Regulatory Compliance with Limited Enforceability: Evidence from Privacy Policies

Bernhard Ganglmair[1] ⓡ **Julia Krämer**[2] ⓡ **Jacopo Gambato**[1]

[1]**ZEW, University of Mannheim, and MaCCI**

[2]**Erasmus University Rotterdam**

ASSA 2024 Annual Meeting
San Antonio, TX
January 5–7, 2024

Download at `https://ssrn.com/abstract=4600876`. The order of authors is randomized using the AEA's Author Randomization Tool (ID: iqI2Orn8rEuo).

Regulatory Compliance with Limited Enforceability

- Firms compliance with regulation is only as good as the rules are clear and the regulator is properly equipped to regulate.
- If firms do not expect strong enforcement, there wont be much compliance!
- Especially relevant with multidimensional regulation and multiple regulatory targets (assume: same value of compliance)
  - The regulator chooses a path of least resistance $\rightarrow$ focus on what is *easy to regulate*.
  - Conversely, the regulator neglects what is *difficult to regulate*.
- We should expect this asymmetric enforcement and compliance pattern to be stronger for budget-constrained regulators!

Many such environments! This paper: data regulation via the GDPR

GDPR: An Update of EU Data Regulation

- In 2018, the GDPR (General Data Protection Regulation) fundamentally transformed the legal requirements of privacy policies
- One of the (many) goals (going back to earlier consumer protection agenda):
  - Enhance transparency and accountability

1. **Art. 13–14 GDPR:** Disclose what data is collected, how, by whom
   - Lists of items and requirements
2. **Art. 12(1) GDPR:** *"concise, transparent, intelligible and easily accessible form, using clear and plain language"*
   - Little to no guidance (e.g., Art. 29 Working Party: "readability testing")

Enforcement and Compliance

- Two aspects of the *transparency requirement* come with stark differences
  - *Disclosure* (Art. 13–14) is clear and objective → **easy enforcement**
  - *Readability* (Art. 12(1)) is vague and subjective → **difficult enforcement**
- We expect budget-constrained regulators to **prioritize** and emphasize disclosure in their enforcement activities.
  - Firms expect *more* enforcement of disclosure → better compliance
  - Firms expect *less* enforcement of readability → little compliance
- Limited enforceability → differential enforcement → differential compliance

How do the differential enforcement dynamics affect compliance with the GDPR's transparency requirements?

Here's How We Tackle This Problem

1. Theoretical framework of asymmetric enforceability
   - A simple two-dimensional audit model provides predictions
2. Panel of privacy policies
   - Novel data set with privacy policies for German firms (2014–2021)
   - Texts of 585,329 privacy policies by 75,683 firms
3. Firm and industry-level information
   - Mannheim Enterprise Panel (MUP) for firm size and HHI
   - Budget information for state data protection authorities
4. Natural language processing of privacy policies to
   - measure **disclosure** and **readability**
   - $\rightarrow$ as measures for compliance!

We ask and answer three questions

Preview

1. **How did firms respond to the GDPR?**
   - Disclosure: Large increase
   - Readability: Little compliance (with what is difficult to enforce!)

2. **How does more scrutiny by regulators affect firms' compliance?**
   - Better readability compliance for firms in concentrated industries (enforcement targets!)
   - Disclosure compliance does not improve in concentrated industries

3. **How do a regulator's resources affect firms' compliance?**
   - 16 state data protection authorities regulate firms in their respective states (same rules, different enforcers)
   - Firms regulated by higher-budget DPAs show better readability compliance
   - No effect on disclosure compliance (maybe even less, as theory predicts)

Data: Privacy Policy Panel

Privacy Policy Panel

- Construction of panel
  - Start with 1.2m privacy policies by more than 500,000 German firms between 2012 and 2021.
  - Data source: Internet Archive via the Wayback Machine
- Take subsample of firms with at least one observation in the following time periods:
  - **Pre-GDPR:** 2014–Q2 2018 (before GDPR's enforcement date: on May 25, 2018)
  - **Post-GDPR:** Q2 2018–Q2 2021 (after GDPR's enforcement)
- Unbalanced quarterly panel:
  - **585,329 privacy policies** posted by **75,683 firms**

