case Outcome (win = 1, lose = 0)				
	(1)	(2)	(3)	(4)
Masculine	-0.00707* (0.00413)	-0.00346* (0.00184)	-0.00975** (0.00440)	-0.00674*** (0.00201)
Confident			0.00417 (0.00312)	0.00248 (0.00178)
Attractive			-0.000821 (0.00328)	0.00198 (0.00167)
Intelligent			0.00634** (0.00292)	0.00116 (0.00155)
Trust			-0.00177 (0.00264)	0.000510 (0.00155)
Aggressive			0.00182 (0.00266)	0.000421 (0.00149)
Likely winner			-0.00270 (0.00315)	0.00265 (0.00211)
Lawyer FE	No	Yes	No	Yes
Observations	35330	35330	35329	35329

- Within-lawyer changes & between-lawyer variance; joint  $p < 0.05$

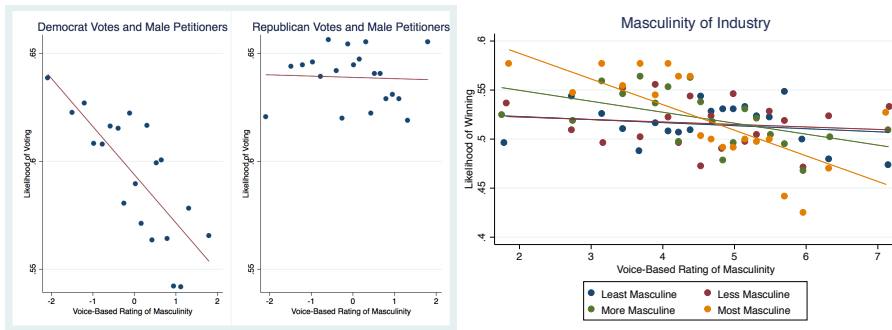
# Attributeable to Between and Within-Variation

**Table:** Advocate Experience at the SCOTUS and Outcomes (raw scores)

Dependent Variable: Case Outcome (win = 1, lose = 0)						
	Petitioner (1)	Respondent (2)	Both (3)	Petitioner (4)	Respondent (5)	Both (6)
	argued only once			argued more than once		
Masculine	-0.00189 (0.00861)	0.00755 (0.00850)	-0.000210 (0.00630)	-0.0190*** (0.00726)	0.00516 (0.00872)	-0.0128** (0.00598)
Confident	-0.000625 (0.00650)	-0.00543 (0.00659)	0.000428 (0.00480)	0.00709 (0.00443)	-0.0111* (0.00569)	0.00599 (0.00392)
Attractive	-0.00295 (0.00728)	-0.00947 (0.00654)	-0.00676 (0.00507)	-0.000407 (0.00531)	0.00260 (0.00653)	0.00167 (0.00442)
Intelligent	0.00297 (0.00604)	0.00822 (0.00646)	0.00805* (0.00454)	-0.00000342 (0.00508)	0.00461 (0.00525)	0.00245 (0.00379)
Trust	0.00315 (0.00573)	-0.00437 (0.00533)	-0.00215 (0.00397)	0.00348 (0.00418)	-0.00474 (0.00574)	-0.00149 (0.00354)
Aggressive	0.0000306 (0.00590)	0.00685 (0.00511)	0.00285 (0.00408)	-0.00134 (0.00409)	0.00445 (0.00507)	0.00235 (0.00348)
Likely winner	-0.00938 (0.00621)	-0.0205*** (0.00650)	-0.0131*** (0.00450)	0.00367 (0.00540)	-0.00228 (0.00622)	0.00440 (0.00444)
Observations	8016	7950	15966	10526	8837	19363

- Half argued only once during the 15 years of our study; half argued multiple times
- Among repeat petitioners, roughly 30% of the association stems from [within-advocate](#) variation, 70% from [between-variation](#)

# Covering



- Votes of Democrats negatively correlated with perceived masculinity
- Stronger negative correlation in more masculine industries
- Consistent with taste differences or misbeliefs in those industries



# De-Biasing Experiment Reduces Misbeliefs

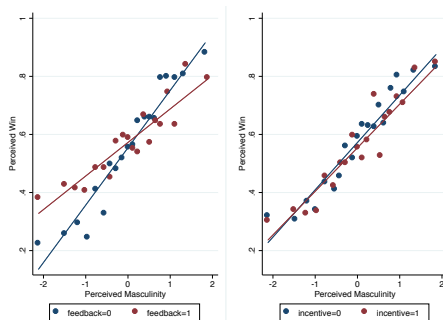


Figure: Feedback ( $p < 0.01$ ), Incentives

- Attitudes precipitate decisions
- Incentives: \$0.5 if get 2/3 or more of outcomes correct.

# Incentives Reveals Taste-Based Discrimination

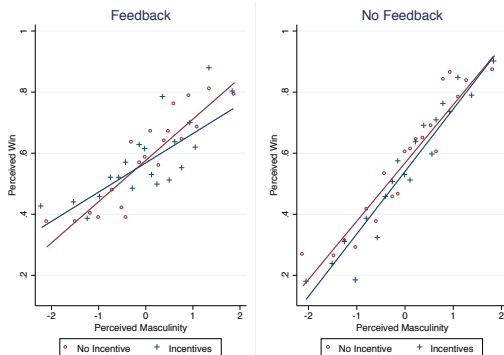


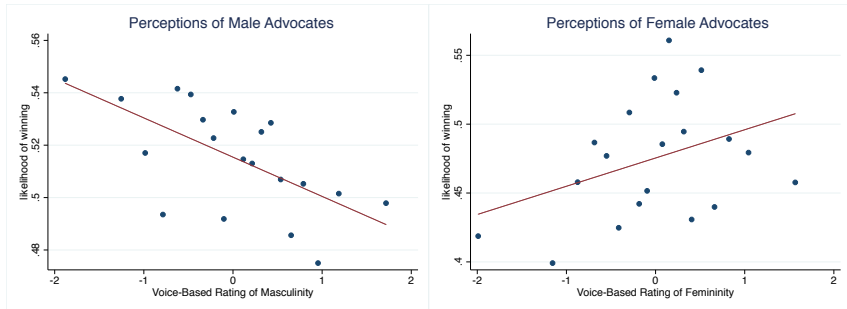
Figure: Incentives ( $p < 0.05$ ) with Feedback

- Incentives to choose correctly erode the effect of taste on choices ( $\pi_F - \pi_M > \frac{d}{\alpha}$ )
- Any changes in behavior are due to preferences ( $d > 0$ )

# Gender

- Female lawyers are also coached to be more masculine (Starecheski 2014)
  - ▶ Deep-voiced men/women perceived competent, persuasive, confident, trustworthy (Apple, et al 1979; Klofstad et al., 2012; Tigue et al., 2011; Jones et al. 2010)
  - ▶ Margaret Thatcher was coached to be less shrill (Kramer 1987)
  - ▶ As women entered the work force and into positions of authority, their voices moved closer to a masculine standard (Pemberton 1993); rewarded for doing so (Case 1995)
- Are our findings restricted to male advocates alone or do they extend?

# Extends to Female Advocates



**Figure:** Voice-Based Perceptions and Court Outcomes by Advocate Gender

- Less masculine males and more feminine females are more likely to win
  - ▶ If masculine = - feminine, pooled results would be stronger

# Magnitudes

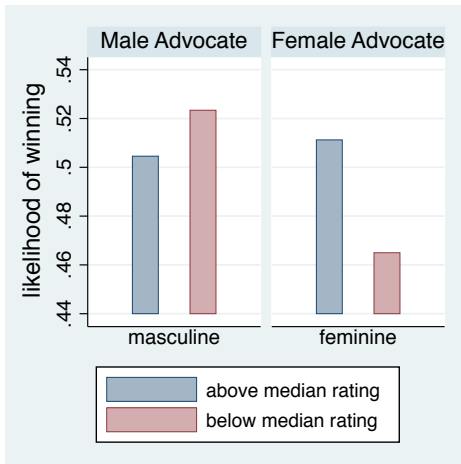


Figure: Perceptions of Masculinity (Femininity) and Advocate Win Rates

- $\leq$  median masculinity rating  $\sim$  2 percentage points greater likelihood of winning for males
- $\leq$  median femininity rating  $\sim$  5 percentage points lower likelihood of winning for females

# Association is Specific to Perceived Femininity

**Figure:** Voice-Based Perceptions and Court Outcomes by Advocate and Participant Gender (raw scores uncollapsed)

Dependent Variable: Case Outcome (win = 1, lose = 0)								
	(1) male advocate		(3) female advocate		(5) male advocate		(7) female advocate	
	petitioner	both	petitioner	both	petitioner	both	petitioner	both
Masculine/ Feminine	-0.0118** (0.00524)	-0.00707* (0.00413)	0.00991 (0.00952)	0.0153** (0.00727)	-0.0149*** (0.00565)	-0.00975** (0.00440)	0.0131 (0.00933)	0.0125* (0.00678)
Confident					0.00508 (0.00387)	0.00417 (0.00312)	-0.00821 (0.00992)	0.00701 (0.00730)
Attractive					0.0000377 (0.00445)	-0.000821 (0.00328)	0.00280 (0.0112)	-0.00587 (0.00743)
Intelligent					0.00244 (0.00385)	0.00634** (0.00292)	-0.00192 (0.0109)	-0.00578 (0.00847)
Trust					0.00356 (0.00344)	-0.00177 (0.00264)	-0.00439 (0.00973)	-0.00175 (0.00647)
Aggressive					-0.00134 (0.00345)	0.00182 (0.00266)	0.00501 (0.0102)	0.00243 (0.00786)
Win					-0.000977 (0.00411)	-0.00270 (0.00315)	0.00251 (0.0108)	0.00361 (0.00785)
Observations	18543	35330	2543	5868	18542	35329	2543	5868

- No other personality dimension consistently predictive

# Reverse voice analysis

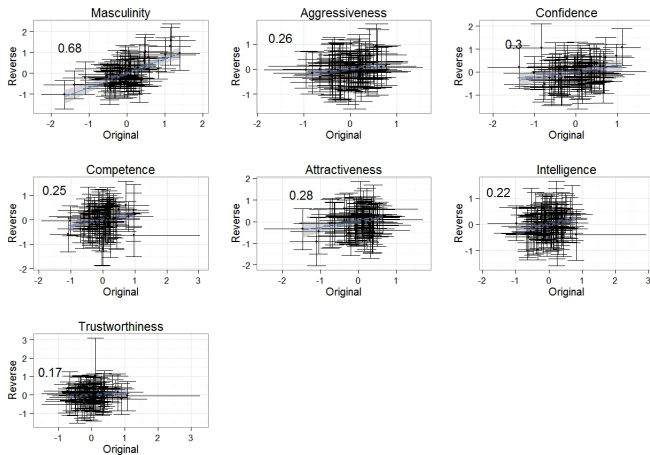


Figure: Correlation in Voice Perceptions across Reversal

# Robust to Lawyer Heterogeneity (and lawyer FE)

Dependent Variable: Case Outcome (win = 1, lose = 0)					
Masculine	-0.0149*** (0.00565)	-0.0145*** (0.00559)	-0.0151*** (0.00558)	-0.0139*** (0.00537)	-0.0864** (0.0340)
Confident	0.00508 (0.00387)	0.00535 (0.00386)	0.00595 (0.00385)	0.00482 (0.00382)	0.0851 (0.0539)
Attractive	0.0000377 (0.00445)	-0.000927 (0.00445)	-0.000399 (0.00441)	0.000460 (0.00431)	-0.00237 (0.0501)
Intelligent	0.00244 (0.00385)	0.00264 (0.00384)	0.00309 (0.00381)	0.00166 (0.00375)	-0.0167 (0.0639)
Trust	0.00356 (0.00344)	0.00336 (0.00343)	0.00330 (0.00345)	0.00305 (0.00338)	0.0644 (0.0618)
Aggressive	-0.00134 (0.00345)	-0.00139 (0.00343)	-0.00145 (0.00343)	-0.00170 (0.00339)	-0.0235 (0.0472)
Likely winner	-0.000977 (0.00411)	-0.00118 (0.00411)	-0.000821 (0.00412)	-0.00152 (0.00405)	-0.0401 (0.0755)
Masculinity of Name	N	Y	Y	N	N
SCOTUS Experience	N	N	Y	N	N
Additional Lawyer Covariates	N	N	N	Y	Y
Collapsed	N	N	N	N	Y
Observations	18542	18542	18542	18542	856
R-squared	0.002	0.006	0.008	0.018	0.026
Sample: Male Petitioners					

