The Czech Constitutional Court Database*

ŠTĚPÁN PAULÍK, Humboldt Universität zu Berlin, stepan.paulik.1@hu-berlin.de

The article at hand introduces a comprehensive foundational database on the Czech Constitutional Court spanning from its inception in 1993 to 2023. The database includes metadata on all decisions, full text corpus, and additional background data on judges and law clerks, filling a gap in high-quality datasets for empirical legal research in the Central and Eastern European region. As one of the first comprehensive court databases in the CEE region, it has the potential to catalyze similar research efforts and contribute to methodologically rigorous empirical legal research in a region of increasing European significance.

Contents

Contents

1 Introduction

It has been traditionally espoused that there has been a divide between the empirically oriented US legal scholarship, stemming from a different perception of the role of courts and judges, and the rest of the world (Hamann 2019, 416). Therein the judges empirically researched whether and to what extent they behave as for example political or strategic actors.(Carrubba et al. 2012; Clark and Lauderdale 2010; Epstein and Knight 1997, 2000; Lauderdale and Clark 2014; Sunstein et al. 2006; Cameron and Kornhauser 2017; Clark, Engst, and Staton 2018; Epstein, Landes, and Posner 2011; Kornhauser 1992b, 1992a; Posner 1993, 2010; Roussey and Soubeyran 2018).

*The database itself is available at a Zenodo Repository (ANONYMIZED) and replication files are available on the author's Github account (ANONYMIZED). **Current version**: April 05, 2024

In contrast to, especially in European legal systems, such as the one at hand - Czechia, judges have been perceived as "proclaimers of law" and the law handed down by them. (Hamann 2019, 417) even claims that such a view had hindered empirical legal research in Europe. The lack of empirical legal research can be partially blamed on lack of high quality data, a prerequisite for any quantitative empirical research. At least so the story goes until recently. The interest in empirical legal studies has picked up in the last years across the whole continent, including studies on plethora of topics within Germany (Arnold, Engst, and Gschwend 2023; Coupette and Fleckner 2018; Benjamin G. Engst et al. 2017; Wittig 2016), Spain and Portugal (Hanretty 2012), the UK (Hanretty 2020) or the EU institutions (Bielen et al. 2018; Brekke, Naurin, et al. 2023; Fjelstul 2023, 2019; Fjelstul, Gabel, and Carrubba 2022).

Publications of new high quality publicly accessible data have gone hand in hand with these developments. To the many released comprehensive datasets and databases in the recent years belong namely the Iuropa project's CJEU database (Brekke, Fjelstul, et al. 2023), the German Federal courts (Hamann 2019) as well as the German Federal Constitutional Court (Benjamin G. Engst, Hönnige, and Gschwend Forthcoming). The mushrooming research proves that there is demand for quality data in Europe as well. The article at hand presents a Czech Constitutional Court ("CCC") database, a comprehensive high-quality multi-user database on the CCC.

The CCC database is foundational in that it encompasses plethora of data that other researches can base their research efforts on, it has the capacity address research questions and it adheres to the tidy data principles. The database encompasses all decisions of the CCC starting from its founding in 1993 until the end of 2023, including plenty of metadata - such as information on the judge rapporteur, subject matter, or concerned legal acts, a complete text corpus, as well additional background information on judges or clerks.

To my best knowledge, the CCC database is one of the first, if not the first, comprehensive databases coming out of the Central and Eastern European ("CEE") region. The CEE region has gotten to the spotlight of the European legal research, for example, thanks to the various rule of law crises, in which the regional constitutional courts and their interplay with the CJEU have played an important role (R. Daniel Kelemen 2020; R. Daniel Kelemen and Pech 2019; Sadurski 2019 and many other articles). Despite that the CEE scholarship has so far produced very little in terms of methodologically rigorous empirical legal research output concerning the role of the judiciary, constitutional courts or judicial politics. The lack of high-quality data is undoubtedly a piece of this puzzle.

To zero in on Czechia, there have been solitary attempts go gather data in some shape or form in the CCC context (Harašta et al. 2018; Novotná and Harašta 2019), mainly thanks to the Institute of Law and Technology based in Brno, as well as isolated attempts to conduct network analysis or research employing natural language processing and alike methods (Chmel 2017; Eliášek, Kól, and Švaňa 2020; Harašta et al. 2021; Vartazaryan 2022). Unfortunately, the former group does not always adhere to the principles of of high-quality infrastructure, namely the principle of foundationality, espoused by Weinshall and Epstein (2020), p. 424, the latter group did not publish data/code at all.

Therefore, the effort to put together and to publish a high-quality database on the CCC is more than warranted, especially to enable empirical legal scholarship to flourish. As the article explains later, the database at hand can serve as an foundation for wide variety of research inquiries, for example, research into the role of gender or the role of law clerks. Moreover, the database offers a blueprint for future efforts to build akin databases. In building the database, I attempted to name

the variables and to structure the data in a transparent, replicable and comparable way so that any efforts from different courts can mimic my approach without steep costs and efforts.

The article proceeds as follow. In the section 2, I introduce the CCC, namely its compositions, its internal organization and its powers to give the reader a little bit of context. In the section 3, I introduce the CCC database. Therein, I briefly discuss its structure, its creation, and describe its variables. The section 4 then discusses the adherence of the CCC database to four principles of a high-quality dataset, including its relevance for research, as well as to the adherence of the tidy data principles. The last section 5 concludes.

2 A Brief Primer on the CCC

The CCC consists of fifteen justices[1], out of which one is the president of the CCC, two are vice presidents and twelve associate justices (following the terminology of Kosař and Vyhnánek 2020). These fifteen justices are appointed by the president of the Czech republic upon approval of the Senate, the upper chamber of the Czech two-chamber Parliament. The justices enjoy 10 years terms with the possibility of re-election; there is no limit on the times a justice can be re-elected. The three CCC functionaries are unilaterally appointed by the Czech president.

The appointment procedure is similar to how the SCOTUS justices are appointed as the procedure lays in the hands of the president of the republic and the upper chamber. The minimal requirement for a CCC nominee are 40 years of age, a clean criminal record, a finished legal education and experience in the legal field. Other than that, the nomination is left to the consideration of the President of the Republic. After a nomination, the nominee is firstly interviewed in the constitutional law committee of the Senate, which produces an unbinding recommendation for the plenary Senate hearing. The final binding decision is then made by simple majority of the Senate plenary hearing. This procedure has lead to a situation, in which there is very little variance as to the nominating background of the justices. First, there is no nominating political party akin to the US context or the Spanish context (Hanretty 2012). Second, because the court was established in 1993 and filled within roughly a year of its establishment and because the term of the Czech president is 5 years and all the 3 Presidents, who'd finished their term at the time of writing this article, have been elected twice (for ten years it total), each president has had the chance to appoint all the fifteen members of "their" CCC. Therefore, the first term of the CCC has been termed the Václav Havel, the second the Václav Klaus and the third Miloš Zeman terms of the CCC.

Regarding the competences, the CCC is a typical Kelsenian court inspired mainly by the German Federal Constitutional Court. The CCC enjoys the power of abstract constitutional review, including constitutional amendments. The abstract review procedure is initiated by political actors (for example MPs) and usually concerns political issues. Moreover, an ordinary court can initiate a concrete review procedure, if that court reaches the conclusion that a legal norm upon which its decision depends is not compatible with the constitution. Individuals can also lodge constitutional complaints before the CCC. Lastly, the CCC can also resolve separation-of-powers disputes, it can *ex ante* review international treaties, decide on impeachment of the president of the republic, and it has additional ancillary powers (for a complete overview, see Kosař and Vyhnánek 2020).

