

The Comprehensive Court of Appeals Database

Codebook — Release 09-2026

Eitan Sapiro-Gheiler
Princeton University
eitans@princeton.edu

Jonathan P. Kastellec
Princeton University
jkastell@princeton.edu

Introduction

This codebook documents the Comprehensive Court of Appeals Database, a case-level dataset covering all published U.S. Courts of Appeals opinions from 1892 through 2025 (approximately 445,000 opinions). A companion paper, “Appealing to Large-Language-Model-as-Judge: Annotating and Analyzing the Universe of U.S. Appellate Opinions” (Sapiro-Gheiler and Kastellec) introduces the dataset, describes the database construction pipeline, and reports the LLM-coding validation results referenced throughout. Each case row combines four blocks of information:

1. **Case metadata** (§1): identifiers, decision date, circuit, citation, per-opinion structure (type, author, and word count for each opinion in the case), per-type opinion counts, and the panel-quality flags.
2. **Judge metadata** (§2): per-judge attributes drawn from the FJC Biographical Directory and standard judge-ideology score files. Non-en-banc panels carry one block per panel judge; en banc rosters carry a single block of parallel semicolon-separated attribute lists. Four convenience blocks pull the same attributes forward for the most-asked-about opinion authors: the majority author and the first dissent, first concurrence, and first mixed-opinion authors.
3. **Panel aggregates** (§3): derived summary variables over the case’s matched judges — composition counts and flags, lone-judge indices, and per-measure summary statistics (min, max, median, mean) for ideology, age, and tenure, with matching judge indices. All apply uniformly to panel and en banc cases.
4. **LLM-coded variables** (§4): substantive case characteristics — parties, issue area, disposition, ideological direction, and many others — generated by a large language model from the full opinion text, in two parallel tracks (SCDB and Songer) and two passes per track. See the accompanying paper for the coding pipeline and validation results.

The dataset also links cases to two external sources: the Songer Courts of Appeals Database (hand-coded ground truth for a subset of cases) and the Federal Judicial Center’s Integrated Database (administrative records for almost all cases since 1971). Data from those sources is included without any corrections, and variable definitions are reproduced in Appendices A and B (respectively) for convenience. The original codebooks remain authoritative for those sources. Appendix C reproduces the corresponding native-code definitions from the Supreme Court Database (SCDB) for the SCDB-track LLM-coded variables, again for users joining against raw SCDB data.

To find a specific variable (either from this database or from any of the three reproduced external codebooks) by column name, see the alphabetical index in Appendix D.

How to use this codebook

Variables are organized by what they describe and where the information comes from, not by the order they appear in the data file. Each entry follows the same template:

Display Name `column_name` *data type* [Source: ...]

A short prose description of what the variable means and how it was constructed, followed by an inline value list when the variable is coded.

For LLM-coded variables, a blue box reports validation accuracy against the relevant source.

Coverage notes (e.g., era differences, sparsity) appear in orange boxes where applicable.

File formats and how to read them

This dataset comes in four formats for ease of use. All four contain exactly the same rows and the same columns — the CSV, Stata and R files are generated from the parquet, so no format is a subset of another — and all use the codebook column names documented below (except where the Stata 32-character variable-name cap requires truncation; see below).

Uniform dtypes across all formats. Every column the codebook tags as *numeric* is stored as a 64-bit floating-point value in every format (`double` in Parquet, `float64` in pandas, `double` in Stata, `numeric` in R); every column tagged as *string* is stored as a string (`string` in Parquet, `string` / `object` in pandas, `str#` or `strL` in Stata, `character` in R). The downside of this uniformity is purely cosmetic: integer-valued columns like `year` and `judge_count` print as “1880.0” in CSV rather than “1880”, and Stata reports their storage type as `double`. The upside is that the parquet partitions all share a single schema (DuckDB, Polars, R’s `arrow`, and any other parquet consumer can read them in any order with no schema-merge step), and missing values flow consistently through every numeric column as `NaN` / `.` / `NA`.

Parquet (chunked directory, ~1 GB): 52 partition files, all sharing a uniform schema. Preferred for Python users, and used to build the other formats below.

CSV (single file, ~4 GB): standard RFC-4180 CSV with a header row. When reading the `caseid` column, pass `dtype={"caseid": str}` to treat as `str` rather than the default `int64`. A few string-valued cells (e.g., long LLM-generated justifications) contain embedded newlines, so use a CSV reader that handles RFC-4180 quoted multi-line fields. Per the uniformity rule above, integer-valued numeric columns print with a trailing `.0` (e.g., `year` reads as 1880.0, not 1880); cast back to integer at read time if needed.

DTA (single file, ~13 GB): every numeric column is stored as Stata `double`; every string column is stored as the smallest fixed-length `str#` that fits its longest value, except for a handful of columns whose strings exceed Stata’s 2045-character `str#` cap (long LLM justifications, JSON blobs) which use the `strL` (long string) storage type that Stata 14+ supports. Every variable carries a label (the codebook display name). Stata caps variable names at 32 characters, so 16 codebook columns are renamed to a shorter prefix-truncated form. The mapping is included in `ccad_08-2026.stata.rename.csv` (`codebook_name` → `stata_name` → `label`); the Stata file has the full codebook name attached via `label variable` on every column.

RData (single file, ~6 GB): one data frame named `case_frame` with the codebook column names verbatim (no 32-character cap). R has only one numeric type (`double`), so integer-valued columns (e.g., `year`, `judge_count`) are stored as `double` rather than integer — the values are identical, just the type tag differs.

Two crosswalk CSVs (`songer_crosswalk.csv`, `idb_crosswalk.csv`) are also included for users who want to inspect the matching to Songer or IDB or reapply it against a fresh release; see §5 for their schemas and join keys.

Ideological direction convention

Every “direction” variable in this dataset uses the Songer convention:

0 = not ascertained
1 = conservative
2 = mixed
3 = liberal

This is *not* the Supreme Court Database (SCDB) convention (which codes 1 = conservative, 2 = liberal). The convention applies to LLM-coded direction variables and reproduced Songer ground-truth direction. Remap when merging with SCDB.

Filtering to clean panels

Most panel-decision analyses want a clean three-judge-panel subset with all judges and opinion authors matched to FJC records. The recommended “clean-panel” filter is:

```
judge_count == 3 & all_judges_matched == 1 & all_authors_matched == 1
```

Each piece addresses a distinct concern:

`judge_count == 3` keeps only complete three-judge panels. Drops two-judge panels (one panel member recused, died, or was otherwise unavailable) and en banc panels in one step — en banc cases all have `judge_count > 3`, so a separate `is_enbanc == 0` filter is redundant.

`all_judges_matched == 1` keeps only cases where every panel judge was matched unambiguously to an FJC identifier.

`all_authors_matched == 1` keeps only cases where every populated opinion author was matched to a panel judge.

The last two flags are independent: dropping `all_authors_matched` keeps cases with unmatched opinion authors but matched panel judges; dropping `all_judges_matched` keeps the symmetric case.

Looser filters work too. `is_enbanc == 0` alone keeps one- or two-judge panels alongside three-judge panels. Use `judge_count >= 2` if you want at least a two-judge panel without requiring exactly three (i.e., including en banc panels).

A note on terminology

A few terms recur throughout the codebook:

Slot A panel-judge position on the case, indexed 1, 2, or 3 for three-judge panels. Slots are arbitrary — they reflect the order in which judges were matched, not seniority or opinion role. For en banc cases the slot blocks (`j1_*/j2_*/j3_*`) are all NA; the full roster is instead carried in `j_enbanc` (IDs, not names) and the parallel `enbanc_*` attribute lists. Judges on en banc cases are 1-indexed by their position in `j_enbanc` (see `opinion_N_author`, §1.4).

Focal party The party whose perspective is used to key a substantive outcome (e.g., “defendant” in a criminal case, “taxpayer” in a tax case). Project-defined for LLM coding, and distinct from appellant or respondent, which are procedural roles.

LLM-coded vs. ground truth Variables in §4 were produced by the LLM coding pipeline described in the paper. Variables in Appendices A and B (`songer_gt_*` and `idb_*`) are external hand-coded or administrative data used for validation; we refer to them as “ground truth.”

Mechanical direction An ideological direction variable derived by a deterministic lookup from other LLM-coded variables (issue area, litigant type, and outcome), rather than coded directly by the LLM. See §4.4.

The paper References to a/the/accompanying paper are pointing to the companion work, “Appealing to Large-Language-Model-as-Judge: Annotating and Analyzing the Universe of U.S. Appellate Opinions” (Sapiro-Gheiler and Kastellec), which contains more details of the data pipeline and thorough validation of LLM-coded variables.

What changed since release 06-2026

This release rebuilds the database from the source opinions rather than patching the previous one, so figures can move slightly on any case. The changes worth knowing about:

1. **Author matching corrected.** Opinion authors are now resolved by judge identifier throughout. A defect in the previous release matched some authors by name where an identifier was already available, which could attribute an opinion to the wrong judge of the same surname and dropped authorships whose names were OCR-damaged or ambiguous. Author-linked variables are more complete and more accurate in this release.
2. **Byline detection improved.** The routines that decide where a separate opinion begins, and who signed it, were substantially reworked. Opinions are separated more reliably, and fewer passages are split off from the opinion they belong to or attributed to a judge who did not write them. Opinion counts, opinion types, and authorship all benefit.
3. **Non-participating judges recorded (`nonpart_*`).** When a court states that a listed judge took no part, that judge is removed from the panel — correct, since they cast no vote, but previously indistinguishable from never having been on the case. The new columns record who they were, with the same attribute set carried for every other judge, plus the court’s own sentence verbatim. See §2.5.
4. **Multi-author opinions recorded (`opinion_N_coauthors`).** Where a court records more than one judge on a single opinion (“BARKETT and MARCUS, Circuit Judges”), the additional signers are now listed rather than only the first. See §1.4.

Note that these changes interact. Items 1, 2 and 4 all feed the panel: a judge recovered from a byline may now be seated who previously was not, and duplicate spellings of one judge collapse into a single seat. `judge_count` and `is_enbanc` therefore change on a small number of cases, which

moves those cases in or out of any three-judge-panel filter. Analyses that condition on panel size should be re-run rather than compared across releases. Full technical detail for every change is in the project’s code history; this note is a summary for data users.

Contents

What changed since 06-2026	4
1 Case metadata	8
1.1 Case identifiers	8
1.2 Court and date	8
1.3 Regex-extracted disposition (<code>regex_disp_*</code>)	9
1.4 Opinion structure	11
1.4.1 Per-opinion schema (<code>opinion_N_*</code>)	12
1.5 Panel-structure indicators	14
1.6 Data availability flags	15
2 Judge metadata	16
2.1 Panel-judge attributes (slots $N \in \{1, 2, 3\}$)	16
2.1.1 Identifying information	17
2.1.2 Party and ideology	17
2.1.3 Demographics and background	18
2.2 Majority-author attributes (<code>author_*</code>)	20
2.3 First-opinion convenience blocks (<code>dis1_*</code> , <code>conc1_*</code> , <code>mix1_*</code>)	21
2.4 En-banc roster (<code>j_enbanc</code> , <code>enbanc_*</code>)	22
2.5 Non-participating judges (<code>nonpart_*</code>)	23
3 Panel metadata (derived aggregates)	25
3.1 Composition counts	25
3.2 Composition flags (0/1)	25
3.3 Lone-judge slots	26
3.4 Panel summary statistics	27
4 LLM-coded variables	28
4.1 SCDB track — litigants	29
4.2 SCDB track — disposition	31
4.3 SCDB track — issue	31
4.4 SCDB track — ideological direction	33
4.5 SCDB track — other	34
4.6 Songer track — litigants	35
4.7 Songer track — disposition	36
4.8 Songer track — issue	36
4.8.1 Threshold issues (first pass)	37
4.8.2 Sub-issues by type of issue (second pass)	38
4.9 Songer track — ideological direction	39
4.10 Songer track — other	40

5	Linked external sources	42
5.1	Songer Court of Appeals Database	42
5.2	FJC Integrated Database	42
A	Songer ground truth (songer_gt_*)	44
A.1	Identifiers and dates	44
A.2	Issue area, type, and authority	47
A.2.1	Three-digit casetyp1 subcodes	47
A.3	Direction, opinion, and voting	56
A.4	Panel judges and votes	57
A.5	Concur and dissent basis (KH-update fields)	59
A.6	Parties: counts and detailed codes	60
A.7	Counsel, attorneys, firms, and amicus briefs	63
A.8	Threshold issues	66
A.9	Substantive sub-issues: criminal (songer_gt_typeiss = 1)	68
A.10	Substantive sub-issues: civil-private and civil-government (songer_gt_typeiss ∈ {2, 3, 4})	70
A.11	Substantive sub-issues: agency and administrative (songer_gt_typeiss = 2)	73
A.12	Constitutional and legal-provision citations	74
B	IDB ground truth (idb_*)	77
B.1	Identifiers and case origin	77
B.2	Dates	81
B.3	Disposition, opinion, and publication	82
B.4	Panel judges and party status	86
B.5	Legacy 1971–2007 fields	88
B.6	Match metadata	91
C	SCDB variables coded by the LLM	92
C.1	Litigants	92
C.2	Issue area	92
C.3	Specific issue (issue)	93
C.3.1	Issue area 1: Criminal Procedure	93
C.3.2	Issue area 2: Civil Rights	95
C.3.3	Issue area 3: First Amendment	96
C.3.4	Issue area 4: Due Process	96
C.3.5	Issue area 5: Privacy	96
C.3.6	Issue area 6: Attorneys	97
C.3.7	Issue area 7: Unions	97
C.3.8	Issue area 8: Economic Activity	97
C.3.9	Issue area 9: Judicial Power	98
C.3.10	Issue area 10: Federalism	100
C.3.11	Issue area 11: Interstate Relations	100
C.3.12	Issue area 12: Federal Taxation	100
C.3.13	Issue area 13: Miscellaneous	101
C.3.14	Issue area 14: Private Action	101
C.4	Disposition and winning party	101
C.5	Direction of decision	102

C.6	Decision type	102
C.7	Declaration of unconstitutionality	102
C.8	Legal provisions	103
D	Alphabetical variable index	104

1. Case metadata

These variables describe the case itself: its identifiers, when and where it was decided, where it was published, and what opinions it contains. All come directly from the published opinion text and bibliographic records.

1.1 Case identifiers

Case ID `caseid` *string*

Unique case identifier and the merge key across all parts of the dataset. Begins with 1 for cases sourced from the Caselaw Access Project and 2 for cases sourced from CourtListener; the remainder is the source's own identifier. Use this column to link to the Songer and IDB crosswalk files (see §5).

Citation `citation` *string*

Primary Federal Reporter citation, encoding the reporter, volume, and first page in a single string (e.g., 135 F. 39, 927 F.3d 1023, 785 F. App'x 421). Users who need the reporter / volume / page as separate fields can parse this column with a small regex; we keep them packed here to avoid schema bloat. NA when no primary citation could be parsed from the source.

Docket number `docket_number` *string*

Docket number as recorded by the source. Format is not standardized across circuits or eras. NA when no docket number was recorded in the source (common in older reporters).

Data source `source` *string*

Origin of the case text: `caselaw` (Caselaw Access Project) or `courtlistener` (CourtListener).

Citation count (placeholder) `cite_count` *numeric*

Number of citations as of December 31, 2025. NA for `source = caselaw` cases (citation counts are currently retrieved only from CourtListener), and for any CourtListener case where the count could not be retrieved.

1.2 Court and date

Circuit `circuit` *numeric*

Federal appellate circuit, as an integer:

0 = D.C. Circuit

1--11 = numbered circuits

12 = Federal Circuit

NA = deciding court does not parse to a known circuit

Year of decision `year` *numeric*

Four-digit year in which the appellate decision was announced. Use this column for time-series analyses. The Songer ground-truth year column (`songer_gt_year`) is retained in Appendix A for reference but may be inconsistent with `year` on a small subset of cases; `year` is the canonical column. NA when the decision date could not be parsed from the source.

Decision date `decision_date` *string*

Full decision date in ISO format (YYYY-MM-DD) where available, else the raw date string as it appeared in the source. NA when no date string was present in the source at all.

1.3 Regex-extracted disposition (`regex_disp_*`)

A small set of disposition codes extracted directly from the majority-opinion text by a rule-based regex scanner, independent of the LLM pipeline. These columns give users a transparent, machine-checkable fallback for the LLM-coded disposition (§4.2); see the companion paper’s appendix on mechanical-disposition validation for accuracy benchmarks.

The scanner runs on every case in the database. To pick which opinion to scan, it parses the case’s stored opinion list and selects the first opinion whose type is treated as *majority*, including the Courtlistener-only alternate types `majority-lead` or `010combined`. If no opinion is tagged as a majority but the case has at least one opinion, the scanner falls back to the first opinion in source order. For the chosen opinion, the scanner runs two matching variants:

Plain. Case-insensitive matching on the disposition stems `affirm`, `revers`, `remand`, `vacat`, `dismiss`, `grant`, `den(y|ied|ies|ial)`, `enforc`, `cert(if|iorari)`.

Strict. All-caps matching on the same stems (decretal convention).

Each variant is tried via two strategies:

Conclusion-block. If a conclusion section header (e.g. CONCLUSION, DISPOSITION, DECISION, or a terminal Roman-numeral subheading) is detected in the last 50% of the opinion, scan the entire conclusion block.

Walkback. Walk back from the last sentence up to five sentences and take the union of stem hits from the anchor sentence to end of opinion.

Conclusion-section header detected `regex_disp_concl_found` *numeric*

Binary flag (0/1). 1 if a conclusion-section header was detected in the last 50% of the majority opinion; 0 otherwise. This is independent of whether any disposition stems were matched inside the conclusion — a header may exist with no bucketable stems. NA only if no opinions with text exist for a given case.

Code values for the four disposition columns below follow the SCDB scheme used by the LLM track — the same code list as `scdb_first_disposition` (§4.2), so a row-wise comparison between regex and LLM disposition requires no remap. To compare against the Songer hand-coded ground

truth in `songer_gt_treat` (§A) instead, remap SCDB code 9 to Songer 0 (stay / petition / motion granted) and SCDB code 10 to Songer 9 (certification); the other codes are identical between the two schemes.

Disposition — plain, conclusion-block `regex_disp_plain_concl` *numeric*

SCDB-coded disposition from scanning the conclusion block with the plain (case-insensitive) variant. NA when no conclusion-section header was detected; otherwise the code reflects the union of stems found in the conclusion block. 0 marks an ambiguous match (stems present but did not resolve to a single bucket — e.g. **affirm** together with **dismiss** and no **revers** or **vacat**, or a lone **grant** / **remand** without a companion stem). Full code list, identical to `scdb_first_disposition` (§4.2):

- NA = no scan attempted (no conclusion-section header detected)
- 0 = no disposition / ambiguous match (see above)
- 1 = affirmed
- 2 = reversed
- 3 = reversed and remanded
- 4 = vacated and remanded
- 5 = affirmed in part, reversed (or vacated) in part
- 6 = affirmed in part, reversed (or vacated) in part, and remanded
- 7 = vacated
- 8 = petition denied or appeal dismissed
- 9 = stay, petition, or motion granted
- 10 = certification to or from a lower court
- 11 = affirmed, vacated, and remanded (without reversal) — rare, introduced in the Kuersten–Haire update; the LLM emits this code in only a handful of cases. Most analyses should remap to 6 (see `songer_gt_treat`, §A).

To convert any value in this column to the Songer `treat` scale used by `songer_gt_treat` (§A), drop 0 and apply the remap {9→0, 10→9, 11→6}.

Disposition — plain, walkback `regex_disp_plain_walkback` *numeric*

SCDB-coded disposition from the walkback strategy with the plain variant: starting from the last sentence and walking back up to five sentences, take the first sentence that contains any disposition stem, then take the union of stem hits from that anchor sentence through end of opinion. NA when no stem matched in any of the last five sentences. Code list and Songer remap: same as `regex_disp_plain_concl`.

Disposition — strict, conclusion-block `regex_disp_strict_concl` *numeric*

As `regex_disp_plain_concl` but using the strict (all-caps) variant; matches only the decretal all-caps convention (AFFIRMED, REVERSED, etc.). Code list and Songer remap: same as `regex_disp_plain_concl`.

Disposition — `strict, walkback` `regex_disp.strict.walkback` *numeric*

As `regex_disp_plain.walkback` but using the `strict` (all-caps) variant. Code list and Songer remap: same as `regex_disp_plain.concl`.

1.4 Opinion structure

These variables describe which opinions the case produced and who authored them. Authorship is recorded as an index: on non-en-banc cases 1--3 refers to the panel-judge slot (see §2); on en banc cases it is the 1-indexed position in `j_enbanc`. 0 is reserved for *per curiam* opinions.

Each opinion in the case falls into exactly one of five mutually exclusive types: *majority*, *concurrence*, *dissent*, *mixed* (concurring in part and dissenting in part), or *rehearing* (the court's order on a petition for rehearing, typically a denial). The five per-type counts sum to the total:

$$\text{n_opinions} = \text{n_majority} + \text{n_concurrence} + \text{n_dissent} + \text{n_mixed} + \text{n_rehearing}.$$

All six count columns are always populated as non-negative integers (never NA); a count of 0 means no opinions of that type appeared.

Total opinions `n_opinions` *numeric*

Total number of opinions in the case, summed across the five types below.

Number of majority opinions `n_majority` *numeric*

Count of majority opinions. Typically 1; 0 when no opinion in the case is explicitly designated a majority; occasionally >1 when the upstream metadata records plurality / split majorities. See the `author` field for the slot index of the majority author. The CourtListener-specific raw type values `majority-lead` (plurality with lead opinion) and `010combined` (combined opinion text) are counted as majority.

Number of concurrences `n_concurrence` *numeric*

Count of *strict* concurring opinions. Opinions designated “concurring in part and dissenting in part” are **not** counted here — they are counted under `n_mixed`.

Number of dissents `n_dissent` *numeric*

Count of *strict* dissenting opinions. Mixed (concur-in-part, dissent-in-part) opinions are **not** counted here.

Number of mixed opinions `n_mixed` *numeric*

Count of opinions designated “concurring in part and dissenting in part.”

Number of rehearing orders `n_rehearing` *numeric*

Count of opinions of type `rehearing` — typically the court's one-paragraph order on a petition

for panel or en banc rehearing (usually denied). This is an *opinion-level* count of orders attached to the case file; it is distinct from the *case-level* question of whether the substantive decision was a first-time panel decision or a rehearing.

1.4.1 Per-opinion schema (opinion_N_*)

The case’s opinions are stored in source order in a uniform four-column block per slot. For each opinion slot $N \in \{1, \dots, K\}$, where K is the maximum number of opinions observed on a single case in the dataset, four columns describe that opinion:

Opinion type `opinion_N_type` *string*

One of **majority**, **concurrence**, **dissent**, **mixed** (concurring in part and dissenting in part), or **rehearing**. CourtListener-specific raw values **majority-lead** and **010combined** are normalized to **majority**. NA for slots beyond the actual number of opinions on the case.