Figure: Case Outcomes and Perceived Masculinity



# Linguistic Profiling: Possible reasons for judicial behavior

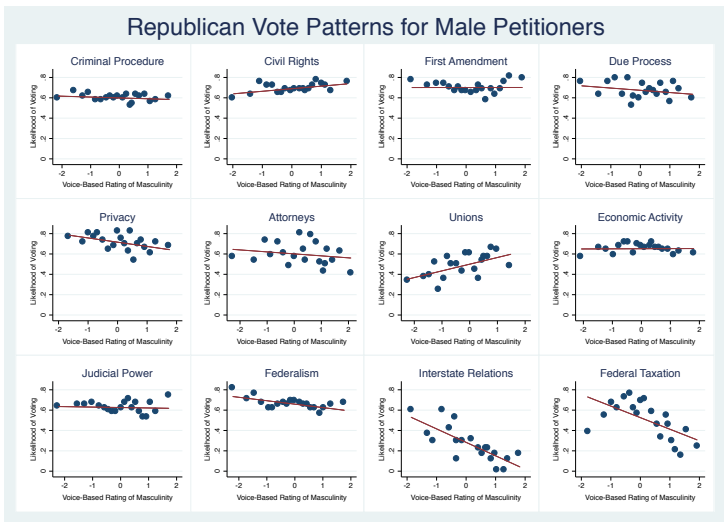
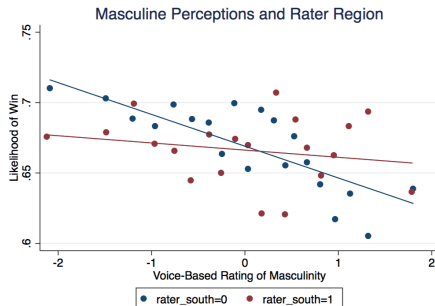
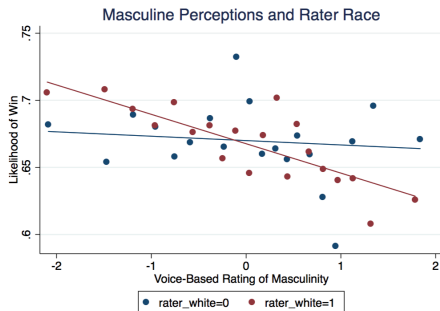


Figure: Republicans vote like Democrats in **less-ideological** cases

- Attorneys, Interstate Relations, Federal Taxation ( $p < 0.1$ )

# Rater Heterogeneity



**Figure:** White ( $p < 0.05$ ) and Non-Southerner ( $p < 0.05$ ) raters' perceptions of masculinity predicted court outcomes

- If White non-Southerners  $\sim$  law firm HR, consistent with firm heterogeneity

# Predicting Behavior of the Supreme Court

Predicted Vote	Judge Votes for Lawyer					
	0.257***		0.258***	0.250***		0.248***
from Random Forest	(0.0486)		(0.0487)	(0.0485)		(0.0489)
Masculine		-0.0223**	-0.0207**		-0.0852**	-0.0780**
		(0.0101)	(0.0101)		(0.0359)	(0.0361)
Cluster			Lawyer and Judge			
Collapsed	No	No	No	Yes	Yes	Yes
Observations	26447	26391	26391	1229	1229	1229
R-squared	0.061	0.002	0.063	0.058	0.008	0.064
Sample: Male Petitioners, Democrat Judges						

Figure: Best Prediction and Perceived Masculinity

- Perceived masculinity basically orthogonal to random forest prediction
- Rater-level: Additional 3% of variance explained
- Lawyer-level: Additional 10% of variance explained
- Random forest also selects perceptions, improves accuracy by 2%
- Katz, Bonmarito, Blackman (Plos-ONE 2017) don't predict close cases well

# Acoustic Data (formant, dispersion, spectral tilt, duration, rate, rhythm, pitch)

Dependent Variable: Case Outcome (win = 1, lose = 0)			
Masculine	-0.0875** (0.0369)	-0.0972** (0.0364)	-0.0858** (0.0348)
Confident		0.0258 (0.0247)	0.0360 (0.0220)
Attractive		-0.0171 (0.0181)	-0.0197 (0.0144)
Educated		0.0158* (0.00878)	0.0146 (0.00932)
Intelligent		0.00549 (0.00893)	0.00635 (0.00783)
Trust		-0.00512 (0.00979)	-0.00528 (0.00786)
Likely winner		-0.00355 (0.00793)	-0.00132 (0.00729)
Acoustic Controls	No	No	Yes
Observations	10920	10080	10080

Figure: Case Outcomes and Perceived Masculinity

- Perceptions matter beyond acoustics
- Results extend with pre-1999 data
  - ▶ Pitch (Dietrich, Enos, Sen, Political Analysis 2018)
  - ▶ ML prediction of masculinity using 15 years of training data (Chen and Kumar 2016)

# Mimicry

- Text-audio alignment for vowel extraction
  - ▶ Eg. AA, AE, UH, etc.
  - ▶ Formants = frequency components: shape/position of tongue
  - ▶ The first two formants typically disambiguate vowels
- ABA triplets
  - ▶ The first segment with speaker A:  $A_1$
  - ▶ The second segment with speaker B's response:  $B$
  - ▶ The third segment with speaker A's response to speaker B:  $A_2$
- Convergence definition:

$$\begin{aligned} & \mathbf{E}[f_j - \bar{f}_j(A_1) | \bar{f}_j(A_1), \bar{f}_j(B)] \\ &= \mathbf{conv} \cdot [\bar{f}_j(B) - \bar{f}_j(A_1)] + \gamma \cdot \bar{f}_j(A_1) \end{aligned}$$

Lawyers converge to judges more than judges do (role of heirarchy)

**Table:** ABA Basic Convergence Parameters

	F1		F2	
	Estimate (S.E.)		Estimate (S.E.)	
	I. Overall (Non Directional)			
Overall	0.175	(0.003)	0.156	(0.003)
	II. Lawyer $\longrightarrow$ Judge			
Overall	0.213	(0.005)	0.187	(0.005)
Winning Lawyer	0.222	(0.006)	0.186	(0.006)
Losing Lawyer	0.205	(0.009)	0.188	(0.006)
	III. Judge $\longrightarrow$ Lawyer			
Overall	0.190	(0.004)	0.151	(0.003)
Winning Lawyer	0.200	(0.006)	0.157	(0.004)
Losing Lawyer	0.181	(0.006)	0.146	(0.004)

Winning lawyers may converge to judges more than losing lawyers do (F1)

Judges converge more when concurring

**Table:** AxBxA Basic Convergence Parameters

	F1		F2	
	Estimate (S.E.)		Estimate (S.E.)	
	I. Overall			
Overall	0.363	(0.007)	0.339	(0.006)
	II. By Decision			
Concurring	0.374	(0.007)	0.359	(0.007)
Not Concurring	0.227	(0.032)	0.159	(0.020)

MIMICRY IS A VERY BASIC HUMAN TENDENCY

# Covering

## Theory

- Mutable characteristics makes everyone “marginal”
  - ▶ for interpreting correlations with productivity

## Data

- Identical introductory sentences in high-stakes setting
  - ▶ to study speech variation beyond word choice

## Method

- De-biasing experiment identifies both misbeliefs and prejudice
  - ▶ Deep-voiced men/women perceived as competent, persuasive
  - ▶ Homophily for masculine-voiced lawyers

## Policy

- Large economic consequences of covering
  - ▶ Stronger effects in industries with more masculine voices

## AI and Rule of Law

- Behavioral economics offer intuitive understanding of feature relevance
  - ▶ Predictive analytics of judges and legal outcomes



## (II) Why Do First Impressions of Lawyers' Voices Predict U.S. Supreme Court Outcomes?

### ● Lawyers

- ▶ Are perceived masculine lawyers different?

- ★ ◀ Is coefficient on perceived masculinity affected by observable lawyer controls?

- ▶ Is voice masculinity a response to case weakness?

- ★ ◀ Is coefficient on perceived masculinity affected by observable pre-trial controls?

### ● Justices

- ▶ Are judges responding differentially to masculine voices?

- ★ ◀ Are masculinity ratings correlated with votes for some judges but not others?

### ● Firms

- ▶ Do firms engage in linguistic profiling in choosing their oral advocates?

- ★ ◀ Are masculinity ratings correlated with court outcomes for some firms but not others?

### ● “Perceived” / performative nature of masculinity

- ▶ Are lawyers performing for another audience (potential client)?

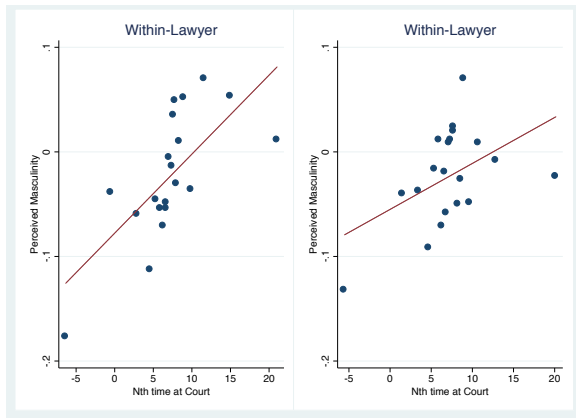
- ★ ◀ Whose ratings of masculinity predict outcomes?

### (III) De-Biasing Experiment

- What are lawyers doing over time?
- Information
- Incentives

# Are lawyers learning?

Figure: Masculine Ratings and Order of Appearance



Within-lawyer graph is residualized on lawyer fixed effects, so negative means early years of experienced advocates.

- No, lawyers' masculinity ratings increase over time within-lawyer & controlling for years since graduation.
- ◀ No significant response to previous wins/losses

## Are lawyers learning?

	Outcome	Masculinity Rating	$\Delta$ Masculinity Rating
Masculinity Rating	-0.0605** (0.0308)		
Previous Loss		0.0445 (0.0551)	-0.0463 (0.0359)
Collapsed	Y	Y	Y
Observations	862	389	389
R-squared	0.004	0.002	0.004
Sample: Male Petitioners			

- We can reject the hypothesis that losses are significantly associated with perceived masculinity at the next appearance
- Whatever drives increase in masculine rating (may be “covering”), they seem not to be taking into account actual win rates

## Diff-in-diff framework

- Attitudes precipitate decisions.
- Let  $y_{ij}$  be attitudes in treatment  $j = 1, 2, 3, 4$  and recording  $i$ ,  $F_j$  be an indicator variable that takes on a value of 1 if raters receive information about whether the attorney on the recording won the case or not and 0 otherwise, and  $I_j$  be an indicator for whether the raters were given an incentive to choose correctly ( $I_j = 1$ ) or not ( $I_j = 0$ ).
- The estimation equation takes the following form:

$$y_{ij} = \alpha + \beta V_i + \gamma' T_J + \delta' T_J \times V_i + e_{ij},$$

where  $V_i \in \{H, L\}$  is advocate  $i$ 's voice type and  $T_J \in \{[F_j], [I_j]\}$  is a treatment vector (2x2 treatment).

