

Rauh, Christian

Article — Published Version

One agenda-setter or many? The varying success of policy initiatives by individual Directorates-General of the European Commission 1994–2016

European Union Politics

Provided in Cooperation with:

WZB Berlin Social Science Center

Suggested Citation: Rauh, Christian (2021) : One agenda-setter or many? The varying success of policy initiatives by individual Directorates-General of the European Commission 1994–2016, European Union Politics, ISSN 1465-1165, Sage, Thousand Oaks, CA, Vol. 22, Iss. 1, pp. 3-24, <https://doi.org/10.1177/1465116520961467>

This Version is available at:

<https://hdl.handle.net/10419/228463>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

ONLINE APPENDIX

accompanying

Rauh, Christian (2020) ‘One agenda-setter or many? The varying success of policy initiatives by individual Directorates-General of the European Commission 1994-2016’, *European Union Politics*

Contents

Appendix A: Logic of the Eur-Lex Scraper.....	2
Appendix B: Mapping treaty policy areas to Directorates-General of the Commission	3
Appendix C: Variables in data set.....	4
Appendix D: Nested regression models of the inverted normalized DL distance of proposal and law text.....	7
Appendix E: Alternative model specifications.....	8

Appendix A: Logic of the Eur-Lex Scraper

Step	Task	Details & Examples
1	Identify universe of document numbers	• CELEX identifiers: [SECTOR][YEAR][TYPE][DOCNUM] • Com proposals for binding secondary EU law: CELEX sector 5 (preparatory documents) of type ‘PC’ (COM – legislative proposals) • Manual search shows that the highest observed DOCNUM in this domain between 1985 and 2016 is 2282 • Proposals in legislative packages might share a CELEX number separated by bracketed numbers - the scraper allows for up to five of such sub-proposals; the highest number retrieved in random manual searches • This defines the possible numerical range: 51985PC0001(01) to 52016PC2282(05)
2	Check whether documents exist	• If EUR-Lex contains a document with these identifiers, it can be accessed via: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:[IDENTIFIER] • The scraper accordingly pings the respective URLs and stores the http response • ‘404’ codes indicate that no such document exists in the database and corresponding identifiers are disregarded in all further steps
3	Download bibliographical information	• For existing identifiers, the scraper stores the html content of the landing URL • These files contain bibliographical information such as the proposal title, the adoption dates and – partially – the full legal text of the proposal • Execution: July 31 2017
4	Download procedural information	• EUR-Lex provides procedural information via a dedicated URL with the structure: http://eur-lex.europa.eu/legal-content/EN/HIS/?uri=CELEX:[IDENTIFIER] • The html content of these URLs is separately stored • These files contain further document information and various details on the inter-institutional process and its outcome (such as Council and EP readings) • Execution: July 31 2017
5	Parse html and extract data	• All downloaded files are then parsed to extract relevant information and store it in a combined data frame centred on the original Commission proposal • Besides procedural information, this also includes references to the finally adopted law, if applicable, which is also uniquely identified by a CELEX number in sector 3 (legislation) of type L (directives), R (regulations), or D (decisions)
6	Remove irrelevant cases	• Some documents have been wrongly classified as original Commission proposals by the database maintainers (as indicated in the document title) • First, I remove documents that are recommendations or reports only • Second, documents that do not initiate a new procedure but present inter-inst. steps only: ‘opinions’ and ‘modified revised reexamined proposals’ are removed
6	Download and parse adopted law	• Similar to steps 3 and 5, the scraper then exploits the CELEX numbers of the adopted laws to download and store bibliographical information and the full legal text (where available) • Execution: August 3 2017
7	Data cleaning	• Infer missing procedures from legal basis, or EP involvement and Treaty in force • Harmonize actor and procedure names that have varied over time • Check and correct false entries by exploiting overlapping data points

Table 1: Logic of the custom EUR-Lex scraper

Appendix B: Mapping treaty policy areas to Directorates-General of the Commission