Opinion author `opinion_N_author` *numeric*

Index of the opinion’s author. Interpretation depends on whether the case is en banc:

0 = *per curiam* (no individual author)

1--3 = for non-en-banc cases (`is_enbanc == 0`): panel-judge slot index, pointing into the `j1_*` / `j2_*` / `j3_*` block in §2.

1+ = for en banc cases (`is_enbanc == 1`): one-indexed position in the `j_enbanc` list (also in §2).
So `opinion_1_author == 5` on an en banc case means the first opinion was authored by the fifth judge in `j_enbanc`.

NA = author could not be matched (or slot beyond `n.opinions`).

Opinion co-authors `opinion_N_coauthors` *string*

Semicolon-separated list of the opinion’s *additional* authors, where a court records more than one (“BARKETT and MARCUS, Circuit Judges”). Each entry is an author index on exactly the same scale as `opinion_N_author` — panel slot 1--3 on non-en-banc cases, position in `j_enbanc` on en banc cases — so the two columns of a slot are read the same way. The primary author in `opinion_N_author` is *not* repeated here.

NA when the opinion has a single author (the overwhelming majority: co-authorship appears on roughly 0.1% of cases). An *empty* entry within a non-empty list (e.g. “3;”) marks a co-author whose FJC identity could not be resolved: the list length still counts the co-authors, but that one cannot be pointed at a judge. Consistent with the rest of the dataset, an unresolvable author is recorded as missing rather than guessed at from the name.

Off-panel co-authors are seated. A co-author who matched the FJC directory but was not among the judges extracted for the case is appended to the roster, exactly as an off-panel *primary* author already was. Both are recovered from opinion text by the same reader, so treating them differently was an inconsistency rather than a safeguard. Two consequences for anyone reading these columns:

- An empty token no longer means “not on this case’s roster” — that reason is gone, because such a judge is now put on the roster. It means only that no FJC identity could be established.

- Seating adds a judge, so it can raise `judge_count` and flip `is_enbanc`. Corpus-wide this moves five cases from a three-judge panel to en banc, and in four of them the opinion says so in its own words (“We concur in the judgment of the en banc court”, two separate opinions on a denial of rehearing *in banc*, and a four-signer byline on a three-judge panel). Those cases were previously inside the `clean_panel` filter (§) while their own text described a larger bench; they now fall outside it, which is the correct treatment.

Co-authorship is overwhelmingly an en banc phenomenon: of the cases carrying any co-author, roughly nine in ten are already `is_enbanc == 1`, where seating changes nothing because the judge is on the roster already.

Opinion word count `opinion_N_words` *numeric*

Word count of the opinion text. NA for slots beyond `n_opinions`.

Working with opinions. To get all dissents on a case, filter slots where `opinion_N_type == "dissent"`; the strict-count `n_dissent` already gives the number of such slots. For author attributes (party, ideology, age, ...), read `opinion_N_author` as a slot index. For non-en-banc cases, the panel-judge attribute columns use the prefix `j1_*` / `j2_*` / `j3_*`, so an `opinion_N_author` of 2 points at the `j2_gbp`, `j2_age`, etc. block. For en banc cases, take the `opinion_N_author`-th position in `j_enbanc` (and in the parallel `enbanc_*` attribute lists; see §2.4).

Four dedicated convenience blocks shortcut the most common queries without requiring the lookup:

- `author_*` — majority-opinion author (§2.2).
- `dis1_*` — first dissent’s author.
- `conc1_*` — first concurrence’s author.
- `mix1_*` — first mixed-opinion (concur-in-part, dissent-in-part) author.

See §2.3 for the dissent, concurrence, and mixed attribute blocks. Authors of *additional* opinions (`dis2`, `conc2`, `mix2`, slots beyond the first of each type) require the manual lookup described above.

Majority opinion author `author` *numeric*

Index of the majority opinion’s author. Same interpretation rules as `opinion_N_author`: 0 = per curiam, 1--3 = panel slot (non-en-banc) or 1+ = position in `j_enbanc` (en banc), NA = no explicit majority or unmatched author. For cases with `n_majority > 1` (split majorities), `author` carries the first majority slot in source order; the full set is recoverable from `opinion_N_author` filtered to `opinion_N_type == "majority"`.

First-dissent author `dis1` *numeric*

Index of the first dissenting opinion’s author, in source order. Same encoding as `author`: 0 = per curiam, 1--3 = panel slot (non-en-banc) or 1+ = position in `j_enbanc` (en banc), NA = the case has no dissent (`n_dissent == 0`) or the dissent author could not be matched to a roster judge. Pairs with the `dis1_*` attribute block in §2.3.

First-concurrence author `conc1` *numeric*

Index of the first concurring opinion's author, in source order. Same encoding as `author` and `dis1`; NA when `n_concurrence == 0` or the author is unmatched. Pairs with `conc1_*`.

First-mixed-opinion author `mix1` *numeric*

Index of the first mixed (concurring-in-part, dissenting-in-part) opinion's author, in source order. Same encoding; NA when `n_mixed == 0` or the author is unmatched. Pairs with `mix1_*`.

1.5 Panel-structure indicators

Panel size `judge_count` *numeric*

The number of judges on the case (1, 2, 3, 5, 9, 11, 15, ...). This is the *structural* panel size: every slot that carries a judge name from the source counts, including slots where the matching pipeline did not resolve the name to an FJC identifier. So `judge_count == 2` is always a legitimate two-judge panel (one member recused, died, or was otherwise unavailable); it is never a three-judge panel where one slot failed to match. Standard three-judge panels have `judge_count == 3`; en banc panels have `judge_count > 3` (typically the circuit's active membership, often 9–15). One-judge panels are rare but do occur. NA *only* when no judge names were parsed from the source for the case — `judge_count` is otherwise always populated as a positive integer (and is never 0).

Pair with `all_judges_matched == 1` when an analysis needs every slot resolved to an FJC identifier.

En banc indicator `is_enbanc` *numeric*

0 = not en banc (`judge_count <= 3`)

1 = en banc (`judge_count > 3`)

NA = `judge_count` missing

Convenience flag derived from `judge_count`. Use this in filters rather than re-deriving from the count each time.

Per curiam indicator `is_percuriam` *numeric*

0 = majority opinion has an identified author

1 = *per curiam* (`author == 0`)

NA = unmatched or no explicit majority opinion

Convenience flag derived from `author`.

All judges matched `all_judges_matched` *numeric*

Indicator (0/1) for whether every panel-judge slot matched unambiguously to a unique FJC identifier. 0 when the matching pipeline could not match one or more judges. NA when no judge names were extracted from the source at all (`judge_count` is NA).

All opinion authors matched `all_authors_matched` *numeric*

Indicator (0/1) for whether every populated `opinion_N_author` (and the `author` convenience column) was either identified as *per curiam* or matched to a panel-judge slot / en banc position. A value of 1 thus includes cases whose only authorship resolution is *per curiam*, alongside cases where every author was resolved to an FJC-matched judge. NA when no author names were extracted from any opinion on the case. Note that if only *per curiam* was extracted as an opinion author, it may be that `all_authors_matched` is not NA (*per curiam* is a valid author), but `all_judges_matched` is NA (*per curiam* authorship does not identify any panel judges).

1.6 Data availability flags

For each case, these 0/1 flags indicate which of the three companion data sources (LLM coding, IDB, Songer ground truth) is available. Variables in the corresponding sections are missing on cases where the flag is 0. All three flags are always populated (never NA).

LLM coding available `llm_present` *numeric*

1 if this case has LLM-coded variables (see §4); 0 otherwise.

IDB record available `idb_present` *numeric*

1 if this case is linked to a record in the FJC Integrated Database (see Appendix B); 0 otherwise. Coverage begins in 1971.

Songer ground truth available `songer_present` *numeric*

1 if this case is in the Songer hand-coded ground-truth subset (see Appendix A); 0 otherwise.

2. Judge metadata

These variables describe the judges associated with each case — both the panel judges and the authors of the case’s opinions. All attributes are drawn from the FJC Biographical Directory and from standard judge-ideology sources; the mapping from raw judges to case slots is the product of the project’s matching pipeline. The variables in this section are administrative or scholarly data about the judges themselves; they are not coded by the LLM and are not part of the LLM-validation results.

The information appears in three families, all using the same underlying set of attributes:

- **Panel-judge attributes:** one block per panel slot $N \in \{1, 2, 3\}$, named `jN_<attr>`. Populated only for cases with a normal panel (`is_enbanc == 0`). See §2.1.
- **En-banc roster:** `j_enbanc` (semicolon-separated list of judge IDs) plus a parallel `enbanc_<attr>` list for every other judge attribute. Populated only for cases with `is_enbanc == 1`. See §2.4.
- **Non-participating judges:** `nonpart_<attr>`, a parallel list per attribute for the judges recorded as taking no part, plus the court’s own `nonpart_statement`. Populated on either panel or en banc cases. See §2.5.
- **Convenience first-opinion blocks:** four parallel attribute blocks for the most-asked-about opinion authors — the majority author (`author_<attr>`), the first dissent (`dis1_<attr>`), the first concurrence (`conc1_<attr>`), and the first mixed-opinion (`mix1_<attr>`). Each is populated for both panel and en banc cases (looked up via panel slot or `j_enbanc` position as appropriate). See §2.2–§2.3. Authors of opinions beyond the first of each type require a manual lookup against the panel block or `enbanc_*` lists.

To avoid repetition, the full set of attributes is documented once under panel-judge variables in §2.1. The `author_*`, `dis1_*`, `conc1_*`, and `mix1_*` blocks share the same attributes with renamed columns; differences are noted in §2.2 and §2.3.

En banc cases. For `is_enbanc == 1`, the per-slot blocks `j1_*`, `j2_*`, `j3_*` are all NA. The full roster lives in `j_enbanc` (as judge IDs, not names) and `enbanc_*` (parallel attribute lists). Per-opinion authors on en banc cases are addressed by their position in `j_enbanc` (see `opinion_N_author` in §1.4); the `author_*`, `dis1_*`, `conc1_*`, and `mix1_*` convenience blocks are populated by looking up that position in the `enbanc_*` lists.

2.1 Panel-judge attributes (slots $N \in \{1, 2, 3\}$)

The following block of variables appears once per panel-judge slot. *Slot indices are arbitrary* — they reflect matching order, not seniority, ordering by name, or opinion role. To identify which slot wrote the majority opinion, use `author` from §1.4.

Populated only when `is_enbanc == 0`. All per-slot columns below are NA on en banc cases; the full roster for en banc cases lives in `j_enbanc` (§2.4).

Shared NA conventions. A column is NA when (i) the slot is unmatched (no FJC record could be assigned to the judge name), (ii) the FJC record or ideology score source has no value for that attribute (e.g., era-specific coverage — ABA ratings begin in the 1950s, DIME scores in 1979, JudJIS scores in 1990), or (iii) the case is en banc (per the rule above). Per-attribute exceptions or refinements are noted in the relevant entry below.

For each variable below, replace N with 1, 2, or 3.

2.1.1 Identifying information

Judge identifier `jN_id` *numeric*

FJC National Identifier (NID) of the judge in slot N . The NID is the canonical key linking to the FJC Biographical Directory and to standard judge-ideology score files. Use this to join external judge-level data to the case-level dataset.

Judge name `jN_name` *string*

Judge's display name, formed by joining the FJC Biographical Directory **First Name**, **Middle Name**, **Last Name**, and **Suffix** fields with single spaces (and dropping any of the four that are missing). Provided as a reader-friendly label for cross-tabs and plots; *not* a stable join key — use `jN_id` (the FJC NID) to merge against external judge-level data.

Unmatched slots. When `jN_id` is NA (the matching pipeline found zero or multiple FJC candidates for the slot, so no single judge could be assigned), `jN_name` is back-filled with the uppercase surname as parsed from the opinion byline — e.g., "WOOD" when the byline reads "Before FLAUM, Chief Judge, and WOOD and WILLIAMS, Circuit Judges." `jN_id = NA` is the tell that this value is an unmatched fallback rather than an FJC display name. All other `jN_*` attributes remain NA in this case (even if they could be theoretically filled because all candidate judge matches have the same value of an attribute).

2.1.2 Party and ideology

Democratic-appointee indicator `jN_dem` *numeric*

0 = Republican appointee

1 = Democratic appointee

NA = non-presidentially-appointed judge sitting by designation

Appointing president `jN_appointing_president` *string*

Name of the U.S. president who appointed the judge to the seat held at the time of the decision (e.g., Ronald Reagan, Joseph R. Biden), in the raw form recorded by the FJC Biographical Directory. The corresponding party is collapsed into the binary `jN_dem`; this column preserves the specific appointer for cohort-style analyses (Reagan vs. Bush appointees, Clinton vs. Obama, etc.).

Giles–Hettinger–Peppers ideology score `jN_ghp` *numeric*

Giles–Hettinger–Peppers (GHP) Judicial Common Space score for the judge, on a DW-NOMINATE-style scale. Negative values are liberal, positive values are conservative. This is the project's canonical panel-ideology measure.

DIME ideology score `jN_dime` *numeric*

Bonica–Sen DIME judicial common-space (CFscore) for the judge. Negative values are liberal, positive values are conservative.

JudJIS ideology score `jN_judjis` *numeric*

JudJIS (Jurist-Derived Judicial Ideology Score) for the judge. Unlike GHP and DIME, the underlying JudJIS series is time-varying: the value stored here is the most recent JudJIS score whose date predates the case’s `decision_date`, so that scores “leak” no information from the future into a regression. Missing if no in-date JudJIS observation exists for the judge.

2.1.3 Demographics and background

Birth date `jN_birth_date` *string*

Judge’s date of birth from the FJC Biographical Directory, in YYYY-MM-DD format. Provided for users who want age in finer units than the year-derived `jN_age`.

Judge’s age at decision `jN_age` *numeric*

Judge’s age in years at the date of the decision.

Commission date `jN_commission_date` *string*

Date the judge received the commission to the seat held at the time of the decision, in YYYY-MM-DD format. `jN_tenure` is computed from this date and the case’s `decision_date`; the raw value is exposed for users who need precise day-level tenure or who want to inspect recess or backdated appointments.

Years on the bench at decision `jN_tenure` *numeric*

Years on the bench at the date of the decision, computed from the commission date in the FJC Biographical Directory. Lower-clipped at zero for backdated and recess appointments.

Senior status date `jN_senior_date` *string*

Date the judge took senior status, in YYYY-MM-DD format, or NA if the judge had not taken senior status as of December 31, 2025 (the FJC build date). The binary indicator `jN_senior` is derived from this column; the raw date lets users compute time-in-senior-status when relevant.

Senior status indicator `jN_senior` *numeric*

0 = active service (judge had not taken senior status by the case’s `decision_date`)
1 = senior status (judge had retired from active service to senior status on or before `decision_date`)

NA = unparsable case date only; we treat missing senior status dates as 0 (had not taken senior status)

Time-dependent: a judge can flip from 0 to 1 during their tenure on the bench. Note: this is *not* a chief-judge indicator; the codebook does not track chief-judge status.

Female indicator `jN_fem` *numeric*

0 = male
1 = female

Race or ethnicity `jN_race` *string*

Race or ethnicity string from the FJC Biographical Directory (e.g., **White**, **African American**).

Nonwhite indicator `jN_nonwhite` *numeric*

0 = White
1 = any other race/ethnicity
NA = `jN_race` not recorded

Birth city `jN_birth_city` *string*

Judge's city of birth from the FJC Biographical Directory. Pairs with `jN_birth_state`.

Birth state `jN_birth_state` *string*

Judge's state of birth.

Court name `jN_court_name` *string*

Full name of the court to which the judge was appointed for the seat held at the time of the decision, as recorded by the FJC Biographical Directory — e.g., **U.S. Court of Appeals for the Ninth Circuit** for a regular circuit judge, or **U.S. District Court for the Northern District of California** for a district judge sitting on the panel by designation. Useful for identifying the home court of visiting / designated judges; for regular circuit judges this column duplicates information already in case-level `circuit` (§1).

Court-of-Appeals appointee indicator `jN_is_appellate` *numeric*

1 = judge's seat at decision is a U.S. Court of Appeals appointment (a regular circuit judge)
0 = judge's seat at decision is on another court (e.g., U.S. District Court, the Supreme Court, the Court of Customs and Patent Appeals) — the judge is sitting on this appellate panel by designation
NA = no court type available for the judge in this slot

Roughly 10% of matched panel-judge slots in the published-cases universe carry `jN_is_appellate == 0`, almost all of them district-judge designations.

Duty state `jN_duty_state string`

State of the judge’s primary duty station (the state in which the seat is physically located). For circuit-court judges, this is typically the home-chambers state rather than the circuit headquarters. This is an *appointment* attribute that locates the judicial seat, not personal demographic information about the judge.

ABA rating at appointment `jN_aba string`

American Bar Association rating at the time of appointment (e.g., **Well Qualified**).

Law school (normalized) `jN_law string`

Law-school name after normalization (lowercased, punctuation stripped). Use this column when comparing law schools across judges in the same case.

Law school (grouped) `jN_law-grouped string`

Canonical institution label for the judge’s law school. The grouping consolidates name variants of the *same* institution (mergers, renamings, “college” vs. “university” vs. “law school” sub-units, predecessor institutions, etc.) into a single canonical name. For example, “Harvard College,” “Harvard Law School,” and “Harvard University” all map to **Harvard**; “Columbia University School of Law (now George Washington University Law School)” maps to **George Washington**. The grouping does **not** bucket schools into prestige tiers (Top 14, Top 50, etc.); it preserves institutional identity. Falls back to the normalized name (`jN_law`) when no grouping rule applies; NA when `jN_law` itself is NA.

2.2 Majority-author attributes (author_*)

For the author of the majority opinion, the same attributes documented in §2.1 are pulled forward into a parallel set of columns with the prefix **author_**. For non-en-banc cases, **author_ghp** is equivalent to `jN_ghp` for whichever slot *N* the author occupies. For en banc cases (where the per-slot blocks are NA), the **author_*** block still carries the majority author’s attributes, looked up from the corresponding NID in `j_enbanc` and joined against the FJC Biographical Directory; **author_role_<score>** is NA for en banc cases because it is defined relative to a three-judge panel.

The full column list is:

`author_id, author_name, author_dem, author_appointing_president, author_dime, author_role_dime, author_judjis, author_role_judjis, author_ghp, author_role_ghp, author_birth_date, author_age, author_commission_date, author_tenure, author_senior_date, author_senior, author_fem, author_law, author_law_grouped, author_race, author_nonwhite, author_aba, author_birth_city, author_birth_state, author_court_name, author_is_appellate, author_duty_state.`

Three of these suffixes do not appear in the panel-judge block and deserve separate definitions:

Author is on a Democratic-appointee panel slot `author_dem` *numeric*

Equivalent to `jN_dem` for the author’s slot N . NA for *per curiam* opinions, or when the majority author is unmatched (or the case has no majority opinion).

Author’s role on the ideology distribution `author_role_dime`, `author_role_judjis`, `author_role_ghp` *numeric*

For each ideology score, the author’s position in the panel ranking of that score:

- 1 = panel median (the middle judge by score)
- 2 = panel second-extreme (the middle in distance from the median)
- 3 = panel extreme (judge furthest from the median)
- NA = *per curiam*, en banc, unmatched author, missing case with no majority opinion, or any tie (score tie, or distance-from-median tie among non-median judges) that makes the role assignment ambiguous

For *per curiam* opinions, the three ideology scores (`author_dime`, `author_judjis`, `author_ghp`) are set to the panel mean across non-missing values; `author_id` is set to the sentinel 999999999; and all other attributes (`author_name`, `author_appointing_president`, `author_birth_date`, `author_age`, `author_commission_date`, `author_tenure`, `author_senior_date`, `author_senior`, `author_fem`, `author_law`, `author_law_grouped`, `author_race`, `author_nonwhite`, `author_aba`, `author_birth_city`, `author_birth_state`, `author_court_name`, `author_is_appellate`, `author_duty_state`, plus the four `role_*`) are NA. When the majority author is unmatched (or the case has no explicit majority), the entire `author_*` block is NA throughout.

To ask “is the author the panel’s minority-party judge?” compare slots directly: `author == minority_slot`.

2.3 First-opinion convenience blocks (`dis1_*`, `conc1_*`, `mix1_*`)

Three parallel blocks carry the attributes of the *first* dissent, the first concurrence, and the first mixed (concur-in-part, dissent-in-part) opinion. Each shortcuts the common case where there is only one opinion of the given type (~80% of cases with any non-majority opinion).

Each block has the same attribute set as `author_*` (see §2.2):

`P_id`, `P_name`, `P_dem`, `P_appointing_president`, `P_dime`, `P_role_dime`, `P_judjis`,
`P_role_judjis`, `P_ghp`, `P_role_ghp`, `P_birth_date`, `P_age`, `P_commission_date`,
`P_tenure`, `P_senior_date`, `P_senior`, `P_fem`, `P_law`, `P_law_grouped`, `P_race`,
`P_nonwhite`, `P_aba`, `P_birth_city`, `P_birth_state`, `P_court_name`, `P_is_appellate`,
`P_duty_state`

for each prefix $P \in \{\text{dis1}, \text{conc1}, \text{mix1}\}$. For example, `dis1_ghp` is the GHP score of the first dissent’s author and `mix1_age` is the age of the first mixed-opinion’s author.

These blocks are populated for both panel and en banc cases: on a non-en-banc panel the attribute is taken from the `j1_*/j2_*/j3_*` block according to the relevant `opinion_N_author`; on an en banc case it is taken from the corresponding position in the `enbanc_*` lists (§2.4).

Missing rules. A block is NA throughout when the case has no opinion of that type (`n.dissent == 0`, `n.concurrence == 0`, or `n.mixed == 0`). The `*role_<score>` columns are also NA on en banc cases (panel-relative position is undefined) and on score ties. For *per curiam* non-majority opinions (rare), `*_id` is the sentinel 999999999 and the ideology scores fall back to the panel mean, as in `author_*`.

For non-first authors. Authors of additional opinions of the same type (`dis2`, `conc2`, `mix2`, ...) require the manual lookup described in §1.4: filter `opinion_N_type`, take the *K*-th match, read `opinion_N_author`, then look up the panel block or `enbanc_*` lists.

2.4 En-banc roster (`j_enbanc`, `enbanc_*`)

For en banc cases (`is_enbanc == 1`), the per-slot panel-judge block (§2.1) is NA; the full roster is instead carried as a set of parallel semicolon-separated lists, one per attribute, all with the same length and same judge order.