- ▶ Incentives: \$0.5 if get 2/3 or more of outcomes correct.

# Masculine voices and perceived outcomes

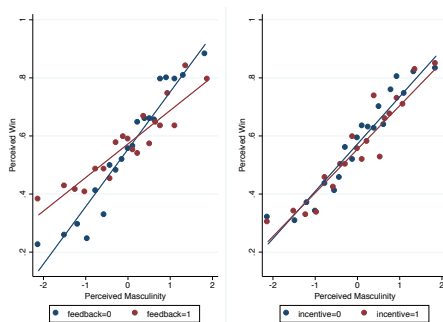


Figure: Feedback ( $p < 0.01$ ), Incentives

# Masculine voices and perceived outcomes

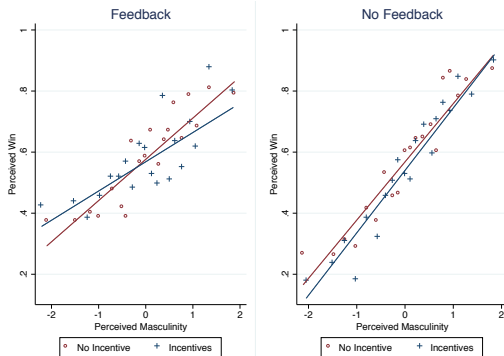


Figure: Incentives ( $p < 0.05$ ) with Feedback

- Recall that incentives to choose correctly erode the effect of taste on choices ( $\pi_F - \pi_M > \frac{d}{\alpha}$ )
- Providing incentives should bring  $d \rightarrow 0$ 
  - The effect ( $p < 0.05$ , 30% of association) also means  $d > 0$

## (IV) Explanation of Variance

- 1 ◀ How does predictive power compare with previous carefully specified multivariate models?
- 2 ◀ How much do perceptions explain beyond objective acoustic data?



# Model

## ① Simple framework to distinguish prejudice vs. statistical discrimination

$$U_v = \alpha\pi_v + V$$

- ▶ Previous literature gives reasons why employers may prefer masculine voices
  - ★ Masculine voices do worse on this job

## ② Information and incentives can reduce selection bias

- ▶ Individuals perceive masculine voices to be winners
  - ★  $\frac{\partial(\pi_F - \pi_M)}{\partial \text{Info}} > 0 \Rightarrow$  statistical discrimination
  - ★  $\frac{\partial^2(\pi_F - \pi_M)}{\partial \text{Info} \partial \text{Incentives}} > 0 \Rightarrow$  prejudice

## ③ Advocates may cover even in high-skill industry

- ▶ Masculine voices do worse in industries with more masculine voices
  - ★ Consistent with differences in  $d > 0$  sacrificing  $\Delta\pi_v$
- ▶ Advocates' voice become more masculine with experience

## ④ Voice is one of many mutable dimensions

# Discussion

## ① If omitted pre-trial characteristics affect voice masculinity,

- ▶ why does perceived masculinity matter beyond best predictive models
- ▶ why do judges respond differently to case weakness
  - ★ why do Republican judges respond to female lawyers' case weakness and Democrats to male lawyers' case weakness
- ▶ why don't lawyers change their voices (since it's mutable)

## ② No direct evidence of employment contingent on covering, but

- ▶ why do masculine lawyers do worse in masculine industries
- ▶ why do lawyers become more masculine with experience
  - ★ ◀ while experience is cross-sectionally associated with less masculine voices

## ③ No data on salaries, but

- ▶ We would need to observe salary differentials of  $\Delta\pi_v = 7\%$  of 1B\$
  - ★ "Billions of \$ at Stake in Supreme Court Power Market Fight" (Bloomberg 1/19/16)

# Results

## ① Do Snap Judgments of Voices Predict SCOTUS Outcomes?

- ▶ Yes, and it is specific to perceived masculinity
  - ★ Petitioner (the first lawyer to argue) is the main driver
  - ★  $\exists$  within- and between- lawyer variation

## ② Why Do Voices Predict SCOTUS Outcomes?

- ▶ Democrats vote for less-masculine ♂; Republicans for feminine ♀
- ▶ Stronger effect in private firms, industries with more masculine voices
  - ★ attribute 75% to linguistic profiling by judges and 25% by firms

## ③ De-biasing Experiment

- ▶ Information reduces 40% of  $\text{Cov}(\text{perceived masculinity}, \text{perceived win})$ 
  - ★ Incentives reduces another 20% when there is information

## ④ Explanation of Variance

- ▶ Explains up to 10% additional variance beyond best prediction models

# Mutable Characteristics

## ● Employment consequences of mutable characteristics

- ▶ Affirmative action targets immutable characteristics
  - ★ women past glass ceiling or masculine women past glass ceiling?
  - ★ “Once that barrier is broken, there may be a little less pressure to keep appointing people from that previous excluded category” (Zweigenhaft et al. 1998)
- ▶ EU refugee policy and laïcité focuses on (assimilation) mutable characteristics

## ● Welfare consequences

- ▶ General equilibrium considerations
- ▶ Need random assignment of laws

## ● Positive not normative (Yoshino 2000)

- ▶ Mutable characteristics gradually entering economic models (Fryer et al. 2010)
- ▶ One sentence (but no data) in two summaries of the empirical literature (Bertrand and Duflo 2016; Neumark 2016)
  - ★ “Recently, the emphasis on ‘fit’ [of] a prospective employee .. as a hiring criterion has raised the spectrum of a new form of subtle discrimination”

## ● Puzzle

- ▶ Economic perspective: Correlations should not persist
  - ★ Complete markets would eliminate misbeliefs, so correlations persist only if  $d > 0$  for the marginal employer (hiring  $M$  or  $F$ ) (Charles and Guryan 2008)

# What's Next?

- Data

- ▶ 1955- audio clips
- ▶ 1955- lawyer covariates & faces & campaign donations & clerked-for
- ▶ 1999- acoustics
- ▶ split "Mr. Chief Justice" from "May it please the Court"
- ▶ "I like to reserve the balance of my time"
- ▶ Identify the cases that moved stock markets (Katz et al 2015)

- Causal mechanisms: manipulate

- ▶ presence of text to separate predictive power of audio
- ▶ presence of face
- ▶ voice itself
- ▶ presence of both clips per case

- Ideological mechanisms

- ▶ rate audio of justices & (male and female) politicians
- ▶ perceived ideology of lawyer

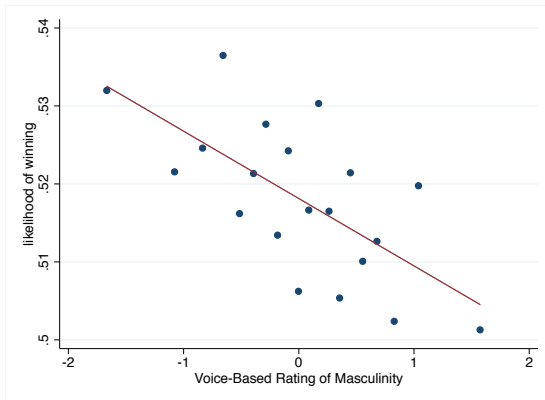
- De-biasing experiment

- ▶ willingness-to-bet on a voice winning

- Separate project on convergence during oral argument

- ▶ Judges converge to each other, convergence is correlated with voting together
- ▶ Advocates and judges converge, but lawyers converging to justices (F1) lose their vote

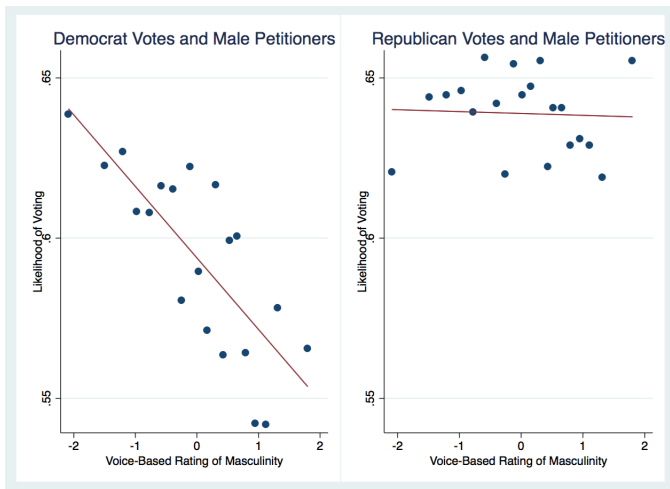
# Results



**Figure:** Males are more likely to win when they are perceived as less masculine

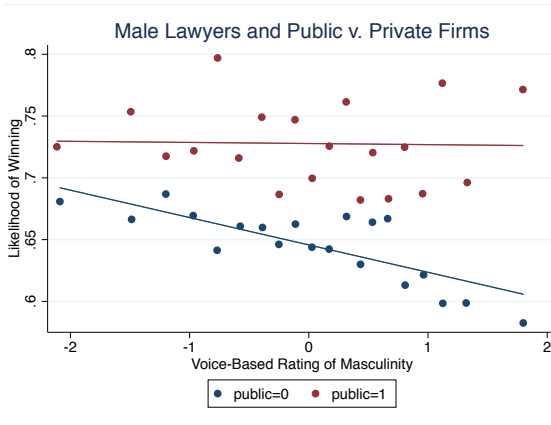
- even controlling for characteristics of: lawyer, case, acoustics, rater and other perceptions

# Judge Heterogeneity



**Figure:** Votes of Democrats, but not Republicans, are negatively associated with perceived masculinity ( $p < 0.01$ )

# Firm Heterogeneity



**Figure:** No effect among public lawyers who argue on behalf of the government

“firm/industry” is the category of the petitioner - the party who petitioned the Supreme Court to review the case (and speaks first) - or the respondent



# Information Reduces Bias

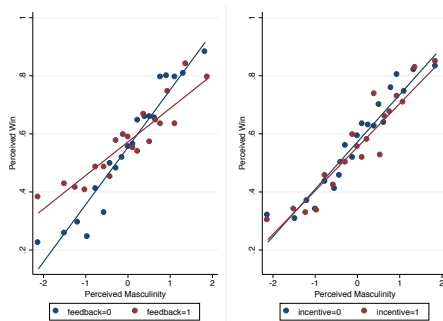





Figure: Feedback ( $p < 0.01$ ), Incentives

40% of the correlation between perceived masculinity and perceived win

# Questionnaire

Recording 1 of 66

  0:00 

1. Please provide your impression of the voice recording in the matrix below:

Very Attractive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Unattractive
Very Masculine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Not At All Masculine
Not Intelligent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Intelligent
Very Unaggressive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Aggressive
Not Trustworthy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Trustworthy
Very Confident	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very Timid

2. Assuming that this is a lawyer arguing a case in front of a panel of judges, how likely do you think this lawyer will win the case?

Will Definitely Lose ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Will Definitely Win

3. How good is the quality of the recording?

Very Bad ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very Good

Next

Figure: Screenshot of Experiment

# Design Perturbations

- 1 To control for the possibility of within-voice modeling by raters; rely on their first impression and to not over think (Ballew and Todorov 2007)
  - ▶ One randomly selected question only for each voice sample
  - ▶ Each voice clip was played aloud only once
- 2 Added perceived age and perceived educatedness, and varied the way we asked raters whether the lawyer *should win* the case versus how likely do they think the lawyer *will win* the case.

