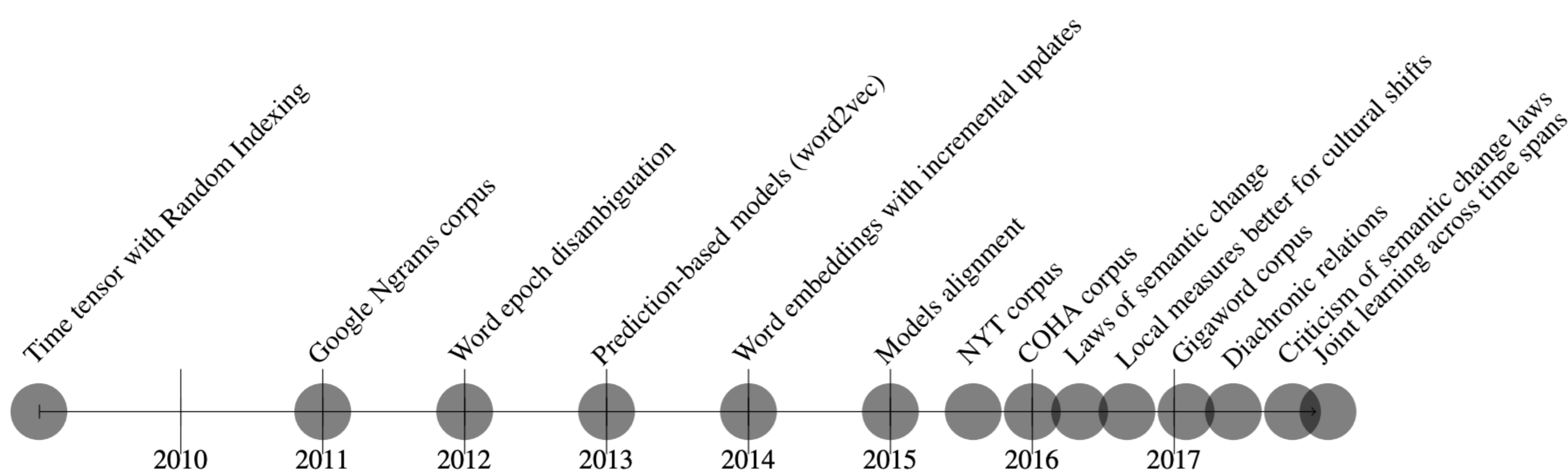


Summary

- The United States of America (U.S.) — much like Europe — has seen declining public support for major political institutions, and a general disengagement with the processes or outcomes of the three branches of government.
- According to Pew’s Public Trust in Government survey, “public trust in the government remains near historic lows” — with only 14% of Americans stating that they can trust the government to do what is right “most of the time”.
- To better understand ideological drifts in the branches of government, we use models of semantic change to track the evolution of issues around reproductive rights and the potential bias inherent in the language used in discussing it.

Semantic Shifts

- Semantic shifts are generally categorised into: (a) linguistic drifts, i.e. slow and regular changes in core meaning of words; and (b) cultural shifts, i.e. culturally determined changes in associations of a given word.
- Historically, much of the theoretical work on semantic shifts had been focused on documenting and categorising types of semantic shifts. More recently, we see a focus on the use of computational/distributional models that are count or prediction-based.



Research timeline for distributional models for tracing semantic shifts (Kutuzov et al., 2018)

Experiment Setup

- To compare the ideological stance of the U.S. Congress and Supreme Court, we use a dataset of opinions and speeches to track changes in the word “abortion”.
- To preprocess the data for our word2vec algorithm, we (i) lower-cased all tokens; (ii) chose a window size of 5; (iii) used a minimum frequency of 200; and (iv) captured case or legislation citations.

Key Characteristics of the Datasets

- In terms of frequencies, we see a more extensive vocabulary used by the U.S. Courts, which does not come as a surprise given the complexity of judicial opinions.

Characteristics	U.S. Court	U.S. Congress
No. of opinions/speeches used:	1,077,605	16,031,032
No. of judges/speakers:	~600	~6,000
Size of corpus (in bytes):	10.7GB	14.3GB
Average length (in words):	1,590	150
No. of unique words:	4,351,232	3,267,619
No. of unique words (>100):	131,632	94,431
No. of unique words (>200):	81,596	65,878