Biesenbender policy classification	Reduction to capture Comm organisation over time	Mapped DGs (PEU nomenclature)
Administrative cooperation	Administrative cooperation	EUROSTAT, HR, OLAF
Agriculture	Agriculture	AGRI
Financial provisions	Budget	BUDG
Economic and Social Cohesion	Cohesion	REGIO
Competition, taxation and approx. of laws	Competition, taxation, customs	COMP, TAXUD
Economic and monetary policy	Economic and monetary policy	ECFIN, FISMA
Employment Social policy	Employment and social affairs	EMPL
Environment	Environment	ENV
External action (incl. development and cooperation with third-countries) Association of the overseas countries and territories	External relations	DEVCO, EEAS, NEAR, TRADE
Public health Consumer protection	Health and consumer protection	SANTE
Free movement of goods	Industry	CONNECT, ENTR, MARE
Industry		
Institutional provisions	Institutional provisions	COM, SecGen
Internal market - general Free movement of persons, services and capital	Internal market	GROW
Citizenship and non-discrimination Justice and home affairs (incl. visas, asylum, immigration) Civil protection	Justice and home affairs	HOME, JUST
Principles and final provisions	Principles and final provisions	n/a
Culture	Research, Education, Culture	DIGIT, EAC, JRC, RTD
Research and technological development		
Transport Trans-European networks Energy Tourism	Transport and Energy	ENER, MOVE

Table 2: Policy-area and DG mapping employed

Appendix C: Variables in data set

Variable	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
anticipation	2,228	0.60	0.25	0.01	0.39	0.60	0.82	1.00
ass	2,225	0.005	0.07	0.00	0.00	0.00	0.00	1.00
cns	2,225	0.34	0.47	0.00	0.00	0.00	1.00	1.00
cod	2,225	0.43	0.49	0.00	0.00	0.00	1.00	1.00
commissioner.experience	2,218	1,352.89	1,070.50	0.00	540.00	1,113.00	1,782.00	5,322.00
commissioner.power	2,218	0.87	0.46	0.24	0.40	0.92	1.01	2.27
council.eu.var	2,228	3.03	0.58	1.41	2.56	2.83	3.40	4.20
council.lr.var	2,228	3.30	0.58	1.69	2.92	3.29	3.72	4.83
council.size	2,228	20.60	5.97	12	15	15	27	28
dg.units	2,218	29.84	11.81	10.00	20.00	26.00	40.00	79.00
dist_norm	2,228	0.40	0.25	0.00	0.18	0.40	0.61	0.99
oral	2,228	0.24	0.43	0	0	0	0	1
policy.age	2,228	32.54	10.23	1	28	34	40	49
prop.entropy	2,228	6.65	0.85	2.65	5.99	6.73	7.38	8.91
qmv	2,228	0.66	0.47	0	0	1	1	1
secgensize	2,228	26.42	3.24	23	23	26	27	32
syn	2,225	0.03	0.17	0.00	0.00	0.00	0.00	1.00
text.length	2,228	2,096.63	3,431.92	34	221	726.5	2,559.8	39,897
year	2,228	2,003.87	5.77	1,994	1,999	2,003	2,008	2,016