The CCC is an example of a collegial court. From the perspective of the inner, the CCC can decide in four bodies: (1) individual justices in the role of judge rapporteur, (2) 3-member chambers

[1] I will proceed with calling the CCC judges as "justices" in the SCOTUS fashion (Boatright 2018) as they technically sit on the highest court in Czechia.

(*senáty*), (3) the plenum (*plénum*), and (4) special disciplinary chamber. The 3-member chambers and the plenum play a crucial role. The plenum is composed of all justices, whereas the four 3-member chambers are composed of the associate justices. Neither the president of the CCC or her vice-presidents are permanents members of the 3-member chambers. Until 2016, the composition of the chambers was static. However, in 2016, a system of regular 2-yearly rotations was introduced, wherein the president of the chamber rotates to a different every 2 years. I am of the view that such a institutional change opens up potential for quasi-experimental research similar to the Gschwend, Sternberg, and Zittlau (2016) study utilizing judge absences within the 3-member chambers of the German federal constitutional court. In general, the plenum is responsible for the abstract review, whereas the 3-member chambers are responsible for the individual constitutional complaints.

In the chamber proceedings, decisions on admissibility must be unanimous, whereas decisions on merits need not be, therefore, a simple majority of two votes is necessary to pass a decision on merits. In the plenum, the general voting quorum is a simple majority and the plenum is quorate when there are ten justices present. The abstract review is one of the exceptions that sets the quorum higher, more specifically to 9 votes.

A judge rapporteur plays a crucial role (Chmel 2017; Hořeňovský and Chmel 2015 study the large influence of the judge rapporteurs at the CCC). Each case of the CCC gets assigned to a judge raporteur. The assignment is regulated by a case allocation plan.[2] They are tasked with drafting the opinion, about which the body then votes. The president of the CCC (in plenary cases) or the president of the chamber (in chamber cases) may re-assign a case to a different judge rapporteur if the draft opinion by the original judge rapporteur did not receive a majority of votes. Unfortunately, the CCC does not keep track of these reassignments.[3]

The act on the CCC allows for separate opinions. They can take two forms: dissenting or concurring opinions. Each justice has the right to author a separate opinion, which then gets published with the CCC decision. It follows that not every anti-majority vote implies a separate opinion, it is up to the justices to decide whether they want to attach a separate opinion with their vote. Vice-versa, not every separate opinion implies an anti-majority vote, as the justices can attach a concurring opinion. In contrast to dissenting opinion, when a judge attaches a concurring opinion, they voted with the majority but disagree with its argumentation.[4]

The CCC justices can hire their clerk teams. Each justice is required to have at least one clerk. The clerk is appointed by the president of the CCC on the nomination of the said justice. The clerk must have a clean criminal record and a finished legal degree. Other than that there are no requirements on the clerks. The term of the clerks may not exceed the term of the nominating justice. The clerks are usually tasked with drafting decisions and, in narrowly defined cases, can be instructed by the justice to decide on their behalf when an application does not meet even the minimal requirements.

It may be concluded that the CCC takes after the american model of selection of justices, with the president of the republic and the upper chamber being in the spotlight, but is also a

[2]The original term is *rozvrh práce*, which is usually translated as a *work schedule*, however, I borrow the term *case allocation plan* from Hamann (2019), p. 673

[3]More specifically, some decisions mention the reassignment. However, not all do, therefore an attempt to retrieve the information from the texts was highly unreliable. According to my internal insight, the CCC should always pass a short procedural decision when a reassignment takes place. I unsuccessfully attempted to retrieve the information with the right to information as the procedural decision is not available in electronic form and retrieving the full information would thus entail manually going through all paper files.

[4]Which makes it difficult to, for example, conduct the same point-estimation with data on dissenting behavior of justices as Hanretty (2012) has done on the Portuguese and Spanish Constitutional Courts.

typical example of a Kelsenian specialized court with concentrated constitutional review. The CCC stands out in how powerful its constitutional review is, having attracted the power to review even constitutional amendments, thus, the CCC is a powerful player in the Czech political system. While the appointment procedure of the justices may be compared to the SCOTUS, its role within the constitutional system is akin to the European constitutional courts, with the German federal constitutional court at its forefront, whose doctrinal approaches and methods, such as the test of proportionality or test of rationality, have often been adopted by the CCC, and its power to review even constitutional amendments may be comparable to the Supreme Court of Israel. The internal organization of the CCC gives room for strategic or policy considerations of its justices. Not only due to the similarities with the constitutional adjudication powerhouses but also due to its own idiosyncrasies, I believe the CCC to be a worthy object of empirical legal research as the conclusions drawn from research of the CCC may be after a careful consideration be extended beyond a mere case study on the CCC.

3 Description of the CCC database

Now that the CCC has been introduced in the previous section, I move on to describing the content and structure of the dataset in this section.

3.1 Case Inclusion

The CCC database includes all publicly available CCC decisions from its foundation until the end of 2023, that is 93826 decisions, as well as background information on its 50 justices and their 221 clerks. All the data was first web scraped from the official CCC website Nalus. Nalus is an official publicly accessible database in the form of website, on which the CCC publishes all its decisions on merits and admissibility. Nalus includes all procedural decisions decided after 1.1.2007, procedural decisions before that date may be missing without a specification as to which.[5]

The web scraping was followed by intense data cleaning and data wrangling processes. A lot of the information was transformed to a more readily form. In the last step, some information was mined from the texts or other metadata of the decisions (such as composition of the court in any given decision). The CCC database is accompanied by a comprehensive codebook, which contains detailed explanation of its structure, parts and all variables contained therein.[6]

3.2 Structure of the CCC Database

The structure of the CCC database can, on a very basic level, be divided into the master decision-level table, variable-decision-level tables and justice-level/clerk-level tables. The decision-variable level tables are linked to the master table by the *doc_id* variable and the justice/clerk-level tables are connected by the *(_)judge_id* and *clerk_id* variables. I now go over each level of the structure. For the clarity of the ensuing description of the database, fig. 1 presents a diagram of the schema of the CCC database, which can be used as a reference point.

[5] The manual to the Nalus database says "All decisions of the Constitutional Court that terminated proceedings from the beginning of the Constitutional Court's activity in 1993 until the end of 2006 (approximately 26,000 decisions) were automatically imported into the NALUS database."