En banc judge NIDs `j_enbanc string`

Semicolon-separated list of FJC National Identifiers (NIDs) of every judge on the en banc panel, in source order as the panel appears in the opinion text. Example: "1383951;1380266;1395396;1389501;1394666". NA on non-en-banc cases.

En banc judge attributes `enbanc_id`, `enbanc_name`, `enbanc_dem`, `enbanc_appointing_president`, `enbanc_ghp`, `enbanc_dime`, `enbanc_judjis`, `enbanc_birth_date`, `enbanc_age`, `enbanc_commission_date`, `enbanc_tenure`, `enbanc_senior_date`, `enbanc_senior`, `enbanc_fem`, `enbanc_race`, `enbanc_nonwhite`, `enbanc_aba`, `enbanc_birth_city`, `enbanc_birth_state`, `enbanc_court_name`, `enbanc_is_appellate`, `enbanc_duty_state`, `enbanc_law`, `enbanc_law_grouped string`

Semicolon-separated lists of judge attributes, with the same length and same ordering as `j_enbanc`. Each `enbanc_<attr>` column carries the value of the corresponding panel-judge attribute (`jN_<attr>`; see §2.1) for each en banc judge in turn.

The list types are stored as strings to keep the schema simple; consumers split on ";" to parse. Missing values for a particular judge's attribute are represented by an empty token, so "0.123;;0.789" means the second judge has a missing value for that attribute. `enbanc_id` mirrors `j_enbanc` exactly (the same NID list); the duplication keeps the convention uniform (`enbanc_<attr>` for every attribute, including the identifier).

Unmatched judges. En banc judges whose FJC match was ambiguous or absent still occupy a position in every list — `j_enbanc` / `enbanc_id` carries an empty token at that position, and so does every other `enbanc_<attr>`, with one exception: `enbanc_name` is back-filled with the uppercase surname as parsed from the opinion byline (e.g., "JOHNSON"). This mirrors the panel-slot rule for `jN_name` (§2.1) and lets users identify the structurally-occupied seat by name even when no FJC ID could be assigned.

NA on non-en-banc cases throughout.

Working with en banc rosters. Per-opinion authors on en banc cases address judges by their one-indexed position in `j_enbanc`: an `opinion_N_author` value of 5 means the opinion was authored by the fifth judge in `j_enbanc`, and that judge's GHP score is the fifth token of `enbanc_ghp` (split on

");). The `author_*` / `dis1_*` / `concl_*` / `mix1_*` convenience blocks already perform this lookup at build time, so most analyses do not need to parse the lists.

2.5 Non-participating judges (`nonpart_*`)

Courts frequently record that a judge who was listed with the case did not take part in deciding it: “MORROW, Circuit Judge, took no part in this decision.” Such a judge cast no vote, so the pipeline removes them from the panel — they do not appear in `j1_*` / `j2_*` / `j3_*` or `j_enbanc`, and they do not count toward `judge_count`, `is_enbanc`, or any panel composition measure (§3). That is the correct treatment of a non-vote, but on its own it is indistinguishable from the judge never having been on the case at all. The columns here are the record of what was removed, who they were, and on what evidence.

They are parallel semicolon-separated lists, one entry per non-participating judge, in the order the statements appear in the opinion. Populated on roughly 0.1% of cases, essentially all of them pre-1950 CAP; NA where no judge is recorded as sitting out.

Attributes. `nonpart_<attr>` carries the *same* attribute set as a panel slot or an en banc judge — `nonpart_id`, `nonpart_name`, `nonpart_dem`, `nonpart_ghp`, `nonpart_dime`, `nonpart_judjis`, `nonpart_age`, `nonpart_tenure`, and the rest of the block documented in §2.1. An unresolved judge contributes an empty token in every list, so the lists stay aligned and length still counts the judges who sat out.

These attributes are *case-relative* and are not recoverable by joining `nonpart_id` back to the FJC directory: `nonpart_age` is the judge’s age *at this decision*, `nonpart_senior` their senior status *as of this case*, and `nonpart_tenure` is measured to the same date. A directory join returns a birth date and a commission date; reproducing these columns from it would mean re-deriving each one against every case’s decision date.

Non-participating judge NIDs `nonpart_id` *string*

Semicolon-separated list of FJC National Identifiers of the judges recorded as not participating. An empty token marks a judge whose statement was found but whose FJC identity could not be resolved (about 5% of entries); their name is recoverable from `nonpart_statement`, which quotes it.

Non-participation statement `nonpart_statement` *string*

The court’s own sentence, verbatim, aligned with `nonpart_id`. Kept because it is the evidence for removing a judge from the panel, and because it names the judge even when no NID could be assigned. Any semicolon inside a statement is rewritten as a comma so the list stays splittable on ";", and internal line breaks are collapsed to single spaces; the text is otherwise unmodified, including OCR damage in the older scans.

Working with non-participation. `nonpart_id` is non-NA exactly when at least one judge sat out, so it doubles as the filter for such cases. A panel of two named judges with one `nonpart_id` entry was a three-judge panel that lost a member — materially different from a two-judge panel, and worth excluding or flagging in any analysis keyed on panel composition. Note that `all_judges_matched` and the clean-panel filter (§) describe only the judges who *remained*, and say nothing about the ones removed here.

New in this release. The whole `nonpart_*` family is new; the previous public release (06-2026) recorded nothing about judges who sat out. We deliberately publish the court's own sentence rather than a classification of it: an earlier draft carried a six-way `nonpart_reason` code, and it was dropped because it was our label for text we now ship verbatim. Classify `nonpart_statement` yourself, to whatever taxonomy your question needs.

3. Panel metadata (derived aggregates)

These variables summarize the composition and characteristics of the deciding panel as a whole. They are derived from the per-judge attributes in §2 to make common analyses easier to specify — they add no information that could not be recomputed from the per-judge columns, but they package it in a form that drops directly into regression specifications.

Aggregates apply to both normal panels and en banc panels. For non-en-banc cases the aggregate runs over the matched slots (`j1_*`, `j2_*`, `j3_*`); for en banc cases it runs over the parsed `enbanc_*` lists. A variable is NA when there are fewer than two judges with a non-missing value of the underlying attribute, or when the variable is by definition restricted to a smaller setting (noted per variable below).

3.1 Composition counts

Democratic appointees `num_dems` *numeric*

Count of Democratic appointees on the panel. 0–3 for non-en-banc; up to the en banc roster size for en banc.

Women on panel `num_women` *numeric*

Count of female judges on the panel.

Nonwhite judges on panel `num_nonwhite` *numeric*

Count of judges coded nonwhite on the panel.

Distinct law schools `num_law_schools` *numeric*

Count of distinct grouped law-school institutions (`jN_law_grouped`; see §2.1) represented on the panel. The grouping consolidates name variants of the same institution into one label (so Harvard College and Harvard Law School both count as “Harvard”).

3.2 Composition flags (0/1)

For each indicator below: 1 if every judge on the panel matches the named attribute; 0 if at least one does not. NA when one or more judges have a missing value for the underlying attribute, so the condition cannot be evaluated.

Mixed-party panel `mixed_panel` *numeric*

1 if the panel mixes Democratic and Republican appointees; 0 if all appointees are from one party.

All-male panel `all_male` *numeric*

1 if every judge on the panel is male.

All-female panel `all_female` *numeric*

1 if every judge on the panel is female.

All-white panel `all_white` *numeric*

1 if every judge on the panel is coded White.

All-nonwhite panel `all_nonwhite` *numeric*

1 if every judge on the panel is coded as a race / ethnicity other than White.

Same law school `all_same_law` *numeric*

1 if every judge on the panel attended the same grouped law-school institution.

3.3 Lone-judge slots

Each variable below is the index of the lone judge with the named attribute, populated only when the panel has exactly one such judge. The slot index follows the same dual convention as `opinion_N_author` (§1.4): for non-en-banc cases 1–3 pointing into `j1_*/j2_*/j3_*`; for en banc cases 1+ pointing into `j_enbanc` and the parallel `enbanc_*` lists (§2.4). NA when the panel has zero or more than one judge with the named attribute.

Lone minority-party judge `minority_slot` *numeric*

Slot of the lone minority-party judge whenever exactly one judge on the panel belongs to the minority party. On three-judge non-en-banc panels this is the 2–1 case. On en banc panels it works the same way — the minority is identified as the lone Democrat-on-Republican-panel (or vice versa) when the partisan split is 1-vs-(N-1). NA on unanimous-party panels and on panels with a more even split (e.g., 4-vs-5 on en banc).

Lone woman `lone_woman_slot` *numeric*

Slot of the lone female judge on a panel with exactly one woman.

Lone man `lone_man_slot` *numeric*

Slot of the lone male judge on a panel with exactly one man.

Lone nonwhite judge `lone_nonwhite_slot` *numeric*

Slot of the lone nonwhite judge on a panel with exactly one nonwhite judge.

Lone white judge `lone_white_slot` *numeric*

Slot of the lone white judge on a panel with exactly one white judge.

Interaction queries. To ask “was the minority-party judge also the lone woman?” compare slots directly: `minority_slot == lone_woman_slot`.

3.4 Panel summary statistics

For each **measure** $M \in \{\text{dime}, \text{judjis}, \text{ghp}, \text{age}, \text{tenure}\}$, the panel aggregate variables follow a uniform pattern:

Value columns (the actual number):

`M_min` minimum value on the panel.

`M_max` maximum value on the panel.

`M_med` median value on the panel.

`M_mean` mean value on the panel.

Slot columns (string; the slot index or indices of the judges with that value, following the dual opinion.*N*.author convention from §1.4):

`M_min_slot` slot(s) of the judge(s) with the minimum value.

`M_max_slot` slot(s) of the judge(s) with the maximum value.

`M_med_slot` slot(s) of the median judge(s) — one slot when the panel size is odd and the median value is unique, two slots when the panel size is even or the median value ties.

(`mean` has no corresponding slot — the mean is not any one judge’s value.)

Tie handling. Slot columns are stored as strings. A unique slot is just the integer as a string (e.g., “2”). Ties are represented as a semicolon-separated list, ordered by slot index (e.g., “1;4” for a four-judge en banc panel whose two middle-ranked judges occupy slots 1 and 4). This matches the `j_enbanc` / `enbanc_*` list convention. Consumers split on “;” when they need to enumerate the tied judges; for the common case of unique slots they can cast the column to integer.

Examples. `ghp_min` is the most-liberal GHP score on the panel; `ghp_min_slot` is the slot index of the judge with that score (or a semicolon-separated list on a tie), so you can read their other attributes via `j{ghp_min_slot}_age`, `j{ghp_min_slot}_law`, etc. (or via the corresponding `enbanc_*` list position for en banc). `age_mean` is the panel’s mean age in years.

Coverage. Each block is computed when at least two judges on the panel have a non-missing value of the underlying attribute. Both the value columns and the slot columns are always well-defined when this condition is met — ties are recorded in the slot lists rather than collapsed to NA.

Total: 7 columns per measure × 5 measures = 35 columns (4 values + 3 slots; `mean` has no slot).

4. LLM-coded variables

These variables describe substantive case characteristics — litigants, issue area, disposition, ideological direction, and others — coded by a large language model from the full opinion text. Full details of the LLM pipeline are in the companion paper, “Appealing to Large-Language-Model-as-Judge: Annotating and Analyzing the Universe of U.S. Appellate Opinions” (Sapiro-Gheiler and Kastellec). We provide a brief summary below:

Two parallel tracks

The coding pipeline runs the LLM over each case twice, in two independent tracks:

- The **SCDB track** (§4.1) adapts the Supreme Court Database codebook to the U.S. Courts of Appeals.
- The **Songer track** (§4.6) follows the Songer Courts of Appeals Database codebook.

The tracks are kept separate: the SCDB prompts never see Songer codes and vice versa. This lets us validate each track against its own ground truth and lets users pick whichever code list their analysis requires.

Two passes per track

Within each track the LLM is queried twice:

- The **first pass** prompts the LLM with the opinion text and a fixed instruction sheet, and produces broad classifications: appellant and respondent identities and categories, disposition, winning party, broad issue area, and a one-shot ideological direction. The first-pass prompt is the same for every case.
- The **second pass** re-prompts the LLM with a case-specific instruction sheet whose contents are customized based on the first-pass output (e.g., the second-pass prompt for a criminal case enumerates only the criminal-procedure issue codes; for an agency case it enumerates only the relevant administrative sub-issues). The second pass produces refined codes: the specific issue within the broad area, the focal party, the substantive outcome from that party’s perspective, and detailed party-type digits.

After both passes complete, a deterministic post-processing step derives the *mechanical* ideological direction from the combination of issue area, focal party, and outcome (see §4.4).

Naming convention

Every LLM-coded column carries the pass it was produced by, so a downstream user can tell at a glance which prompt generated the value:

`<track>_<pass>_<varname>`

where `<track>` is `scdb` or `songer`, `<pass>` is `first` or `second`, and `<varname>` is the substantive column name (e.g., `appellant_category`, `broad_area`, `focal`). Examples:

- `scdb_first_disposition` — the SCDB first-pass disposition.
- `scdb_second_focal` — the SCDB second-pass focal party (the party whose perspective the outcome is keyed to).

- `songer_first_threshold_jurisdiction` — a threshold flag set by the Songer first pass.
- `songer_second_typeiss_search` — the search-and-seizure sub-issue set by the Songer second pass.

The only LLM-output column without a track/pass prefix is `parse_count` (a per-case count of LLM attempts; see below). Composite codes (`songer_second_appellant_subcategory`, `songer_second_respondent_subcategory`) and the mechanical-direction outputs (`*_second_direction_mech`, `*_second_direction_supermech`, `*_second_direction_flag`) are deterministic transformations of LLM output, but they carry the `second` pass marker because they require the second pass to be produced.

Retry-count column

LLM parse attempts `parse_count` *numeric*

Number of LLM attempts (initial pass plus retries) needed to obtain parseable output for this case. 1 is the typical value (~95% of cases parse on the first attempt); higher values correspond to retries triggered by parse-format failures on the initial pass. The pipeline retried until parsing succeeded, so every LLM-matched case has `parse_count` ≥ 1 and the per-track, per-pass parse-success flags that earlier releases stored (e.g. `scdb_first_parsed`, `songer_second_parsed`) are no longer needed — they would all be 1. The full retry distribution is reported in the paper.

Validation results

Where possible, each LLM-coded variable below carries a small validation box reporting accuracy against Songer hand-coded data, IDB administrative data, or both. Headline accuracies against the Songer hand-coded ground truth: 87.0% on disposition, 88.7% on broad issue area (Songer track) / 86.0% (SCDB track, via crosswalk), and 80.7% on ideological direction (Songer track, super-mechanical). Against the FJC Integrated Database as an independent administrative ground truth, the LLM matches or exceeds Songer human-coder accuracy on disposition (88.9% vs. 88.4%) and on broad issue area (86.9% vs. 83.7%). See the paper for full validation, including confusion matrices and error decomposition for the headline variables.

Direction convention reminder. All direction variables below use the Songer convention: 0 = not ascertained, 1 = conservative, 2 = mixed, 3 = liberal. This is not the SCDB convention.

4.1 SCDB track — litigants

Appellant name `scdb_first_appellant_name` *string*

Free-text name of the appellant as identified by the LLM.

Appellant description `scdb_first_appellant.description` *string*

One-line free-text description of the appellant (e.g., “employee of defendant company alleging wrongful termination”).

Appellant category `scdb_first_appellant_category` *numeric*

Validation vs. Songer GT `genapel1` (via SCDB→Songer crosswalk): 84.7% accuracy ($N = 22,910$ matched cases).

Seven-category litigant typology applied to the appellant. Adapted from SCDB **petitioner**; the LLM’s code list (below) differs from the SCDB native 311-code party typology — the SCDB-native scheme is summarized in Appendix C under **petitioner**; remap when joining against raw SCDB.

- 1 = private business or corporation (including partnerships and sole proprietorships)
- 2 = nonprofit organization, interest group, or union (as an institutional litigant)
- 3 = federal government (U.S. as party, federal agencies)
- 4 = sub-state government (city, county, municipality, school district, local official)
- 5 = state government (state, governor, state agency, state official in official capacity)
- 6 = tribal government or Indian nation
- 7 = natural person (individual, including pro se litigants and individual officials sued personally)

Respondent name `scdb_first_respondent_name` *string*

Free-text name of the respondent.

Respondent description `scdb_first_respondent_description` *string*

One-line free-text description of the respondent.

Respondent category `scdb_first_respondent_category` *numeric*

Validation vs. Songer GT `genresp1` (via SCDB→Songer crosswalk): 83.9% accuracy ($N = 22,902$ matched cases).

Seven-category litigant typology applied to the respondent. Adapted from SCDB **respondent**; coding is identical to **scdb_first_appellant_category** above. The SCDB-native scheme is summarized in Appendix C under **respondent**; remap when joining against raw SCDB.

Who appealed `scdb_first_appealed` *numeric*

Who initiated the appeal, from the perspective of the lower-court proceeding. Project-derived: the SCDB codebook has no direct analog (SCDB cases all reach the Supreme Court, so the SCDB-relevant “who appealed” question is which lower court the case came from, not which party initiated the appeal).

- 1 = original-trial plaintiff (or equivalent, e.g. Title VII complainant or habeas petitioner)
- 2 = original-trial defendant (or equivalent, e.g. the employer in a Title VII suit, or the criminal accused)
- 3 = federal government on behalf of the original plaintiff
- 4 = federal government on behalf of the original defendant

4.2 SCDB track — disposition

Disposition `scdb.first.disposition` *numeric*

Validation vs. Songer GT treat: 87.0% accuracy ($N = 22,792$ matched cases).

Validation vs. IDB outcome: 88.9% accuracy ($N = 279,005$ matched cases).

How the Court of Appeals disposed of the case. Original SCDB codebook name: `caseDisposition`. The LLM’s code list (below) differs from the SCDB native scale — the SCDB-native scheme is summarized in Appendix C under `caseDisposition`; remap when joining against raw SCDB.

- 0 = no disposition
- 1 = affirmed (includes modified and affirmed)
- 2 = reversed
- 3 = reversed and remanded
- 4 = vacated and remanded
- 5 = affirmed in part, reversed (or vacated) in part
- 6 = affirmed in part, reversed (or vacated) in part, and remanded
- 7 = vacated
- 8 = petition denied or appeal dismissed
- 9 = stay, petition, or motion granted
- 10 = certification to or from a lower court
- 11 = affirmed, vacated, and remanded (without reversal)

In practice, the LLM only outputs code 11 in a handful of cases. Furthermore, that code was only added to the Songer hand-coded data in the Kuersten-Haire update (see `songer_gt_treat`). Most users will therefore want to map both the LLM-output 11 and Songer hand-coded cases with `songer_gt_treat = 11` to code 6 instead.

Winning party `scdb.first.appellant.win` *numeric*

Whether the appellant (the “petitioning party” in SCDB terms) received a favorable disposition. Original SCDB codebook name: `partyWinning`. Code list matches the SCDB convention exactly.

- 0 = no favorable disposition for the appellant (i.e., the respondent prevailed)
- 1 = appellant received a favorable disposition
- 2 = favorable disposition for the appellant unclear (e.g., a mixed affirm-in-part / reverse-in-part outcome)

4.3 SCDB track — issue

Broad issue area `scdb.first.broad.area` *numeric*

Validation vs. Songer GT geniss (via SCDB→Songer crosswalk): 86.0% accuracy ($N = 22,668$ matched cases).

Primary issue area of the case. Original SCDB codebook name: `issueArea`. Code list matches the SCDB convention exactly.

- 1 = Criminal Procedure
- 2 = Civil Rights
- 3 = First Amendment
- 4 = Due Process
- 5 = Privacy
- 6 = Attorneys
- 7 = Unions
- 8 = Economic Activity
- 9 = Judicial Power
- 10 = Federalism
- 11 = Interstate Relations
- 12 = Federal Taxation
- 13 = Miscellaneous
- 14 = Private Action

Coverage note. When the second-pass detects a government party, an issue-area code of 14 (Private Action) is overridden to 8 (Economic Activity). The override is recorded in `scdb_first_justification`.

Alternate broad issue area `scdb_first_broad_area_alt` *numeric*

A secondary issue area when the case spans two areas; missing otherwise. Same code list as `scdb_first_broad_area`. Project-derived: the SCDB codebook has no direct analog (SCDB records only a single `issueArea` per docket).

Specific issue area `scdb_second_sub_area` *numeric*

Validation vs. Songer GT `casetyp1` (set-membership via SCDB→Songer crosswalk): 40.1% accuracy ($N = 15,530$ matched cases).

Refined SCDB issue code within the chosen issue area (`scdb_first_broad_area`). Produced by the second-pass coder. Original SCDB codebook name: `issue`. Code list matches the SCDB convention exactly — there are 278 codes, which we include in Appendix C under `issue`.

Procedural holding `scdb_first_procedural` *numeric*

1 if the case's holding was purely procedural (no merits ruling); 0 otherwise. Project-derived: the SCDB codebook has no direct analog (SCDB encodes procedural posture indirectly via `decisionType` and the procedural sub-areas under `issue`).

4.4 SCDB track — ideological direction

Direction convention. All three SCDB direction variables in this subsection use the project-standard 0/1/2/3 scale: 0 = not ascertained, 1 = conservative, 2 = mixed, 3 = liberal. (Not the raw SCDB `decisionDirection` convention; remap if merging.)

Mechanical and super-mechanical direction variables are derived deterministically from second-pass variables by applying SCDB codebook rules for each issue area; see the paper for the exact lookup tables. They are typically the preferred direction variable for analysis because they correct for a small set of one-shot coding errors the LLM makes when the issue area is unusual.

Direction (one-shot) `scdb_first_direction_oneshot` *numeric*

Validation vs. Songer GT `direct1`: 80.4% accuracy ($N = 19,097$ matched cases; Songer GT “mixed” dropped, precision-credit correction applied for codebook-hardcoded “unspecifiable” SCDB issue areas).

Ideological direction of the decision as coded directly by the first-pass LLM, on the project-standard 0/1/2/3 scale (see direction convention at the top of this section). Original SCDB codebook name: `decisionDirection`. The LLM’s code list differs from the SCDB native 3-code scale (1 = conservative, 2 = liberal, 3 = unspecifiable); the SCDB-native scheme is summarized in Appendix C under `decisionDirection`.

Focal party `scdb_second_focal` *string*

Categorical description of relevant focal party (e.g., “criminal defendant” or “underdog”). Produced by the second-pass coder; used as input to the mechanical direction (see §4.4).

Focal party is appellant `scdb_second_focal_appellant` *numeric*

- 0 = focal party was the respondent
- 1 = focal party was the appellant

Focal party prevailed `scdb_second_focal_won` *numeric*

- 0 = focal party did not prevail
- 1 = focal party prevailed in the appellate decision
- 2 = mixed outcome (e.g., affirmed in part and reversed in part such that no party clearly won)

Other focal parties considered `scdb_second_focal_alt` *string*

JSON-encoded list of other parties the LLM considered as possible focal parties.