Table 3: Descriptive statistics of model variables

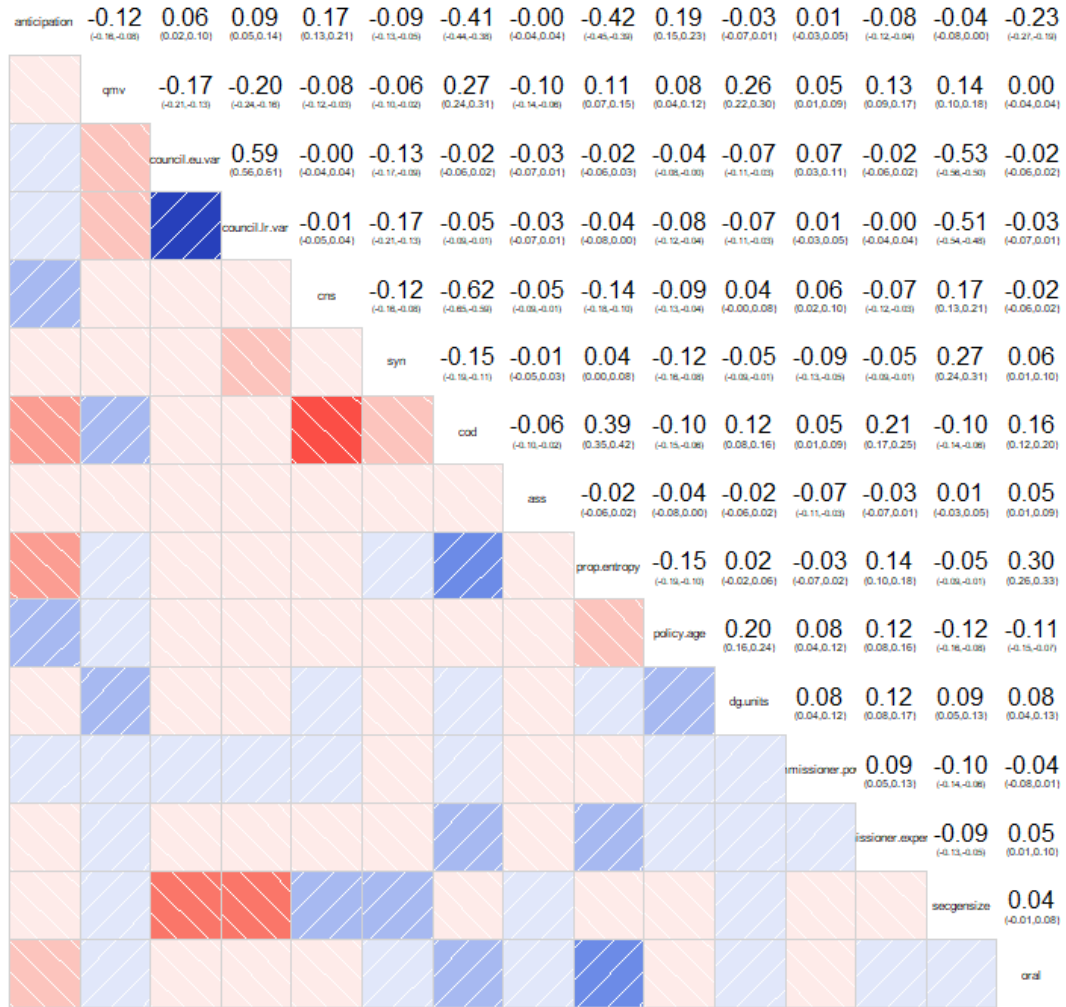


Figure 1: Correlations among model variables

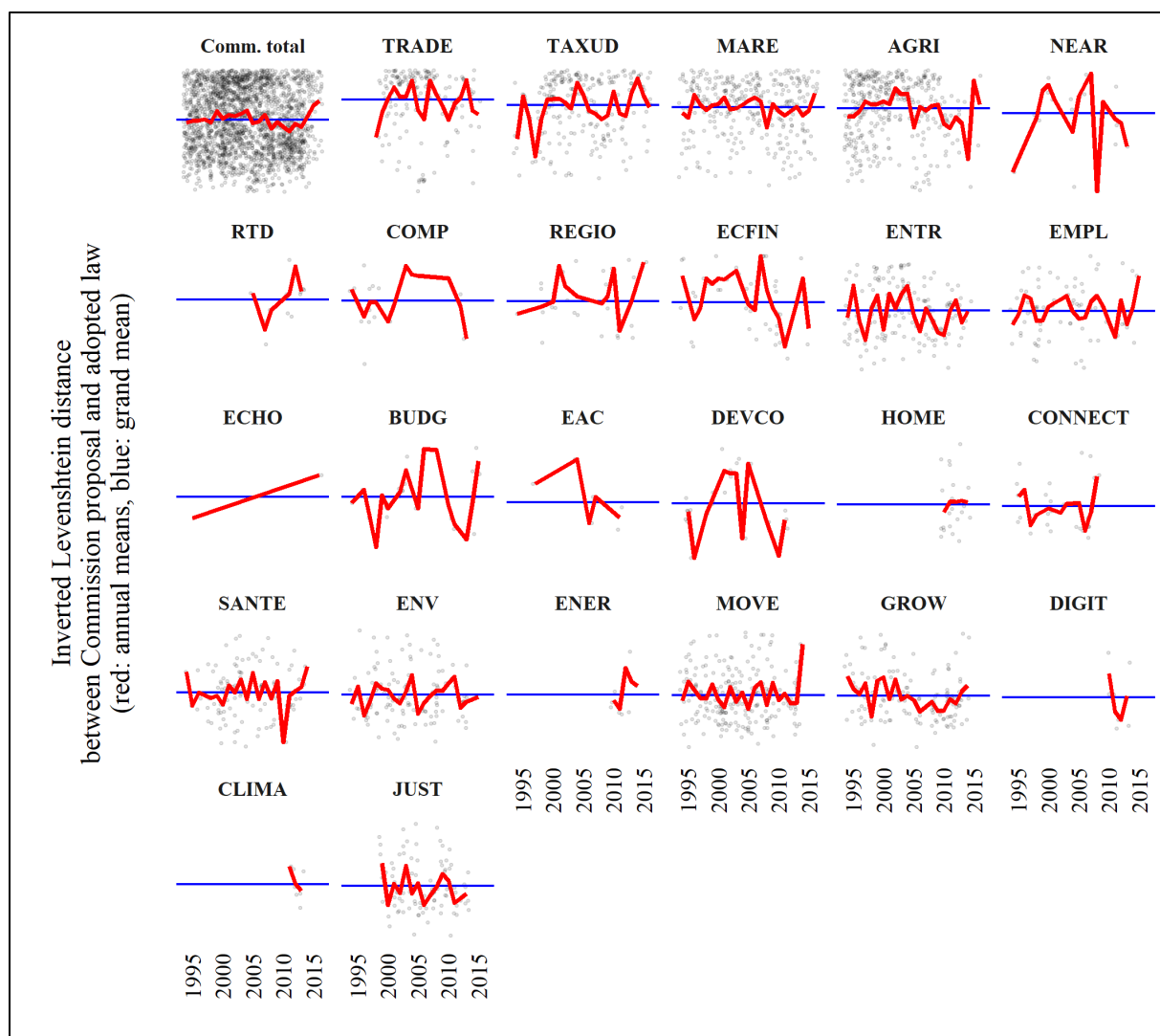


Figure 2: Dependent variable across DGs and time

Appendix D: Nested regression models of the inverted normalized DL distance of proposal and law text