[6] Available at: LINK ANONYMIZED

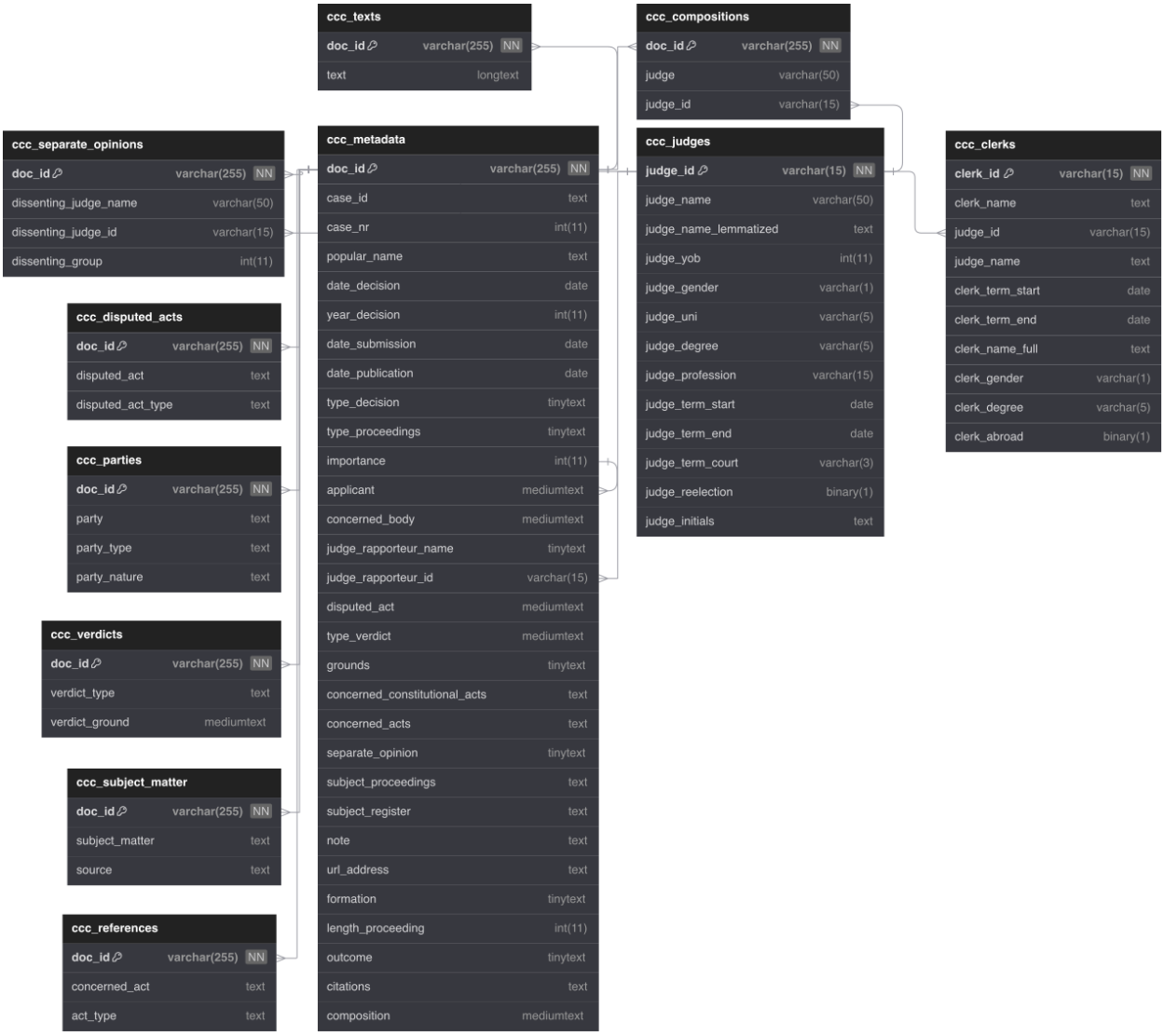

Figure 1: A diagram of the CCC database schema. The main master table is the ccc_metadata table, to which all the variable-decision level tables are linked via the doc_id unique identifier. The ccc_judges table is directly connected to the ccc_metadata via the judge_rapporteur_id variable and indirectly linked via the ccc_composition table.

Table 1: Data summary

Name	data_metadata
Number of rows	93826
Number of columns	31
_____	
Column type frequency:	
character	15
Date	3
list	11
numeric	2
_____	
Group variables	None

3.2.1 Master Table

The whole database is guided and revolves around a master *ccc_metadata* table, as seen on the fig. 1. The master table contains multiple types of information. Identifying variables contain information on each and every case that the CCC has decided. There are variables uniquely identifying the decisions (ECLI renamed to *doc_id* or *case_id* identifying the whole case, which can contain more than one decisions, and *case_nr* identifying each decision within a case) or the dates at which they were lodged and decided (*date_submission* and *date decision* respectively). Procedural variables concern whether the decision was a *usnesení* or *nález*[7](*type_decision*), what type of procedure the decision was made in, such as abstract review or constitutional complaint procedure (*type proceedings*), or on what type of grounds the decision was based (*grounds*). Background variables concern among others parties before the CCC, which are identified (a natural person, a legal person, a court etc.), the body whose decision was under review (typically which court), the type of decision being reviewed and alike. Moreover, the data on the subject proceedings (relates to the area of constitutional law) and subject register (the pertaining area of general law such as criminal-proof, civil damages, or administrative proceedings) variables are included. Such variables are especially useful for controlling for specific features of cases that may have confounding potential. Lastly, miscellaneous variables contain for example an URL address to the decision in the Nalus database (*url_address*) or a note, which typically contains a link to the press release (*note*).

Variable type: character

[7]roughly speaking *nález* is a decision on merits, whereas *usnesení* is a decision on admissibility or a procedural decision

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
doc_id	0	1.00	25	34	0	93826	0
case_id	43	1.00	9	14	0	85596	0
case_nr	27002	0.71	1	1	0	10	0
popular_name	86777	0.08	9	256	0	6395	0
type_decision	0	1.00	5	16	0	3	0
type_proceedings	0	1.00	23	94	0	12	0
importance	0	1.00	1	1	0	4	0
judge_rapporteur_name	0	1.00	9	19	0	48	0
judge_rapporteur_id	0	1.00	3	4	0	48	0
grounds	0	1.00	6	13	0	3	0
note	91848	0.02	3	506	0	1791	0
url_address	0	1.00	56	62	0	93824	0
formation	0	1.00	6	14	0	5	0
outcome	0	1.00	7	8	0	2	0
presence_dissent	0	1.00	4	10	0	2	0

Variable type: Date

skim_variable	n_missing	complete_rate	min	max	median	n_unique
date_decision	0	1.00	1993-09-14	2023-12-06	2012-07-17	6949
date_publication	88739	0.05	1993-12-21	2023-12-06	2012-03-21	2472
date_submission	0	1.00	1992-05-07	2023-11-14	2012-02-28	8856

Variable type: list

skim_variable	n_missing	complete_rate	n_unique	min_length	max_length
applicant	8	1.00	1219	1	286
concerned_body	32373	0.65	7364	1	15
disputed_act	4566	0.95	3140	1	97
type_verdict	9	1.00	638	1	6
concerned_constitutional_acts	41609	0.56	8078	1	8
concerned_acts	16053	0.83	54287	1	47
separate_opinion	93292	0.01	229	1	9
subject_proceedings	14047	0.85	11251	1	28
subject_register	19688	0.79	40834	1	28
citations	0	1.00	52190	0	220
composition	0	1.00	1026	2	2

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	his
year_decision	0	1	2011.16	7.58	1993	2005	2012	2018	2023	
length_proceeding	0	1	157.64	200.34	-322	38	83	195	3809	

3.2.2 Variable-Decision-level Tables

Some of the aforementioned variables may contain more observations per decision. In effect, to keep all information in one table would entail breaking the tidy data principle that each row contains one observation as the observation in the *ccc_table* is one decision. To resolve that the variables in the master table, which potentially contained more than one row of information per observation

(decision), are stored as a nested list in the master table and then unnested into separate tables, in which the observation is a variable-decision level. The variable-decision tables are connected to the main table by the *doc_id* key. The unnested tables include, to name a few, *ccc_references* (contains references to CCC caselaw found in the texts of the decisions), *ccc_subject_matter* (contains subject matters of a decision), *ccc_parties* (contains information on the parties, both the applicant and the concerned body, before the CCC), or *ccc_compositions* (contains the bench composition with a link to the *ccc_judges* table via the *judge_id* variable).