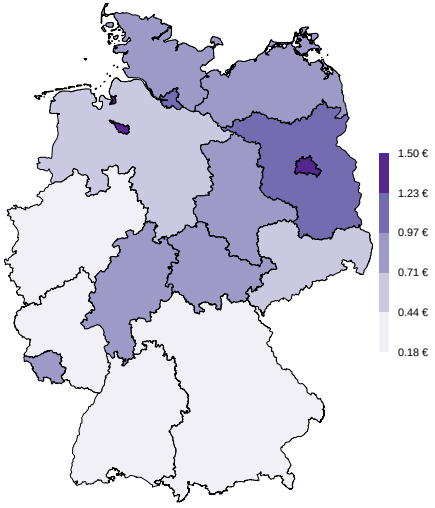
Unbalanced panel

Add: Other Data Sources

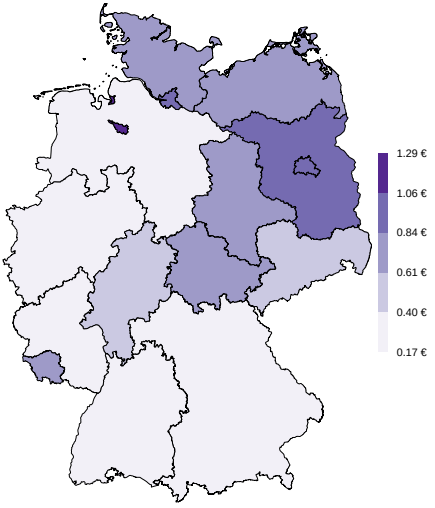
- **Firm-level** information: Mannheim Enterprise Panel (MUP)
  - Firm size information (employment and revenue/sales)
  - Industry classifiers (NACE codes)
  - NACE 4-digit Herfindahl-Hirschman Index (national markets)
  - Firm HQ address ($\rightarrow$ state)
- **State** government websites
  - Budget information for a firm's state data protection authorities (total, labor)
- German Federal Statistical Office
  - Population by state

Sample statistics

State DPA Budgets (2017): Variation Across States

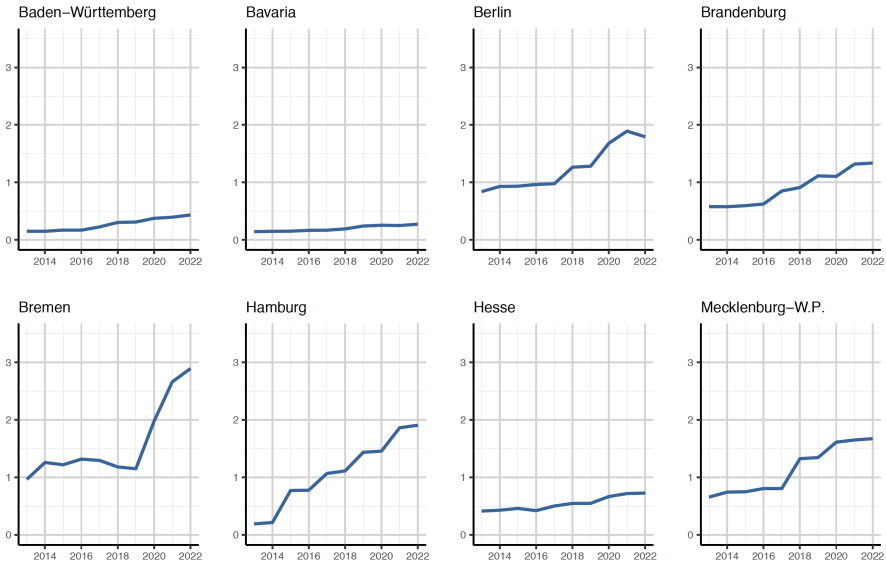


Total DPA Budget Per Capita



Personnel Budget per Capita

Variation Over Time (Example: Labor Budget per Capita)



How do we measure regulatory compliance?

How do we measure *disclosure* and *readability*?

for disclosure:	the standard NLP approach
for readability:	digging into the linguists' toolbox

Disclosure (Art. 13–14 GDPR)

- Simple: Volume of information (length and breadth of policies)
  - Number of words (total and unique)
  - Number of sentences
  - Number of distinct topics from *LDA topic models* (with paragraph-level corpus)
- **Disclosure:** Policy contains GDPR relevant information
  - Number of words of disclosing paragraphs:
    - Identify paragraphs with terms expected in Art. 13–14 related parts of policies
  - Number of words of topic-weighted paragraphs
    - Higher weight for disclosing paragraphs

Topic weights

Note: Not an assessment of GDPR-compliance of the policies!