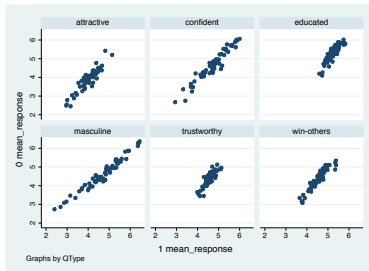


Figure: Correlation in Average Voice Perceptions across Experiments

# Design Perturbations

- Asked more ratings per voice

Table: Robustness to Experimental Variation (collapsed)

Dependent Variable: Case Outcome (win = 1. lose = 0)						
	(1)	(2)	(3)	(4)	(5)	(6)
masculine	-0.224** (0.0846)	-0.324** (0.131)	-0.219* (0.124)	-0.317 (0.197)	-0.197 (0.140)	-0.434** (0.200)
attractive		-0.0654 (0.257)		-0.0616 (0.242)		0.111 (0.221)
confident		0.404 (0.239)		0.346 (0.342)		0.167 (0.253)
trustworthy		-0.436 (0.574)		-0.347 (0.488)		0.285 (0.288)
Ratings per voice	60	60	60	60	20	20
Cross-lawyer ratings	Yes	Yes	No	No	Yes	Yes
Observations	60	60	60	60	58	58

# More Attributes

## What is Perceived Masculinity?

- Performative nature of “masculinity” → existence of non-masculine men and masculine women (Kiesling 2007; Eckert and McGonnell-Ginet 2003)
  - ▶ (1) Gender difference; (2) Heterosexism; **(3) Dominance**; (4) Male solidarity
  - ▶ Evolutionary basis for gendered differentiation (Locke 2011)
    - ★ Males selected to be aggressive and dominant, but aggressive and dominant behaviors would lead to lethal confrontation
    - ★ Invite aversive response from SCOTUS (even controlling perceived confidence)

# More Attributes

Linear regression

Number of obs = 60  
F(7, 29) = 2.83  
Prob > F = 0.0226  
R-squared = 0.1788  
Root MSE = .4867

(Std. Err. adjusted for 30 clusters in casenum)

outcome	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
zratemascul	-.5549709	.1794386	-3.09	0.004	-.9219641	-.1879777
zrateconfi	.3785674	.3000131	1.26	0.217	-.2350282	.9921631
zrateattitude	-.0203677	.2829106	-0.07	0.943	-.5989848	.5582495
zratecompetent	.5881734	.4047179	1.45	0.157	-.2395676	1.415914
zrateempathy	.2124795	.2859498	0.74	0.463	-.3723536	.7973126
zraterespect	-.3886403	.3201185	-1.21	0.235	-1.043356	.2660756
zratewin	-.3668454	.2585223	-1.42	0.167	-.8955829	.1618921
_cons	.5	.0266405	18.77	0.000	.4455142	.5544859

Figure: Collapsed

- R-square is 17%

# More Attributes

Linear regression

Number of obs = 1,260  
F(7, 29) = 1.56  
Prob > F = 0.1862  
R-squared = 0.0207  
Root MSE = .49639

(Std. Err. adjusted for 30 clusters in casenum)

outcome	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
zratemascul	-.0800075	.0312074	-2.56	0.016	-.1438338	-.0161811
zrateconfi	.019719	.026995	0.73	0.471	-.035492	.07493
zrateattitude	.0049058	.0188143	0.26	0.796	-.0335737	.0433853
zratecompetent	.0173319	.0190688	0.91	0.371	-.0216681	.0563319
zrateempathy	.0060675	.0135979	0.45	0.659	-.0217433	.0338783
zraterespect	-.0030413	.0155372	-0.20	0.846	-.0348184	.0287358
zratewin	-.0078778	.0212576	-0.37	0.714	-.0513545	.0355988
_cons	.5	.0033015	151.45	0.000	.4932477	.5067523

Figure: Uncollapsed

# Model Perturbations

Court decision outcome was modeled in terms of mixed effects logistic regression ◀

- Fixed effects:

- ▶ The role of the attorney: Petitioner vs. Respondent
- ▶ By-subject normalized ratings of
  - ★ Masculinity
  - ★ Confidence
  - ★ Trustworthiness
  - ★ Attractiveness
  - ★ Educatedness
  - ★ Intelligence
  - ★ Likelihood of winning

- Random effects

- ▶ Case



## Model Perturbations

	Coef ( SE )	z-value
Intercept	0.002 ( 0.031 )	0.07
Attorney Role	0.702 ( 0.022 )	32.45 ***
Masculinity	-0.434 ( 0.028 )	-15.51 ***
Confidence	0.087 ( 0.030 )	2.91 **
Attractiveness	-0.049 ( 0.027 )	-1.81 .
Educatedness	0.050 ( 0.034 )	1.49
Trustworthiness	-0.035 ( 0.027 )	-1.29
Intelligence	0.028 ( 0.035 )	0.79
Winnability	-0.042 ( 0.031 )	-1.38

# Results

Table: Baseline Results (raw scores collapsed)

Dependent Variable: Case Outcome (win = 1, lose = 0)				
	(1)	(2)	(3)	(4)
Masculine	-0.0279* (0.0169)	-0.0767 (0.0646)	-0.0369 (0.0239)	-0.166** (0.0778)
Confident			0.0502 (0.0431)	0.0358 (0.0970)
Attractive			-0.0167 (0.0349)	0.0562 (0.0933)
Intelligent			0.0907* (0.0464)	0.0398 (0.105)
Trust			-0.0185 (0.0461)	0.0305 (0.0948)
Aggressive			0.0160 (0.0378)	-0.00136 (0.0836)
Likely winner			-0.0796 (0.0568)	0.0551 (0.119)
Lawyer FE	No	Yes	No	Yes
Observations	1634	1634	1634	1634

- Magnitudes are larger at an analysis at the level of advocates
- One s.d. of masculine measure  $\sim$  5.2 percentage points in col (1) and 31 in col (4)

# Correlation Matrix

Table: Correlation in Trait Judgements (uncollapsed)

variable	masculine	confident	attractive	intelligent	trustworthy	aggressive
masculine	1.0000					
confident	0.4418	1.0000				
attractive	0.3381	0.3604	1.0000			
intelligent	0.2332	0.4006	0.3477	1.0000		
trustworthy	0.2002	0.2658	0.3546	0.3678	1.0000	
aggressive	0.3448	0.4974	0.2296	0.2349	0.1017	1.0000
likely winner	0.4133	0.5593	0.4387	0.4770	0.3968	0.3917

- ◀ Principal components show that voice judgements cluster in one factor explaining 66 percent of variation
- Regression tree yields Masculinity and Confidence in the final pruned tree
- ◀ Masculinity and Confidence in reverse voice analysis

◀ Substantial Variation

# Principal components

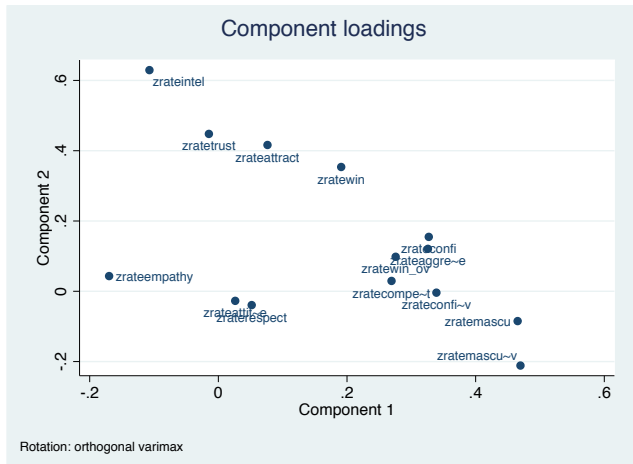


Figure: First two components

# Reverse voice analysis

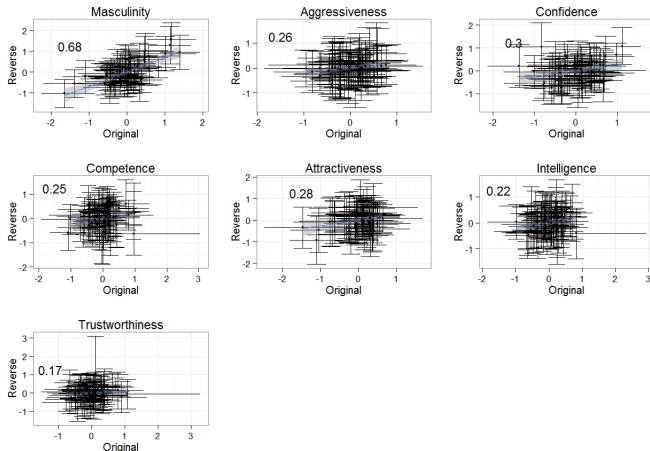


Figure: Correlation in Voice Perceptions across Reversal

- Forward version
- Reversed version
- ◀ Correlation matrix

# Results

**Table:** Rater Characteristics and Voice Perceptions (raw scores)

Dependent Variable: Case Outcome (win = 1, lose = 0)						
	Petitioner (1)	Respondent (2)	Both (3)	Petitioner (4)	Respondent (5)	Both (6)
Masculine	-0.0144** (0.00571)	0.00491 (0.00630)	-0.00963** (0.00444)	-0.00569** (0.00241)	0.000420 (0.00232)	-0.00668*** (0.00203)
Confident	0.00457 (0.00390)	-0.00924** (0.00438)	0.00379 (0.00314)	0.00314 (0.00194)	-0.00520** (0.00219)	0.00233 (0.00179)
Attractive	-0.000428 (0.00447)	-0.00271 (0.00479)	-0.00137 (0.00332)	0.000228 (0.00185)	0.00148 (0.00235)	0.00152 (0.00169)
Intelligent	0.00323 (0.00390)	0.00652 (0.00425)	0.00662** (0.00298)	0.00187 (0.00177)	0.00108 (0.00204)	0.00201 (0.00157)
Trust	0.00279 (0.00351)	-0.00475 (0.00399)	-0.00221 (0.00267)	0.00105 (0.00189)	-0.000211 (0.00213)	0.000673 (0.00156)
Aggressive	-0.00156 (0.00347)	0.00522 (0.00372)	0.00182 (0.00272)	0.00174 (0.00179)	-0.000667 (0.00199)	0.000554 (0.00151)
Likely winner	-0.000989 (0.00415)	-0.0101** (0.00452)	-0.00258 (0.00320)	0.00172 (0.00228)	-0.00112 (0.00261)	0.00235 (0.00216)
MTurker controls	Yes	Yes	Yes	Yes	Yes	Yes
Lawyer FE	No	No	No	Yes	Yes	Yes
Observations	18542	16787	35329	18542	16787	35329

- More precision when controlling for MTurk rater characteristics

Dependent Variable: Case Outcome (win = 1, lose = 0)

	(1)	(2)	(3)	(4)
masculine	-0.197 (0.140)	-0.434** (0.200)	-0.719*** (0.206)	-1.203* (0.618)
attractive		0.111 (0.221)	0.281 (0.294)	0.384 (0.850)
confident		0.167 (0.253)	0.277 (0.202)	0.545 (0.554)
trustworthy		0.285 (0.288)	0.279 (0.317)	0.559 (1.419)
SCOTUS clerk			-0.0257 (0.227)	0.134 (0.817)
Lawyer's age			0.0121* (0.00646)	0.0203 (0.0261)
Number of clerkships			-0.114 (0.116)	-0.182 (0.465)
Law review			0.415*** (0.131)	0.815** (0.363)
Other graduate degree			0.115 (0.158)	-0.00364 (0.499)
Top-5 law school			0.0687 (0.153)	0.0956 (0.397)
Case fixed effects	No	No	No	Yes
Observations	58	58	58	58