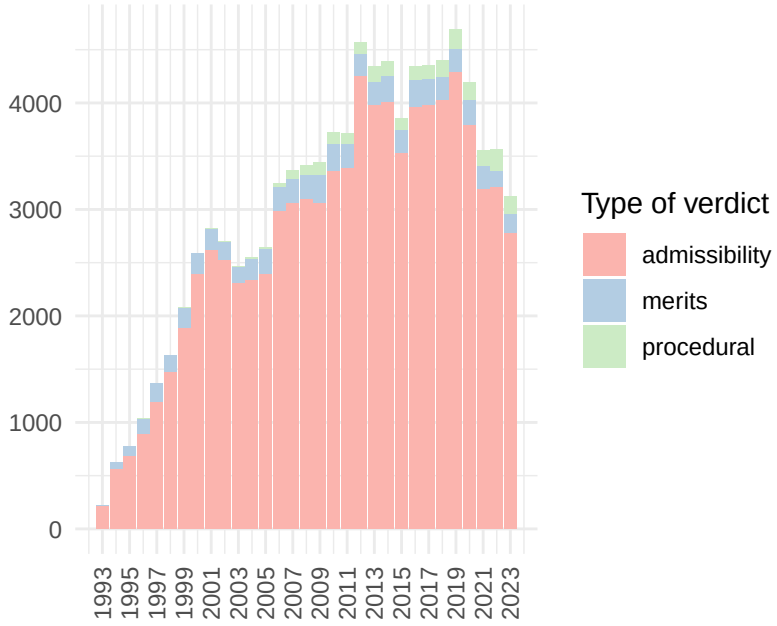

Figure 2: We can see the developing caseload overtime with the colors refering to the type of the decision. It may be observed that while the caseload of the CCC has grown, the proportions between the admissibility and merits decisions have not changed, despite the efforts of the CCC to administer more decisions more quickly with the decisions on admissibility (especially the ill-foundness).

Lastly, the *ccc_texts* contains the full text of the decision, which unlocks plethora of potential research utilizing quantitative text analysis or various machine learning endeavor. The texts have underwent a very little pre-processing as the texts in the Nalus database are in a good state. Most of the html tags have been removed apart from paragraph tag (mostly in the form of one or more $\backslash n$ tags). As the decisions have no clear structure, the texts have been kept as a whole[8] and any researcher intending to run any NLP task can simply split them up into a unit they deem fit (tokens, sentences, paragraphs etc.). A number of variables have already been mined from these texts. To name two, the compositions of sitting benchs have been mined using various regex variations of the justices' names and dissenting opinions as well as their relationships to each other (whether more judges signed one dissenting opinion or whether they dissented separately) have been mined from

[8]I have been able to partition the text into an implicit structure such as the heading, procedure history, parties' arguments, or court arguments using a supervised machine learning algorithm. However, the decisions do not lend themselves to a simple and reliable regex partitioning. The only clear general rule is that the first paragraph most of the time contains the heading of the decision, in which the composition as well as the parties of the case can be located.

Table 2: Data summary

Name	data_judges
Number of rows	55
Number of columns	14
_____	
Column type frequency:	
character	7
Date	2
factor	3
numeric	2
_____	
Group variables	None

the texts.

3.3 Justice-level and Clerk-level Tables

Justice-level and clerk-level variables contain information on the individual justices and clerks respectively. Most of the information was collected partly automatically and partly manually from the official profiles of current justices, former justices, and clerks at the CCC website, as well as Wikipedia profile pages of the justices. The CCC database includes information on the terms of the justices, their age and gender, their alma mater, highest reached degree, as titles play an especially important "ceremonial" role in the Czech legal environment, their professional background before they became a judge[9], or information on whether the justice ran for a reelection as the Czech Constitution and the act on the CCC allow for reelection of justices after their 10 year term runs out. The table 2 shows an overview of the variables included in the *ccc_judges* table.

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
judge_id	0	1	3	4	0	50	0
judge_name	0	1	9	24	0	50	0
judge_name_lemmatized	0	1	29	73	0	50	0
judge_profession	0	1	5	10	0	4	0
judge_term_court	0	1	3	3	0	4	0
judge_term_president	0	1	10	12	0	4	0
judge_initials	0	1	7	7	0	41	0

Variable type: Date

skim_variable	n_missing	complete_rate	min	max	median	n_unique
judge_term_start	0	1	1993-07-15	2023-12-19	2004-06-16	33
judge_term_end	0	1	1999-12-08	2033-12-19	2014-06-16	39

Variable type: factor

[9]Coded as the last profession before they started their CCC term.

Table 3: Data summary

Name	data_clerks
Number of rows	376
Number of columns	10
_____	
Column type frequency:	
character	7
Date	2
numeric	1
_____	
Group variables	None

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
judge_gender	0	1	FALSE	2	M: 41, F: 14
judge_uni	0	1	FALSE	3	CUN: 37, MUN: 15, KOM: 3
judge_degree	0	1	TRUE	5	jud: 29, pro: 16, phd: 6, doc: 3

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
judge_yob	0	1	1950.84	13.13	1924	1945	1951	1957.5	1978	
judge_reelection	0	1	0.22	0.42	0	0	0	0.0	1	

Secondly, the *ccc_clerks* table includes information on all 221 clerks that have served in the CCC's history. The information on the clerks contains under which judge they deserved, what was their term, what is their gender, education, or whether they studied abroad. To this end, I believe inclusion of such an information makes the dataset quite unique and opens up a lot of avenues for research. The table 3 shows an overview of the variables included in the *ccc_clerks* table.

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
judge_id	0	1	3	4	0	50	0
judge_name	0	1	9	24	0	50	0
clerk_id	0	1	3	5	0	221	0
clerk_name	0	1	8	21	0	221	0
clerk_name_full	0	1	11	41	0	236	0
clerk_gender	0	1	1	1	0	2	0
clerk_degree	1	1	3	4	0	5	0

Variable type: Date

skim_variable	n_missing	complete_rate	min	max	median	n_unique
clerk_term_start	0	1	1993-08-02	2024-01-21	2013-05-21	250
clerk_term_end	0	1	1993-11-30	2034-01-21	2014-10-30	208

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
clerk_abroad	0	1	0.06	0.24	0	0	0	0	1	

4 Principles Guiding the CCC database

The Czech Constitutional Court database is a "multi-user dataset"[10] created in a principled manner. Epstein et al. (2014), p. 14 define a multi-user dataset as a dataset created with the purpose of that "[r]ather than collect data to answer particular research questions [...] the idea is to amass a dataset so rich in content that multiple users, even those with distinct projects, can draw on it."

Accordingly, the CCC database upholds the principles of a high-quality datasets espoused by Weinshall and Epstein (2020), p. 424, namely that the database is (1) capable of addressing real-world problems, (2) accessible, (3) reproducible and reliable, and (4) foundational.[11] The data structure also follows the principles of tidy data. According to Wickham (2014), tidy data are data with such a tabular structure, i.e. data with a column and row structure, that stick to the following principles

(1) every column is a variable,
(2) every row is an observation,
(3) every cell is a single value.

I now go over and discuss the Weinshall and Epstein (2020) principles one by one and describe them in detail.

4.1 Capacity to Address Real-World Problems

In the words of Weinshall and Epstein (2020), "By definition, data infrastructure should promote innovation, inventions, and insights. Although no product can guarantee these ends, infrastructure aimed at solving (or developing implications for) real-world problems increases the odds of success." With the database at hand I hope to enable data- and evidence- based research on the CCC, which has so far not been an object of thorough empirical legal research. I now present two examples that corroborate the capacity of the CCC database to address real-world problems and research concerns.

4.1.1 Clerks

The first brief example concerns clerks of justices. Kosař and Vyhnánek (2020) argue that the clerks at the CCC play an especially vital and unappreciated role: "The initial idea of the legislature was to grant each Justice one law clerk who would take administrative burdens unrelated to substantive decision-making off the Justices' shoulders. Yet the reality is different. First, due to the growing caseload, the number of law clerks per Justice increased gradually; today, each Justice has three law clerks. Moreover, law clerks de facto prepare drafts of most CCC judgments and decisions, and the real administrative burden has been 'outsourced' to secretaries of the cabinets." The difficulty of studying the role of clerks was highlighted in the Clark, Engst, and Staton (2018) study on the effects of leisure on judicial performance. In the existing studies on clerks, their influence on the final decision as "an information source" (Kromphardt 2015) or the influence of their gender on their career choice to become a clerk have been researched (Badas and Stauffer 2023).