Readability (Art. 12(1) GDPR)

- EU advisory body proposes *readability testing* (Art. 29 Working Party, 2018)
- Linguists have proposed a long list of indices and scores to measure readability
- Some popular readability scores
  - Flesch Reading Ease Score (Flesch, 1948)
    [e.g., used for regulation of insurance contract language in FL, MA, MI, and TX]
  - Gunning Fog Index (Gunning, 1952) [also in the literature: obfuscation measure]
  - Flesch-Kincaid Grade Level Score (Kincaid et al., 1975)
  - Neue Wiener Sachtextformeln (Bamberger and Vanecek, 1984)
- Often weighted average of various factors: higher readability with
  - ... more common words
  - ... shorter words
  - ... shorter sentences

We Use 2 Readability Indices

1. German version (Amstad, 1978) of the **Flesch Reading Ease Score (German FRE)** (Flesch, 1948)

$$180 - ASL - 58.5 \times \frac{n_{sy}}{n_w}$$

   ... because of its regulatory history

2. **Läsbarhetsindex (LIW)** (Björnsson, 1968)

$$ASL + \frac{100 \times n_{sy \geq 7}}{n_w}$$

   ... because it best predicts 4,000 pairwise comparisons of text snippets (assessed by "users") (following Benoit et al., 2019)

Results

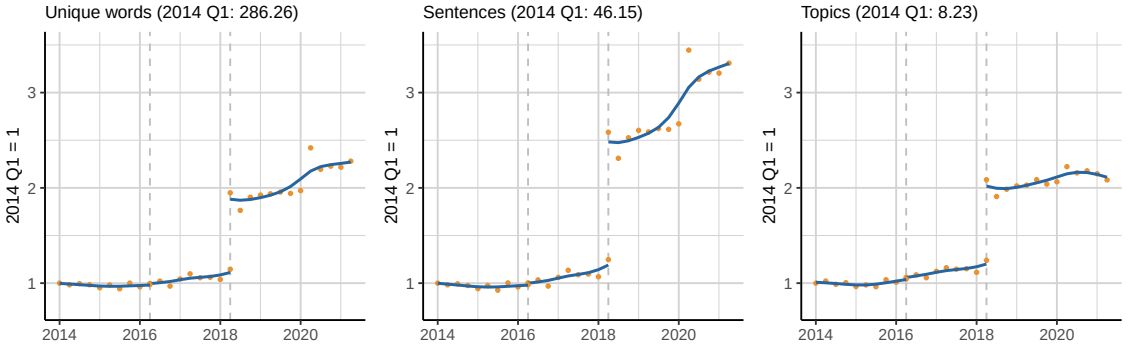
- Before-and-After Results
- Regulatory **Exposure**
- Regulatory **Scrutiny**
- Regulatory **Capacity**

Before-and-After Results

Prediction:

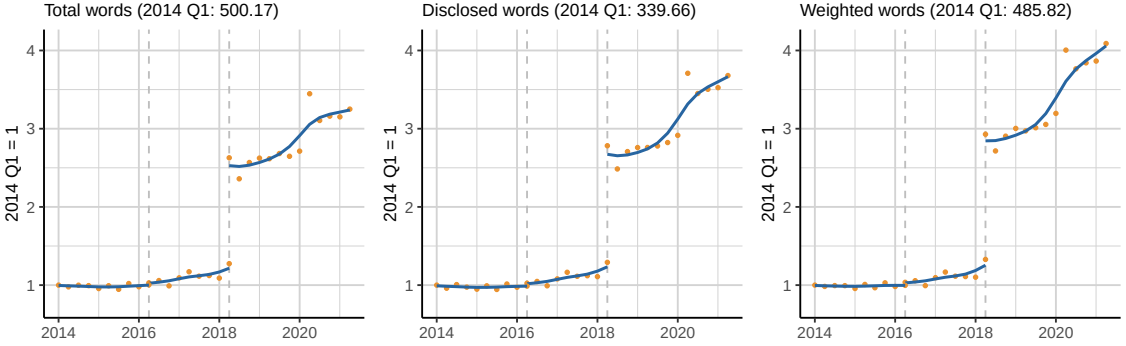
GDPR $\rightarrow$ more disclosure, maybe more readability

Simple Informational Volume



Policies double to triple in length and continue to increase; breadth (distinct topics) increases to a constant!

Disclosure



Disclosure triples/quadruples!