	(1)	(2)	(3)	(4)	(5)
CONSTANT	0.390*** (0.043)	0.086 (0.295)	0.932*** (0.331)	2.679*** (0.539)	2.524*** (0.931)
QMV		-0.224** (0.094)	0.045 (0.102)	0.013 (0.104)	-0.008 (0.108)
COUNCIL.EU.VAR		0.003 (0.092)	0.018 (0.095)	0.020 (0.096)	0.022 (0.103)
COUNCIL.LR.VAR		0.135 (0.092)	0.057 (0.096)	0.083 (0.097)	0.078 (0.101)
CNS			-0.549*** (0.138)	-0.343** (0.145)	-0.337** (0.151)
SYN			-1.317*** (0.281)	-0.928*** (0.295)	-0.900*** (0.307)
COD			-1.303*** (0.135)	-0.912*** (0.150)	-0.904*** (0.156)
ASS			-0.767 (0.628)	-0.564 (0.639)	-0.491 (0.645)
PROP.ENTROPY				-0.352*** (0.059)	-0.330*** (0.062)
POLICY.AGE				0.009* (0.005)	0.008 (0.005)
DG.UNITS					0.001 (0.004)
COMMISSIONER.POWER					0.017 (0.100)
COMMISSIONER.EXPERIENCE					0.00002 (0.00004)
SECGENSIZ					0.001 (0.019)
ORAL					-0.186* (0.109)
Pseudo R2 (McFadden)	0	0.01	0.101	0.131	0.139
Observations	2,228	2,228	2,225	2,225	2,215
Log likelihood	-1,493.597	-1,479.290	-1,342.142	-1,297.782	-1,286.546
Akaike information criterion	2,989.194	2,966.580	2,700.283	2,615.565	2,603.092
Notes:	*** p < .01; ** p < .05; * p < .1				

Table 4: Nested regression models

Appendix E: Alternative model specifications

	<i>Logit (main model)</i>	<i>Tobit</i>	<i>OLS</i>	<i>ML OLS (random DG intercepts)</i>
CONSTANT	2.524*** (0.931)	1.076*** (0.095)	1.053*** (0.093)	1.131*** (0.103)
QMV	-0.008 (0.108)	-0.003 (0.011)	-0.002 (0.011)	-0.013 (0.012)
COUNCIL.EU.VAR	0.022 (0.103)	0.005 (0.011)	0.005 (0.010)	0.001 (0.010)
COUNCIL.LR.VAR	0.078 (0.101)	0.017* (0.010)	0.017* (0.010)	0.017* (0.010)
CNS	-0.337** (0.151)	-0.062*** (0.015)	-0.061*** (0.014)	-0.062*** (0.016)
SYN	-0.900*** (0.307)	-0.200*** (0.031)	-0.198*** (0.032)	-0.151*** (0.035)
COD	-0.904*** (0.156)	-0.199*** (0.016)	-0.198*** (0.015)	-0.159*** (0.018)
ASS	-0.491 (0.645)	-0.101* (0.053)	-0.097 (0.066)	-0.115* (0.068)
PROP.ENTROPY	-0.330*** (0.062)	-0.076*** (0.007)	-0.073*** (0.006)	-0.068*** (0.006)
POLICY.AGE	0.008 (0.005)	0.002*** (0.0005)	0.002*** (0.001)	0.001 (0.001)
DG.UNITS	0.001 (0.004)	0.0003 (0.0004)	0.0003 (0.0004)	-0.001 (0.001)
COMMISSIONER.POWER	0.017 (0.100)	0.002 (0.010)	0.003 (0.010)	0.007 (0.011)
COMMISSIONER.EXPERIENCE	0.00002 (0.00004)	0.00000 (0.00000)	0.00000 (0.00000)	0.00001 (0.00000)
SECGENSIZE	0.001 (0.019)	0.0002 (0.002)	0.0003 (0.002)	-0.002 (0.002)
ORAL	-0.186* (0.109)	-0.044*** (0.011)	-0.046*** (0.011)	-0.037*** (0.012)
<i>Observations</i>	2,215	2,215	2,215	2,215
<i>R-squared</i>			0.290	
<i>Adjusted R-squared</i>			0.285	
<i>Log likelihood</i>	-1,286.546	232.592		226.734
<i>Residual standard error</i>			0.214 (df = 2200)	
<i>F statistic</i>			64.123*** (df = 14; 2200)	
<i>Wald test</i>		969.580*** (df = 14)		
<i>Akaike information criterion</i>	2,603.092			-419.467
<i>Bayesian information criterion</i>				-322.516
<i>Notes:</i>	*** p < .01; ** p < .05; * p < .1			

Table 5: Alternative estimation approaches

The main model discussed in the article relies on a logit link, primarily justified by interpreting the dependent variable as the likelihood of textual change. Yet, while the DV is bound to the [0,1] interval it is continuous within this interval. Thus, by enforcing the logit estimation I may give away information on the mid-range of the DV. To see whether this affects my results models 2 and 3 in Table 4 estimate censored and uncensored linear specifications instead. The main interpretations are identical to the logit model. Using more information from the DV, furthermore, increases the levels of statistical significance for some of the variables (left/right heterogeneity in Council, EP assent, oral procedure).