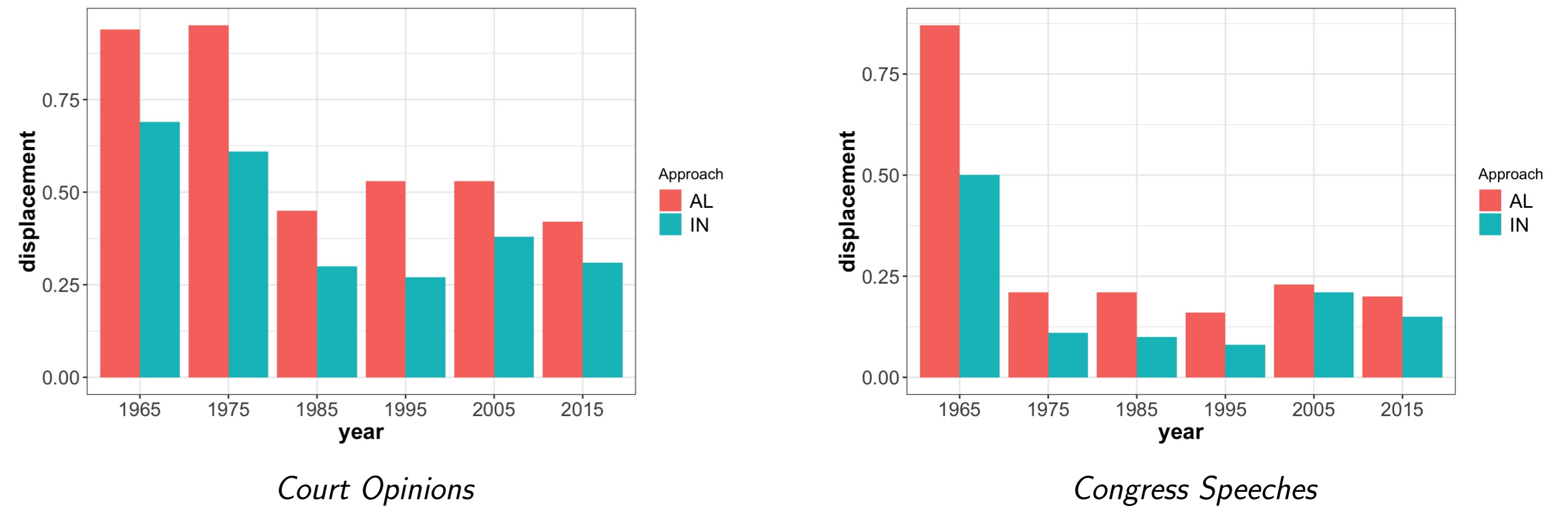
Interpreting Results

- To contextualise the results we obtain, we categorise the nearest neighbours using the assessment from an expert interview (see table key).
- We also note that the results confirm our intuition that it is harder to detect how the ideology of a judge might feed into the words used.

Category	Colour Code	Description
Medical	Yellow	Medical or scientific terminology
Legal	Green	Legal terminology
User-Chosen	Orange	Potentially biased words

Semantic Displacement of Target Word

- To formally measure semantic displacement, we compute the cosine distance between a word’s representation for different time-periods, i.e. $dist(w_t, w_{t+\delta})$, as a measure of semantic change. This shows the semantic displacement from one time period to the next for our target word (“abortion”).



U.S. Court Opinions

Local Changes in U.S. Federal Court Opinions. The top 10 nearest neighbours to the target word “abortion” ranked by cosine similarity for each model.

INCREMENTAL UPDATES					
1965	1975	1985	1995	2005	2015
sodomy	pregnant	parent	fetus	clinic	clinic
chemically	unborn	roe v. wade	unborn	unborn	unborn
pharmaceutic	griswold v. conn.	school-age	clinic	fetus	abort
polyurethane	parent	doe v. bolton	medically	partial-birth	fetus
prolix	therapeutic	immature	saline	i.u.d.	i.u.d.
carnal	medicaid	anonymity	medicaid	pregnant	pregnant
trafficker	sterilization	sterilization	pregnant	woman	partial-birth
smallest	doe v. bolton	fetus	woman	reproductive	late-term
transmits	physician	unborn	hyde amendment	physician	woman
steal	sodomy	life-saving	demonstration	abort	physician

DIACHRONIC ALIGNMENT					
1965	1975	1985	1995	2005	2015
prolix	pregnant	unemancipated	clinic	sterilize	women
polyurethane	physician	parent	medicaid-eligible	i.u.d.	clinic
sodomy	trimester	immature	fetus	contraception	reproductive
chemically	saline	school-age	roe v. wade	clinic	mcallen
trafficker	parent	anonymity	parent	partial-birth	partial-birth
carnal	medical	unborn	reproductive	f.g.m.	a.s.c’s
conspiracy	roe v. wade	fetus	unborn	unborn	parent
steal	unborn	life-saving	protester	pregnancy	fetus
mayhem	doe v. bolton	well-informed	pregnancy	fetus	roe v. wade
purloin	munday	hyde amendment	hyde amendment	reproductive	i.u.d.

U.S. Congress Speeches

Local Changes in U.S. Congress Speeches. The top 10 nearest neighbours to the target word “abortion” ranked by cosine similarity for each model.

INCREMENTAL UPDATES					
1965	1975	1985	1995	2005	2015
sterilization	anti-abortion	anti-abortion	anti-abortion	partial-birth	partial-birth
unborn	sterilization	contraception	partial-birth	pro-life	pro-life
contraception	contraception	pregnancy	pregnancy	pregnancy	contraception
therapeutic	unborn	reproductive	infanticide	infanticide	pregnancy
fetus	fetus	pro-life	third-trimester	rape	infanticide
pregnant	rape	infanticide	rape	trimester	abortionist
pro-life	sodomy	parent	abortionist	reproductive	trimester
anti-abortion	infanticide	clinic	clinic	woman	woman
back-alley	pro-life	trimester	reproductive	contraception	late-term
pro-abortion	abortionist	ectopic	woman	abortionist	self-induced

DIACHRONIC ALIGNMENT					
1965	1975	1985	1995	2005	2015
sheri	sterilization	anti-abortion	abortion-related	partial-birth	late-term
harrassment	unborn	non-therapeutic	pregnancy	late-term	partial-birth
preemptively	contraceptive	unborn	anti-abortion	third-trimester	sterilization
disingenuous	non-therapeutic	pregnancy	reproductive	pregnancy	abortionist
soviet-style	legalized	pro-abortion	pro-life	pro-life	unborn
beheading	fetus	abortion-related	clinics	abortionist	infanticide
doublespeak	pregnancy	infanticide	roe v. wade	gruesome	pro-life
theatrics	abortion-related	abortionists	infanticide	procedure	gosnell
soviet-backed	fetal	rape	gruesome	infanticide	elective
annan	alcoholic	hyde amendment	abortionist	roe v. wade	dismembering