Badas and Stauffer (2023) have discovered that women are in general under-represented among law clerks and that one of the reasons behind under-representation of female clerks is that "female

[10] I refer to it as a datase

[11] I decided to skip the principle of sustainability as I do not see the benefit of discussing it separately, most of its issues are answered in the other sections

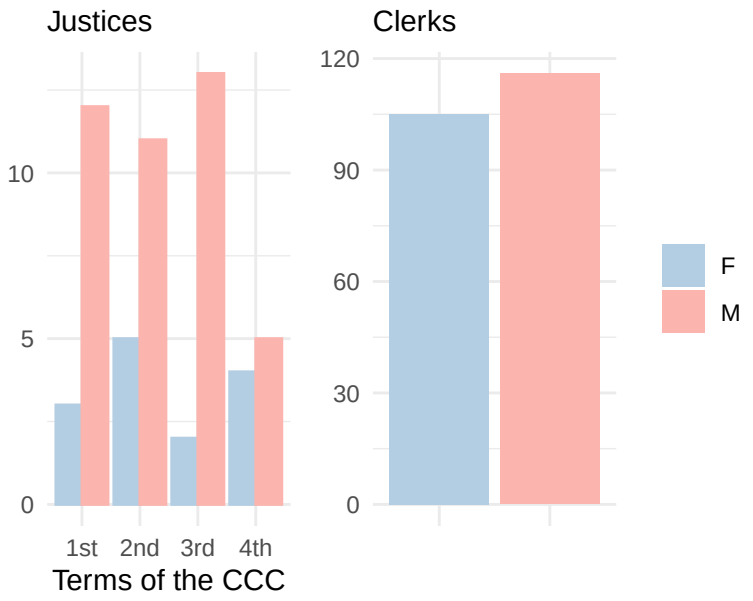

Figure 3: Comparison of proportions of genders among justices (on the left) and among their clerks (on the right). The X axis of the left bar chart signifies the terms of the CCC, which roughly correspond to decades.

Table 4: A table showing the representation of women among CCC justices, their clerks, and law graduates in Czechia.

	% of Women
Graduates	58.6
Clerks	47.5
Judges	26.0

law students may have lower levels of ambition compared to men. (…) Examining potential sources of this difference, we find that while women view themselves to be just as qualified for these positions as men, men are more willing to apply with lower feelings of qualification. Likewise, while women and men report similar levels of encouragement, more encouragement is required before women express ambition to hold these posts." In two studies on the gender equality in the Czech judiciary, Havelková (2017) and Urbániková, Havelková, and Kosař (2023) have revealed that at first glance the representation of women is rather high. However, structurally, the distribution is vertically unequal: female judges dwell on the first-instance court and take care of the run-of-the-mill decision-making, whereas male judges are over-represented in the upper echelons of the judiciary, which exert higher influence over doctrinal development, and occupy the judicial functionary positions, which mainly take care of court administration. That raises two questions: (1) is there a discrepancy between the proportion of women among graduates and law clerks and (2) is the representation of women similarly vertically unequally distributed as that of the justices?

The CCC database contains the information on all clerks that have served on the CCC, including their gender, education, beginning and end of mandates. For the purpose of showing the capability of solving real-world problems, I briefly test the theoretical expectations generated above

by presenting descriptive statistics, namely whether there is a discrepancy between the representation of women among clerks, graduates and justices. To do so, I collected data by Eurostat on the gender distribution among law graduates in Czechia between 2015-2021 (Eurostat 2024) and appended it to the CCC database data. Table 4 compares the proportion of women among CCC justices, clerks and the graduates. The discrepancy between clerks and graduates less so but it is still rather pronounced. Tab. 3 confirms the discrepancy at the CCC.

It can further inquired whether male or female justice have different preferences regarding their clerks. Surprisingly, given the theoretical expectations, the judges at the CCC seem to prefer clerks equally whereas female judges seem to hire more male clerks.

4.1.2 Dissenting behavior of justices

The research on judicial coalitions at the CCC has revealed that the third period of CCC between 2013-2023 is rather polarized and that there are two big coalitions of judges that clash against each other (Chmel 2021; Smekal et al. 2021; Vartazaryan 2022). The articles rely primarily on network analysis of the dissenting opinions in the plenary proceedings and make inferential conclusions based on a rather superficial descriptive analysis. I predict that should the relationship from the plenum indeed exist, they should also carry over to the 3-member chamber hearings. The hypothesis is that chambers composed of judges from both coalitions will be more likely to show disagreement in the form of dissenting opinion; i.e. whether having a 3-member chamber composed of members of both judicial coalitions increases judges' likelihood of a dissent. If this shows to be true, it would provide further evidence to the two coalition theory of the CCC (Chmel 2021; Vartazaryan 2022; Smekal et al. 2021).

I very concisely tested whether the presumable existence of the coalitions carries over to and has any effect on the dissenting behavior of judges in the chambers building a very simple logistic regression. I manually annotated which justices of the 3rd term were from which coalition, thus further proving that the CCC database adheres to the below discussed principle of foundationality as it can be additionally built upon.

The model predicts the likelihood of a judge attaching a separate opinion occurring in the 3-member chamber proceedings depending on whether the chamber is fully composed of members from the 1st or 2nd coalition and on the composition being a mix of one minority justice from one coalition and the remaining two from the other. The dependent variable is thus a dummy variable containing the information whether a judge attached a separate opinion or not and the independent variable is a categorical variable of whether the chamber was composed of members of one coalitoin or whether the chamber was mixed.

The very simple model reveals that there does seem to be a negative association between the likelihood of a judge to dissent and the chamber being composed fully of members from one or the other coalition. Such a conclusion is in line with the Czech legal scholarship and warrants further more in-depth inquiry. The example moreover proves (as will be discussed bellow in the section 4.1.2) that the database is foundational in the sense that for the aforementioned model, the CCC database was used as a basis with majority of the data stemming directly out of it (such as the information on the individual decisions, on the compositions etc.) and the remaining information being added and adjusted according to the specific research goal, in this case verifying the theory on the coalitions posited by Czech legal scholars.

Table 5: Results from the Logit Model

	(1)
(Intercept)	−3.783***
	(0.125)
coalitionfull_coal	−1.214***
	(0.316)
Num.Obs.	4710
AIC	771.3
BIC	784.2
Log.Lik.	−383.645
RMSE	0.13

4.1.3 Natural language processing

The last potential usage of the CCC database that immediately springs into mind is the application of various natural language processing methods. For example, one could replicate the research on vagueness of language of the CCC (Sternberg 2019) or measure the readibility of the CCC decisions (Crossley, Skalicky, and Dascalu 2019; Fix and Fairbanks 2020) and link those measures to interesting research question. For example, do the better readable CCC decisions get cited by the CCC more than the less readable (Fix and Fairbanks 2020; Crossley, Skalicky, and Dascalu 2019)? Does the CCC use vague language in certain areas/for certain reasons more than for others (Sternberg 2019)?

Practically every dataset introduction article contains a disclaimer that the goal of presenting these simplified models is not to draw any inference but rather to show the potential of using the dataset "to develop real-world implications and contribute to public and academic discourse on pressing legal-political issues". (Weinshall and Epstein 2020, 427) This section has attempted to do exactly that.

4.2 Accessibility

The principle of accessibility demands that "in the creation of high-quality infrastructure is that members of the community should be able to access it with no barriers to entry or use." (Weinshall and Epstein 2020, 427)

As I have shown in the introduction with specific examples, not all research is reproducible and not all data are made available. That goes against the principle of accessibility. Weinshall and Epstein refer to studies, according to which majority of psychological research data stays under embargo or never gets released at all (Houtkoop et al. 2018) or that only minority of papers published in journals requiring a data availability statement actually do publish their data (Federer et al. 2018). There is no need to prove why such a practice is wrong, why it hinders reproducibility of science.