I have also tested for theoretically plausible hierarchical structures in my data. Yet, the intra-class correlation coefficients for Commission term / presidency (.046), EU treaty in force (.012), or year of adoption (.029) do not immediately warrant caution in this regard. Yet, the ICC for the responsible Commission DG is more meaningful with a value of .17. But this is exactly the variation that I discuss and want to show in the main text, rather than modelling it away. However, there might be a concern that not accounting for this data structure affects the multivariate results. Thus model 4 in Table 4 estimates a multi-level OLS with random intercepts for drafting DGs. Unsurprisingly these intercepts absorb the effects of the factors that vary on DG level. Yet, the interpretation of the other variables in the model is by and large identical to the main model discussed in the article.

Finally, the correlation matrix (Figure 1 above) raises concerns about possible multi-collinearity of the two measures for preference heterogeneity in the Council. Thus, Table 6 below replicates the analysis for each of these two measures separately and compares it to the main model discussed in the article. Neither the conclusions on Council preference heterogeneity nor on the other variables of interest have to be substantially altered in these alternative specifications.

	Main model	Heterogeneity on EU dim only	Heterogeneity on L-R dim only	W/out Council heterogeneity
CONSTANT	2.524** (0.931)	2.826*** (0.844)	2.601*** (0.856)	3.148*** (0.623)
QMV	-0.008 (0.108)	-0.014 (0.107)	-0.009 (0.108)	-0.019 (0.107)
COUNCIL.EU.VAR	0.022 (0.103)	0.053 (0.095)		
COUNCIL.LR.VAR	0.078 (0.101)		0.086 (0.093)	
CNS	-0.337** (0.151)	-0.341** (0.151)	-0.335** (0.151)	-0.338** (0.151)
SYN	-0.900*** (0.307)	-0.920*** (0.305)	-0.898*** (0.306)	-0.918*** (0.305)
COD	-0.904*** (0.156)	-0.913*** (0.155)	-0.903*** (0.156)	-0.914*** (0.155)
ASS	-0.491 (0.645)	-0.514 (0.645)	-0.494 (0.645)	-0.528 (0.645)
PROP.ENTROPY	-0.330*** (0.062)	-0.331*** (0.062)	-0.330*** (0.062)	-0.333*** (0.062)
POLICY.AGE	0.008 (0.005)	0.007 (0.005)	0.008 (0.005)	0.007 (0.005)
DG.UNITS	0.001 (0.004)	0.001 (0.004)	0.001 (0.004)	0.001 (0.004)
COMMISSIONER.POWER	0.017 (0.100)	0.014 (0.100)	0.018 (0.100)	0.015 (0.099)
COMMISSIONER.EXPERIENCE	0.00002 (0.00004)	0.00002 (0.00004)	0.00002 (0.00004)	0.00002 (0.00004)
SECGENSIZE	0.001 (0.019)	-0.003 (0.018)	-0.0003 (0.018)	-0.008 (0.015)
ORAL	-0.186* (0.109)	-0.186* (0.109)	-0.186* (0.109)	-0.185* (0.109)
<i>Observations</i>	2,215	2,215	2,215	2,215
<i>Log likelihood</i>	-1,286.546	-1,288.089	-1,286.419	-1,288.203
<i>Akaike information criterion</i>	2,603.092	2,604.178	2,600.838	2,602.407

Notes:

***p < .01; **p < .05; *p < .1

Table 6: Alternative estimation approaches