Regression Results: Disclosure

Panel (a): Informational Volume

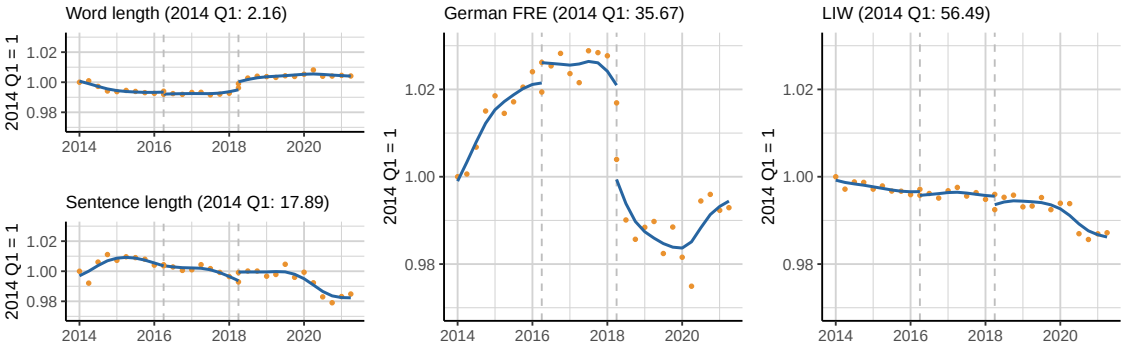
Dependent variable (in log):	**Unique words** (1)	**Sentences** (2)	**Topics** (3)
Post GDPR (=1)	0.5182***	0.7219***	0.4886***
	(0.0048)	(0.0064)	(0.0074)
# Firm FE	64,605	64,605	64,596
R^2	0.789	0.793	0.696
Observations	409,221	409,221	409,071

Panel (b): Disclosure

Dependent variable (in log):	**Total words** (4)	**Disclosed words** (5)	**Weighted words** (6)
Post GDPR (=1)	0.7078***	0.7929***	0.7775***
	(0.0065)	(0.0084)	(0.0073)
# Firm FE	64,605	64,605	64,605
R^2	0.792	0.779	0.782
Observations	409,221	409,221	409,221

+ additional controls: firm size (log employees) and market concentration (HHI)

Readability



Readability factors and scores show mixed results!

Regression Results: Readability

Dependent variable (in log):	**Word length** (5)	**Sentence length** (6)	**German FRE** (7)	**LIW** (8)
Post GDPR (=1)	0.0081***	-0.0102***	-0.0418***	-0.0041***
	(0.0003)	(0.0014)	(0.0018)	(0.0006)
log Employees	0.0007***	0.0011	-0.0030**	0.0004
	(0.0003)	(0.0014)	(0.0015)	(0.0006)
Concentration (HHI in '00)	-0.00001**	0.0000002	0.0001***	-0.00002*
	(0.000007)	(0.00003)	(0.00004)	(0.00001)
Firm FE, Year FE	Yes	Yes	Yes	Yes
# Firm FE	64,605	64,605	64,602	64,605
R^2	0.678	0.642	0.624	0.648
Observations	409,221	409,221	409,131	409,221

Add fixed effects, firm size, HHI: Readability factors and scores show mixed results!

So What?

- Is disclosure compliance much cheaper than readability compliance?
  - "This is not about a response to enforcement but a response to compliance costs."
- Do either firms or data protection authorities know what readability means?
  - "Of course, we see little to no readability compliance."

Three sets of results to address these concerns ⇒
- it is not (only) about costs
- the readability requirement is binding

Regulatory Exposure

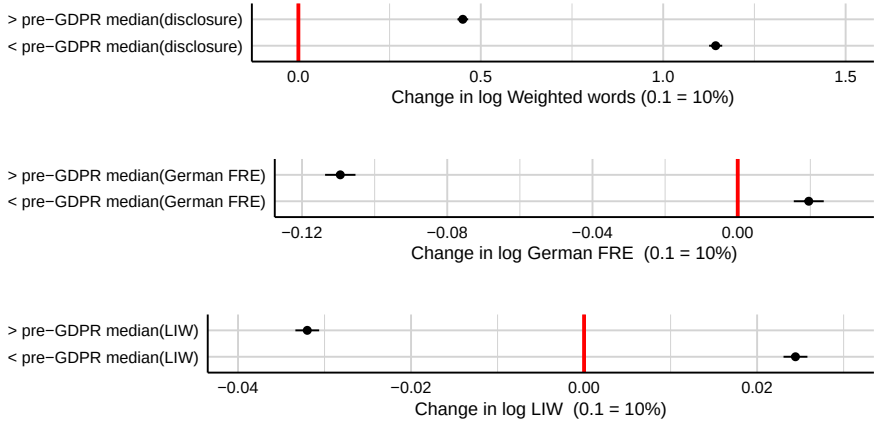
Prediction:

Higher treatment intensity, better compliance

Treatment Intensity: GDPR Exposure

- Treatment intensity:
  - Firms with *lower pre-GDPR disclosure* are *more exposed* to GDPR
  - Firms with *lower pre-GDPR readability* are *more exposed* to GDPR
- $\rightarrow$ think: treatment-intensity DiD
- We expect firms with higher pre-GDPR disclosure to exhibit better compliance with the transparency requirement (both disclosure and readability)
- *How:* We interact the post-GDPR coefficient with above/below median dummy variables (disclosure and readability)

GDPR Exposure: Readability Requirement Bites!



Compliance cost differences or ¯_(ツ)_/¯ do not explain this!

Regulatory Scrutiny

Prediction:
More scrutiny, better compliance

Effect of Market Concentration on Compliance

	Disclosure	Readability			
Dependent variable (in log):	Weighted words	Word length	Sentence length	German FRE	LIW
Post GDPR (=1)	0.7509***	0.0081***	-0.0113***	-0.0401***	-0.0039***
	(0.0062)	(0.0002)	(0.0012)	(0.0015)	(0.0005)
Concentration (HHI in '00)	0.0006***	0.00002**	-0.0002***	0.000002	-0.00003
	(0.0002)	(0.000008)	(0.00004)	(0.00005)	(0.00002)
× Concentration (HHI in '00)	0.00001	-0.00005***	0.0003***	0.0002**	0.00002
	(0.0003)	(0.00001)	(0.00006)	(0.00007)	(0.00002)
Firm FE, Year FE	Yes	Yes	Yes	Yes	Yes
# Firm FE	73,493	73,493	73,493	73,490	73,493
R^2	0.757	0.643	0.606	0.592	0.612
Observations	567,166	567,166	567,166	566,997	567,166

Better readability compliance in concentrated industries (targets!). Firm size?

Regulatory Capacity

Prediction:
Higher-budget DPAs induce better readability compliance but not disclosure compliance

We exploit budget variation
(across states and time)
to proxy for enforcement intensity

Give regulators more money,
and firms will fall in line!

Compliance by Budget

Dependent variable (in log):	Disclosure	Readability			
	Weighted words	**Word length**	**Sentence length**	**German FRE**	**LIW**
Panel (a): DPA Budget – Total Budget Per Capita					
× Total budget (per capita, lagged)	-0.0087	-0.0022***	0.0087***	0.0048*	0.0008
	(0.0117)	(0.0005)	(0.0024)	(0.0029)	(0.0010)
Panel (b): DPA Budget – Personnel Budget Per Capita					
× Staff budget (per capita, lagged)	-0.0165	-0.0027***	0.0109***	0.0053	0.0005
	(0.0138)	(0.0006)	(0.0029)	(0.0034)	(0.0012)

Compliance by state

Let's Wrap Up!

Track Record of the GDPR is Mixed

- Ambiguous effect of GDPR on firms' privacy policies pointing at enforcement dynamics
  - **Readability** is a vague concept and subjective → difficult enforcement
  - **Disclosure** requirement is clear and objective → easy enforcement
- Choice of the appropriate readability test (enforcement of Art. 12(1)) is crucial as indices often disagree
  - Comparison data suggests LIW for German legal texts
  - Results also for German FRE (used in regulatory context)
- Better compliance for firms in concentrated industries (→ primary enforcement targets?)
- Regulator capacity matters!
  - Higher-budget DPAs induce better readability compliance
  - Higher-budget DPAs do not affect disclosure compliance

Literature: Where Do We Fit In?