Following the principle of accessibility, the CCC database is freely and publicly available in full, with the handbook as well as this article attached to it. **That data are downloadable at the**

Zenodo Repository.[12] The data are published out of my own accord, the publication is not funded by any grant or national science foundation.

4.3 Reliability and Reproducibility

Weinshall and Epstein (2020) define the principles as follows: "[r]eproducibility means that users and developers alike must understand how to duplicate the data housed in the infrastructure. Reliability is related: it is the extent to which encoded data can be replicated, producing the same value using the same standard for the same subject at the same time, regardless of who or what is doing the replicating" The heart of the matter of reliability and reproducibility is internal consistency of the dataset, not necessarily its external validity.

The data must have been reliably generated. In my case, I did not narrow down the selection of cases; all cases of the CCC that have been made publicly available from its history have been web scraped from its website, including all the available information as well as the texts of the decisions. Reproducibility also demands that anyone with sufficient skill should be able to reproduce the database on their own based on the provided information. All the code has been made available on Github, the code is written in a clean manner and is well commented.

Bound to both principles is the issue of coding the variables. To this end, the amount of human input has been minimized. Vast majority of the information provided has either been directly (or with minimal input) collected from the CCC website or it has been automated to the maximal possible extent (including the full information on the clerks). Only the biographic information on judges has been imputed via human input using the official profiles of justices at the CCC website as well as Wikipedia as sources. The rest is product of the published code.

There are two potential sources of unreliability. One is coming from the Nalus database, the other is coming from the data mining process that was to great extent automated. Regarding the former source of unreliability, it is hard to estimate the unreliability. According to my internal insight, some of the information (such as the subject matter) is inserted manually, mainly by the justices' clerks and the court's analytic unit. It is easy to imagine that it is hard to maintain consistency across time and between different chambers and justices. To verify the validity and reliability of the data mining process, I check two variables that have been mined from the texts of the decisions: the compositions and the information about separate opinion.

The compositions have been retrieved using regex search of the first couple of paragraphs of the decisions using lemmatized names of the justices. After some trial and error, a couple of patterns emerged. Many chamber decisions contained four names. The issue was that those decisions were decisions on the independence on one of the justices deciding a case, which always occurred last in the decision. Therefore, in the case of chamber decisions with 4 justices, the last name found in the text has been removed as it always came up last. chamber decisions with either 1 or 3 found names are deemed correct as simple cases can be expedited by one justice, whereas chamber decisions with 0 or 2 found names are deemed incorrect.

Plenum decisions are harder to verify as it is impossible to determine the correct number of justices as a benchmark. To name an example as to why, at one point of CCC's history, there was as few as 10 justices sitting on it when the president Václav Klaus hesitated with nominating justices after a feud with the Senate and the number fluctuated within a short period of time.

In any case, it can be determined when a number of justices found in the text of a decision is

[12] The link is currently ANONYMIZED

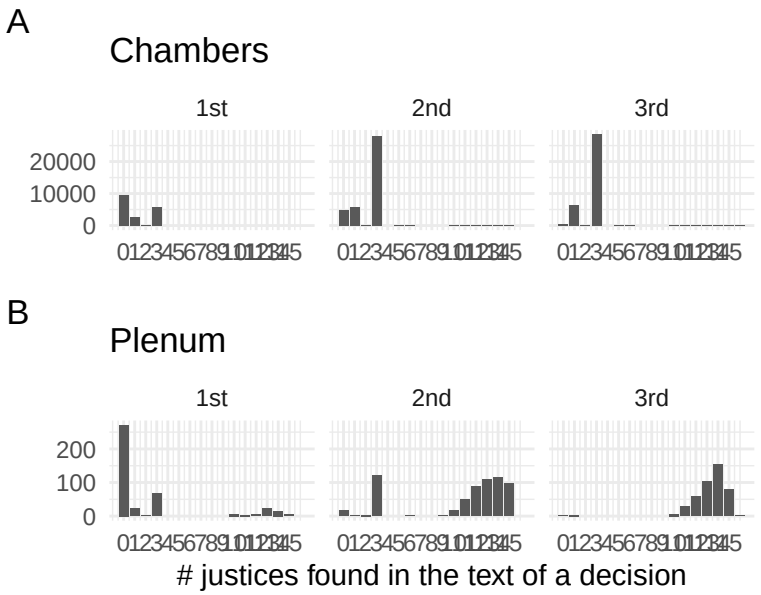

Figure 4: The number of justices retrieved from the text of each decision faceted by the formation of the court and the term. The data are uncorrected.

undoubtedly faulty. The clear mistake is when there is either 0 or 2 justices found in the text. The 0 name found is typical for the first term of the CCC, in which the composition of the bench was not always enumerated in the text of the decision. As such, the is nearly impossible to retrieve the information using the case allocation plan and the identification of the chamber as any justice could have been sidelined due to illness, lack of independence, or replaced by one of the functionaries, none of which is captured in the original database. The error with 2 names being found typically included a judge rapporteur decision with another former or future justice being included in different role such as the legal representative of one of the parties or more rarely a hard to generalize typo in one of the names of the justices. To prevent this type of error inasmuch possible, the regex search was limited only to the first two paragraphs of a decision.

Fig. 4 that the accuracy develops over time. The first term is rather unreliable, especially the plenum decisions barely ever contain the names of at least 10 justices. The second term is rather reliable and the third term is practically completely reliable. Table 6 shows the ratio of correctly to incorrectly retrieved compositions.

The accuracy of data extraction of the information on separate opinions underwent verification too. While the information whether a judge attached a separate opinion or not is generated by the Nalus database and, therefore, is presumed to be accurate, the information on whether the judge dissented alone or in a group with others was retrieved using regex search.

Therefore, I can conclude that the CCC database is reliable to the extent that the data generating process is reliable and consistent. Insofar the decisions of the first decade of the CCC were plagued with a degree of inconsistency and missing information, so is the database. To some extent I attempted to capture and correct the errors, to the remaining extent the database does not attempt to conceal the incomplete data. Partly because the database easily reproducible, that is, at least with enough time to run the whole web scraping process and data cleaning process, anyone can verify the data themselves.

Table 6: A table showing the absolute number of decisions as well as the proportion of incorrectly extracted compositions from the texts of the CCC decisions. The numbers clearly show that while the first term is rather inaccurate, the consistency with whitch the CCC includes information in its decisions greatly increased over time to the point that the 3rd term is practically completely accurate.

Formation	Term	Incorrect	Correct	Proportion of Incorrect
Panel	1st	9514	8385	53.15%
Plenum	1st	273	146	65.16%
Panel	2nd	4979	33810	12.84%
Plenum	2nd	19	600	3.07%
Panel	3rd	573	35092	1.61%
Plenum	3rd	1	432	0.23%

Table 7: A table showing the number of correctly and incorrectly extracted information about the dissenting group from the texts of the CCC decisions. The numbers clearly show that while the first term is rather inaccurate, the consistency with whitch the CCC includes information in its decisions greatly increased over time to the point that the 3rd term is practically completely accurate.