Mechanical direction `scdb_second_direction_mech` *numeric*

Validation vs. Songer GT direct1: 82.1% accuracy ($N = 19,097$ matched cases; same conditioning as `scdb_first_direction_oneshot`).

Ideological direction derived from the second-pass focal party (`scdb_second_focal`) and whether that focal party won (`scdb_second_focal_won`) by applying SCDB codebook rules for each issue area (`scdb_first_broad_area` and `scdb_second_sub_area` as appropriate). Original SCDB codebook name: `decisionDirection` (mechanical reconstruction). Uses the project-standard 0/1/2/3 direction convention, which differs from the SCDB native 3-code scale; see Appendix C under `decisionDirection`.

Super-mechanical direction `scdb_second_direction_supermech` *numeric*

Validation vs. Songer GT direct1: 79.7% accuracy ($N = 19,097$ matched cases; same conditioning as `scdb_first_direction_oneshot`).

Ideological direction derived similarly to `scdb_second_direction_mech` but replacing the LLM’s focal-party-win judgment (`scdb_second_focal_won`) with the case disposition (`scdb_first_disposition`) and whether the focal party was the appellant (`scdb_second_focal_appellant`). Original SCDB codebook name: `decisionDirection` (mechanical reconstruction). Uses the project-standard 0/1/2/3 direction convention, which differs from the SCDB native 3-code scale; see Appendix C under `decisionDirection`.

Direction edge-case flag `scdb_second_direction_flag` *string*

Free-text annotation describing the rule that was applied (e.g., “criminal forced to defendant”, “area-14 overridden”). Useful for diagnosing disagreements between the one-shot and mechanical directions.

4.5 SCDB track — other

Method of review `scdb_first_method` *numeric*

How the appellate court reviewed the case. This field codes features like en banc and per curiam as mutually exclusive, even though they may not be so in practice, and should be used with caution. Adapted from SCDB `decisionType`; the LLM’s code list (below) is recast for the Courts of Appeals (panel / panel rehearing / en banc / per curiam) and differs from the SCDB native scale (signed opinion / per curiam / decree / equally divided vote / etc.). The SCDB-native scheme is summarized in Appendix C under `decisionType`.

- 1 = standard panel decision (typically three judges, first hearing)
- 2 = panel rehearing (same panel reconsidering its prior decision)
- 3 = en banc (full court rehearing the case)
- 4 = per curiam (unsigned opinion; may overlap with codes 1 or 3, with en banc taking priority)

Declaration of unconstitutionality `scdb_first_unconst` *numeric*

Whether and how the court struck down a statute, regulation, or government action as uncon-

stitutional. Adapted from SCDB `declarationUncon`; the LLM’s 3-code list (below) distinguishes facially-invalid from as-applied strike-downs, where the SCDB native 4-code scale instead distinguishes by type of law struck down (federal / state / municipal). The SCDB-native scheme is summarized in Appendix C under `declarationUncon`; remap when joining against raw SCDB.

- 0 = no strike-down (either no constitutional challenge was raised, or the challenge was rejected)
- 1 = facially invalid (the law itself is struck down)
- 2 = unconstitutional as-applied (struck down only as applied to this case)

Cited legal provisions `scdb_first_legal_provisions` *string*

JSON-encoded list of statutes and constitutional sections cited as controlling in the case. Parse with a standard JSON parser; each entry is a free-text citation string. Adapted from SCDB `lawType` / `lawSupp`; the LLM emits free-text citations rather than the SCDB native broad-type code (Constitution / Federal Statute / etc.) plus the 206-code specific-provision code. The SCDB-native scheme is summarized in Appendix C under `lawType`.

First-pass justification `scdb_first_justification` *string*

Free-text justification produced by the first-pass (party) coder, including postprocessing annotations (e.g., “area overridden 14→8”).

Second-pass justification `scdb_second_justification` *string*

Free-text justification produced by the second-pass (issue and outcome) coder.

4.6 Songer track — litigants

The Songer track does not separately code free-text appellant and respondent names; it produces only broad-category codes and inputs to reconstruct 5-digit subcategory codes. Subcategory digit accuracies are reported separately in the paper.

Appellant category (broad) `songer_first_appellant_category` *numeric*

Validation vs. Songer GT `genape11`: 85.7% accuracy ($N = 22,911$ matched cases).

Broad category of the first listed appellant. Original Songer codebook name: `genape11`. Code list matches the Songer convention exactly.

- 1 = private business
- 2 = private organization or association
- 3 = federal government
- 4 = sub-state government
- 5 = state government
- 6 = government level not ascertained
- 7 = natural person
- 8 = miscellaneous
- 9 = not ascertained

Respondent category (broad) `songer_first_respondent_category` *numeric*

Validation vs. Songer GT `genresp1`: 84.4% accuracy ($N = 22,904$ matched cases).

Broad category of the first listed respondent. Original Songer codebook name: `genresp1`. Same code list as `songer_first_appellant_category` plus the value 0 (not applicable, e.g., ex parte cases), matching the Songer convention exactly.

Appellant detail code `songer_second_appellant_subcategory` *numeric*

Five-digit subcategory code for the first listed appellant. Original Songer codebook name: `appel1`. Reconstructed by the LLM from the first-pass broad category and second pass detail-digit outputs on the standard Songer 5-digit scale. For natural-person appellants (broad category 7), the race and income detail digits are fixed at 0 (not coded by the LLM). For broad categories 6 and 9 (fiduciary and unidentified), the value is fixed at 69999 and 99999 respectively, and no detail variables are populated.

Respondent detail code `songer_second_respondent_subcategory` *numeric*

Five-digit subcategory code for the first listed respondent. Original Songer codebook name: `respond1`. Reconstructed analogously to `songer_second_appellant_subcategory` on the standard Songer 5-digit scale.

4.7 Songer track — disposition

The Songer track does not code disposition separately. Use `scdb_first_disposition` (§4.2); the disposition is coded only once per case (in the SCDB first pass) and shared across tracks.

4.8 Songer track — issue

Broad issue area `songer_first_broad_area` *numeric*

Validation vs. Songer GT `geniss`: 88.7% accuracy ($N = 22,667$ matched cases).

Primary issue area, on the Songer `geniss` scale. Original Songer codebook name: `geniss`. Code list matches the Songer convention exactly — code 8 is intentionally not used; miscellaneous is 9.

- 1 = criminal
- 2 = civil rights
- 3 = First Amendment
- 4 = due process
- 5 = privacy
- 6 = labor relations
- 7 = economic activity and regulation
- 9 = miscellaneous

Specific issue area `songer_second_sub_area` *numeric*

Validation vs. Songer GT `casetyp1`: 51.7% accuracy ($N = 22,657$ matched cases).

Refined Songer issue code (3-digit subject-matter classification) within the chosen issue area (`songer_first_broad_area`). Original Songer codebook name: `casetyp1`. Code list follows the Songer convention; the full list (organized by issue area), as well as two minor adjustments for LLM validation, are documented in Appendix A under `songer_gt.casetyp1` (§A.2.1).

Procedural holding `songer_first_procedural` *numeric*

1 if the holding was purely procedural; 0 otherwise. Project-derived: the Songer codebook has no direct analog (Songer encodes procedural status only indirectly via case-type sub-codes).

Type of issue `songer_first_typeiss` *numeric*

General type of issues discussed in the opinion; the same name and code list as `typeiss` in the Songer codebook.

- 0 = not ascertained
- 1 = criminal or prisoner
- 2 = civil-government
- 3 = diversity
- 4 = civil-private
- 5 = other or not applicable

The value of `songer_first_typeiss` controls which set of sub-issue flags is populated in the second-pass output (see §4.8.2).

4.8.1 Threshold issues (first pass)

The Songer first pass generates a free-response list of threshold issues (jurisdiction, standing, mootness, etc.) present in the case. For each threshold issue the LLM identifies, it codes 1, 2, or 9 for claims that favored the respondent, appellant, or were mixed (respectively). In post-processing, we add a code 0 for threshold issues *not included* by the LLM. The full four-code set and the variable names are the same as the Songer codebook’s approach; see Appendix A for a full description of each threshold issue’s meaning. Note that for approximately 75% of Songer hand-coded cases, all threshold issues have code 0.

Per-flag accuracy against the Songer GT threshold-issue variables is too granular to reproduce per-variable here; see the paper for the full per-threshold-issue performance table. Names are included below:

`songer_first_threshold_jurisdiction`, `songer_first_threshold_standing`,
`songer_first_threshold_mootness`, `songer_first_threshold_ripeness_exhaustion`,
`songer_first_threshold_timeliness`, `songer_first_threshold_immunity`,
`songer_first_threshold_failure_to_state_claim`,
`songer_first_threshold_frivolous_case`, `songer_first_threshold_political_question`,
`songer_first_threshold_other_trial_threshold`, `songer_first_threshold_late_appeal`,

`songer_first_threshold_frivolous_appeal,`
`songer_first_threshold_other_appellate_threshold`

Threshold issues *numeric*

4.8.2 Sub-issues by type of issue (second pass)

These sub-issue flags are populated according to the value of `songer_first_typeiss` coded by the first pass. Similar to threshold issues, the Songer second pass generates a free-response list of sub-issues present in the case. For each sub-issue the LLM identifies, it codes 1, 2, or 9 for claims that favored the respondent, appellant, or were mixed (respectively). In post-processing, we add a code 0 for sub-issues *not included* by the LLM. The variable names are the same as the Songer codebook's approach. The four-code set is largely the same, though Songer uses category-specific reference parties (pro-defendant for criminal sub-issues, pro-government for agency sub-issues) where the LLM uses a uniform pro-appellant convention, so codes 1 and 2 swap meanings in those families whenever the reference party is not the appellant. For some sub-issues codes 3 and 4 are also used. See Appendix A for a full description of each sub-issue's meaning.

Per-flag accuracy against the Songer GT sub-issue variables is too granular to reproduce per-variable here; see the paper for the full per-sub-issue performance table. Names are grouped by first-pass value and included below:

Criminal-procedure sub-issues (`songer_first_typeiss = 1`):

`songer_second_typeiss_prejud`, `songer_second_typeiss_insane`,
`songer_second_typeiss_improper`, `songer_second_typeiss_juryinst`,
`songer_second_typeiss_othjury`, `songer_second_typeiss_deathpen`,
`songer_second_typeiss_sentence`, `songer_second_typeiss_indict`,
`songer_second_typeiss_confess`, `songer_second_typeiss_search`,
`songer_second_typeiss_othadmis`, `songer_second_typeiss_plea`,
`songer_second_typeiss_counsel`, `songer_second_typeiss_rtcouns`,
`songer_second_typeiss_suffic`, `songer_second_typeiss_indigent`,
`songer_second_typeiss_entrap`, `songer_second_typeiss_procdis`,
`songer_second_typeiss_othcrim`.

Civil-law sub-issues (`songer_first_typeiss ∈ {2, 3, 4}`):

`songer_second_typeiss_dueproc`, `songer_second_typeiss_execord`,
`songer_second_typeiss_stpolicy`, `songer_second_typeiss_weightev`,
`songer_second_typeiss_pretrial`, `songer_second_typeiss_trialpro`,
`songer_second_typeiss_post_trl`, `songer_second_typeiss_attyfee`,
`songer_second_typeiss_judgdisc`, `songer_second_typeiss_altdisp`,
`songer_second_typeiss_injunct`, `songer_second_typeiss_summary`,
`songer_second_typeiss_fedvst`, `songer_second_typeiss_foreign`,
`songer_second_typeiss_int_law`, `songer_second_typeiss_st_v_st`,
`songer_second_typeiss_discover`, `songer_second_typeiss_othcivil`.

Government / administrative sub-issues (`songer_first_typeiss = 2`):

songer_second_typeiss_subavid, songer_second_typeiss_denovo,
songer_second_typeiss_erron, songer_second_typeiss_capric,
songer_second_typeiss_abusedis, songer_second_typeiss_judrev,
songer_second_typeiss_genstand, songer_second_typeiss_notice,
songer_second_typeiss_alj, songer_second_typeiss_agen_acq,
songer_second_typeiss_freeinfo, songer_second_typeiss_comment,
songer_second_typeiss_record.

Diversity-jurisdiction sub-issues (songer_first_typeiss = 3):
songer_second_typeiss_diverse, songer_second_typeiss_whlaws.

Sub-issues *numeric*

4.9 Songer track — ideological direction

Direction convention. All three Songer direction variables in this subsection use the project-standard 0/1/2/3 scale: 0 = not ascertained, 1 = conservative, 2 = mixed, 3 = liberal. (Not the raw SCDB decisionDirection convention; remap if merging.)

As in the SCDB track, mechanical and super-mechanical direction are derived deterministically from second-pass codings, which are now applied through Songer codebook lookup tables (see the paper for full details).

Direction (one-shot LLM) songer_first_direction_oneshot *numeric*

Validation vs. Songer GT direct1: 71.6% accuracy ($N = 20,749$ matched cases).

Ideological direction of the decision as coded directly by the first-pass LLM, on the 0/1/2/3 scale. Original Songer codebook name: **direct1** (per-case-type direction). Code list matches the Songer convention exactly.

Focal party songer_second_focal *string*

The party (e.g., **defendant**, **taxpayer**, **business**, **government**) the LLM identified as the focal party whose substantive interests the court's outcome is keyed to. Used as input to the mechanical direction.

Focal party is appellant songer_second_focal_appellant *numeric*

- 0 = focal party is the respondent
- 1 = focal party is the appellant

Focal party prevailed songer_second_focal_won *numeric*

- 0 = focal party did not prevail
- 1 = focal party prevailed in the appellate decision
- 2 = mixed outcome (e.g., affirmed in part and reversed in part such that no party clearly won)

Other focal parties considered `songer_second_focal_alt` *string*

JSON-encoded list of other parties the LLM considered as possible focal parties.

Mechanical direction `songer_second_direction_mech` *numeric*

Validation vs. Songer GT direct1: 78.9% accuracy ($N = 20,749$ matched cases).

Ideological direction derived from the second-pass focal party (`songer_second_focal`) and whether that focal party won (`songer_second_focal_won`) by applying Songer codebook rules for each issue area (`songer_first_broad_area` and `songer_second_sub_area` as appropriate). Uses the project-standard 0/1/2/3 direction convention.

Super-mechanical direction `songer_second_direction_supermech` *numeric*

Validation vs. Songer GT direct1: 80.7% accuracy ($N = 20,749$ matched cases).

Ideological direction derived similarly to `songer_second_direction_mech` but replacing the LLM's focal-party-win judgment (`songer_second_focal_won`) with the case disposition (`scdb_first_disposition`, shared across tracks) and whether the focal party was the appellant (`songer_second_focal_appellant`). Uses the project-standard 0/1/2/3 direction convention.

Direction edge-case flag `songer_second_direction_flag` *string*

Free-text annotation describing the rule applied.

4.10 Songer track — other

Basis of decision (authority) `songer_first_authority` *numeric*

The legal authority on which the decision primarily rests. The closest Songer analogs are three separate present/absent flags: `constit` (constitutional issue), `fedlaw` (federal-statute interpretation), and `procedur` (precedent / case-law interpretation, non-statutory). However, these are an imperfect comparison since they are broader and coded individually (allowing multiple flags per case), compared to these narrower and mutually exclusive codes. For validation, the LLM's authority code is binarized to match these flags; see the paper for per-flag results.

- 1 = constitutional interpretation
- 2 = statutory construction (federal statute interpretation)
- 3 = judicial review of administrative action (Chevron, arbitrary-and-capricious, substantial-evidence review, APA review)

- 4 = court rules / supervisory authority (Federal Rules of Civil / Criminal / Appellate Procedure or Evidence)
- 5 = common law or equity (judge-made law, equitable principles)
- 6 = other

Type of lower-court judgment `songer_first_appealed_from` *numeric*

Type of lower-court judgment or proceeding from which the appeal was taken. Original Songer codebook name: **applfrom**. The LLM’s code list (below) differs from the Songer codebook **applfrom** scale — the GT scale is in Appendix A under **songer_gt_applfrom**; remap when joining against ground truth.

- 1 = trial (jury or bench trial)
- 2 = injunction (grant, denial, modification, or dissolution of injunction or TRO)
- 3 = summary judgment
- 4 = guilty plea (criminal case resolved by plea)
- 5 = dismissal (district-court dismissal before trial, including habeas dismissals)
- 6 = post-judgment order (attorneys’ fees, costs, Rule 59/60(b) motions, JNOV, reconsideration)
- 7 = post-settlement order (related to a consent decree or settlement agreement)
- 8 = interlocutory appeal (non-final judgment; 28 U.S.C. § 1292, collateral order doctrine)
- 9 = mandamus (petition for writ of mandamus)
- 10 = other (pre-trial orders, motions in limine, directed verdicts, nature of judgment unclear)
- 11 = does not fit any of the above, but opinion mentions a trial judge (last-resort code)
- 12 = not applicable (direct appeal from agency / Tax Court / Board of Immigration Appeals; no district-court proceeding)

First-pass justification `songer_first_justification` *string*

Free-text justification produced by the first-pass (party) coder.

Second-pass justification `songer_second_justification` *string*

Free-text justification produced by the second-pass (issue + focal party) coder.

5. Linked external sources

In addition to the project-defined variables documented above, the database bundles in two external case-level data sources — the Songer Court of Appeals Database and the FJC Integrated Database — so that every matched case carries the external data alongside the project’s own columns. Users do not need to download either source separately to use it; the merged columns are documented in Appendices A and B, with availability per case indicated by the `songer_present` / `idb_present` flags.

The per-case crosswalk tables are also distributed as standalone CSV files alongside the main dataset, both as documentation of how each `caseid` was matched and to let users update the bundled values themselves when an external source is revised. The two crosswalks are summarised below.

5.1 Songer Court of Appeals Database

Bundled columns: `songer_gt_*` (see Appendix A)

Coverage: approximately 23,000 cases linked (99.45% of Songer cases successfully matched)

Crosswalk file: `public/songer_crosswalk.csv`

Columns: `caseid`, `songer_casenum`, `reporter`, `vol`, `beginpg`, `phase`

Join keys: `caseid` is unique within this dataset. `songer_casenum` is not strictly unique in the Songer dataset (2 duplicate pairs); join using the natural key, `reporter` + `vol` + `beginpg`.

The Songer Courts of Appeals Database contains hand-coded information for a stratified random sample of approximately 18,000 cases from 1925–1996, with an extension by Kuersten and Haire (approximately 2,000 cases; 1997–2002) and a Phase 2 SCOTUS-bridge subset (approximately 3,000 cases; every Court of Appeals cases reviewed by the Supreme Court). Every matched case in this database carries the corresponding Songer hand-coded row inline, prefixed `songer_gt_`. The `songer_present` flag identifies matched cases. The companion crosswalk file is included for users who want to inspect the matching or reapply it after a Songer release update.

For variable definitions, see Appendix A. The original and KH update codebooks (available from the project website) are the authoritative sources for these definitions.

5.2 FJC Integrated Database

Bundled columns: `idb_*` (see Appendix B)

Coverage: approximately 285,000 cases linked (97.9% of post-1970 published appellate cases successfully matched)

Crosswalk file: `public/idb_crosswalk.csv`

Columns: `caseid`, `idb_idx`, `circuit`, `docket`, `dktdate`, `reopen`, `tapeyear`, `idb.alt_rows`

Join keys: `caseid` is unique within this dataset. `idb_idx` is a unique file and row identifier; in case of file changes, use the natural key, `circuit` + `docket` + `dktdate` + `reopen` + `tapeyear`. `idb.alt_rows` is a semicolon-joined list of additional `idb_idx` values for the approximately 1,000 caseids whose match intentionally produced more than one candidate IDB row (consolidated cases in this database that are split by IDB). Two important caveats:

Intentional many-to-one. For about 11,000 of our `caseids` (about 4% of matched cases), our database splits a case that IDB consolidates, so multiple `caseids` intentionally point to the same `idb_idx`.

IDB-internal duplicates. The 2008–present IDB file has no natural key duplicates; the 1971–2007 IDB file has eight duplicated natural keys. Five of those eight natural key groups are

not used by our database. Two groups represent consolidated cases in our database that IDB split; the duplicate rows are recorded in `idb_alt_rows`. The one remaining group contains two records of a single case with different recorded outcomes; `idb_idx` resolves to exactly one record even though the natural key does not.

The Federal Judicial Center’s Integrated Database (IDB) provides administrative records on all federal appellate cases since 1971. Every matched case in this database carries the corresponding IDB row inline, prefixed `idb_`. The `idb_present` flag identifies matched cases. The companion crosswalk file maps each `caseid` to its IDB docket-number/circuit pair and is included for users who want to inspect the matching or reapply it after an IDB release update.

For variable definitions, see Appendix B. The IDB codebooks (separate codebooks for 1971–2007 and 2008-forward; available from the project website) are the authoritative sources.

A. Songer ground truth (`songer_gt_*`)

This appendix documents the variables in the Songer Courts of Appeals Database that are merged into the dataset under the prefix `songer_gt_`. Definitions follow the Songer codebook (Songer 1996), reproduced here for convenience; era differences between the original database (1925–1996), the Kuersten–Haire update (1997–2002), and the Phase 2 SCOTUS-bridge cases (approximately 2,700 cases) are flagged in coverage notes where they affect interpretation.

A.1 Identifiers and dates

Federal Reporter volume `songer_gt_vol` *numeric* [Source: Songer codebook]

Federal Reporter volume number in which the case is published.

Beginning page `songer_gt_beginpg` *numeric* [Source: Songer codebook]

Page on which the case begins in the Federal Reporter.

End page of opinion `songer_gt_endopin` *numeric* [Source: Songer codebook]

Last page of the majority opinion in the Federal Reporter.

End page of case `songer_gt_endpage` *numeric* [Source: Songer codebook]

Last page of the case, including any dissent, concurrence, or appendix.

Page length `songer_gt_pagelgth` *numeric* [Source: Project-derived]

Page count of the case, computed as `endpage – beginpg`.

Year of decision `songer_gt_year` *numeric* [Source: Songer codebook]

Four-digit year in which the appellate decision was announced.

Month of decision `songer_gt_month` *numeric* [Source: Songer codebook]

Month of decision (1–12).

Day of decision `songer_gt_day` *numeric* [Source: Songer codebook]

Day of month on which the decision was announced.

Decision date `songer_gt_decisiondate` *string* [Source: Project-derived]

Decision date in ISO format, assembled from year, month, and day.

Circuit `songer_gt_circuit` *numeric* [Source: Songer codebook]

Circuit of the deciding court (00 = DC, 01–13 = by circuit number).

Citation `songer_gt_cite` *string* [Source: Songer codebook]

Nine-character Federal Reporter citation (4-digit volume, slash, 4-digit page).