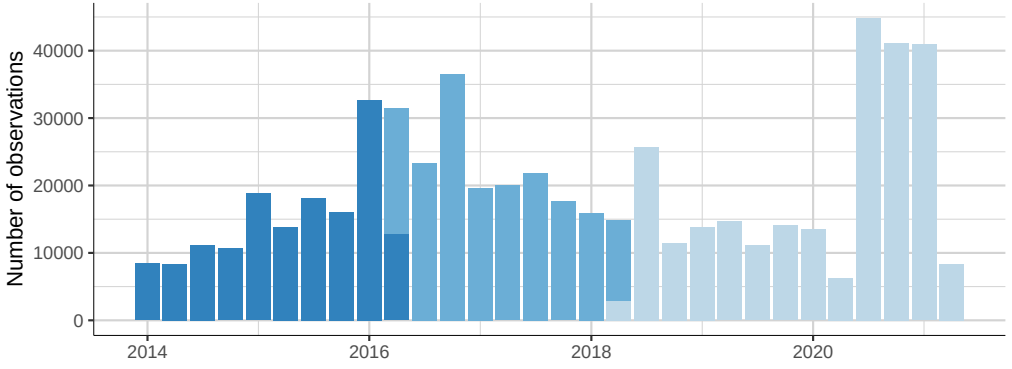
- Effects of GDPR on firm outcomes:
  - Yuan and Li (2019); Goldberg et al. (2019); Koski and Valmari (2020); Peukert et al. (2022); Johnson et al. (forthcoming)
  - **Contribution: more nuanced picture of the effectiveness of GDPR**
- Privacy policies (content and readability)
  - Jensen and Potts (2004); Degeling et al. (2019); Sanchez-Rola et al. (2019); Becher and Benoliel (2021); Milne et al. (2021)
  - Larger scale: Frankenreiter (2022) [60,000] or Amos et al. (2021) [>1m]
  - **Contribution: larger sample size/firm-level information to explain; guidance for readability testing (data-driven approach)**
- Fine print in contracts and terms of use
  - Bakos et al. (2014); Marotta-Wurgler (2008) [shows striking homogeneity in contract terms]; Marotta-Wurgler (2016)
  - **Contribution: Privacy policies exhibit high level of heterogeneity**
- Audits, tax evasion
  - Heyes (1994); Barsley et al. (1994); Macho-Stalder and Perez-Castrillo (2006) [all on imperfect audits]
  - **Contribution: two-dimensional regulation goals**

Thank you!

Find the paper

- internet search engine of your choice
- on our websites
- `https://ssrn.com/abstract=4600876`

b.ganglmair@gmail.com | @ganglmair
j.k.kramer@law.eur.nl | @Julia_Kraemer
jacopo.gambato@zew.de | @JGambato_econ



Shades of blue highlight three phases (before passage; between; after enforcement) Back

Estimation Sample

	Obs.	Mean	Std.	Min	Max
Number of observations per firm	75683	7.734	4.67	2	30
... in pre-GDPR enforcement phase	75683	4.446	3.69	1	18
... in post-GDPR enforcement phase	75683	3.288	2.17	1	13
Employees (firm-level means)	65863	36.446	408.48	1	48300
... Micro	40578	3.72	2.54	1	10
... Small and medium-sized (SME)	23920	39.222	42.13	10	249.6
... Large	1365	960.678	2671.81	250	48300
Sales (in million; firm-level means)	55656	14.942	351.78	0	62379.6
Herfindahl-Hirschman Index (HHI; in 2017)	44883	551.131	1178.23	1.5	10000
Agriculture/Mining	688	1.03%	1.96%		
Manufacturing	6387	9.56%	6.72%		
Utilities	1028	1.54%	0.92%		
Construction	4679	7.01%	10.69%		
Trade	14907	22.32%	23.89%		
Services	39105	58.55%	55.82%		
	66794	**(Sample)**	**(MUP)**		

Back

Topic-Weighted Paragraphs: An Illustration

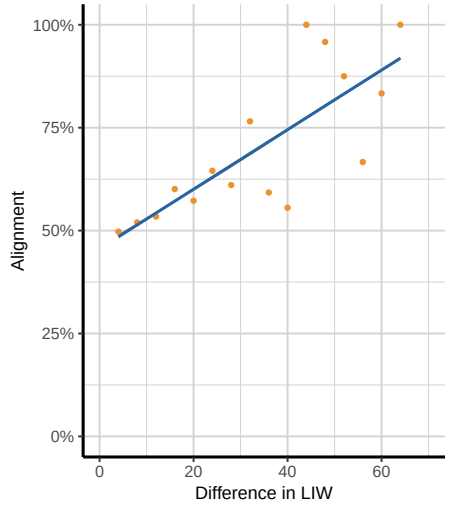
		Example 1			Example 2		
	Words	**Topic**	**Factor ϕ_k**	$\phi_k w_{c\|k}$	**Topic**	**Factor ϕ_k**	$\phi_k w_{c\|k}$
Paragraph 1	10	A	2.0	20	A	2.0	20
Paragraph 2	20	B	1.0	20	B	1.0	20
Paragraph 3	30	C	0.5	15	C	0.5	15
Paragraph 4	40	C	0.5	20	A	2.0	80
Total word count	**100**	**Disclosure (Ex. 1)**		**75**	**Disclosure (Ex. 2)**		**135**