# Lawyer Heterogeneity

Dependent Variable: Case Outcome (win = 1, lose = 0)					
Masculine	-0.0149*** (0.00565)	-0.0145*** (0.00559)	-0.0151*** (0.00558)	-0.0139*** (0.00537)	-0.0864** (0.0340)
Confident	0.00508 (0.00387)	0.00535 (0.00386)	0.00595 (0.00385)	0.00482 (0.00382)	0.0851 (0.0539)
Attractive	0.0000377 (0.00445)	-0.000927 (0.00445)	-0.000399 (0.00441)	0.000460 (0.00431)	-0.00237 (0.0501)
Intelligent	0.00244 (0.00385)	0.00264 (0.00384)	0.00309 (0.00381)	0.00166 (0.00375)	-0.0167 (0.0639)
Trust	0.00356 (0.00344)	0.00336 (0.00343)	0.00330 (0.00345)	0.00305 (0.00338)	0.0644 (0.0618)
Aggressive	-0.00134 (0.00345)	-0.00139 (0.00343)	-0.00145 (0.00343)	-0.00170 (0.00339)	-0.0235 (0.0472)
Likely winner	-0.000977 (0.00411)	-0.00118 (0.00411)	-0.000821 (0.00412)	-0.00152 (0.00405)	-0.0401 (0.0755)
Masculinity of Name	N	Y	Y	N	N
SCOTUS Experience	N	N	Y	N	N
Additional Lawyer Covariates	N	N	N	Y	Y
Collapsed	N	N	N	N	Y
Observations	18542	18542	18542	18542	856
R-squared	0.002	0.006	0.008	0.018	0.026
Sample: Male Petitioners					

**Figure:** Case Outcomes and Perceived Masculinity

- Masculinity of first name (male-female ratio in census), # of previous SCOTUS oral arguments
- Years since graduation, # of admitted courts, # practice areas, firm size, office size, partner, top 5 law school
- No robust association between voice pitch and body size (Fitch 1997; Künzel, 1989)



# Lawyer Heterogeneity

Dependent Variable: Case Outcome (win = 1, lose = 0)					
Feminine	0.0125*	0.0135**	0.0104	0.00890	0.151
	(0.00678)	(0.00668)	(0.00647)	(0.00647)	(0.117)
Confident	0.00701	0.00910	0.00690	0.00402	0.0239
	(0.00730)	(0.00724)	(0.00714)	(0.00689)	(0.104)
Attractive	-0.00587	-0.00645	-0.00508	-0.00412	-0.0537
	(0.00743)	(0.00728)	(0.00715)	(0.00673)	(0.103)
Intelligent	-0.00578	-0.00716	-0.00408	-0.00122	0.0162
	(0.00847)	(0.00852)	(0.00838)	(0.00797)	(0.0756)
Trust	-0.00175	-0.00266	-0.00140	-0.000920	-0.0536
	(0.00647)	(0.00641)	(0.00629)	(0.00607)	(0.124)
Aggressive	0.00243	0.000763	-0.00182	-0.00112	-0.0524
	(0.00786)	(0.00774)	(0.00771)	(0.00758)	(0.0955)
Likely winner	0.00361	0.00294	0.000882	0.00296	0.00768
	(0.00785)	(0.00781)	(0.00758)	(0.00746)	(0.129)
Masculinity of Name	N	Y	Y	N	N
SCOTUS Experience	N	N	Y	N	N
Additional Lawyer Covariates	N	N	N	Y	Y
Collapsed	N	N	N	N	Y
Observations	5868	5868	5868	5868	267
R-squared	0.004	0.023	0.039	0.079	0.088
Sample: Female Lawyers					

Figure: Case Outcomes and Perceived Femininity

- Masculinity of first name (male-female ratio in census), # of previous SCOTUS oral arguments
- Years since graduation, # of admitted courts, # practice areas, firm size, office size, partner, top 5 law school
- Feminine voices perceived as more attractive (Puts et al. 2011; Feinberg et al. 2008)
- Why does perceived masculinity matter?

# Judge Heterogeneity

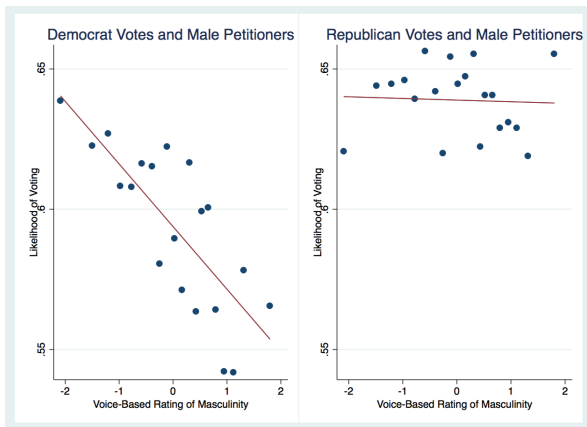
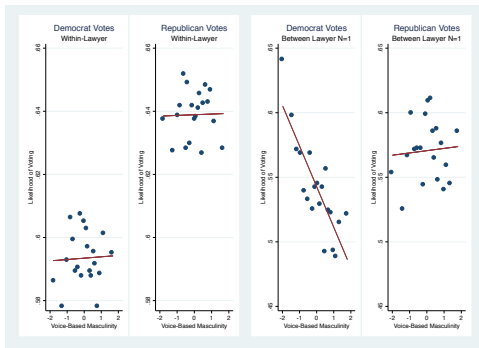


Figure: Political Party and Response to Masculinity ( $p < 0.01$ )

- Republicans favor males (Szmer et al. 2010)
- Democrats disfavor masculine males
- ◀ Not due to judge gender

# Judge Heterogeneity



**Figure:** Political Party and Response to Male Petitioner Masculinity (Within & Between Lawyers)

- Something other than judge party would need to explain within-lawyer effects
- Heterogeneity of between-lawyer effect is quite strong
  - ▶ Between-lawyer variance for non-repeat advocates may have been masked

# Judge Heterogeneity

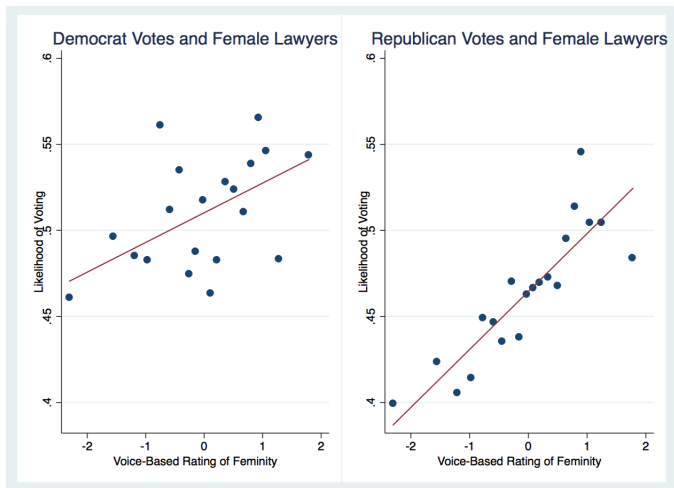


Figure: Political Party and Response to Femininity ( $p < 0.1$ )

- Republicans favor feminine females
- ◀ Also not due to judge gender

# Linguistic Profiling: Possible reasons for judicial behavior

## Strategic voting

- Judges could signal their attitudes toward gender role issues
  - ▶ Democrats want to signal they are progressive
    - ★ Support gender neutral sounding men.
  - ▶ Republicans want to signal their support for traditional gender roles.
    - ★ Females should sound female, men and women have a distinct role.

## Sincere voting

- Judges value information from advocates who are less masculine since they must be excellent to jump over market hurdles to get this far.
  - ▶ For both men and women this means not sounding masculine.

## Homophily

- Perceived masculinity of voice may signal ideology
  - ▶ Advocacy groups put forth attorneys who share their ideology
  - ▶ Republican women in Congress look feminine (Carpinella and Johnson 2012)
  - ▶ Rep Senators look powerful (Dem look warm) (Rule and Ambady 2010)

# Linguistic Profiling: Possible reasons for judicial behavior

## Strategic voting

- Judges could signal their attitudes toward gender role issues
  - ▶  $H_0$  : no variation across cases

## Sincere voting

- Judges value information from advocates who are less masculine
  - ▶  $H_1$  : particularly in difficult or technical cases

## Homophily

- Perceived masculinity of voice may signal ideology
  - ▶  $H_2$  : particularly in ideological cases

# Linguistic Profiling: Possible reasons for judicial behavior

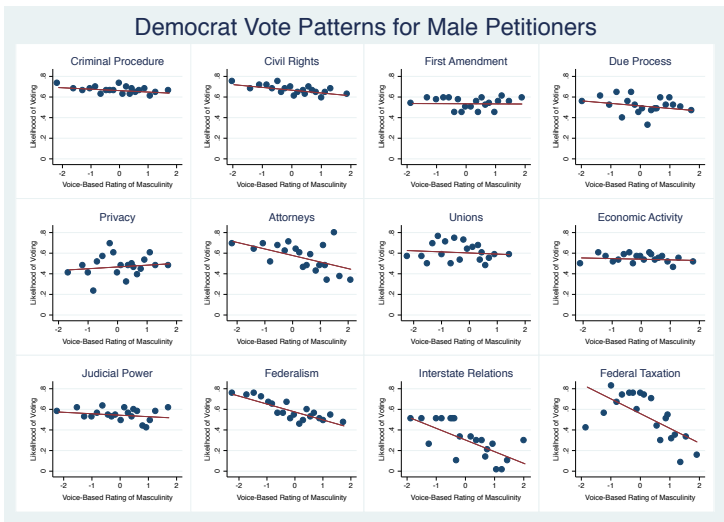


Figure: Substantial variation by case, stronger in less-ideological cases

- Attorneys, Interstate Relations, Federal Taxation ( $p < 0.05$ ) (also Federalism)

# Linguistic Profiling: Possible reasons for judicial behavior

Legal area ranked by polarization (% 5-4 decisions)

- Unions: 34%
- Criminal Procedure: 31%
- **Federalism**: 28%
- Civil Rights: 27%
- First Amendment: 26%
- Judicial Power: 22%
- **Interstate Relations**: 22%
- Due Process: 21%
- Privacy: 21%
- **Attorneys**: 21%
- Economic Activity: 12%
- **Federal Taxation**: 10%



# Linguistic Profiling: Possible reasons for judicial behavior

- Both Democrats and Republicans vote for less-masculine lawyers in Federal Taxation, Interstate Relations, Federalism, and Attorneys cases (which are less ideological & maybe more technical) => Sincere voting
  - ▶ But differ in ideological cases like Unions and Civil Rights => Homophilly
  - ▶ ◀ Similar pattern for female lawyers

# Linguistic Profiling: Possible reasons for judicial behavior

$$\begin{aligned} \text{judge votes for advocate}_{itw} = & \alpha + \widehat{\text{masc}}'_{itw}\beta_0 + \widehat{\text{masc}}'_{itw} * \text{Dem}_w\beta_1 \\ & + \widehat{\text{masc}}'_{itw} * \text{NonIdeo}_t\beta_2 + \mathbf{x}'_{itw}\gamma + \varepsilon_{itw} \end{aligned} \quad (3)$$

Suggests the overall effect can be decomposed into

- Ideology - 75% ( $H_1$ )
- Information - 25% ( $H_2$ )

Regardless of the source, lawyers should adjust - unless another audience matters

◀ Why does perceived masculinity matter?