Case number `songer_gt_casenum` *numeric* [Source: Songer codebook]

Sequential case identifier assigned by the Songer coders. Phase 1 (1925–1996) uses values 1–15315; Phase 2 (SCOTUS-bridge) uses 50121+; Phase 3 (Kuersten–Haire update, 1997–2002) starts at 20001.

Not strictly unique. The KH update contains two casenum collisions — 21006 and 21140 each carry two unrelated cases (different circuits, dockets, decision dates, citations, and substantive coding for both pairs), apparently a KH coder error copying the previous row’s casenum when starting a new case. Both rows survive in the combined dataset for fidelity to the source. To get a guaranteed-unique join key on `songer_gt_*` data, use (`songer_gt_casenum`, `songer_gt_vol`, `songer_gt_beginpg`) or the project’s own `caseid`.

Docket number `songer_gt_docnum` *string* [Source: Songer codebook]

Docket number of the first case decided by the opinion (2-digit year, zero, up to 5-digit number).

Docket number (legacy alias) `songer_gt_docknum` *string* [Source: Songer codebook]

Legacy alias of `songer_gt_docnum`, always blank. Use the canonical column for new analyses.

Multiple docket count `songer_gt_multdoc` *numeric* [Source: Songer codebook]

Number of docket numbers consolidated in the case coded. KH-update field.

Coverage note. This field was added in the Kuersten–Haire update and is missing in Phase 1 cases.

Multi-citation flag `songer_gt_multiple` *numeric* [Source: Project metadata]

Project-derived flag indicating that the case has multiple citations. Not part of the Songer codebook.

Prior published opinion `songer_gt_priorpub` *string* [Source: Songer codebook]

Citation of the most recent prior published opinion (from another court, or from a prior CoA decision in the same case). Missing if none.

File identifier `songer_gt_fileid` *string* [Source: Project metadata]

Internal file or source identifier. Not part of the Songer codebook.

Sampling stratum `songer_gt_sample` *string* [Source: Songer codebook]

Songer's original sampling-stratum indicator. The column is empty in every row of the released Songer files (original, KH update, and Phase 2 alike), so it carries no information here; we keep it for fidelity to the Songer schema. Use `songer_gt_phase` (below) for the project-derived source-provenance label, or `songer_gt_source` (next entry) for the original Songer codebook's pre-appeals-forum field.

Source forum `songer_gt_source` *numeric* [Source: Songer codebook]

Identifies the forum that heard the case immediately before it came to the court of appeals (Songer codebook Field 16). Often the same as `songer_gt_origin`.

- 1 = federal district court (single judge)
- 2 = three-judge district court
- 3 = state court
- 4 = bankruptcy court or referee in bankruptcy
- 5 = federal magistrate
- 6 = federal administrative agency
- 7 = Court of Customs & Patent Appeals
- 8 = Court of Claims
- 9 = Court of Military Appeals
- 10 = Tax Court or Tax Board
- 11 = administrative law judge
- 12 = U.S. Supreme Court (remand)
- 13 = special DC court (not the U.S. District Court for D.C.)
- 14 = earlier appeals-court panel
- 15 = other
- 16 = not ascertained
- 17 = bankruptcy appellate panel (BAP) — KH-update addition

Coverage note. Code 17 (BAP) was added in the Kuersten-Haire 1997–2002 update; absent from earlier vintages.

Source phase `songer_gt_phase` *string* [Source: Project metadata]

The Songer dataset phase(s) from which this row was assembled. Single-source rows take the bare phase name (`original`, `kh`, or `phase2`); rows whose (`vol`, `beginpg`) key appears in more than one source list every contributor, alphabetized and joined with `+`: `original+phase2`, `kh+original`, `kh+original+phase2`. Use `regexm(songer_gt_phase, "phase2")` or `split` on `+` to filter on a specific contributor. Distinct from `songer_gt_source` (above), which is the Songer codebook's pre-appeals-forum field.

Merge audit log `songer_gt_merge_log` *string* [Source: Project metadata]

Audit log noting any field-level edits applied when the project assembled the combined Songer file. Useful for diagnosing discrepancies between the original Songer release and the merged version used here.

A.2 Issue area, type, and authority

Primary case-type code `songer_gt_casetyp1` *numeric* [Source: Songer codebook]

Three-digit subject-matter classification within eight general categories (criminal, civil rights, First Amendment, due process, privacy, labor, economic activity, miscellaneous). 220 specific codes; the full list (organized by issue area) appears in §A.2.1 below.

Second case-type code `songer_gt_casetyp2` *numeric* [Source: Songer codebook]

A second case-type classification when a case has two case types. Coded on the same scale as `songer_gt_casetyp1`; not necessarily less important than the primary code. See §A.2.1 for the full code list.

A.2.1 Three-digit `casetyp1` subcodes

The 3-digit `songer_gt_casetyp1` (and analogously `songer_gt_casetyp2`) codes are organized by the first digit, which encodes the same broad issue area as `songer_gt_geniss`. Note that the codes below are the ones used by the LLM prompt for `songer_second_sub_area`; the original Songer codebook includes a handful of extension codes not enumerated here.

Issue area 1: Criminal

- 101 murder
- 102 rape (the KH-update guidance instructs coders to record sexual assault under this code as well)
- 103 arson
- 104 aggravated assault
- 105 robbery
- 106 burglary
- 107 auto theft
- 108 larceny (over \$50)
- 109 other violent crimes
- 110 narcotics
- 111 alcohol related crimes
- 112 tax fraud
- 113 firearm violations
- 114 morals charges (gambling, prostitution, obscenity)
- 115 criminal violations of government regulations of business
- 116 other white collar crime (embezzlement, computer fraud, bribery)
- 117 other federal crimes

- 118 federal offense, specific crime not ascertained
- 119 RICO

Issue area 2: Civil Rights

- 201 suit for damages for false arrest, false confinement, malicious prosecution
- 202 cruel and unusual punishment
- 203 due process rights in prison
- 204 denial of other rights of prisoners (42 USC 1983)
- 205 denial or revocation of parole (due process grounds)
- 206 other denial or revocation of parole
- 207 other prisoner petitions
- 208 excessive force used in arrest
- 209 other civil rights violations alleged by criminal defendants
- 210 voting rights - reapportionment & districting
- 211 participation rights - ballot access
- 212 voting rights - other (includes race discrimination in voting)
- 213 desegregation of schools
- 214 other desegregation
- 221 employment race discrimination - alleged by minority
- 222 other race discrimination - alleged by minority
- 223 employment race discrimination - alleged by Caucasian (reverse discrimination)
- 224 other reverse race discrimination claims
- 231 employment sex discrimination - alleged by woman
- 232 pregnancy discrimination
- 233 other sex discrimination - alleged by woman
- 234 employment sex discrimination - alleged by man (reverse sex discrimination)
- 235 other sex discrimination - alleged by man
- 239 42 USC 1983 claims based on race or sex discrimination
- 241 alien petitions (includes deportation disputes)
- 251 Indian rights (civil rights claims under Indian law only)
- 261 juveniles
- 271 poverty law, rights of indigents
- 281 rights of handicapped / ADA
- 282 age discrimination
- 283 discrimination based on religion or nationality
- 284 discrimination based on sexual preference
- 285 AIDS/HIV discrimination
- 290 challenge to federal government hiring/firing/promotion (not above categories)
- 291 other 14th Amendment and civil rights act cases
- 299 other civil rights

Issue area 3: First Amendment

- 301 commercial speech
- 302 libel, slander, defamation
- 303 free exercise of religion
- 304 establishment of religion (other than aid to parochial schools)

305 aid to parochial schools
306 press
307 obscenity
308 association
309 federal internal security / communist control acts / loyalty oaths / security risks
310 expression protesting race discrimination (speeches, parades, picketing)
311 overt acts opposing war and the military
312 conscientious objection to military service
313 expression of political/social beliefs conflicting with regulation of physical activity
314 threats to peace, safety, order (fighting words, incitement)
315 challenges to campaign spending limits
399 other (includes tests of belief)

Issue area 4: Due Process

410 denial of fair hearing or notice - government employees
411 denial of hearing or notice in non-employment context
412 taking clause (5th or 14th Amendment)
413 freedom of information act and other access-to-information claims
499 other due process issues

Issue area 5: Privacy

501 abortion rights
502 homosexual rights where privacy claim raised
503 contraception and other marital/sexual privacy
504 suits for compensation for violation of privacy rights (e.g., 1983 suits)
505 mandatory testing (drugs, AIDS, etc.)
506 mandatory sterilization
507 right to die or right to refuse medical help
599 other

Issue area 6: Labor Relations

601 union organizing
602 unfair labor practices
603 Fair Labor Standards Act (FLSA) — minimum wage, overtime, child labor,
603 Fair Labor Standards Act issues
604 Occupational Safety and Health Act (OSHA)
604 Occupational Safety and Health Act (OSHA) — workplace safety violations
605 collective bargaining
606 conditions of employment
606 conditions of employment (non-FLSA) — use for workplace conditions
607 employment of aliens
608 which union has right to represent workers
609 non-civil-rights grievances by worker against union
610 other labor relations
611 Family Medical Leave Act (FMLA)
612 ERISA

Issue area 7: Economic Activity and Regulation

- 701 state or local tax
- 702 federal taxation - individual income tax
- 703 federal tax - business income tax
- 704 federal tax - excess profits
- 705 federal estate and gift tax
- 706 federal tax - other
- 710 patents
- 711 copyrights
- 712 trademarks
- 713 trade secrets, personal intellectual property
- 720 motor vehicle
- 721 airplane
- 722 product liability
- 723 Injuries to dockworkers and longshoremen under FELA or the Longshore
- 723 federal employer liability; injuries to dockworkers/longshoremen
- 724 other government tort liability
- 725 workers compensation
- 726 medical malpractice
- 727 other personal injury
- 728 fraud
- 729 other property damage
- 730 other torts
- 731 contract disputes - general (private parties)
- 732 disputes over government contracts
- 733 insurance disputes
- 734 debt collection, disputes over loans
- 735 consumer disputes with retail business or service providers
- 736 breach of fiduciary duty; franchise agreements
- 737 contract disputes - validity of contract
- 738 commerce clause challenges to state/local government
- 739 other contract disputes
- 740 private economic disputes (other than contract)
- 741 Individual bankruptcy — Chapter 7 for individuals, personal bankruptcy
- 741 bankruptcy - private individual (e.g., chapter 7)
- 742 Business reorganization — Chapter 11, corporate restructuring, business
- 742 bankruptcy - business reorganization (e.g., chapter 11)
- 743 Other bankruptcy — use ONLY when neither 741 nor 742 applies. Examples:
- 743 other bankruptcy
- 744 antitrust - brought by individual or private business
- 745 antitrust - brought by government
- 746 regulation of mergers (non-antitrust grounds)
- 747 securities - private party conflicts
- 748 government regulation of securities
- 750 social security benefits (including SS disability)
- 751 other government benefit programs (welfare, veterans, food stamps)
- 752 state or local economic regulation (non-pollution)

753 federal environmental regulation (non-pollution)
 754 federal consumer protection regulation
 755 rent control; government price controls
 756 federal regulation of transportation
 757 oil, gas, mineral regulation
 758 federal regulation of utilities (telephone, radio, TV, power)
 759 other commercial regulation by federal government
 760 civil RICO suits
 761 Admiralty personal injury under general maritime law — Jones Act seamen,
 761 admiralty - personal injury
 762 admiralty - seamen's wage disputes
 763 admiralty - maritime contracts, charter contracts
 764 admiralty - other
 765 state or local pollution control
 766 federal pollution control
 770 disputes over real property (private)
 771 eminent domain and disputes with government over real property
 772 landlord-tenant disputes
 773 government seizure of property - criminal enforcement
 774 government seizure of property - civil (taxes, liens)
 780 legal malpractice
 799 other economic activity

Issue area 9: Miscellaneous

901 miscellaneous interstate conflict
 902 federalism (preemption, state vs. federal power)
 903 attorneys (disbarment, discipline)
 904 selective service or draft issues
 905 challenge to authority of magistrates, special masters
 906 challenge to authority of bankruptcy judges/referees
 910 Indian law - criminal verdict under tribal law
 911 Indian law - commercial disputes under Indian law/treaties
 912 Indian law - claims acts and real property disputes
 913 Indian law - federal regulation of Indian land
 914 Indian law - state/local authority over Indian land
 915 Indian law - tribal regulation of economic activities
 916 other Indian law
 920 international law
 921 immigration (except civil rights claims)
 922 Patriot Act
 923 other national security
 924 14th Amendment - congressional power to enforce
 925 executive privilege
 999 other

Federal / state / jurisdiction-unclear collapse for validation In the criminal area, the Songer codebook uses parallel ranges for federal versus state offenses: codes 121–138 are state-

offense parallels of 101–118 (121 = state murder, 132 = state tax fraud, etc.), and codes 141–158 are jurisdiction-unclear parallels of the same. The LLM prompt enumerates all three ranges and instructs the model to pick the federal / state / unclear sub-digit; in practice the model most often falls back to the federal code. For validation against Songer GT, we collapse all three ranges to the federal base (1xx) on both the LLM-predicted side and the GT side, so the model is credited for getting the offense type right regardless of jurisdiction sub-digit.

Kuersten–Haire-update casetyp1 codes (era-aware validation) Eleven codes in the LLM prompt were added to the Songer casetyp1 list in the Kuersten–Haire (KH) update — they are not in the original Phase-1 (1925–1996) codebook but are in the KH-update codebook. In the project’s Songer ground-truth data, all 56 cases carrying any of these codes have `songer_gt_phase = "kh"` (year 1997–2002); Phase-1 and Phase-2 cases never carry them, because the human coder used the older codebook.

Validation handles this era split asymmetrically. For KH-era GT cases, the LLM and the human coder share the same codebook, so we expect direct exact match and apply no remap. For pre-KH cases (Phase-1 *original* and Phase-2 SCOTUS-bridge), the LLM is following the KH codebook but the GT cannot use a KH-only code; we remap the LLM-emitted KH-only code to its closest in-Phase-1-codebook analogue — the value a Phase-1 human coder would most plausibly have used — before comparison. The map is:

119	RICO — pre-KH analogue: 117 (other federal crimes)
285	AIDS/HIV discrimination — pre-KH analogue: 281 (handicapped / ADA)
611	Family Medical Leave Act (FMLA) — pre-KH analogue: 603 (Fair Labor Standards Act issues)
612	ERISA — pre-KH analogue: 610 (other labor relations)
765	state or local pollution control — pre-KH analogue: 752 (state or local economic regulation)
766	federal pollution control — pre-KH analogue: 753 (federal environmental regulation)
780	legal malpractice — pre-KH analogue: 730 (other torts)
922	Patriot Act — pre-KH analogue: 920 (international law)
923	other national security — pre-KH analogue: 920 (international law)
924	14th Amendment, congressional power to enforce — pre-KH analogue: 902 (federalism)
925	executive privilege — pre-KH analogue: 902 (federalism)

Rare codes in the Songer ground truth not enumerated in the codebook A small number of Songer GT cases (1–3 cases each) carry case-type codes that do not appear in either the original Songer codebook or the Kuersten–Haire update — likely legacy codes, typos, or unrecorded extensions: 10, 16, 99, 242, 294, 392, 449, 693, 695, 716, 749, 1110, 1223, 1703, 2999, 7500. The four-digit codes are especially likely to be data-entry artifacts (the `casetyp1` field is defined as three columns wide; each four-digit value resembles a valid 3-digit code with one extra digit smeared on, e.g. 1110 $\approx$ 111, 7500 $\approx$ 750). Rows carrying any of these codes are dropped from `casetyp1` validation denominators because the LLM cannot emit them by construction.

General issue summary `songer_gt_geniss` *numeric* [Source: Songer codebook]

Eight-category general issue summary of `songer_gt_casetyp1`.

1 = criminal

- 2 = civil rights
- 3 = First Amendment
- 4 = due process
- 5 = privacy
- 6 = labor relations
- 7 = economic activity and regulation
- 9 = miscellaneous

Note: code 8 is not used in this variable.

Type of issue `songer_gt_typeiss` *numeric* [Source: Songer codebook]

General category of issues discussed in the opinion.

- 0 = not ascertained
- 1 = criminal or prisoner
- 2 = civil-government
- 3 = diversity
- 4 = civil-private
- 5 = other or not applicable

Coverage note. The Kuersten–Haire update extended code 5 to cover cases that involve both civil and criminal issues; users analyzing KH-update cases should expect code 5 to be more inclusive than in the original Phase 1 data.

Type of issue (legacy alias) `songer_gt_isstyp` *numeric* [Source: Project metadata]

Legacy alias of `songer_gt_typeiss`. Use `songer_gt_typeiss` for new analyses.

Administrative review `songer_gt_adminrev` *numeric* [Source: Songer codebook]

Identifier of the federal regulatory agency (if any) whose decision was reviewed. Codes 1–14 are specific agencies (e.g., 1 = Benefits Review Board, 14 = other federal agency); 15 indicates the case did not involve agency review. See `value_codes_songer_gt_adminrev.csv` for the full agency list.

Court of origin `songer_gt_origin` *numeric* [Source: Songer codebook]

Type of court that made the original decision being appealed. Cases removed from state court are coded as originating in federal district court.

- 1 = federal district court (single judge)
- 2 = three-judge district court
- 3 = state court (includes habeas corpus petitions after conviction in state court; also petitions from territorial courts other than U.S. District Courts)
- 4 = bankruptcy court, referee in bankruptcy, or special master
- 5 = federal magistrate
- 6 = federal administrative agency

- 7 = special DC court (not the U.S. District Court for D.C.)
- 8 = other (e.g., Tax Court, court martial)
- 9 = not ascertained
- 10 = Tax Court (KH-update addition; split out from the original code 8)
- 11 = Administrative Law Judge (KH-update addition)

Coverage note. Codes 10 (Tax Court) and 11 (ALJ) were added in the Kuersten–Haire update to break out cases that the original codebook absorbed into code 8; they do not appear in Phase 1 cases.

District-court decision appealed from `songer_gt_applfrom` *numeric* [Source: Songer codebook]

Type of district-court decision or judgment from which the appeal was taken. Coded as “not applicable” when there was no prior district-court action.

- 1 = trial (jury or bench trial)
- 2 = injunction, denial of injunction, or stay of injunction
- 3 = summary judgment or denial of summary judgment
- 4 = guilty plea or denial of motion to withdraw plea
- 5 = dismissal (includes dismissal of a habeas-corpus petition)
- 6 = post-judgment orders (e.g., attorneys’ fees, costs, damages, JNOV); also recorded for criminal cases whose sole appellate issue is the sentence (KH-update convention)
- 7 = appeal of post-settlement orders
- 8 = interlocutory appeal (not a final judgment)
- 9 = mandamus (not a final judgment)
- 10 = other (e.g., pre-trial orders, rulings on motions, directed verdicts) or could not determine
- 11 = does not fit any of the above, but opinion mentions a “trial judge”
- 12 = not applicable (e.g., decision below was from an agency or the Tax Court)

Coverage note. The KH-update codebook extends code 6 to cover criminal cases whose only appellate issue is the sentence; in earlier vintages those cases would typically be coded elsewhere.

Nature of the proceeding `songer_gt_method` *numeric* [Source: Songer codebook]

Nature of the appellate proceeding for the citation coded — indicating whether the case has prior appellate history.

- 1 = decided by panel for first time (no indication of rehearing or remand)
- 2 = decided by panel after rehearing
- 3 = decided by panel after Supreme Court remand
- 4 = en banc after a single panel decision
- 5 = en banc after multiple panel decisions
- 6 = en banc with no prior panel decisions
- 7 = decided by panel after remand to a lower court
- 8 = other
- 9 = not applicable

Court of appeals disposition `songer_gt_treat` *numeric* [Source: Songer codebook]

Disposition by the court of appeals.

- 0 = stay, petition, or motion granted
- 1 = affirmed
- 2 = reversed
- 3 = reversed and remanded
- 4 = vacated and remanded
- 5 = affirmed in part and reversed in part (or vice versa)
- 6 = affirmed in part, reversed in part, and remanded
- 7 = vacated
- 8 = petition denied or appeal dismissed
- 9 = certification to another court
- 10 = not ascertained
- 11 = affirmed, vacated, and remanded (KH-update only)

Coverage note. Code 11 (affirmed, vacated, and remanded) was added in the Kuersten–Haire update and does not appear in Phase 1 cases.

Party initiating the appeal `songer_gt_initiate` *numeric* [Source: Songer codebook]

Which party initiated the appeal.

- 1 = original plaintiff
- 2 = original defendant
- 3 = federal agency representing plaintiff
- 4 = federal agency representing defendant
- 5 = intervenor
- 8 = not applicable
- 9 = not ascertained

Class action `songer_gt_classact` *numeric* [Source: Songer codebook]

- 0 = not a class action
- 1 = case described as a class action

Habeas corpus `songer_gt_habeas` *numeric* [Source: Songer codebook]

Whether the case was an appeal of a habeas corpus decision.

- 0 = not a habeas appeal
- 1 = state habeas (criminal)
- 2 = federal habeas (criminal)
- 3 = federal habeas concerning deportation

Cross-appeals `songer_gt_crossapp` *numeric* [Source: Songer codebook]

Whether cross-appeals were filed and consolidated in the present case.

- 0 = no cross-appeals
- 1 = cross-appeals filed and consolidated
- 2 = not ascertained

Intervenor `songer_gt_interven` *numeric* [Source: Songer codebook]

- 0 = no intervenor
- 1 = intervenor on the appellant's side
- 2 = intervenor on the respondent's side
- 3 = intervenors on both sides
- 9 = not applicable

A.3 Direction, opinion, and voting

Direction of decision on casetyp1 `songer_gt_direct1` *numeric* [Source: Songer codebook]

Ideological direction of the decision on the primary case type (`songer_gt_casetyp1`). Uses the standard 0/1/2/3 convention; see the direction-convention callout in the front matter.

Direction of decision on casetyp2 `songer_gt_direct2` *numeric* [Source: Songer codebook]

Ideological direction on the secondary case type (`songer_gt_casetyp2`), where applicable. Same coding as `songer_gt_direct1`.

Majority votes `songer_gt_majvotes` *numeric* [Source: Songer codebook]

Number of judges voting with the majority disposition. Typically 2 or 3 on a panel; up to 15 in en banc decisions. Judges who concurred in result only are counted with the majority.

Number of dissents `songer_gt_dissent` *numeric* [Source: Songer codebook]

Number of judges who dissented from the majority disposition, with or without a written opinion. Judges who dissented in part are counted as dissenting.

Number of concurrences `songer_gt_concur` *numeric* [Source: Songer codebook]

Number of judges who wrote, joined, or filed a concurring opinion, or concurred in result only.