- For the overall topic distribution, assume (0.25, 0.25, 0.50).
- Topic A is the most disclosing! Topics B and C are less relevant.
  $\rightarrow$ For the topic distribution of disclosing paragraphs, assume (0.50, 0.25, 0.25).
- The topic factors are therefore $(\phi_A, \phi_B, \phi_C) = (2, 1, 0.5)$.

Back

Are They Useful? Let the Data Speak! ($\rightarrow$ Benoit et al., 2019)

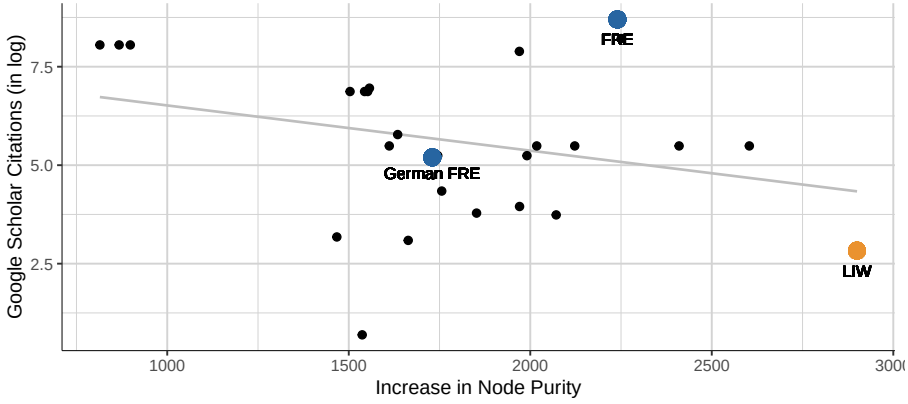
1. Collect human assessments that serve as a "gold-standard" for determining text readability
   - 700 text pairs; 14 respondents; 4,000 comparisons (data points)
   - pairs of similar length and similar content (topic: justification of data processing)
2. Fit unstructured Bradley-Terry model for pairwise comparisons (Bradley and Terry, 1952) to the data
   - originally developed for sports competition to rank the best contestant
   - estimates odds that a contestant will outperform another in a competition
3. Select best predictors (i.e., indices/scores) of readable texts by using a random forest algorithm
   - Result: Läsbarhetsindex (LIW) by Björnsson (1968)



Around 4,000 pairs of paragraphs / alignment: human ranking = text-based ranking

10 point increase of LIW (about 20% in our sample) increases alignment by 7 p.p.

Best Does Not Mean Most Popular



Performance of readability indices in predicting pairwise comparisons and their popularity

Do They Make Sense? Putting Readability in Perspective

	Obs.	Word length	Sentence length	Big words	German FRE	LIW
Privacy policy panel	585329	2.16	17.84	0.21	35.98	56.13
		(0.07)	(3.26)	(0.04)	(5.64)	(3.94)
Simple-language news (nachrichtenleicht.de)	1594	1.74	10.74	0.04	67.5	39.11
		(0.12)	(1.8)	(0.03)	(7.28)	(5.42)
Speeches and statements: Angela Merkel	1128	1.83	18.16	0.3	54.84	48.05
		(0.07)	(2.3)	(0.03)	(4.47)	(3.1)
Decisions by German Constitutional Court (BVerfG)	9358	1.96	16.35	0.15	49.27	50.17
		(0.09)	(2.91)	(0.03)	(6.75)	(4.91)
Wikipedia (German)	10000	1.9	20.63	0.12	48.48	53.51
		(0.2)	(14.48)	(0.04)	(18.23)	(15.48)
Wikipedia (English)	10000	1.71	19.78	0.05	60.33	47.8
		(0.16)	(6.57)	(0.03)	(11.58)	(9.31)
GDPR/DS-GVO (Wikipedia)	1	2.1	18.63	0.12	38.35	57.1
GDPR/DS-GVO (official)	1	2.24	40.39	0.18	8.83	81.39