# Linguistic Profiling: Possible reasons for judicial behavior

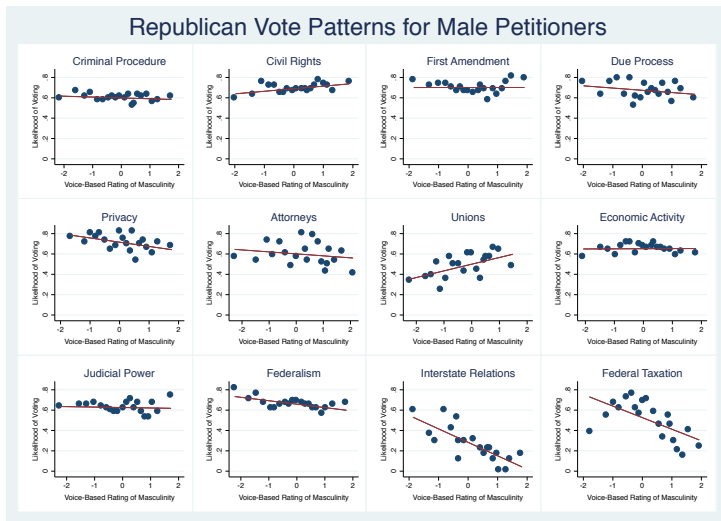


Figure: Even Republicans vote like Democrats in less-ideological cases

# Linguistic Profiling: Possible reasons for judicial behavior

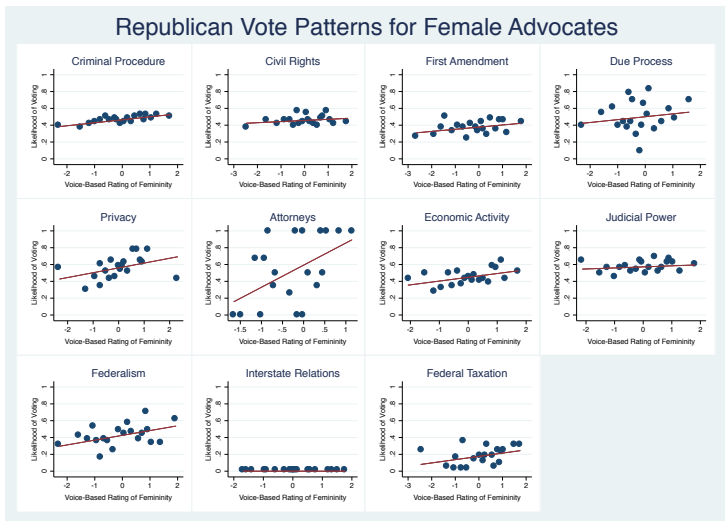


Figure: Substantial variation by case, stronger in less-ideological cases

● Attorneys ( $p < 0.01$ ) (also Federalism, Federal Taxation, Due Process and Privacy)

# Linguistic Profiling: Possible reasons for judicial behavior

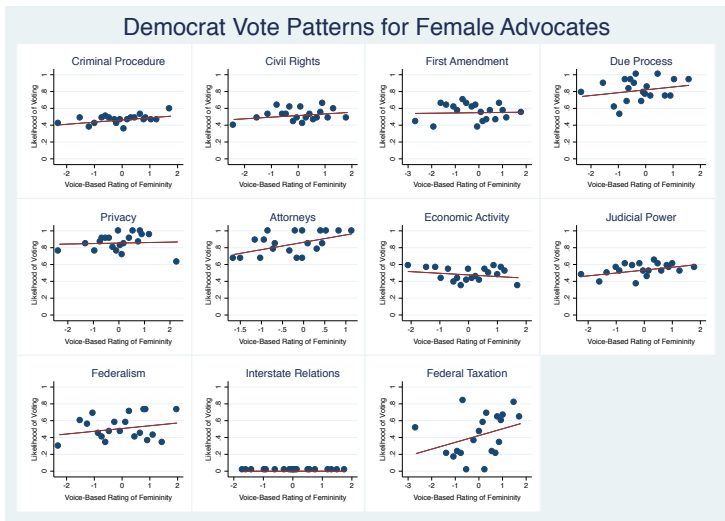


Figure: Even Democrats vote like Republicans in less-ideological cases ( $p < 0.05$ )

● Attorneys ( $p < 0.05$ ) (also Federal Taxation) ◀ Sincere or homophily?

# Judge Heterogeneity

	(1)
	Judge Votes for Lawyer
Masculinity	-0.000460 (0.00806)
Female Judge	-0.00554 (0.0194)
Masculinity * Female Judge	-0.00236 (0.00266)
Democrat	-0.0426* (0.0246)
Masculinity * Democrat Judge	-0.0203*** (0.00756)
Other Ratings	Y
Cluster	Lawyer and Judge
N	150304
R-sq	0.003

**Figure:** Political Party and Response to Male Petitioner Masculinity ( $p < 0.01$ )

- Nor due to experience or age of judge

# Judge Heterogeneity

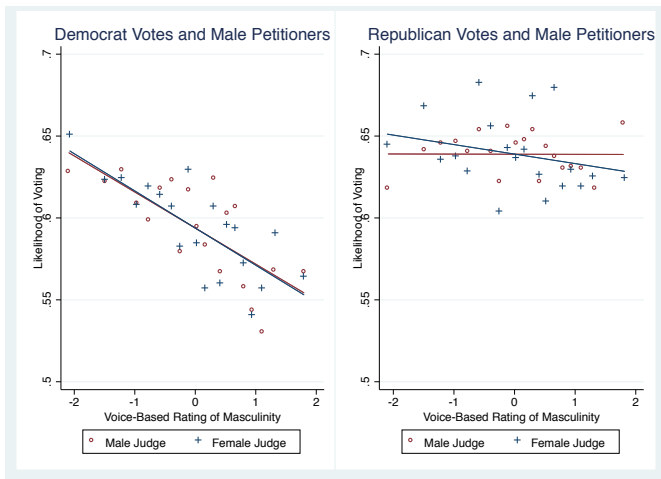


Figure: Political Party and Response to Male Petitioner Masculinity ( $p < 0.01$ )

- Fully interacting gender and party shows that there is a slight gender effect.

# Judge Heterogeneity

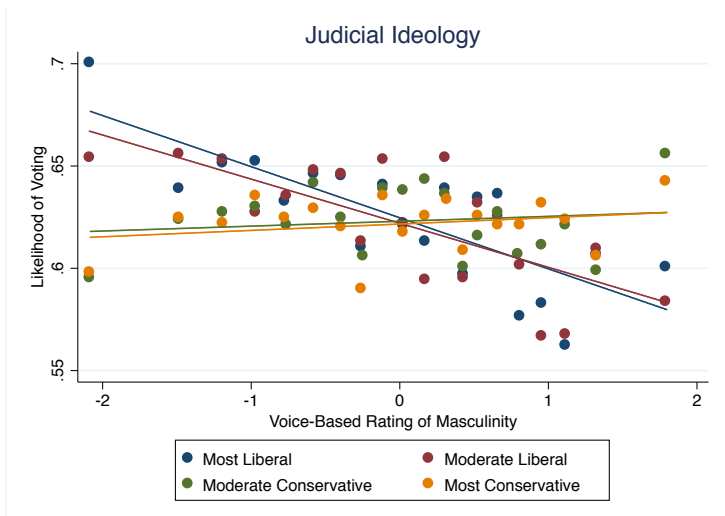


Figure: Judicial Ideology and Response to Male Petitioner Masculinity ( $p < 0.01$ )



# Judge Heterogeneity

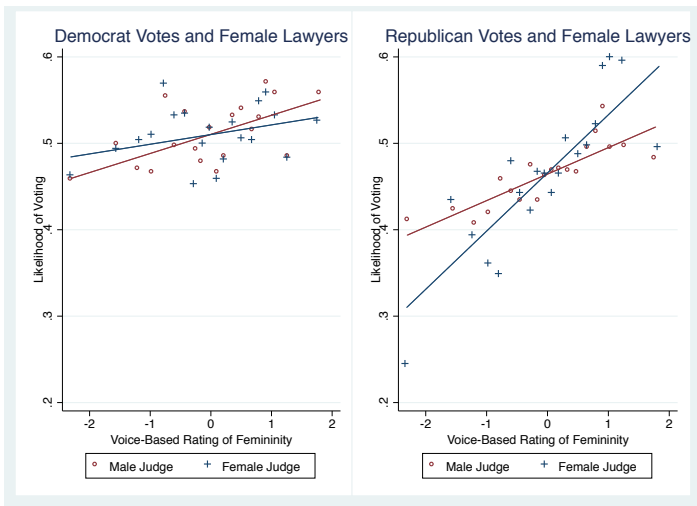


Figure: Political Party and Response to Femininity ( $p < 0.01$ )

- Seems to be an ideology effect among female judges

◀ Judge heterogeneity for female lawyers

Dependent Variable: Case Outcome (win = 1, lose = 0)						
Masculine	-0.0153*** (0.00576)	-0.0125** (0.00576)	-0.0166** (0.00756)	-0.00546** (0.00225)	-0.00554** (0.00228)	-0.00389* (0.00235)
Aggressive	-0.000563 (0.00352)	0.0000108 (0.00350)	0.00181 (0.00424)	0.00228 (0.00173)	0.00263 (0.00175)	0.00217 (0.00164)
Attractive	0.00189 (0.00454)	-0.000299 (0.00443)	0.00189 (0.00570)	0.00202 (0.00179)	0.00172 (0.00172)	0.00101 (0.00191)
Confident	0.00554 (0.00396)	0.00390 (0.00384)	0.00199 (0.00461)	0.00284 (0.00191)	0.00258 (0.00196)	-0.00221 (0.00192)
Intelligent	0.00250 (0.00391)	0.00260 (0.00389)	0.00191 (0.00536)	0.000816 (0.00175)	0.000983 (0.00178)	0.00250 (0.00179)
Trust	0.00391 (0.00352)	0.00560 (0.00345)	0.00421 (0.00428)	0.00232 (0.00179)	0.00242 (0.00178)	0.000647 (0.00172)
Likely winner	-0.00194 (0.00421)	-0.00230 (0.00426)	-0.00435 (0.00518)	0.000736 (0.00219)	0.000825 (0.00225)	0.000916 (0.00204)
Lower Court Disagreement		0.107*** (0.0366)	0.142*** (0.0518)		0.0323 (0.0537)	-0.0659 (0.0786)
Lower Court Reversing Trial Court		-0.0218 (0.0385)	-0.00438 (0.0483)		-0.0909* (0.0543)	-0.0964 (0.0703)
Lower Court Liberal Decision		0.0750** (0.0355)	0.0406 (0.0471)		0.0155 (0.0592)	0.0127 (0.0771)
Lower Court Opinion Word Count			-0.00612 (0.0134)			0.0203 (0.0256)
Lower Court Opinion Self-Certainty Words			-0.00130 (0.402)			-0.647 (0.710)
Lower Court Politically Divided			0.0977* (0.0559)			0.139* (0.0801)
Circuit and topic fixed effects	No	No	Yes	No	No	Yes
Lawyer fixed effects	No	No	No	Yes	Yes	Yes
Observations	17782	17358	9542	17782	17358	9542
Sample: Male Petitioners						

# Firm Heterogeneity

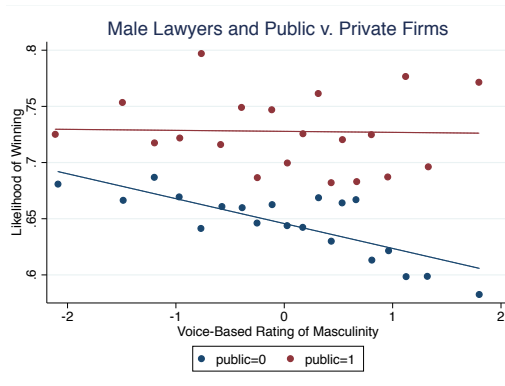


Figure: Public vs. Private Firms and Response to Masculinity

- Not-for-profit (e.g., government) vs. private firms based on petitioner status
- ◀ Not-for-profit firms' lawyers are slightly less masculine than private firm lawyers ( $p < 0.1$ )
- ◀ Firm heterogeneity for females
- ◀ Why does perceived masculinity matter?