Unanimous decision indicator `songer_gt_unan` *numeric* [Source: Project-derived]

1 if the decision was unanimous (no dissents and no concurrences); 0 otherwise. Project-derived

from `songer_gt_dissent` and `songer_gt_concur`.

Opinion type `songer_gt_opinstat` *numeric* [Source: Songer codebook]

Whether the opinion was signed by an individual author or issued *per curiam*.

- 1 = signed opinion with reasons
- 2 = per curiam opinion with reasons
- 9 = not ascertained

Signed-opinion indicator `songer_gt_signed` *numeric* [Source: Project-derived]

1 if the opinion was signed (`songer_gt_opinstat == 1`); 0 otherwise.

Supreme Court review indicator `songer_gt_supct` *numeric* [Source: Project-derived]

1 if the case was subsequently reviewed by the Supreme Court; 0 otherwise. Derived from membership in the Songer SCOTUS-bridge Phase 2 file, which by construction enumerates every Courts-of-Appeals case subsequently reviewed by the Supreme Court. Equivalent to `regexm(songer_gt_phase, "phase2")`.

Sanctions `songer_gt_sanction` *numeric* [Source: Songer codebook]

Whether sanctions were imposed by the court of appeals.

- 0 = no sanctions
- 1 = sanctions on the appellant
- 2 = sanctions on the respondent
- 3 = sanctions on both parties
- 4 = not ascertained

A.4 Panel judges and votes

Panel judge codes `songer_gt_codejN` *numeric* [Source: Songer codebook]

For each $N \in \{1, \dots, 15\}$: Songer judge identification code for panel-judge slot N . Slots 4–15 accommodate en banc panels. Empty for unfilled slots. The full code list is provided in `value_codes_songer_judge.csv`; the Songer codebook includes the corresponding alphabetical list as Appendix 3.

District-court judge code `songer_gt_distjudg` *numeric* [Source: Songer codebook]

Five- or six-digit code for the federal district judge who heard the case below. 99999 if unknown. The full code list is provided in `value_codes_songer_district_judge.csv`; the Songer codebook includes the corresponding alphabetical list as Appendix 4.

State of origin `songer_gt_state` *numeric* [Source: Songer codebook]

Two-digit state code (00–57) for the state in which the case was first heard. For federal cases, the state of the district court; for habeas cases, the state of the state court that first heard the case. Not applicable if the case originated in a federal administrative agency.

See `value_codes_songer_gt_state.csv` for the full list of state codes.

Coverage note. Code 57 = Northern Mariana Islands is a KH-update addition; absent from Phase 1 cases.

State code (legacy alias) `songer_gt_statecode` *numeric* [Source: Songer codebook]

Legacy alias of `songer_gt_state`. Use the canonical column for new analyses.

District within state `songer_gt_district` *numeric* [Source: Songer codebook]

District within the state from which the case originated.

- 0 = not applicable
- 1 = eastern
- 2 = western
- 3 = central
- 4 = middle
- 5 = southern
- 6 = northern
- 7 = whole-state district
- 8 = not ascertained

Judge majority indicators on casetyp1 `songer_gt_jNmaj1` *numeric* [Source: Songer codebook]

For each $N \in \{2, \dots, 16\}$: whether the judge in slot N voted with the majority on the primary case type (`songer_gt_casetyp1`).

- 1 = voted with the majority
- 2 = dissented
- 3 = voted with majority but authored a concurrence on this issue (KH-update extension)
- 4 = voted with majority but joined a concurrence on this issue (KH-update extension)
- 5 = concurred in result only (KH-update extension)

Coverage note. Codes 3, 4, and 5 were added in the Kuersten–Haire update. Phase 1 cases use only codes 1 and 2.

Judge majority indicators on casetyp2 `songer_gt_jNmaj2` *numeric* [Source: Songer codebook]

For each $N \in \{2, \dots, 15\}$: same as `jNmaj1` but for the secondary case type (`songer_gt_casetyp2`).

Judge vote direction on casetyp1 `songer_gt_jNvote1` *numeric* [Source: Songer codebook]

For each $N \in \{2, \dots, 16\}$: vote of judge N on the primary case type, using the project-wide direction convention (0 = not ascertained, 1 = conservative, 2 = mixed, 3 = liberal). See the direction-convention callout in the front matter.

Judge vote direction on casetyp2 `songer_gt_jNvote2` *numeric* [Source: Songer codebook]

For each $N \in \{2, \dots, 15\}$: same as `jNvote1` but for the secondary case type.

A.5 Concur and dissent basis (KH-update fields)

The following fields characterize the substantive basis of concurring and dissenting opinions. All were added in the Kuersten-Haire update; they are missing in Phase 1 cases.

Concur-merits basis (first concurrence) `songer_gt_cmbasis` *numeric* [Source: Songer codebook]

Basis of the first concurring opinion as it relates to the merits of the majority's reasoning.

- 0 = did not address the merits
- 1 = addressed the merits and disagreed with the majority's reasoning
- 2 = addressed the merits and agreed with the majority's reasoning
- 3 = addressed merits not discussed in the majority opinion

Coverage note. KH-update field; missing in Phase 1 cases.

Concur-merits basis (second concurrence) `songer_gt_cmbasis2` *numeric* [Source: Songer codebook]

Same as `songer_gt_cmbasis` but for the second concurring opinion, where one exists.

Coverage note. KH-update field; missing in Phase 1 cases.

Concur-procedural basis (first concurrence) `songer_gt_cpbasis` *numeric* [Source: Songer codebook]

Basis of the first concurring opinion as it relates to the procedural aspects of the case. Coding parallels `songer_gt_cmbasis`.

Coverage note. KH-update field; missing in Phase 1 cases.

Concur-procedural basis (second concurrence) `songer_gt_cpbasis2` *numeric* [Source: Songer codebook]

Same as `songer_gt_cpbasis` but for the second concurring opinion.

Coverage note. KH-update field; missing in Phase 1 cases.

Dissent-merits basis `songer_gt_dmbasis` *numeric* [Source: Songer codebook]

Basis of the dissenting opinion as it relates to the merits.

- 0 = did not address the merits
- 1 = addressed the merits and disagreed with the majority's reasoning
- 2 = addressed the merits and agreed with the majority's reasoning
- 3 = addressed merits not discussed in the majority opinion

Coverage note. KH-update field; missing in Phase 1 cases.

Dissent-procedural basis `songer_gt_dpbasis` *numeric* [Source: Songer codebook]

Basis of the dissenting opinion as it relates to procedure. Coding parallels `songer_gt_dmbasis`.

Coverage note. KH-update field; missing in Phase 1 cases.

A.6 Parties: counts and detailed codes

Appellant category (first listed) `songer_gt_genapel1` *numeric* [Source: Songer codebook]

Broad classification of the first listed appellant.

- 1 = private business
- 2 = private organization or association
- 3 = federal government
- 4 = sub-state government
- 5 = state government
- 6 = government level not ascertained
- 7 = natural person
- 8 = miscellaneous
- 9 = not ascertained

Appellant category (second listed) `songer_gt_genapel2` *numeric* [Source: Songer codebook]

Broad classification of the second listed appellant, populated only when distinct from `songer_gt_genapel1`. Same coding.

Respondent category (first listed) `songer_gt_genresp1` *numeric* [Source: Songer codebook]

Broad classification of the first listed respondent. Same nine categories as `songer_gt_genapel1`, plus 0 = not applicable (e.g., ex parte cases).

Respondent category (second listed) `songer_gt_genresp2` *numeric* [Source: Songer codebook]

Broad classification of the second listed respondent, populated only when distinct from `songer_gt_genresp1`. Same coding.

Appellant detail code (first listed) `songer_gt_appell1` *numeric* [Source: Songer codebook]

Five-digit detailed code for the first listed appellant, packing the broad category from `songer_gt_genappell1` with sub-category detail digits. See `value_codes_songer_gt_appell1.csv` for the full five-digit code list.

Appellant detail code (second listed) `songer_gt_appell2` *numeric* [Source: Songer codebook]

Five-digit detailed code for the second listed appellant, populated only when distinct from `songer_gt_appell1`.

Respondent detail code (first listed) `songer_gt_respond1` *numeric* [Source: Songer codebook]

Five-digit detailed code for the first listed respondent. Uses the same scheme as `songer_gt_appell1`.

Respondent detail code (second listed) `songer_gt_respond2` *numeric* [Source: Songer codebook]

Five-digit detailed code for the second listed respondent, populated only when distinct from `songer_gt_respond1`.

Real appellant indicator `songer_gt_realapp` *numeric* [Source: Songer codebook]

Whether the formally listed appellants are the real parties in interest.

- 0 = both listed appellants are real (or only one is listed and is real)
- 1 = the first listed appellant is not the real party
- 2 = the second listed appellant is not the real party
- 3 = neither listed appellant is the real party
- 4 = not ascertained

Real respondent indicator `songer_gt_realresp` *numeric* [Source: Songer codebook]

Same as `songer_gt_realapp` but for respondents.

Number of appellants `songer_gt_numappel` *numeric* [Source: Songer codebook]

Total number of appellants in the case. Coded 99 if the case lists “et al.” without specifying.

Number of respondents `songer_gt_numresp` *numeric* [Source: Songer codebook]

Total number of respondents in the case. Coded 99 if unspecified.

Appellant counts by category `songer_gt_appbus, songer_gt_appfed, songer_gt_appfiduc, songer_gt_appnatpr, songer_gt_appnonp, songer_gt_appstate, songer_gt_appsubst` *numeric* [Source: Songer codebook]

Counts of appellants by broad category: private businesses (**appbus**), federal-government agencies or officials (**appfed**), fiduciaries or trustees (**appfiduc**), natural persons (**appnatpr**), non-profit groups or associations (**appnonp**), state-government agencies or officials (**appstate**), and sub-state governments (**appsubst**). Each is a count, not an indicator.

Respondent counts by category `songer_gt_r_bus, songer_gt_r_fed, songer_gt_r_fiduc, songer_gt_r_natpr, songer_gt_r_nonp, songer_gt_r_state, songer_gt_r_subst` *numeric* [Source: Songer codebook]

Counts of respondents by broad category, paralleling the appellant counts above.

State of first state/local appellant `songer_gt_app_stid` *numeric* [Source: Songer codebook]

State code (per the Songer state list) of the first listed state-government or local-government appellant. 0 if no state or local appellant is listed. The Songer source .dta files abbreviate this field name as `ap_stid` (single “p”); we rename it on load to match the codebook’s documented `APP_STID`.

Appellant state code (legacy alias) `songer_gt_astate_code` *numeric* [Source: Songer codebook]

Legacy alias of `songer_gt_app_stid`. Use the canonical column for new analyses.

State of first state/local respondent `songer_gt_r_stid` *numeric* [Source: Songer codebook]

State code of the first listed state-government or local-government respondent.

Respondent state code (legacy alias) `songer_gt_resstate` *numeric* [Source: Songer codebook]

Legacy alias of `songer_gt_r_stid`. Use the canonical column for new analyses.

Appellant bankruptcy indicators `songer_gt_bank_ap1, songer_gt_bank_ap2` *numeric* [Source: Songer codebook]

Whether the first and second listed appellants were bankrupt at the time of the appeal.

- 1 = bankrupt
- 2 = not bankrupt (presumed if not mentioned)

Respondent bankruptcy indicators `songer_gt_bank_r1, songer_gt_bank_r2` *numeric* [Source: Songer codebook]

Same as `bank_ap1/bank_ap2` but for the first and second listed respondents.

Government-as-party indicator `songer_gt_govt` *numeric* [Source: Project-derived]

Project-derived indicator for whether the case involves the government as a party. Not part of the Songer codebook.

A.7 Counsel, attorneys, firms, and amicus briefs

Most fields in this subsection were added in the Kuersten–Haire update. Expect them missing in Phase 1 (1925–1996) rows unless noted otherwise.

Nature of appellant counsel `songer_gt_counsel1` *numeric* [Source: Songer codebook]

Type of counsel representing the appellant.

- 1 = pro se
- 2 = court-appointed
- 3 = legal aid or public defender
- 4 = private
- 5 = U.S. government
- 6 = state or local government
- 7 = interest group, union, or professional group
- 8 = other, not ascertained, or some combination of the categories above (the KH update extended this code to absorb hybrid representations)

Per the original Songer codebook: if an attorney’s name appears with no other indication of affiliation, the coder assumed “private” (4) unless a government agency was the party.

Nature of respondent counsel `songer_gt_counsel2` *numeric* [Source: Songer codebook]

Type of counsel representing the respondent. Same coding as `songer_gt_counsel1`.

Appellant counsel (legacy duplicate) `songer_gt_acounsel1` *numeric* [Source: Songer codebook]

Legacy/duplicate of `songer_gt_counsel1`. Use the canonical column for new analyses.

Respondent counsel (legacy duplicate) `songer_gt_rcounselcode` *numeric* [Source: Songer codebook]

Legacy/duplicate of `songer_gt_counsel2`. Use the canonical column for new analyses.

Counsel (project metadata) `songer_gt_counsel` *numeric* [Source: Project metadata]

Project-derived field; likely an alias of `songer_gt_counsel1`. Not part of the Songer codebook.

Number of appellant attorneys `songer_gt_aatty` *numeric* [Source: Songer codebook]

Number of attorneys appearing for the appellant. 99 if unknown.

Coverage note. KH-update field.

First appellant attorney name `songer_gt_aattyfirstname, songer_gt_aattylastname` *string* [Source: Songer codebook]

First and last name of the first listed attorney for the appellant.

Coverage note. KH-update fields.

Number of respondent attorneys `songer_gt_ratty` *numeric* [Source: Songer codebook]

Number of attorneys appearing for the respondent. 99 if unknown.

Coverage note. KH-update field.

First respondent attorney name `songer_gt_rattyfirstname, songer_gt_rattylastname` *string* [Source: Songer codebook]

First and last name of the first listed attorney for the respondent.

Coverage note. KH-update fields.

Number of appellant law firms `songer_gt_afirms` *numeric* [Source: Songer codebook]

Number of law firms appearing for the appellants. Each governmental legal office (e.g., U.S. Attorney's office) counts as a firm. 99 if unknown.

Coverage note. KH-update field.

First appellant firm name `songer_gt_afirmname` *string* [Source: Songer codebook]

Name of the first listed law firm representing the appellant.

Coverage note. KH-update field.

Number of respondent law firms `songer_gt_rfirms` *numeric* [Source: Songer codebook]

Number of law firms appearing for the respondents, using the same conventions as `songer_gt_afirms`.

Coverage note. KH-update field.

First respondent firm name `songer_gt_rfirmname` *string* [Source: Songer codebook]

Name of the first listed law firm representing the respondent.

Coverage note. KH-update field.

Appellant interest groups `songer_gt_agroups` *numeric* [Source: Songer codebook]

Number of organized groups participating on behalf of the appellant.

Coverage note. KH-update field.

Respondent interest groups `songer_gt_rgroups` *numeric* [Source: Songer codebook]

Number of organized groups participating on behalf of the respondent.

Coverage note. KH-update field.

Amicus participation `songer_gt_amicus` *numeric* [Source: Songer codebook]

Number of separate amicus briefs filed in the case.

0 = no amicus participation

1--7 = number of separate amicus briefs filed

8 = eight or more briefs

9 = not ascertained

Amicus briefs for appellant `songer_gt_aamicusbriefs` *numeric* [Source: Songer codebook]

Number of amicus briefs filed in support of the appellant. 99 if briefs were filed but the side could not be determined.

Coverage note. KH-update field.

Amicus briefs for respondent `songer_gt_ramicusbrfs` *numeric* [Source: Songer codebook]

Number of amicus briefs filed in support of the respondent. 99 if side unknown.

Coverage note. KH-update field.

First amicus group for appellant `songer_gt_afirstlistedamicusgroup` *numeric* [Source: Songer codebook]

Corresponds to Songer codebook variable IDENTAMICAP, which the source codebook describes as “Name of first listed group that filed amicus for appellant.” This column is stored as

a numeric code field rather than a free-text name — effectively unpopulated, with 1,655 of 1,657 non-null values equal to 0 (no first-listed appellant amicus identified). The two non-zero values are isolated codes (7 and 99) rather than names. The asymmetric companion `songer_gt_rnameoffirstlistedamicusgroup` *does* carry name strings, so the two columns are not directly comparable.

Coverage note. KH-update field.

First amicus group for respondent `songer_gt_rnameoffirstlistedamicusgroup` *string* [Source: Songer codebook]

Name of the first listed group that filed an amicus brief for the respondent. Corresponds to Songer codebook variable IDENTAMICRP.

Coverage note. KH-update field.

A.8 Threshold issues

Threshold issues are doctrines that may prevent a court from reaching the merits of a case (jurisdiction, standing, mootness, etc.). Each variable below uses the same four-code scheme:

- 0 = threshold issue not raised or not discussed in the opinion (the typical value, since most cases do not contest a threshold)
- 1 = threshold raised, court ruled in favor of the respondent (the barrier is accepted; the appeal cannot proceed)
- 2 = threshold raised, court ruled in favor of the appellant (the barrier is rejected; the appeal proceeds)
- 9 = mixed result (e.g., jurisdiction over some claims but not others) or not ascertained

The per-threshold descriptions below cover the substantive content (what triggers a non-zero value); the code values follow the scheme above uniformly. A handful of cases carry codes 3 or 4 on individual threshold variables (`juris`, `mootness`, `othappth`). These are not defined in the project's working copy of the Songer codebook and are likely legacy codes or typos.

Jurisdiction (threshold) `songer_gt_juris` *numeric* [Source: Songer codebook]

Whether the court determined it had jurisdiction to hear the case.

Standing (threshold) `songer_gt_standing` *numeric* [Source: Songer codebook]

Whether the court determined the parties had standing to bring the case.

Mootness (threshold) `songer_gt_mootness` *numeric* [Source: Songer codebook]

Whether the court concluded the issue was moot.

Exhaustion / ripeness (threshold) `songer_gt_exhaust` *numeric* [Source: Songer codebook]

Whether the court refused to hear the appeal because administrative remedies were unexhausted or the case was not ripe.

Governmental immunity (threshold) `songer_gt_immunity` *numeric* [Source: Songer codebook]

Whether a governmental-immunity issue caused the court to refuse to reach the merits.

Failure to state a claim (threshold, legacy alias) `songer_gt_statecl` *numeric* [Source: Songer codebook]

Legacy alias of STATECLAIM: whether the court dismissed the case for failure to state a claim.

Frivolous case (threshold) `songer_gt_frivol` *numeric* [Source: Songer codebook]

Whether the court concluded the original case was frivolous or raised only trivial issues.

Political question (threshold) `songer_gt_polques`, `songer_gt_polquest` *numeric* [Source: Songer codebook]

Whether the court refused to rule on the merits because the case was a nonjusticiable political question. `songer_gt_polques` is a legacy alias of `songer_gt_polquest`; use the canonical column for new analyses.

Other trial-level threshold `songer_gt_oththres` *numeric* [Source: Songer codebook]

Catch-all for trial-level threshold issues not covered by the named fields above (e.g., collateral estoppel).

Late appeal (threshold) `songer_gt_late` *numeric* [Source: Songer codebook]

Whether the court refused to decide the appeal because of a timeliness, filing-fee, or missed-deadline failure on appeal.

Frivolous appeal (threshold) `songer_gt_frivapp` *numeric* [Source: Songer codebook]

Whether the court concluded it could not reach the merits because the appeal itself was frivolous or trivial.

Other appellate-level threshold `songer_gt_othappth` *numeric* [Source: Songer codebook]

Catch-all for appellate-level threshold issues not covered above (e.g., case became moot post-trial).

Timeliness (threshold) `songer_gt_timely` *numeric* [Source: Songer codebook]

Whether the court refused the merits because of timeliness rules, filing-fee rules, or statute of limitations.

General threshold flag `songer_gt_threshgen` *numeric* [Source: Songer codebook]

KH-update summary threshold-issue flag.

- 0 = no threshold issue
- 1 = threshold issue mentioned but not discussed
- 2 = discussed; favored the appellant
- 3 = discussed; favored the respondent
- 4 = discussed; favored neither or both

Coverage note. KH-update field; missing in Phase 1 cases.

Case turned on threshold issue `songer_gt_turnthresh` *numeric* [Source: Songer codebook]

When `songer_gt_threshgen` ≥ 2 , whether the case turned on the resolution of the threshold issue.

- 1 = yes
- 2 = no

Coverage note. KH-update field (TURNONTHRESH); missing in Phase 1 cases.

A.9 Substantive sub-issues: criminal (`songer_gt_typeiss` = 1)

Each variable below is a codebook flag indicating whether the named issue was raised and how the court ruled. The shared four-code scheme is:

- 0 = sub-issue not raised or not discussed
- 1 = court's ruling favored the prosecution (against the defendant)
- 2 = court's ruling favored the defendant (against the prosecution)
- 9 = mixed or not ascertained

Note that the LLM coder `songer_second_typeiss_*` uses a uniform pro-appellant convention rather than this pro-defendant convention; codes 1 and 2 swap meanings whenever the prosecution (rather than the defendant) is the appellant.

The Songer codebook provides an additional code 3 = “yes, but error was harmless” for most criminal sub-issues (the error was identified but did not affect the outcome). This code appears with non-trivial frequency on `prejud` (57 cases), `juryinst` (62), `othadmis` (72), `search` (14), `confess` (13), `improper` (15), `othcrim` (15), `sentence` (11), and is rare elsewhere. Code 4 appears in two cases on `othappth` and is not enumerated in the project's working copy of the codebook; it is likely a legacy code or typo.

Prosecutor misconduct `songer_gt_prejud` *numeric* [Source: Songer codebook]

Whether there was prejudicial conduct by the prosecutor.

Insanity defense `songer_gt_insane` *numeric* [Source: Songer codebook]

Whether the court below erred in not permitting an insanity defense, or there was error in handling the defendant's mental competence.

Improper jury influence `songer_gt_improper` *numeric* [Source: Songer codebook]

Whether the court concluded there was improper influence on the jury, other than prosecutor conduct.

Jury instructions `songer_gt_juryinst` *numeric* [Source: Songer codebook]

Whether the jury instructions were improper.

Other jury issues `songer_gt_othjury` *numeric* [Source: Songer codebook]

Whether jury composition or selection was invalid, or the jury was biased or tampered with.

Death penalty `songer_gt_deathpen` *numeric* [Source: Songer codebook]

Whether the death penalty was improperly imposed. (Validity of the sentence only.)

Sentencing `songer_gt_sentence` *numeric* [Source: Songer codebook]

Whether some other (non-death) penalty was improperly imposed.

Indictment `songer_gt_indict` *numeric* [Source: Songer codebook]

Whether the indictment was defective.

Confession admissibility `songer_gt_confess` *numeric* [Source: Songer codebook]

Whether a confession or incriminating statement was improperly admitted.