Formation	Term	Incorrect	Correct	Proportion of Incorrect
Panel	1st	22	4	15.38%
Plenum	1st	108	60	35.71%
Panel	2nd	92	1	1.08%
Plenum	2nd	369	9	2.38%
Panel	3rd	125	0	0.00%
Plenum	3rd	360	4	1.10%

4.4 Foundational

The principle that a dataset be foundational requires that it should serve "as a foundation upon which researchers can build by adding content, backdating, updating, or otherwise adapting it to their own needs; it should not be the be-all, end-all." In other words, the principle promotes a generally usable data over one-off solutions to particular research questions. The CCC database is foundational. The dataset includes comprehensive background data on each and every case, bibliographic data on the justices, quite unique data on the clerks, as well as a full text corpus of all the decisions. As I have shown in the example of coalitions, to answer a real research question raised by Czech legal scholarship, the database, used as a foundation, was supplemented with additional variable. The coalitions variable was again based on the CCC database's information on compositions and the manually annotated information on which justice belonged to which coalition. The CCC database can, to name a few examples, serve researches interested in judicial behavior, in institutional studies, or anyone wanting to leverage natural language processing on large amount of text data.

5 Conclusion

The present article introduced a database on the Czech Constitutional Court while aiming at bridging the gap between the traditionally doctrine oriented European scholarship and the more empirically methodologically more rigorous US scholarship. The database, in my view, enables empirical research in the CEE region that has been lacking in the past on methodologically rigorous empirical research. The database unlocks research on the decision-making of judges and judicial politics, such as their dissenting behavior, strategic acting, or the influence of their clerk teams, on the institutional set up of the CCC, such as the introduction of rotations or the various ways to expedite the CCC caseload, and lastly on the texts of the decisions themselves, for example studying various linguistic features of the decisions such as readability or vagueness. Therefore, the present article makes a valuable contribution to the (European) empirical legal research scholarship. To this end, the database contains rich information on the background of the cases, background of the justices, their clerk teams as well as the complete text corpus of all decisions until the end of 2023. The database adheres to the principles of tidy data as well as to the principles of accessibility, reliability, foundationality, and capacity to address real-world problems.

Literature

Arnold, Christian, Benjamin G. Engst, and Thomas Gschwend. 2023. "Scaling Court Decisions with Citation Networks." *Journal of Law and Courts* 11 (1): 25–44. https://doi.org/10.1086/717420.

Badas, Alex, and Katelyn E. Stauffer. 2023. "Gender and Ambition Among Potential Law Clerks." *Journal of Law and Courts* 11 (1): 116–40. https://doi.org/10.1017/jlc.2022.8.

Bielen, Samantha, Ludo Peeters, Wim Marneffe, and Lode Vereeck. 2018. "Backlogs and Litigation Rates: Testing Congestion Equilibrium Across European Judiciaries." *International Review of Law and Economics* 53 (March): 9–22. https://doi.org/10.1016/j.irle.2017.09.002.

Boatright, Jason. 2018. "The History, Meaning, and Use of the Words Justice and Judge." *St. Mary's Law Journal* 49 (4): 727–48. https://commons.stmarytx.edu/thestmaryslawjournal/vol49/iss4/1.

Brekke, Stein Arne, Joshua C. Fjelstul, Silje Synnøve Lyder Hermansen, and Daniel Naurin. 2023. "The CJEU Database Platform: Decisions and Decision-Makers." *Journal of Law and Courts*, January, 1–22. https://doi.org/10.1017/jlc.2022.3.

Brekke, Stein Arne, Daniel Naurin, Urška Šadl, and Lucía López-Zurita. 2023. "That's an Order! How the Quest for Efficiency Is Transforming Judicial Cooperation in Europe." *JCMS: Journal of Common Market Studies* 61 (1): 58–75. https://doi.org/10.1111/jcms.13346.

Cameron, Charles M., and Lewis A. Kornhauser. 2017. "Chapter 3: What Do Judges Want? How to Model Judicial Preferences." SSRN Scholarly Paper. Rochester, NY. June 2, 2017. https://doi.org/10.2139/ssrn.2979419.

Carrubba, Cliff, Barry Friedman, Andrew D. Martin, and Georg Vanberg. 2012. "Who Controls the Content of Supreme Court Opinions?" *American Journal of Political Science* 56 (2): 400–412. https://doi.org/10.1111/j.1540-5907.2011.00557.x.

Chmel, Jan. 2017. "Zpravodajové a Senáty: Vliv Složení Senátu Na Rozhodování Ústavního Soudu České Republiky o Ústavních Stížnostech." *Časopis Pro Právní Vědu a Praxi* 25 (4): 739. https://doi.org/10.5817/CPVP2017-4-9.

———. 2021. *Co Ovlivňuje Ústavní Soud a Jeho Soudce? /*. Vydání první. Teoretik (Leges). Leges,.

Clark, Tom S., Benjamin G. Engst, and Jeffrey K. Staton. 2018. "Estimating the Effect of Leisure on Judicial Performance." *The Journal of Legal Studies* 47 (2): 349–90. https://doi.org/10.1086/699150.

Clark, Tom S., and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54 (4): 871–90. https://doi.org/10.1111/j.1540-5907.2010.00470.x.

Coupette, Corinna, and Andreas M. Fleckner. 2018. "Quantitative Rechtswissenschaft." *Juristen-Zeitung (JZ)* 73 (8): 379–89. https://doi.org/10.1628/jz-2018-0020.

Crossley, Scott A., Stephen Skalicky, and Mihai Dascalu. 2019. "Moving Beyond Classic Readability Formulas: New Methods and New Models." *Journal of Research in Reading* 42 (3-4): 541–61. https://doi.org/10.1111/1467-9817.12283.

Eliášek, Martin, Jakub Kól, and Miloš Švaňa. 2020. "Automatická Klasifikace Významových Celků v Judikatuře." *Revue Pro Právo a Technologie* 11 (21): 3–20. https://doi.org/10.5817/RPT2020-1-1.

Engst, Benjamin G., Thomas Gschwend, Nils Schaks, Sebastian Sternberg, and Caroline Wittig. 2017. "Zum Einfluss Der Parteinähe Auf Das Abstimmungsverhalten Der Bundesverfassungsrichter – Eine Quantitative Untersuchung." *JuristenZeitung* 72 (17): 816–26. https://www.jstor.org/stable/44867374.

Engst, Benjamin G, Christoph Hönnige, and Thomas Gschwend. Forthcoming. "The Constitutional Court Database." *Working Paper*, Forthcoming, 39.

Epstein, Lee, and Jack Knight. 1997. *The Choices Justices Make*. SAGE. https://books.google.com?id=hSnom2h2_zUC.

———. 2000. "Toward a Strategic Revolution in Judicial Politics: A Look Back, A Look Ahead." *Political Research Quarterly* 53 (3): 625–61. https://doi.org/10.1177/106591290005300309.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2011. "Why (And When) Judges Dissent: A Theoretical And Empirical Analysis." *Journal of Legal Analysis* 3 (1): 101–37. https://doi.org/10.1093/jla/3.1.101.

Epstein, Lee, Andrew D. Martin, Lee Epstein, and Andrew D. Martin. 2014. *An Introduction to Empirical Legal Research*. Oxford, New York: Oxford University Press.

Eurostat. 2024. "Distribution of Male and Female Graduates in Different Fields of Education, by Education Level and Programme Orientation." https://doi.org/10.2908/EDUC_UOE_GRAD10.

Federer, Lisa M., Christopher W. Belter, Douglas J. Joubert, Alicia Livinski, Ya-Ling Lu, Lissa N. Snyders, and Holly Thompson. 2018. "Data Sharing in PLOS ONE: An Analysis of Data Availability Statements." *PLOS ONE* 13 (5): e0194768. https://doi.org/10.1371/journal.pone.0194768.

Fix, Michael P., and Bailey R. Fairbanks. 2020. "The Effect of Opinion Readability on the Impact of U.S. Supreme Court Precedents in State High Courts." *Social Science Quarterly* 101 (2): 811–24. https://doi.org/10.1111/ssqu.12752.