# Firm Heterogeneity

	Masculine Rating
Public	-0.0480* (0.0287)
Observations	1580

Figure: Public vs. Private Firms and Response to Masculinity

- Perhaps private firms prefer masculine lawyers, selecting “too many” masculine lawyers
  - ▶ CEOs with deeper voices manage larger companies (Mayew, Venkatachalam, Parsons 2012)
  - ▶ Deep-voiced individuals perceived to have greater leadership capacity (Klofstad et al., 2012; Tighe et al., 2011; Apple, Streeter & Krauss 1979; Burgoon, Buller & Woodall, 1996)

# Law Firm Voicemails

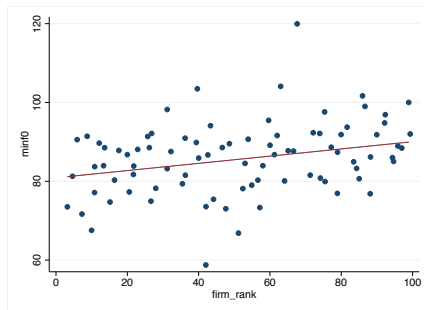


Figure: Higher Ranked Firms have low-pitched voicemails

- Perhaps private firms prefer masculine lawyers, selecting “too many” masculine lawyers
  - ▶ CEOs with deeper voices manage larger companies (Mayew, Venkatachalam, Parsons 2012)
  - ▶ Deep-voiced individuals perceived to have greater leadership capacity (Klofstad et al., 2012; Tigue et al., 2011; Apple, Streeter & Krauss 1979; Burgoon, Buller & Woodall, 1996)

# Firm Heterogeneity

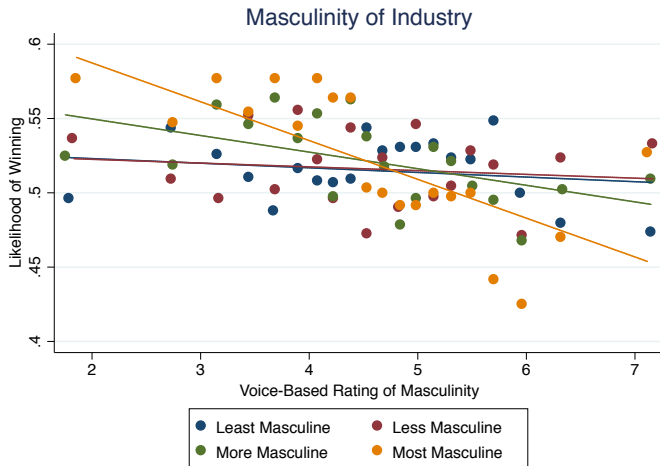


Figure: Masculinity of Industry and Response to Masculinity ( $p < 0.05$ )

- Consistent with differences in  $d > 0$  and industries with high  $d$  sacrificing more ( $\pi_F - \pi_M$ )
- Are masculinity ratings correlated with court outcomes for some firms but not others?

# Firm Heterogeneity

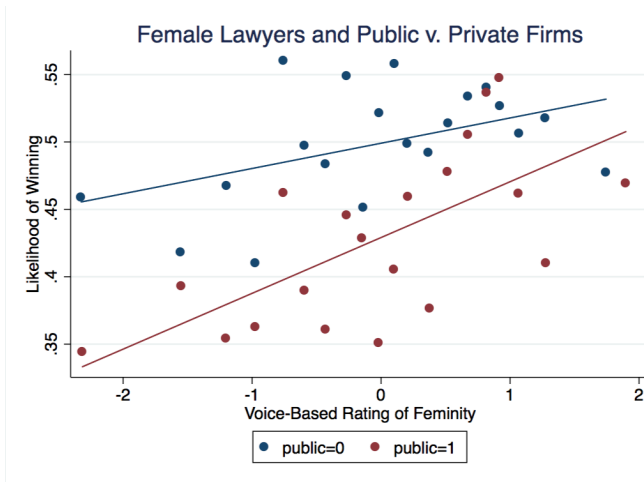


Figure: Public vs. Private Firms and Response to Femininity

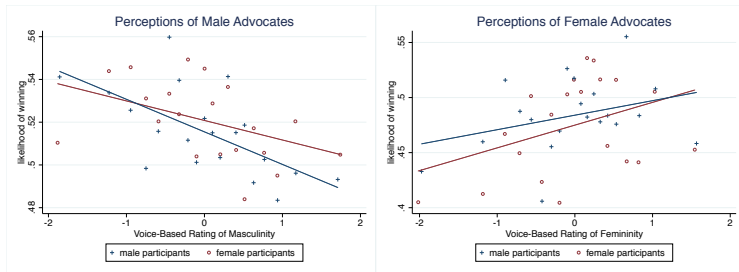
- No significant differences. ◀ Why does perceived masculinity matter?

## Rater Heterogeneity

- Half (382) of the 748 distinct raters were female.
- Two thirds of the raters were aged between 20 and 35 years old and one third were older than 35.
- One third indicated they had some college education, whereas one third claimed to have a bachelor's degree.
- Median income of those who completed the survey roughly \$40,000.
- Racial and geographical distribution of the raters broadly reflected that of the US population.
  - ▶ 74% White, 30% Southern, 20%  $\leq$  \$20,000 income, 11%  $\leq$  HS educ.
  - ▶ Generated 5 dummies and only tested significant interactions for them.



# Rater Heterogeneity



**Figure:** Voice-Based Perceptions and Court Outcomes by Advocate and Participant Gender

- Correlations between perceptions and outcomes depend on the gender of the participant rating the voices
- Gender-matching phenomenon echoes previous reports that members of the same sex may use different criteria (Babel et. al 2014)
  - ▶ But not significantly different

# Rater Heterogeneity

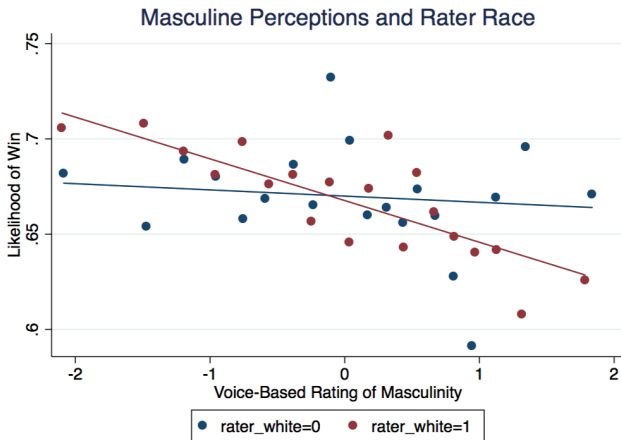


Figure: White raters' perceptions of masculinity predicted court outcomes ( $p < 0.05$ )

# Rater Heterogeneity

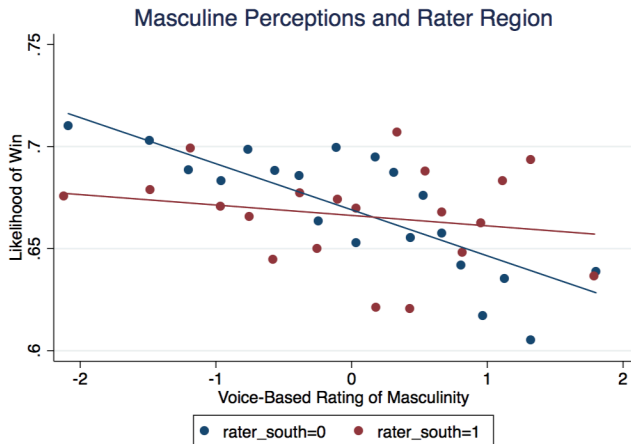
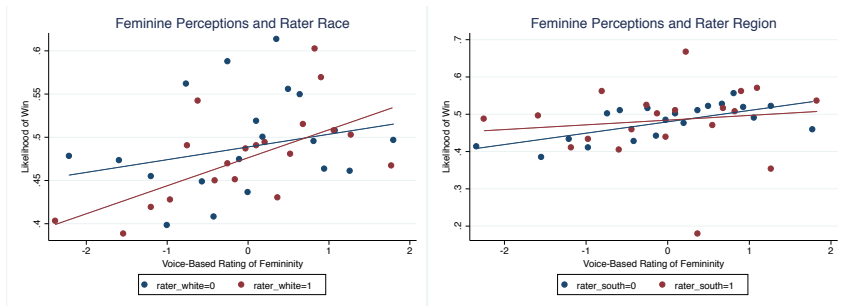


Figure: Non-Southerners' perceptions of masculinity predicted court outcomes ( $p < 0.05$ )

- If White non-Southerners  $\sim$  law firm HR, consistent with firm heterogeneity

# Rater Heterogeneity



**Figure:** White non-Southerners' perceptions of femininity also predict court outcomes ( $p > 0.1$ )

- If White non-Southerners  $\sim$  law firm HR, consistent with firm heterogeneity

# Rater Heterogeneity

- Poor ( $p < 0.05$ ), non-whites ( $p < 0.05$ ) are less likely to correlate masculinity with winning.
- Poor ( $p < 0.05$ ), male ( $p < 0.1$ ), with less education ( $p < 0.05$ ) are less likely to correlate femininity with winning.
- ▶ If these are Democrats, consistent with judicial heterogeneity

◀ Why does perceived masculinity matter?

# Predicting Behavior of the Supreme Court

Random forest tree: Katz et al. (2014), Guimera and Sales-Pardo (2011), Ruger et al. (2004), and Martin et al. (2004)

- Correctly identifies 70% of the Court's overall affirm / reverse decisions and 71% of the votes of individual justices across 1953-2013
- Uses pre-trial characteristics (25%) as well as court and judge historical trends specific to issue, parties, and lower courts (75%)
- We use prediction of justice' votes
  - ▶ Intuition: if predictive model was perfect (i.e., control for the correct decision or true case quality), then no extraneous factor should matter
  - ▶ Does perceived masculinity have an explanatory effect above and beyond this best predictor?
    - ★ If so, how much R-square improvement?
    - ★ How does perceived masculinity correlate with observable measures of case weakness?