Search and seizure `songer_gt_search` *numeric* [Source: Songer codebook]

Whether the court below improperly ruled for the prosecution on an alleged illegal search and seizure.

Other evidence admissibility `songer_gt_othadmis` *numeric* [Source: Songer codebook]

Whether some other (non-confession, non-search) evidence was inadmissible.

Plea bargaining `songer_gt_plea` *numeric* [Source: Songer codebook]

Whether the court ruled for the defendant on a plea-bargaining issue.

Counsel (cross-reference) `songer_gt_counsel` *numeric* [Source: Songer codebook]

See the “Counsel, attorneys, firms, and amicus briefs” subsection above. Legacy duplicate.

Right to counsel `songer_gt_rtcouns` *numeric* [Source: Songer codebook]

Whether the defendant’s right to counsel was violated, other than through inadequate counsel.

Sufficiency of evidence `songer_gt_suffic` *numeric* [Source: Songer codebook]

Whether there was insufficient evidence for conviction.

Indigent-defendant rights `songer_gt_indigent` *numeric* [Source: Songer codebook]

Whether the defendant’s rights as an indigent were violated.

Entrapment `songer_gt_entrap` *numeric* [Source: Songer codebook]

Whether the defendant was the victim of illegal entrapment.

Procedural dismissal `songer_gt_procdis` *numeric* [Source: Songer codebook]

Whether the court upheld dismissal by the district court on procedural grounds.

Other criminal grounds `songer_gt_othcrim` *numeric* [Source: Songer codebook]

Whether the court ruled for the defendant on other grounds (speedy trial, double jeopardy, retroactivity, self-defense, waiver).

A.10 Substantive sub-issues: civil-private and civil-government
(`songer_gt_typeiss` $\in \{2, 3, 4\}$)

Each variable below is a flag for whether the named civil-law issue was raised and how the court ruled. The shared four-code scheme matches the LLM coder’s pro-appellant convention:

- 0 = sub-issue not raised or not discussed
- 1 = court’s ruling favored the respondent
- 2 = court’s ruling favored the appellant
- 9 = mixed or not ascertained

A small number of civil sub-issue variables also show stray codes that are not enumerated in the project’s working copy of the codebook, representing likely legacy codes or typos: **summary** carries codes 3, 4, and 10 on one case each; **foreign** carries code 7 on one case; and **whlaws** carries code 920 (likely a typo for code 2) on one case.

Due process `songer_gt_dueproc` *numeric* [Source: Songer codebook]

Whether the court's interpretation of due-process requirements favored the appellant (in non-prisoner civil cases).

Executive order / agency regulation `songer_gt_execord` *numeric* [Source: Songer codebook]

Whether the interpretation of an executive order or administrative regulation favored the appellant.

State policy `songer_gt_stpolicy` *numeric* [Source: Songer codebook]

Whether the interpretation of state or local law, executive order, administrative regulation, doctrine, or rule favored the appellant.

Weight of evidence `songer_gt_weightev` *numeric* [Source: Songer codebook]

Whether the court's factual interpretation (weight of evidence, sufficiency, burden of proof) favored the appellant.

Pretrial procedure `songer_gt_pretrial` *numeric* [Source: Songer codebook]

Whether the court's pretrial-procedure ruling favored the appellant. Excludes summary judgment and discovery (separate fields).

Trial procedure `songer_gt_trialpro` *numeric* [Source: Songer codebook]

Whether the court's ruling on trial-procedure issues (jury instructions, directed verdicts) favored the appellant.

Post-trial procedure `songer_gt_post_trl` *numeric* [Source: Songer codebook]

Whether the court's ruling on post-trial procedure or motions (court costs, motions to set aside verdict) favored the appellant.

Attorneys' fees `songer_gt_attyfee` *numeric* [Source: Songer codebook]

Whether the attorneys'-fees ruling favored the appellant.

Judicial discretion `songer_gt_judgdisc` *numeric* [Source: Songer codebook]

Whether the court's ruling on abuse of discretion by the trial judge favored the appellant.

Alternative dispute resolution `songer_gt_altdisp` *numeric* [Source: Songer codebook]

Whether the court's ADR ruling (settlement conference, mediator, arbitrator) favored the ap-

pellant.

Injunction `songer_gt_injunct` *numeric* [Source: Songer codebook]

Whether the court's injunction-related ruling (validity, grant, denial, stay) favored the appellant.

Summary judgment `songer_gt_summary` *numeric* [Source: Songer codebook]

Whether the court's summary-judgment ruling (or denial thereof) favored the appellant.

Federal vs. state law `songer_gt_fedvst` *numeric* [Source: Songer codebook]

Whether the court ruled that federal law should take precedence over state or local law in a conflict-of-laws case.

Domestic vs. foreign law `songer_gt_foreign` *numeric* [Source: Songer codebook]

Whether the court ruled that domestic law should take precedence over foreign law in a conflict-of-laws case.

International law `songer_gt_int_law` *numeric* [Source: Songer codebook]

Whether the court ruled for the appellant on a treaty or international-law issue.

State vs. state `songer_gt_st_v_st` *numeric* [Source: Songer codebook]

Whether the court ruled for the appellant on a state-versus-state conflict-of-laws issue.

Discovery `songer_gt_discover` *numeric* [Source: Songer codebook]

Whether the court's interpretation of discovery rules favored the appellant.

Other civil issue `songer_gt_othcivil` *numeric* [Source: Songer codebook]

Whether the court resolved some other significant civil-law issue, not in the enumerated categories above, in favor of the appellant.

Coverage note. KH-update field; missing in Phase 1 cases.

Diversity sufficiency `songer_gt_diverse` *numeric* [Source: Songer codebook]

Whether the court concluded the parties were truly diverse (in diversity-jurisdiction cases).

Which-laws control `songer_gt_whlaws` *numeric* [Source: Songer codebook]

Whether the court's discussion of which state's laws should control supported the appellant's

position (in diversity cases).

A.11 Substantive sub-issues: agency and administrative (`songer_gt_typeiss` = 2)

These flags ask whether the standard or doctrine in question “supported the government” on the issue raised. The shared four-code scheme is:

- 0 = sub-issue not raised or not discussed
- 1 = court’s ruling supported the government
- 2 = court’s ruling went against the government
- 9 = mixed or not ascertained

Note that the LLM coder `songer_second_typeiss_*` uses a uniform pro-appellant convention rather than this pro-government convention; codes 1 and 2 swap meanings whenever the government is the appellant (rather than the agency action being challenged from outside). A single case each carries code 3 on `alj`, `comment`, and `record`; these are not enumerated in the project’s working copy of the codebook and are likely legacy codes or typos.

Substantial evidence `songer_gt_subevid` *numeric* [Source: Songer codebook]

Whether the court’s use of the substantial-evidence rule supported the government.

De novo review `songer_gt_denovo` *numeric* [Source: Songer codebook]

Whether the court’s use of de-novo-on-facts review supported the government.

Clearly erroneous `songer_gt_erron` *numeric* [Source: Songer codebook]

Whether the court’s use of the clearly-erroneous standard of review supported the government.

Arbitrary and capricious `songer_gt_capric` *numeric* [Source: Songer codebook]

Whether the court’s use or interpretation of the arbitrary-and-capricious standard supported the government.

Deference to agency discretion `songer_gt_abusedis` *numeric* [Source: Songer codebook]

Whether the court concluded it should defer to agency discretion.

Judicial review of agency action `songer_gt_judrev` *numeric* [Source: Songer codebook]

Whether the court concluded the agency decision was subject to judicial review.

General standard `songer_gt_genstand` *numeric* [Source: Songer codebook]

Whether the agency articulated the appropriate general standard.

Notice `songer_gt_notice` *numeric* [Source: Songer codebook]

Whether the agency gave proper notice.

ALJ decision `songer_gt_alj` *numeric* [Source: Songer codebook]

Whether the court supported the decision of an administrative law judge.

Agency acquisition of information `songer_gt_agen_acq` *numeric* [Source: Songer codebook]

Whether the court ruled for the government on an issue related to agency acquisition of information (physical inspections, subpoenas).

Freedom of information `songer_gt_freeinfo` *numeric* [Source: Songer codebook]

Whether the court ruled for the government on an issue involving agency provision of information (FOIA, sunshine laws).

Notice-and-comment rulemaking `songer_gt_comment` *numeric* [Source: Songer codebook]

Whether the agency gave proper opportunity to comment.

Adequacy of agency record `songer_gt_record` *numeric* [Source: Songer codebook]

Whether the agency failed to develop an adequate record.

A.12 Constitutional and legal-provision citations

Constitutionality issue `songer_gt_constit` *numeric* [Source: Songer codebook]

Whether there was an issue about the constitutionality of a law or administrative action, and if so how it was resolved.

- 0 = not discussed
- 1 = favored respondent
- 2 = favored appellant
- 9 = mixed

Type of unconstitutionality declaration `songer_gt_decuncon` *numeric* [Source: Songer codebook]

Type of judicial review with a declaration of unconstitutionality. Only explicit statements in the opinion that some provision is unconstitutional were counted; procedural constitutional violations in the courts below (e.g., evidence thrown out under the 4th Amendment) do not count as judicial review for this field.

- 0 = no declaration of unconstitutionality
- 1 = Act of Congress declared unconstitutional, facial invalidity

- 2 = interpretation or application of federal law invalid
- 3 = federal administrative action or regulation unconstitutional on its face
- 4 = interpretation or application of administrative regulations unconstitutional
- 5 = state constitution declared unconstitutional on its face
- 6 = interpretation or application of state constitution unconstitutional
- 7 = state law declared unconstitutional on its face
- 8 = interpretation or application of state law unconstitutional
- 9 = sub-state law declared unconstitutional on its face
- 10 = interpretation or application of sub-state law unconstitutional

Most-cited constitutional provision `songer_gt_const1` *numeric* [Source: Songer codebook]

The U.S. Constitution provision most frequently cited in the headnotes. Codes 001–007 are the seven articles of the original Constitution (with their sub-section indicators in higher codes per the source codebook); codes 101–114 are the First through Fourteenth Amendments. 0 if none. See `value_codes_songer_gt_const1.csv` for the full list.

Second most-cited constitutional provision `songer_gt_const2` *numeric* [Source: Songer codebook]

Second most-cited constitutional provision; same coding as `songer_gt_const1`.

Federal statute interpretation `songer_gt_fedlaw` *numeric* [Source: Songer codebook]

Whether the court’s interpretation of a federal statute favored the appellant.

- 0 = not discussed
- 1 = favored respondent
- 2 = favored appellant
- 9 = mixed

Federal-law case-law application `songer_gt_fedlaw2` *numeric* [Source: Songer codebook]

Whether the application of case law (relating to a federal statute) favored the appellant. Coding parallels `songer_gt_fedlaw`.

Coverage note. KH-update field; missing in Phase 1 cases.

Precedent interpretation (non-statutory) `songer_gt_procedur` *numeric* [Source: Songer codebook]

Whether there was an interpretation of precedent that did not involve statutory or constitutional interpretation.

- 0 = not discussed
- 1 = favored respondent
- 2 = favored appellant
- 9 = mixed

Most-cited Federal Rule of Civil Procedure `songer_gt_civproc1` *numeric* [Source: Songer codebook]

Number of the Federal Rule of Civil Procedure most frequently cited in the headnotes. 0 if none.

Second most-cited Federal Rule of Civil Procedure `songer_gt_civproc2` *numeric* [Source: Songer codebook]

Number of the second most frequently cited Federal Rule of Civil Procedure.

Most-cited Federal Rule of Criminal Procedure `songer_gt_crmproc1` *numeric* [Source: Songer codebook]

Number of the Federal Rule of Criminal Procedure most frequently cited in the headnotes. 0 if none.

Second most-cited Federal Rule of Criminal Procedure `songer_gt_crmproc2` *numeric* [Source: Songer codebook]

Number of the second most frequently cited Federal Rule of Criminal Procedure.

Most-cited U.S. Code title `songer_gt_usc1` *numeric* [Source: Songer codebook]

Title of the U.S. Code most frequently cited in the headnotes. 0 if none.

Section of most-cited U.S. Code title `songer_gt_usc1sect` *numeric* [Source: Songer codebook]

Section number (up to 4 digits) of the title in `usc1` most frequently cited.

Second most-cited U.S. Code title `songer_gt_usc2` *numeric* [Source: Songer codebook]

Title of the U.S. Code second most frequently cited in headnotes. Same coding as `songer_gt_usc1`.

Section of second most-cited U.S. Code title `songer_gt_usc2sect` *numeric* [Source: Songer codebook]

Section number of the title in `usc2` most frequently cited.

Total West key numbers `songer_gt_keynotes` *numeric* [Source: Songer codebook]

Total number of West Topic key numbers in the opinion. Corresponds to Songer codebook variable `KEYTOTAL`.

Coverage note. KH-update field; missing in Phase 1 cases.

B. IDB ground truth (idb_*)

This appendix documents the variables from the FJC Integrated Database that are merged into the dataset under the prefix `idb_`. Definitions follow the two IDB appellate codebooks (Federal Judicial Center; *Appeals Codebook 1971-2007* and *Appeals Codebook 2008 Forward*). Era differences between the 1971–2007 and 2008-forward codebooks are flagged in coverage notes where they affect interpretation. The 1971–2007 codebook references “statistical year” (SY) dating, where SY_n runs from July of year $n - 1$ through June of year n ; the 2008-forward codebook uses fiscal-year dating throughout.

B.1 Identifiers and case origin

Circuit `idb_circuit` *numeric* [Source: FJC IDB codebook]

Circuit court of appeals where the appeal was filed. 0 denotes the D.C. Circuit; 1–11 denote the numbered circuits; -8 denotes missing.

Type of appeal `idb_apptype` *numeric* [Source: FJC IDB codebook]

Type of appeal at filing.

- 1 = administrative agency review
- 2 = administrative agency enforcement
- 3 = civil, U.S. involved
- 4 = civil, private
- 5 = criminal (used 1971–2007; coexisted with sentencing-guideline subtypes 13–17 after 1988; retired in the 2008-forward codebook)
- 6 = original proceeding
- 7 = emergency court / civil (pre-SY81; retired in the 2008-forward codebook)
- 10--12 = bankruptcy variants
- 13--17 = sentencing-guideline criminal subtypes (added 1988 with the sentencing reform)
- 18 = criminal, other (2008-forward addition)
- 19 = direct criminal appeal (2008-forward addition)
- 20 = interlocutory (2008-forward addition)
- 21 = post-conviction (2008-forward addition)
- 22 = miscellaneous case (2008-forward addition)

Coverage note. Codes 19–22 were introduced in the 2008-forward codebook to break the old 5/8/18 criminal lumping into finer subtypes. Codes 5, 7, 8, and 9 are flagged “No longer in use” in the 2008-forward codebook but appear on pre-2008 cases. Consult the case’s filing date before interpreting.

Reviewing agency `idb_agency` *string* [Source: FJC IDB codebook]

U.S. agency or organization whose action is being reviewed in administrative appeals (coded

only when `idb_apptype` $\in \{1, 2\}$). Alphanumeric agency codes (e.g., EPA, NLRB, IRS); -9 for non-agency appeals. See `value_codes_idb_agency.csv` for the full agency code list.

Jurisdictional basis `idb_juris` *numeric* [Source: FJC IDB codebook]

Basis of federal jurisdiction for civil appeals. In the 1971–2007 codebook this is coded when `idb_apptype` $\in \{3, 4, 7\}$; the 2008-forward codebook drops `idb_apptype` = 7 entirely (see the `idb_apptype` coverage note), so only $\{3, 4\}$ remain valid for post-2007 cases.

- 1 = U.S. plaintiff
- 2 = U.S. defendant
- 3 = federal question
- 4 = diversity
- 5 = local (territorial)
- 9 = non-civil

Coverage note. For prisoner-petition Nature-of-Suit values (`idb_nos` $\in \{510, 530, 535, 540, 550\}$), the codes have different meanings: 2 = federal prisoner, 3 = state prisoner, 5 = local prisoner. Check `idb_nos` before interpreting.

Nature of suit `idb_nos` *numeric* [Source: FJC IDB codebook]

Three-digit Administrative Office of the U.S. Courts (AOUSC) code identifying the underlying civil cause of action (coded only for civil `idb_apptype`). Common values include 440 = civil rights, 530 = habeas corpus, and 870 = tax. -9 for non-civil cases. See `value_codes_idb_nos.csv` for the full Nature-of-Suit code list.

Statutory offense `idb_offense` *numeric* [Source: FJC IDB codebook]

Statutory offense charged against the defendant in criminal appeals.

Coverage note. Coded only when `idb_apptype` $\in \{5, 8\}$ (and 13–17 in the 1971–2007 codebook); the 2008-forward codebook restricts coding to `idb_apptype` $\in \{5, 8\}$. -9 for non-criminal cases. See `value_codes_idb_offense.csv` for the full Offense Code list.

Appellant name `idb_appellan` *string* [Source: FJC IDB codebook]

Name of the party initiating the appeal.

Coverage note. Redacted to missing on public-use files for bankruptcy and criminal appeals.

Appellee name `idb_appellee` *string* [Source: FJC IDB codebook]

Name of the party against whom the appeal is taken.

Coverage note. Subject to the same public-use redaction as `idb_appellan`.

U.S. as appellant `idb_usapt` *numeric* [Source: FJC IDB codebook]

- 0 = U.S. was not the appellant
- 1 = U.S. was the appellant
- 8 = missing

U.S. as appellee `idb_usape` *numeric* [Source: FJC IDB codebook]

- 0 = U.S. was not the appellee
- 1 = U.S. was the appellee
- 8 = missing

Nature of original proceeding `idb_orgproc` *numeric* [Source: FJC IDB codebook]

Nature of the original proceeding (coded only when `idb_apptype` = 6).

- 1 = writ of prohibition
- 2 = writ of mandamus (pre-FY99); writ of mandamus, fee paid in full at time of filing (since FY99)
- 3 = habeas corpus (pre-SY81); writ of mandamus, fee not paid in full at time of filing (since FY99)
- 4 = certiorari (pre-SY81; “No longer in use” in the 2008-forward codebook)
- 5 = other extraordinary writ
- 6 = second or successive habeas filing (since FY98)

Coverage note. Code meanings shift at two transition points: SY81 (code 3 stops meaning habeas corpus; code 4 stops being assigned to new cases) and FY99 (the mandamus fee-paid / fee-not-paid split is encoded into codes 2 and 3). Consult the case’s filing date before interpreting code values.

Originating district `idb_ddist` *numeric* [Source: FJC IDB codebook]

District or bankruptcy court in which the case was originally filed. Numeric district code (e.g., 00 = Maine, 71 = Cal. Northern, 90 = D.C.). -8 = missing. Coded for `idb_apptype` ∈ {3, 4, 5, 7} and the criminal sentencing subtypes 8–17. See `value_codes_idb_ddist.csv` for the full district code list.

Originating district circuit `idb_dcirc` *numeric* [Source: FJC IDB codebook]

Circuit of the originating district or bankruptcy court. 0–11 with the same coding as `idb_circuit`. -8 = missing.

Divisional office `idb_doffice` *numeric* [Source: FJC IDB codebook]

Divisional office within the district or bankruptcy court where the case was originally filed. Numeric office code per the AOUSC divisional-office appendix. -8 = missing.

District-court docket number `idb_ddocket` *string* [Source: FJC IDB codebook]

Docket number assigned in the district or bankruptcy court. Format: YYNNNNN. -8 = missing.

District-court defendant number `idb_ddefnum` *numeric* [Source: FJC IDB codebook]

Defendant number assigned in the district court for criminal cases.

Coverage note. 2008-forward field only.

District judge code `idb_djudge` *string* [Source: FJC IDB codebook]

Four-character code identifying the district court, magistrate, or bankruptcy judge whose decision is being appealed.

Coverage note. 2008-forward field only. Blank or missing on public-use files.

Appellate docket number `idb_docket` *string* [Source: FJC IDB codebook]

Court of appeals docket number assigned by the Clerk upon docketing. Standard format since SY87: YYONNNNN (7 chars). Some circuits now use YYNNNNN with a 5-digit serial. -8 = missing.

Consolidated/cross-appeal docket `idb_consdkt` *string* [Source: FJC IDB codebook]

Appellate docket number of the initial cross-appeal or lead consolidated appeal (when `idb_joinapp` $\in \{1, 2\}$). Same format as `idb_docket`. -8 = missing.

Joined-appeal type `idb_joinapp` *numeric* [Source: FJC IDB codebook]

Type of joined appeal.

- 0 = none (single appeal; 2008-forward codebook only)
- 1 = cross appeal
- 2 = consolidated appeal
- 3 = lead case in a consolidated group (since SY81)
- 8 = missing (1971–2007 codebook: also used for single appeals)

Coverage note. The 1971–2007 codebook records single appeals as -8 (no distinct “none” code); the 2008-forward codebook adds an explicit 0 = none code. Filter on `idb_joinapp` $\in \{1, 2, 3\}$ if you want any joined appeal regardless of vintage.

Reopened or remanded indicator `idb_reopen` *numeric* [Source: FJC IDB codebook]

Indicator that the appeal is a reopened or remanded case.

- 1 = reopened

2 = remanded (since SY81)
 3 = second reopening / remand (since SY82)
 4 = reinstated after a procedural default (since SY85)
 5 = third reopening / remand (since SY88)
 6 = fourth reopening / remand (since SY97)
 7 = fifth reopening / remand (since SY97)
 M = remand from the Supreme Court (2008-forward only)
 S = second remand from the Supreme Court (2008-forward only)
 -8 = missing or initial filing

B.2 Dates

All IDB dates use 01/01/1900 as the missing sentinel.

Docketed in court of appeals `idb_dktdate date` [Source: FJC IDB codebook]

Date the case was docketed in the Court of Appeals.

Docketed in lower court `idb_ddktdate date` [Source: FJC IDB codebook]

Date the case was docketed in the district or bankruptcy court below.

Notice of appeal filed `idb_appdate date` [Source: FJC IDB codebook]

Date the notice of appeal was filed in the lower court. For interlocutory 1292(b) appeals, the date the application was granted.

Briefing schedule issued `idb_credate date` [Source: FJC IDB codebook]

Date the original briefing notice or schedule is issued by the court of appeals.

Coverage note. Post-SY85, this field is used only for merits terminations under Rule 28 FRAP. Pre-SY85, this field held the date the complete record was filed in the court.

Last brief filed `idb_brfiled date` [Source: FJC IDB codebook]

Date the last formal written presentation (briefs or memoranda) was received in the court of appeals prior to oral hearing or submission.

Coverage note. Pre-SY85, this field was limited to Rule 28 FRAP briefs only.

Submission date `idb_subdate date` [Source: FJC IDB codebook]

For appeals terminated on the merits without a hearing, the date the appeal was submitted on its merits to the first judge on the panel.

Oral argument date `idb.heardate` *date* [Source: FJC IDB codebook]

Date the appeal was orally argued on its merits before a panel of the court.

