

# Style & Substance:

## *Boring in a New Way*

by Spirling, Huang, & Perry

Juraj Medzihorsky



PolMeth 2018

*Boring*  
is  
**Interesting**

*[T]he ability to listen to bores requires stamina, and such stamina is the essence of politics. It is from the bores that you really find things out.*

Robert Harris, Cicero's character in Imperium

## *Boring in a New Way*

1. *A measure* of rhetorical distinctiveness
2. *Trends* in distinctiveness in the Commons
3. *Effects* of seniority and demotion on distinctiveness in the Commons

# 1. Measuring Rhetorical Distinctiveness

The proposed measure of rhetorical distinctiveness

$$\mathbb{D}_t = d(\mathbf{X}_t, \mathbf{X}_c) = d(\mathbf{f}_t, \mathbf{f}_c) = \sum_{v \in V_c} \bar{f}_{tv} (\eta_{tv} - \bar{\eta}_{cv})$$

# Distributional comparisons

- Many different distance measures capturing different concepts of distinction
- Possible trade-offs between interpretation and computation

## $\mathbb{D}$ 's appealing features

- Related to classical work
- Easy to compute
- Easy to modify
- Straightforward Bayesian interpretation
- 'Robust' in the Soviet sense



## $\mathbb{D}$ and Kullback-Leibler Divergence

- $\mathbb{D}$  related to  $D_{KL}(\mathbf{f}_t \| \mathbf{f}_s)$ ,

$$\sum_v f_{tv}(\eta_{tv} - \eta_{sv}) = \sum_v f_{tv} \ln \frac{f_{tv}}{f_{sv}}$$

- Information loss from using  $\mathbf{f}_s$  instead of  $\mathbf{f}_t$

# KL Divergence

## On information and sufficiency

**S Kullback, RA Leibler** - The annals of mathematical statistics, 1951 - JSTOR

1. Introduction. This note generalizes to the abstract case Shannon's definition of information [115],[161]. Wiener's information (p. 75 of [18]) is essentially the same as Shannon's although their motivation was different (cf. footnote 1, p. 95 of [161]) and Shannon apparently has ...

☆  Cited by 13609 Related articles All 8 versions

# KL Divergence



# Individuals, institutions, and innovation in the debates of the French Revolution

Alexander T. J. Barron<sup>a</sup>, Jenny Huang<sup>a,b</sup>, Rebecca L. Spang<sup>c</sup>, and Simon DeDeo<sup>b,d,1</sup>

<sup>a</sup>School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408; <sup>b</sup>Santa Fe Institute, Santa Fe, NM 87501; <sup>c</sup>Department of History, College of Arts and Sciences, Indiana University, Bloomington, IN 47405; and <sup>d</sup>Department of Social and Decision Sciences, Dietrich College, Carnegie Mellon University, Pittsburgh, PA 15213

Edited by Danielle S. Bassett, University of Pennsylvania, Philadelphia, PA, and accepted by Editorial Board Member Michael S. Gazzaniga March 19, 2018  
(received for review October 9, 2017)

Measuring innovation and influence in legislative speech with  $D_{KL}$

$$\text{Novelty} \quad \mathcal{N} = D_{KL}(\mathbf{s}_t \| \mathbf{s}_{t-1})$$

$$\text{Transience} \quad \mathcal{T} = D_{KL}(\mathbf{s}_{t-1} \| \mathbf{s}_t)$$

$$\text{Resonance} \quad \mathcal{R} = \mathcal{N} - \mathcal{T}$$

# Style, Substance, & $\mathbb{D}$

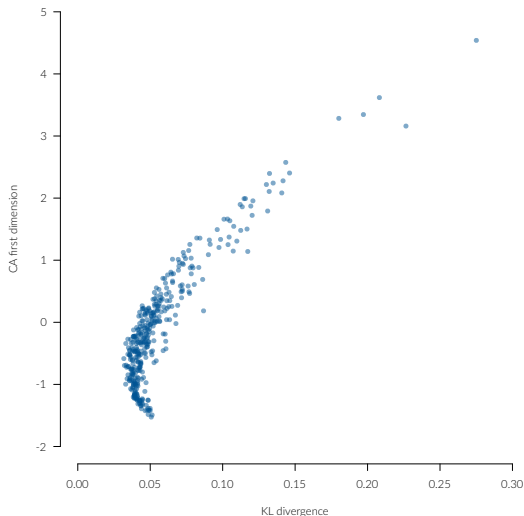
- What concept of style underlies  $\mathbb{D}$ ?
- Is it separate from substance such as policy preferences or ideology?