Justice and Court Background Information		Current Supreme Court Trends	
Justice [S]	0.00781	Mean Agreement Level of Current Court [FE]	0.00955
Justice Gender [FE]	0.00205	Std. Dev. of Agreement Level of Current Court [FE]	0.00936
Is Chief [FE]	0.00283	Mean Current Court Direction Circuit Origin [FE]	0.00789
Party President [FE]	0.00604	Std. Dev. Current Court Direction Circuit Origin [FE]	0.00850
Natural Court [S]	0.00764	Mean Current Court Direction Circuit Source [FE]	0.00945
Seal Cover Score [SC]	0.00971	Std. Dev. Current Court Direction Circuit Source [FE]	0.01021
Year of Birth [FE]	0.00793	Mean Current Court Direction Issue [FE]	0.01469
<b>TOTAL</b>	<b>0.04403</b>	Z-Score Current Court Direction Issue [FE]	0.00832
Case Information		Std. Dev. Current Court Direction Issue [FE]	0.01266
Admin Action [S]	0.00978	Mean Current Court Direction [FE]	0.00918
Case Origin [S]	0.00971	Std. Dev. Current Court Direction [FE]	0.00942
Case Origin Circuit [S]	0.00845	Mean Current Court Direction Petitioner [FE]	0.00863
Case Source [S]	0.00953	Std. Dev. Current Court Direction Petitioner [FE]	0.00894
Case Source Circuit [S]	0.01015	Mean Current Court Direction Respondent [FE]	0.00882
Law Type [S]	0.01370	Std. Dev. Current Court Direction Respondent [FE]	0.00888
Lower Court Disposition Direction [S]	0.01190	<b>TOTAL</b>	<b>0.14456</b>
Lower Court Disposition [S]	0.01125	Individual Supreme Court Justice Trends	
Lower Court Disagreement [S]	0.00706	Mean Justice Direction [FE]	0.01248
Issue [S]	0.01541	Mean Justice Direction 10 [FE]	0.01530
Issue Area [S]	0.01469	Mean Justice Direction Z Score [FE]	0.00826
Jurisdiction Manner [S]	0.00595	Mean Justice Direction Petitioner [FE]	0.00732
Month Argument [FE]	0.02014	Mean Justice Direction Petitioner 10 [FE]	0.01027
Month Decision [FE]	0.01349	Mean Justice Direction Respondent [FE]	0.00724
Petitioner [S]	0.01406	Mean Justice Direction Respondent 10 [FE]	0.01030
Petitioner Binned [FE]	0.01199	Mean Justice Direction for Circuit Origin [FE]	0.00792
Respondent [S]	0.01490	Mean Justice Direction for Circuit Origin 10 [FE]	0.00945
Respondent Binned [FE]	0.01179	Mean Justice Direction for Circuit Source [FE]	0.00891
Cert Reason [S]	0.01408	Mean Justice Direction for Circuit Source 10 [FE]	0.00970
<b>TOTAL</b>	<b>0.22814</b>	Mean Justice Direction by Issue [FE]	0.01881
Overall Historic Supreme Court Trends		Mean Justice Direction by Issue 10 [FE]	0.00950
Mean Court Direction [FE]	0.00988	Mean Justice Direction by Issue Z Score [FE]	0.00771
Mean Court Direction 10 [FE]	0.01997	<b>TOTAL</b>	<b>0.14323</b>
Mean Court Direction Issue [FE]	0.01546	Differences in Trends	
Mean Court Direction Issue 10 [FE]	0.00938	Difference Justice Court Direction [FE]	0.01210
Mean Court Direction Petitioner [FE]	0.00863	Abs. Difference Justice Court Direction [FE]	0.00929
Mean Court Direction Petitioner 10 [FE]	0.00904	Difference Justice Court Direction Issue [FE]	0.01167
Mean Court Direction Respondent [FE]	0.00875	Abs. Difference Justice Court Direction Issue [FE]	0.00968
Mean Court Direction Respondent 10 [FE]	0.00925	Z Score Difference Justice Court Direction Issue [FE]	0.01055
Mean Court Direction Circuit Origin [FE]	0.00791	Difference Justice Court Direction Petitioner [FE]	0.00705
Mean Court Direction Circuit Origin 10 [FE]	0.00864	Abs. Difference Justice Court Direction Petitioner [FE]	0.00708
Mean Court Direction Circuit Source [FE]	0.00951	Difference Justice Court Direction Respondent [FE]	0.00690
Mean Court Direction Circuit Source 10 [FE]	0.01017	Abs. Difference Justice Court Direction Respondent [FE]	0.00699
<b>TOTAL</b>	<b>0.12663</b>	Z Score Justice Court Direction Difference [FE]	0.01280
Lower Court Trends		Justice Lower Court Direction Difference [FE]	0.01922
Mean Lower Court Direction Circuit Source [FE]	0.00962	Justice Lower Court Direction Abs. Difference [FE]	0.02494
Mean Lower Court Direction Circuit Source 10 [FE]	0.01017	Justice Lower Court Direction Z Score [FE]	0.01126
Mean Lower Court Direction Issue [FE]	0.01334	Z Score Justice Lower Court Direction Difference [FE]	0.00992
Mean Lower Court Direction Issue 10 [FE]	0.00933	Agreement of Justice with Majority [FE]	0.00866
Mean Lower Court Direction Petitioner [FE]	0.00949	Agreement of Justice with Majority 10 [FE]	0.01483
Mean Lower Court Direction Petitioner 10 [FE]	0.00874	Difference Court and Lower Ct Direction [FE]	0.01522
Mean Lower Court Direction Respondent [FE]	0.00973	Abs. Difference Court and Lower Ct Direction [FE]	0.01199
Mean Lower Court Direction Respondent 10 [FE]	0.00900	Z Score Difference Court and Lower Ct Direction [FE]	0.01217
<b>TOTAL</b>	<b>0.07946</b>	Z-Score Abs. Difference Court and Lower Ct Direction [FE]	0.01150
		<b>TOTAL</b>	<b>0.23391</b>

Figure: Final Feature Weights (Katz et. al 2014)

# Predicting Behavior of the Supreme Court

Predicted Vote from Random Forest	Judge Votes for Lawyer					
	0.257*** (0.0486)	0.258*** (0.0487)	0.250*** (0.0485)			0.248*** (0.0489)
Masculine	-0.0223** (0.0101)	-0.0207** (0.0101)		-0.0852** (0.0359)	-0.0780** (0.0361)	
Cluster	Lawyer and Judge					
Collapsed	No	No	No	Yes	Yes	Yes
Observations	26447	26391	26391	1229	1229	1229
R-squared	0.061	0.002	0.063	0.058	0.008	0.064
Sample: Male Petitioners, Democrat Judges						

Figure: Best Prediction and Perceived Masculinity

- Perceived masculinity basically orthogonal to random forest prediction
- Uncollapsed: Additional 3% of variance explained
- Collapsed: Additional 10% of variance explained
- Random forest also selects perceptions, improves accuracy by 2%
  - ▶ Katz et al. (2014) doesn't predict close cases well
  - ▶ Explanation of Variance



# Acoustic Data

The following phonetic attributes were measured

- Formant frequencies (F1, F2) for ve stressed vowels (/i, I, O, ej, 2/)
- Formant dispersion (average vowel distance from a central point per talker per sound clip)
  - ▶ pitch range is correlated with ideology (Lieberman 2015; Reeve 2015)
- Spectral tilt (H1-H2, H1-A1, H1-A2, H1-A3)
- Sibilant's duration, four spectral moments, and peak frequency
- Speaking rate (phonemes per second)
- Rhythm (Pairwise Variability Index) (Low et al. 2000)
- F0 mean and standard deviation

Perceived masculinity is correlated with many acoustic variables (Puts et al. 2011)

# Acoustic Data

Dependent Variable: Case Outcome (win = 1, lose = 0)			
Masculine	-0.0875** (0.0369)	-0.0972** (0.0364)	-0.0858** (0.0348)
Confident		0.0258 (0.0247)	0.0360 (0.0220)
Attractive		-0.0171 (0.0181)	-0.0197 (0.0144)
Educated		0.0158* (0.00878)	0.0146 (0.00932)
Intelligent		0.00549 (0.00893)	0.00635 (0.00783)
Trust		-0.00512 (0.00979)	-0.00528 (0.00786)
Likely winner		-0.00355 (0.00793)	-0.00132 (0.00729)
Acoustic Controls	No	No	Yes
Observations	10920	10080	10080

Figure: Case Outcomes and Perceived Masculinity

- Perceptions matter beyond the acoustics

- ▶ Consistent with human perceptions of faces: humans can detect things that computers cannot (White et al. 2015; Adler et al. 2007)
- ▶ Acoustic variables are also correlated with many perceptions

◀ Explanation of Variance

## Other robustness checks

- Drop lawyers with large caseloads
- Permutation test
- Use other subjects' ratings as instrumental variables
- Day of week, seasonality, year controls
- Case fixed effects not relevant because we focus on petitioners
  - ▶ A regression with case fixed effects suggest respondent is perhaps imitating petitioner. [◀ Discussion](#)

# Swing voters

	(1)	(2)	(3)	(4)
	Outcome	Judge Votes for Lawyer	# Votes for Lawyer	Outcome
Masculine	-0.0211** (0.00928)	-0.00829 (0.00676)	-0.0722 (0.0605)	0.0523 (0.0793)
Number of Democrats				0.0215 (0.0380)
Masculine *				-0.0249 (0.0237)
Number of Democrats				
Other Ratings	Y	Y	Y	Y
Collapsed	N	N	N	Y
N	153545	150304	153545	3689
R-sq	0.002	0.001	0.001	0.006

**Figure:** Other Outcomes and Perceived Masculinity

- Results appear at outcome level but not at judge vote level, nor for the number of judge votes, nor for the number of Democrats

# Perceived Masculinity and Experience

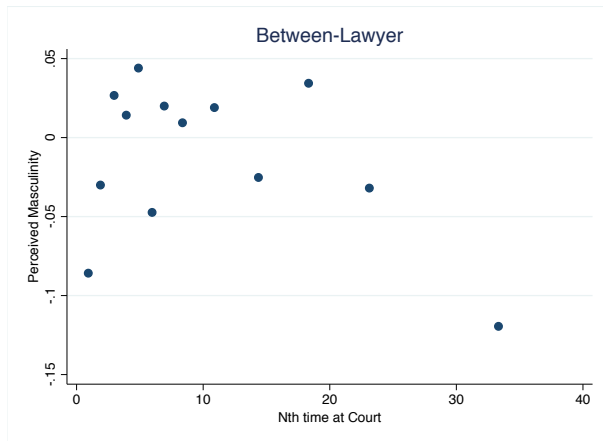
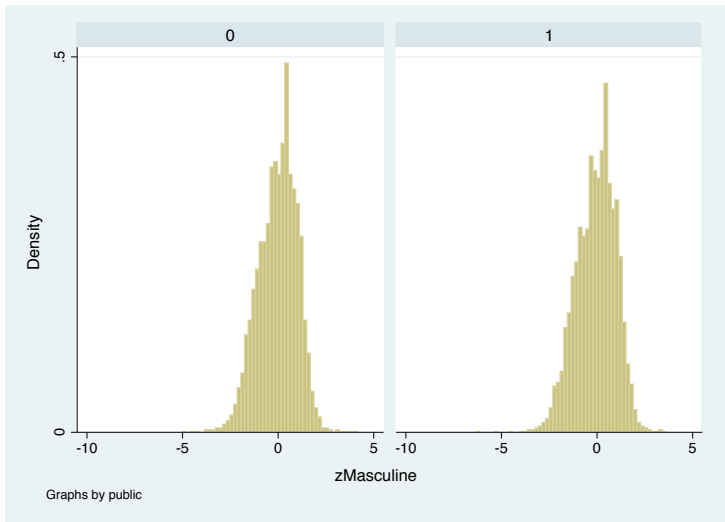


Figure: Cross-sectional variation

- More experienced lawyers have a *less* masculine voice
- Within-lawyer variation (selection of lawyers who repeatedly win) biases towards finding less masculine voice, but we found the opposite

# Firm Heterogeneity



**Figure:** Public vs. Private Firms and Density of Masculinity