Fjelstul, Joshua C. 2019. "The Evolution of European Union Law: A New Data Set on the *Acquis Communautaire*." *European Union Politics* 20 (4): 670–91. https://doi.org/10.1177/1465116519842947.

———. 2023. "How the Chamber System at the CJEU Undermines the Consistency of the Court's Application of EU Law." *Journal of Law and Courts*, 717422. https://doi.org/10.1086/717422.

Fjelstul, Joshua C., Matthew Gabel, and Clifford J. Carrubba. 2022. "The Timely Administration of Justice: Using Computational Simulations to Evaluate Institutional Reforms at the CJEU." *Journal of European Public Policy*, August, 1–22. https://doi.org/10.1080/13501763.2022.2113115.

Gschwend, Thomas, Sebastian Sternberg, and Steffen Zittlau. 2016. "Are Judges Political Animals After All? Quasi-Experimental Evidence from the German Federal Constitutional Court." SSRN Scholarly Paper. Rochester, NY. February 26, 2016. https://doi.org/10.2139/ssrn.2738512.

Hamann, Hanjo. 2019. "The German Federal Courts Dataset 1950–2019: From Paper Archives to Linked Open Data." *Journal of Empirical Legal Studies* 16 (3): 671–88. https://doi.org/10.1111/jels.12230.

Hanretty, Chris. 2012. "Dissent in Iberia: The Ideal Points of Justices on the Spanish and Portuguese Constitutional Tribunals." *European Journal of Political Research* 51 (5): 671–92. https://doi.org/10.1111/j.1475-6765.2012.02056.x.

———. 2020. *A Court of Specialists: Judicial Behavior on the UK Supreme Court*. Oxford University Press. https://doi.org/10.1093/oso/9780197509234.001.0001.

Harašta, Jakub, Jaromír Šavelka, František Kasl, Adéla Kotková, Pavel Loutocký, Jakub Míšek, Daniela Procházková, et al. 2018. "Annotated Corpus of Czech Case Law for Reference Recognition Tasks." In *Text, Speech, and Dialogue*, edited by Petr Sojka, Aleš Horák, Ivan Kopeček, and Karel Pala, 11107:239–50. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-00794-2_26.

Harašta, Jakub, Terezie Smejkalová, Jaromír Šavelka, and Radim Polčák. 2021. *Citační analýza judikatury*. Vydání první. Právní monografie. Praha: Wolters Kluwer.

Havelková, Barbara. 2017. *Gender Equality in Law: Uncovering the Legacies of Czech State Socialism*. Oxford ; Portland, Oregon: Hart Publishing.

Hořeňovský, Jan, and Jan Chmel. 2015. "The Process of making the Constitutional Court Judgements." *Časopis pro právní vědu a praxi* 23 (3): 302–11. https://www.ceeol.com/search/article-detail?id=780150.

Houtkoop, Bobby Lee, Chris Chambers, Malcolm Macleod, Dorothy V. M. Bishop, Thomas E. Nichols, and Eric-Jan Wagenmakers. 2018. "Data Sharing in Psychology: A Survey on Barriers and Preconditions." *Advances in Methods and Practices in Psychological Science* 1 (1): 70–85. https://doi.org/10.1177/2515245917751886.

Kelemen, R Daniel, and Laurent Pech. 2019. "The Uses and Abuses of Constitutional Pluralism: Undermining the Rule of Law in the Name of Constitutional Identity in Hungary and Poland." *Cambridge Yearbook of European Legal Studies* 21 (December): 59–74. https://doi.org/10.1017/cel.2019.11.

Kelemen, R. Daniel. 2020. "The European Union's Authoritarian Equilibrium." *Journal of European Public Policy* 27 (3): 481–99. https://doi.org/10.1080/13501763.2020.1712455.

Kornhauser, Lewis A. 1992a. "Modeling Collegial Courts. II. Legal Doctrine." *Journal of Law, Economics and Organization* 8: 441. https://heinonline.org/HOL/Page?handle=hein.journals/jleo8&id=449&div=&collection=.

———. 1992b. "Modeling Collegial Courts I: Path-dependence." *International Review of Law and Economics* 12 (2): 169–85. https://doi.org/10.1016/0144-8188(92)90034-O.

Kosař, David, and Ladislav Vyhnánek. 2020. "The Constitutional Court of Czechia." In *The Max Planck Handbooks in European Public Law: Volume III: Constitutional Adjudication: Institutions*, edited by Armin von Bogdandy, Peter Huber, and Christoph Grabenwarter, 0. Oxford University Press. https://doi.org/10.1093/oso/9780198726418.003.0004.

Kromphardt, Christopher D. 2015. "US Supreme Court Law Clerks as Information Sources." *Journal of Law and Courts* 3 (2): 277–304. https://doi.org/10.1086/682136.

Lauderdale, Benjamin E., and Tom S. Clark. 2014. "Scaling Politically Meaningful Dimensions Using Texts and Votes: SCALING POLITICALLY MEANINGFUL DIMENSIONS." *American Journal of Political Science* 58 (3): 754–71. https://doi.org/10.1111/ajps.12085.

Novotná, Tereza, and Jakub Harašta. 2019. "The Czech Court Decisions Corpus (CzCDC): Availability as the First Step." October 21, 2019. http://arxiv.org/abs/1910.09513.

Posner, Richard A. 1993. *What Do Judges and Justices Maximize?: (The Same Thing Everyone Else Does)*. Law School, University of Chicago. https://books.google.com?id=ciFUHQAACAAJ.

———. 2010. *How Judges Think*. Harvard University Press. https://books.google.com?id=ZVUC8riEVPQC.

Roussey, Ludivine, and Raphael Soubeyran. 2018. "Overburdened Judges." *International Review of Law and Economics* 55 (September): 21–32. https://doi.org/10.1016/j.irle.2018.02.003.

Sadurski, Wojciech. 2019. *Poland's Constitutional Breakdown*. Oxford, New York: Oxford University Press.

Smekal, Hubert, Jaroslav Benák, Monika Hanych, Ladislav Vyhnánek, and Štěpán Janků. 2021. *Mimoprávní Vlivy Na Rozhodování Českého Ústavního Soudu:* Brno: Masaryk University Press. https://doi.org/10.5817/CZ.MUNI.M210-9884-2021.

Sternberg, Sebastian. 2019. "No Public, No Power? Analyzing the Importance of Public Support for Constitutional Review with Novel Data and Machine Learning Methods." PhD thesis, Mannheim: Universität Mannheim. https://madoc.bib.uni-mannheim.de/52329.

Sunstein, Cass R., David Schkade, Lisa M. Ellman, and Andres Sawicki. 2006. *Are Judges Political? An Empirical Analysis of the Federal Judiciary*. Brookings Institution Press. https://www.jstor.org/stable/10.7864/j.ctt12879t7.

Urbániková, Marína, Barbara Havelková, and David Kosař. 2023. "The Art of Waiting Humbly: Women Judges Reflect on Vertical Gender Segregation." *Feminist Legal Studies*, August. https://doi.org/10.1007/s10691-023-09533-w.

Vartazaryan, Gor. 2022. "Sít'ová Analỳza Disentujících Ústavních Soudců." *Pravnik*, no. 12.

Weinshall, Keren, and Lee Epstein. 2020. "Developing High-Quality Data Infrastructure for Legal Analytics: Introducing the Israeli Supreme Court Database." *Journal of Empirical Legal Studies* 17 (2): 416–34. https://doi.org/10.1111/jels.12250.

Wickham, Hadley. 2014. "Tidy Data." *Journal of Statistical Software* 59 (10). https://doi.org/10.18637/jss.v059.i10.

Wittig, Caroline. 2016. *The Occurrence of Separate Opinions at the Federal Constitutional Court*. Logos Verlag Berlin. https://doi.org/10.30819/4411.