Back

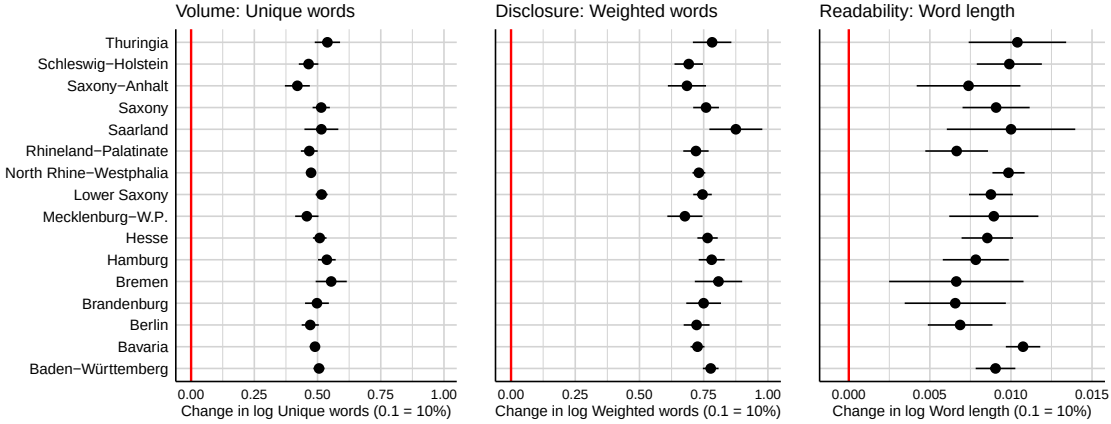
Effect of Firm Size on Compliance

Dependent variable (in log):	Disclosure	Readability			
	Weighted words	**Word length**	**Sentence length**	**German FRE**	**LIW**
Post GDPR (=1)	0.7269***	0.0090***	-0.0136***	-0.0454***	-0.0037***
	(0.0101)	(0.0004)	(0.0021)	(0.0024)	(0.0008)
log Employees	0.0033	0.0010***	0.00005	-0.0042***	0.0005
	(0.0066)	(0.0003)	(0.0015)	(0.0016)	(0.0006)
Post GDPR (=1) × log Employees	0.0216***	-0.0004***	0.0015**	0.0017**	-0.0002
	(0.0032)	(0.0001)	(0.0007)	(0.0008)	(0.0003)
Firm FE, Year FE	Yes	Yes	Yes	Yes	Yes
# Firm FE	65,863	65,863	65,863	65,859	65,863
R^2	0.784	0.680	0.645	0.627	0.650
Observations	413,249	413,249	413,249	413,154	413,249

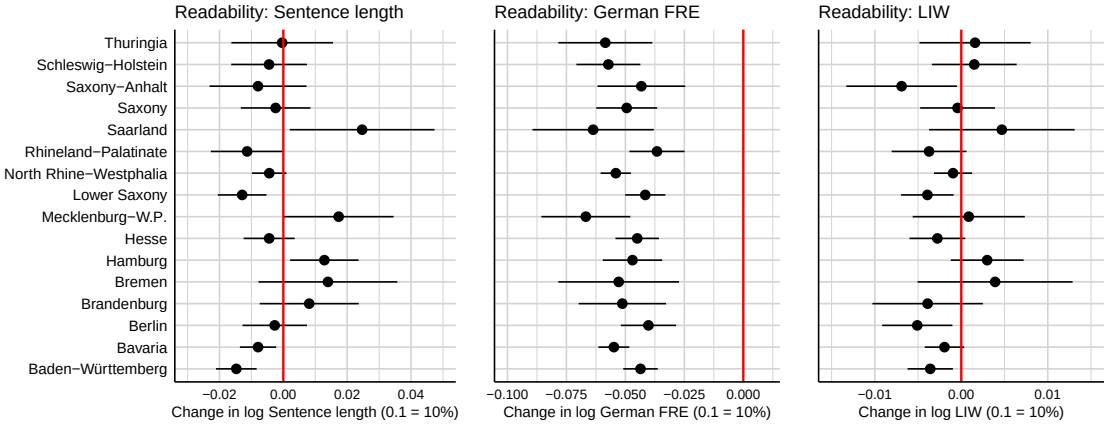
Better readability compliance by larger firms (targets!)

Back

Compliance by State



Compliance by State



Back