## Mosteller & Wallace on the disputed *Federalist Papers*

- Similar political positions
- Shared *Spectator* style: oratorical, complicated
- Measure stylistic elements not easy to observe directly
- Useful signal in preferences over a subset of common words

# Style & Substance, $D_{KL}$ & CA

If word usage reflects both style and substance,  $\mathbb{D}$  captures both.



# Style, Substance, & ID

- Association of style with substance can vary
- Different styles for different audiences
- Flashy style to obscure position
- Boring style to obscure position



## $\mathbb{D}$ and boredom

- What is the concept of boredom motivating  $\mathbb{D}$ ?
- $D_{KL}$  and surprise: Are predictability, repetitiveness, routine, reliability, boring? Varies over persons and domains.
- Is  $\mathbf{f}_c = \mathbf{f}_s$  really the most boring one can get? How about limited vocabularies?
- Boredom is easier to experience than to define

From  $\{\mathbf{X}_s, \mathbf{X}_t\}$  to  $\{\mathbf{f}_s, \mathbf{f}_t\}$

- Weighting speeches:

$$\bar{f}_{tv} = \frac{1}{|I_t|} \sum_{i \in I_t} \frac{x_{iv}}{n_i}; \quad \bar{\eta}_{cv} = \frac{1}{|S_c|} \sum_{s \in S_c} \eta_{sv}$$

insensitive to verbosity

- Smoothing document-feature matrices
- Selecting (weighting) types (now CV)
- Reducing # types with matrix factorization

## 2. Rhetorical Distinctiveness in the UK House of Commons

# $\mathbb{D}$ in the House of Commons

1. External validation
2. Trends in rhetorical distinctiveness
3. Causes of rhetorical distinctiveness

## External validation of $\mathbb{D}$

- Compare distinctiveness with media attention
- Several significance tests reported, but how strong is the association?
- What does the joint distribution look like?

# Causes of rhetorical distinctiveness

1. Socialization, recruitment, career considerations etc. thought to drive propensity to rebel in different directions
2. Rebels speak differently
  - Logic of measurement preferable to logic of discovery: how *large* and *variable* are the effects?

Table 1: Effect of experience and demotion on distinctiveness of MPs. The first model pools all observations. The second adopts a panel structure with MP-fixed effects. The third uses MP-fixed and session-fixed effects.

	<i>Dependent variable: distinctiveness</i>		
	(1)	(2)	(3)
experience	−0.0003*** (0.0001)	−0.001*** (0.0001)	−0.002*** (0.0002)
demoted	0.004 (0.003)	0.006 (0.004)	0.003 (0.004)
constant	0.443*** (0.001)		
MP-fixed effects	✗	✓	✓
Session-fixed effects	✗	✗	✓
Observations	16,022	16,022	16,022
R <sup>2</sup>	0.001	0.006	0.069
F Statistic	4.995*** (df = 2; 16019)	39.602*** (df = 2; 13396)	11.889*** (df = 83; 13315)

Note: \*\*\* $p < 0.01$

# Causes of rhetorical distinctiveness

## Substantive significance

- Experience and demotion capture 0.1% of  $\text{var}(y)$
- with  $>2,500$  MP-fixed effects a further 0.5%
- with  $\sim 80$  session-FEs a further 6.3%
- Implications for
  - link between rebellion and parliamentary speech
  - the focus of the substantive literature
  - type engineering



# Speculations

- Rebelling as a high-risk move with highly variable effects
- Rebelling with the opposition vs. outflanking
- Domains of distinction (vocabulary, accents, clothes etc.) vary in importance across contexts
- Changing audiences (journalists, voters) and their consumption patterns of MP speeches

## *The Secret of Our Success* by Henrich

- Prestige and attention link strongly. High-prestige in-group members are more interesting.
- Imitation of high-prestige individuals explains a lot of what we modern humans do. This imitation is often unintentional and unreflective, and includes irrelevant elements such as speech mannerisms.



<https://www.geograph.org.uk/photo/128371>

*Any fool can appreciate a hill, but it takes a [person] of real discernment to appreciate a fen.*

Tony Robinson, quoting an unnamed farmer