Final judgment date `idb.judgdate` *date* [Source: FJC IDB codebook]

Date the final judgment, decree, or order was entered to announce the court's decision. The canonical **year** column in the case-metadata section is derived from `idb.judgdate` when present.

Last transaction date `idb.transdate` *date* [Source: FJC IDB codebook]

System-generated timestamp of the last action taken on the record.

Coverage note. 2008-forward field only.

Last transaction code `idb.transcode` *string* [Source: FJC IDB codebook]

One-character system-generated code indicating the type of transaction.

E = new filing

J = modification (including termination) of an existing record

Coverage note. 2008-forward field only.

Statistical / fiscal year of record `idb.tapeyear` *numeric* [Source: FJC IDB codebook]

Statistical or fiscal year label of the data file from the AOUSC. Through SY91, this is the statistical year of termination; from FY92 forward, it is the fiscal year of termination. The value 2099 flags pending case records.

B.3 Disposition, opinion, and publication

Type of disposition action `idb.disp` *numeric* [Source: FJC IDB codebook]

Type of disposition action by which the court terminated the appeal.

1 = after oral hearing (merits termination)

2 = after submission without hearing (merits termination)

3 = after submission without hearing under a court rule (merits termination; valid through SY85, retired thereafter — in SY86 onward all submissions without hearing are recorded as code 2)

4 = other judicial action (procedural termination)

5 = without judicial action (clerk-only termination)

Codes 1–3 are merits terminations; codes 4–5 are procedural.

Coverage note. Code 3 (submission under court rule) appears only in pre-SY86 cases.

Merits disposition outcome `idb_outcome` *numeric* [Source: FJC IDB codebook]

Type of merits disposition when `idb_disp` $\in \{1, 2, 3\}$.

- 1 = affirmed or enforced
- 2 = reversed or vacated
- 3 = affirmed in part and reversed in part
- 4 = dismissed for want of jurisdiction (pre-January 1976; retired thereafter)
- 5 = dismissed (frivolous, moot, or for lack of merit)
- 6 = remanded
- 7 = other
- 9 = certificate of appealability denied (post-AEDPA, i.e. 1996 onward)

Coverage note. Pre-SY85 code meanings differ; consult the case's date before interpreting code values. Code 4 was retired in January 1976 and is absent in post-1975 cases.

Procedural-termination type `idb_procterm` *numeric* [Source: FJC IDB codebook]

Type of disposition when `idb_disp` = 4 (procedural termination after other judicial action).

- 1 = jurisdictional defects
- 2 = FRAP 42 (voluntary dismissal or settlement)
- 3 = settlement program (pre-SY87; retired thereafter)
- 5 = default
- 6 = certificate of probable cause denied
- 7 = transferred
- 8 = dismissed or other
- 9 = certificate of appealability denied (post-SY85)

Coverage note. Pre-SY85, this field followed a 1–8 “without judicial action” coding scheme with different meanings. Code 3 was retired in SY87 but appears on pre-SY87 cases. Treat pre-SY85 values carefully and consult the original 1971–2007 codebook for the historical scheme.

Clerk-only termination method `idb_method` *numeric* [Source: FJC IDB codebook]

Type of disposition when `idb_disp` = 5 (no judicial involvement, clerk-only action).

- 1 = FRAP 42 (voluntary dismissal)
- 2 = retired SY87 (“No longer in use” beginning SY87)
- 3 = default
- 4 = other

Coverage note. The codes above apply to post-SY85 cases. Pre-SY85, codes 1–6 covered “written opinion / memorandum / from the bench / court order / consent / other”; treat pre-SY85 values carefully and consult the 1971–2007 codebook.

Original-proceeding determination `idb_opdisp` *numeric* [Source: FJC IDB codebook]

Determination of an original proceeding (coded only when `idb_apptype = 6`).

- 1 = granted
- 2 = denied
- 3 = dismissed or other
- 8 = missing or not an original proceeding

Original-proceeding subtype `idb_torprosub` *string* [Source: FJC IDB codebook]

One-character code reported at termination identifying the nature of an original-proceeding action when the case did not arise from a lower court or agency. Uses the same code list as `idb_orgproc`, restricted to codes 1, 2, 3, 5, and 6 (the 2008-forward codebook does not use code 4 for this subtype field).

Coverage note. 2008-forward field only.

Type of separate opinions `idb_opinion` *numeric* [Source: FJC IDB codebook]

Type of concurring or dissenting opinion produced by the deciding panel.

- 6 = dissent only
- 7 = concurrence only
- 8 = both (not en banc)
- 9 = both (en banc)
- 8 = missing

Valid since SY79.

Publication status `idb_pubstat` *numeric* [Source: FJC IDB codebook]

Publication status or form of the opinion or order disposing of the appeal. (Named `OPINION/ORDER` in the 1971–2007 codebook.) The codes have shifted three times; the scheme that applies to a given case depends on its termination date. The current (SY85-onward) scheme is the seven-value list below.

SY85 onward (opinions/orders, terminations on the merits):

- 1 = unpublished, oral
- 2 = published, written, signed
- 3 = unpublished, written, signed
- 4 = published, written, unsigned, reasoned
- 5 = unpublished, written, unsigned, reasoned
- 6 = published, written, unsigned, without comment
- 7 = unpublished, written, unsigned, without comment
- 8 = missing or invalid

SY81–SY84 (opinions/orders, all terminations):

- 1 = published, other

- 2 = published, signed
- 3 = published, unsigned
- 4 = unpublished, other
- 5 = unpublished, signed
- 6 = unpublished, unsigned

January 1976 – SY80 (opinions only, all terminations):

- 1 = published, other
- 2 = published, signed
- 3 = published, reasons stated
- 4 = unpublished, other
- 5 = unpublished, signed
- 6 = unpublished, reasons stated

SY71 – December 1975 (opinions only, all terminations):

- 1 = none
- 2 = by judge
- 3 = per curiam

Coverage note. The codes above apply to post-SY85 cases. Pre-SY85, this field used a different six-value scheme; consult the 1971–2007 codebook for the historical values.

En banc indicator `idb_enbanc` *string* [Source: FJC IDB codebook]

Whether the appeal was heard en banc rather than by a 3-judge panel.

- Y = en banc
- N = not en banc
- 8 = missing or pending

Coverage note. 2008-forward field only. Pre-2008 en-banc cases were inferred from `idb_jdgcode2 = 0000` starting SY88. See the en banc filtering callout in the front matter.

Federal capital case `idb_fedcap` *string* [Source: FJC IDB codebook]

Flag identifying federal capital cases (e.g., criminal appeals where at least one defendant received the death penalty).

- Y = capital case
- 8 = not a capital case

Coverage note. 2008-forward field only.

Termination-time type of appeal `idb_trmtype` *numeric* [Source: FJC IDB codebook]

Type of appeal as reported by the clerk at termination. Same code list as `idb_apptype`. Useful when an appeal is reclassified between filing and termination.

Coverage note. 2008-forward field only.

B.4 Panel judges and party status

Panel judge codes `idb_jdgcode1, idb_jdgcode2, idb_jdgcode3` *string* [Source: FJC IDB codebook]

Statistical judge codes for the three panel judges. Per IDB conventions: `idb_jdgcode1` is a non-majority-author when `idb_opinion` $\in \{6, 8\}$ (the dissent author); `idb_jdgcode2` is the concurrence author when `idb_opinion` = 7, or carries the en-banc sentinel value 0000 since SY88; `idb_jdgcode3` is by convention the author of the majority opinion.

Coverage note. Judge codes are set to missing on public-use files, except for the 0000 en-banc sentinel in `idb_jdgcode2`.

Judge 1 involvement `idb_jdg1inv` *string* [Source: FJC IDB codebook]

One-character code describing the involvement of judge 1 on the panel.

A = authored the majority opinion
B = co-authored the majority opinion and concurred
C = co-authored the majority opinion and dissented
D = co-authored the majority opinion, concurred, and dissented
E = concurred only
F = dissented only
G = both concurred and dissented (no majority authorship)
H = co-authored the majority opinion
P = panel participant with no separate opinion
blank = not a 3-judge panel

Coverage note. 2008-forward field only.

Judge 2 involvement `idb_jdg2inv` *string* [Source: FJC IDB codebook]

Involvement code for judge 2; same A–H/P scheme as `idb_jdg1inv`.

Coverage note. 2008-forward field only.

Judge 3 involvement `idb_jdg3inv` *string* [Source: FJC IDB codebook]

Involvement code for judge 3; same A–H/P scheme.

Coverage note. 2008-forward field only.

Pro-se status at filing `idb_prosefle` *numeric* [Source: FJC IDB codebook]

Pro-se status of the parties at the time the appeal was filed.

- 0 = neither party pro se
- 1 = appellant pro se
- 2 = appellee pro se
- 3 = both pro se

Coverage note. Data collection began August 1991. Y/N values are valid August 1991–FY92; numeric values from FY93 forward.

Pro-se status at termination `idb_prosetrm` *numeric* [Source: FJC IDB codebook]

Pro-se status of the parties at termination of the appeal; same code list as `idb_prosefle`.

Coverage note. Data collection began August 1991.

In forma pauperis status at filing `idb_filefee` *string* [Source: FJC IDB codebook]

Two-character code indicating in forma pauperis status at filing.

- FP = IFP granted
- blank = not granted

Coverage note. 2008-forward field only. In 1971–2007, the analogous field was IFP coded 0/1.

In forma pauperis status at termination `idb_trmfee` *string* [Source: FJC IDB codebook]

Two-character IFP-status code at termination.

- FP = in forma pauperis status granted
- N = not granted
- 8 = missing or pending

Coverage note. 2008-forward field only.

Record status `idb_statuscd` *string* [Source: FJC IDB codebook]

Status code identifying the type of record.

- S = pending
- L = terminated

Coverage note. Valid since FY01 in the 1971–2007 codebook; always present in the 2008-forward codebook.

B.5 Legacy 1971–2007 fields

These IDB fields appear only in the 1971–2007 Appeals codebook — they were retired before the 2008-forward release. Each is blank (empty string) on cases sourced from the 2008-forward file, and the date ranges in each entry indicate the narrower window during which the field was actually populated. Modern analyses should consult the corresponding successor field where one exists.

In forma pauperis (pre-SY81) `idb.ifp` *numeric* [Source: FJC IDB codebook]

Pre-SY81 in forma pauperis indicator. Successor field: `idb.filefee` (2008-forward).

0 = not an in forma pauperis proceeding

1 = in forma pauperis proceeding

-8 or null = missing or invalid

Coverage note. Data collected prior to SY81.

Divisional office (legacy) `idb.divoff` *numeric* [Source: FJC IDB codebook]

Pre-October 1981 divisional-office indicator for the Ninth and Fifth Circuits only; -8 for missing/invalid. Successor field (now circuit-agnostic): `idb.doffice`.

Coverage note. Only populated on Ninth Circuit and Fifth Circuit appeals docketed before October 1, 1981.

Other type (bankruptcy) `idb.othtype` *numeric* [Source: FJC IDB codebook]

Bankruptcy-appeal subtype, valid only when `idb.apptype` $\in \{9, 10, 11, 12\}$. The 1971–2007 codebook documents the field but does not enumerate its code values inline; consult the codebook directly. -8 = missing/invalid; -9 / 0 / null = non-bankruptcy appeal.

AOUSC filing month `idb.aofilemo` *numeric* [Source: FJC IDB codebook]

AOUSC-assigned filing-month label (1–12). 0 represents missing data.

Coverage note. Valid beginning in SY81. Used with `idb.aofileyr` as a coarse filing-date pair before the daily `idb.apdate` convention was adopted.

AOUSC filing year `idb.aofileyr` *numeric* [Source: FJC IDB codebook]

AOUSC-assigned filing-year label (YYYY). 0 represents missing data.

Coverage note. Valid beginning in SY81. See `idb.aofilemo` for the companion month field.

AOUSC termination month `idb.trusedmo` *numeric* [Source: FJC IDB codebook]

AOUSC-assigned termination-month label (1–12). -8 represents missing/invalid data.

Coverage note. Valid beginning in FY96.

AOUSC termination year `idb.trusedyr` *numeric* [Source: FJC IDB codebook]

AOUSC-assigned termination-year label (YYYY). -8 represents missing/invalid data.

Coverage note. Valid beginning in FY96.

Termination date (legacy) `idb.termdate` *date* [Source: FJC IDB codebook]

Termination date used during two narrow windows when termination differed from final-judgment date (e.g., when a petition for rehearing terminated within the same month as the judgment, or where a court order made termination later than the judgment date). 01/01/1900 or null represents missing/invalid data.

Coverage note. Valid SY71–SY81 and FY96–FY97 only. Successor: `idb.judgdate`.

Certificate of probable cause decision `idb.cpc` *numeric* [Source: FJC IDB codebook]

Outcome of a Certificate of Probable Cause request.

- 1 = granted
- 2 = denied
- 3 = both granted and denied
- 8 = missing or invalid
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Certificate of probable cause decision-maker `idb.cpcdisc` *numeric* [Source: FJC IDB codebook]

Court that ruled on the Certificate of Probable Cause request.

- 1 = circuit
- 2 = district
- 3 = both circuit and district
- 8 = missing or invalid
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Full judge panel indicator `idb_jdgpanel` *numeric* [Source: FJC IDB codebook]

Indicator of whether the appeal was heard by a single judge or a full panel.

- 1 = single judge
- 2 = full panel
- 8 = missing or invalid (within the valid window, missing was treated as “disposed of without judicial action”)
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Counsel appointed `idb_cnslapp` *numeric* [Source: FJC IDB codebook]

Indicator that counsel was appointed for the appellant.

- 1 = yes
- 2 = no
- 8 = missing or invalid
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Counsel continued `idb_cnslcont` *numeric* [Source: FJC IDB codebook]

Indicator that counsel from below continued at the appellate stage.

- 1 = yes
- 2 = no
- 8 = missing or invalid
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Counsel source — district `idb_cnsldis` *numeric* [Source: FJC IDB codebook]

Source of counsel at the district court level.

- 1 = CJA (Criminal Justice Act)
- 2 = private
- 3 = waived / none
- 4 = other / appointment
- 6 = public / community defender
- 8 = missing or invalid
- 9 or null = not collected

Coverage note. Valid SY73–SY80.

Counsel source — **circuit** `idb_cnslcir` *numeric* [Source: FJC IDB codebook]

Source of counsel at the appellate court level; same code list as `idb_cnsldis`.

Coverage note. Valid SY73–SY80.

Statistical year of termination (legacy) `idb_newsytrm` *numeric* [Source: FJC IDB codebook]

Statistical-year label for the case’s termination. Null represents missing data. Successor: `idb_tapeyear`.

Coverage note. Valid through FY95.

B.6 Match metadata

These project-derived fields document how each case in the dataset was matched to its IDB record.

Match stage `idb_match_stage` *string* [Source: Project-derived]

The stage at which the IDB matching pipeline succeeded for this case (e.g., docket-based, party-based, citation-based). Useful for diagnosing match-quality concerns.

Alternate IDB matches `idb_alts_json` *string* [Source: Project-derived]

Sparse JSON-encoded blob of alternate IDB rows in cases where an IDB match is one-to-many, e.g., a single consolidated case in the project’s database is split into individual appeals by IDB. Populated for fewer than 1,000 cases, 0.3% of matched cases; missing otherwise. Parse with a standard JSON parser to inspect individual IDB entries.

C. SCDB variables coded by the LLM

This appendix documents the variables in the Supreme Court Database (SCDB; Spaeth et al. 2025) that the project’s SCDB-track LLM coder targets, together with the SCDB-native code lists. The LLM outputs are described variable-by-variable in §4.1–§4.5; this appendix complements those entries by spelling out the underlying SCDB definitions, which a user accessing raw SCDB data alongside this project’s outputs may want to consult. Where the LLM’s coding scheme deviates from raw SCDB (because it follows the Songer codebook convention, or applies a simplified custom typology), the deviation is flagged on the relevant entry. SCDB variable names appear in the column-name slot of each entry below as they are spelled in the SCDB codebook (i.e., **petitioner**, not a project-prefixed column).

C.1 Litigants

Petitioner `petitioner` *numeric* [Source: SCDB codebook]

The party that petitioned the Supreme Court to review the case (also known as the appellant or moving party). The SCDB native code list comprises 311 inductively created party-type codes (e.g., specific federal agencies, business types, classes of individual litigants), too long to enumerate here; see the SCDB codebook for the full **varParties** list.

Coverage note. The project’s LLM coder collapses the 311-code SCDB typology into a 7-category custom scheme (`scdb_first_appellant_category`; see §4.1); the LLM output is not directly comparable to raw SCDB **petitioner** codes.

Respondent `respondent` *numeric* [Source: SCDB codebook]

The party being sued or tried (also known as the appellee or non-moving party). Same 311-code list as **petitioner**; see the SCDB codebook.

Coverage note. As with **petitioner**, the project’s LLM coder collapses the SCDB typology into a 7-category custom scheme (`scdb_first_respondent_category`; see §4.1).

C.2 Issue area

Issue area `issueArea` *numeric* [Source: SCDB codebook]

Broad issue category, summarizing the more specific 278-code **issue** variable (see below). Coded identically to the project’s `scdb_first_broad_area` (see §4.3).

- 1 = Criminal Procedure
- 2 = Civil Rights
- 3 = First Amendment
- 4 = Due Process
- 5 = Privacy

- 6 = Attorneys
- 7 = Unions
- 8 = Economic Activity
- 9 = Judicial Power
- 10 = Federalism
- 11 = Interstate Relations
- 12 = Federal Taxation
- 13 = Miscellaneous
- 14 = Private Action

C.3 Specific issue (issue)

Specific issue `issue numeric` [Source: SCDB codebook]

The SCDB `issue` variable refines the issue area into 278 specific sub-issue codes (*varIssues* in the SCDB codebook). Codes are 5- or 6-digit integers; the first one or two digits indicate the broad area, and the remaining digits index the specific sub-issue. Coded identically to the project's `scdb_second_sub_area` (see §4.3); definitions below are drawn from the Supreme Court Database codebook, with phrasing simplified for brevity. Consult the original codebook for the most authoritative wording.

C.3.1 Issue area 1: Criminal Procedure

- 10010 involuntary confession
- 10020 habeas corpus
- 10030 plea bargaining: the constitutionality of and/or the circumstances of its exercise
- 10040 retroactivity (of newly announced or newly enacted constitutional or statutory rights)
- 10050 search and seizure (other than as pertains to vehicles or Crime Control Act)
- 10060 search and seizure, vehicles
- 10070 search and seizure, Crime Control Act
- 10080 contempt of court or congress
- 10090 self-incrimination (other than as pertains to Miranda or immunity from prosecution)
- 10100 Miranda warnings
- 10110 self-incrimination, immunity from prosecution
- 10120 right to counsel (cf. indigents appointment of counsel or inadequate representation)
- 10130 cruel and unusual punishment, death penalty (cf. extra legal jury influence, death penalty)
- 10140 cruel and unusual punishment, non-death penalty (cf. liability, civil rights acts)
- 10150 line-up
- 10160 discovery and inspection (in the context of criminal litigation only, otherwise Freedom of Information Act and related federal or state statutes or regulations)
- 10170 double jeopardy
- 10180 ex post facto (state)
- 10190 extra-legal jury influences: miscellaneous
- 10200 extra-legal jury influences: prejudicial statements or evidence
- 10210 extra-legal jury influences: contact with jurors outside courtroom
- 10220 extra-legal jury influences: jury instructions (not necessarily in criminal cases)

10230 extra-legal jury influences: voir dire (not necessarily a criminal case)
 10240 extra-legal jury influences: prison garb or appearance
 10250 extra-legal jury influences: jurors and death penalty (cf. cruel and unusual punishment)
 10260 extra-legal jury influences: pretrial publicity
 10270 confrontation (right to confront accuser, call and cross-examine witnesses)
 10280 subconstitutional fair procedure: confession of error
 10290 subconstitutional fair procedure: conspiracy (cf. Federal Rules of Criminal Procedure: conspiracy)
 10300 subconstitutional fair procedure: entrapment
 10310 subconstitutional fair procedure: exhaustion of remedies
 10320 subconstitutional fair procedure: fugitive from justice
 10330 subconstitutional fair procedure: presentation, admissibility, or sufficiency of evidence (not necessarily a criminal case)
 10340 subconstitutional fair procedure: stay of execution
 10350 subconstitutional fair procedure: timeliness
 10360 subconstitutional fair procedure: miscellaneous
 10370 Federal Rules of Criminal Procedure
 10380 statutory construction of criminal laws: assault
 10390 statutory construction of criminal laws: bank robbery
 10400 statutory construction of criminal laws: conspiracy (cf. subconstitutional fair procedure: conspiracy)
 10410 statutory construction of criminal laws: escape from custody
 10420 statutory construction of criminal laws: false statements (cf. statutory construction of criminal laws: perjury)
 10430 statutory construction of criminal laws: financial (other than in fraud or internal revenue)
 10440 statutory construction of criminal laws: firearms
 10450 statutory construction of criminal laws: fraud
 10460 statutory construction of criminal laws: gambling
 10470 statutory construction of criminal laws: Hobbs Act; i.e., 18 USC 1951
 10480 statutory construction of criminal laws: immigration (cf. immigration and naturalization)
 10490 statutory construction of criminal laws: internal revenue (cf. Federal Taxation)
 10500 statutory construction of criminal laws: Mann Act and related statutes
 10510 statutory construction of criminal laws: narcotics includes regulation and prohibition of alcohol
 10520 statutory construction of criminal laws: obstruction of justice
 10530 statutory construction of criminal laws: perjury (other than as pertains to statutory construction of criminal laws: false statements)
 10540 statutory construction of criminal laws: Travel Act, 18 USC 1952
 10550 statutory construction of criminal laws: war crimes
 10560 statutory construction of criminal laws: sentencing guidelines
 10570 statutory construction of criminal laws: miscellaneous
 10580 jury trial (right to, as distinct from extra-legal jury influences)
 10590 speedy trial
 10600 miscellaneous criminal procedure (cf. due process, prisoners' rights, comity: criminal procedure)

C.3.2 Issue area 2: Civil Rights

- 20010 voting
- 20020 Voting Rights Act of 1965, plus amendments
- 20030 ballot access (of candidates and political parties)
- 20040 desegregation (other than as pertains to school desegregation, employment discrimination, and affirmative action)
- 20050 desegregation, schools
- 20060 employment discrimination: on basis of race, age, religion, illegitimacy, national origin, or working conditions.
- 20070 affirmative action
- 20075 slavery or indenture
- 20080 sit-in demonstrations (protests against racial discrimination in places of public accommodation)
- 20090 reapportionment: other than plans governed by the Voting Rights Act
- 20100 debtors' rights
- 20110 deportation (cf. immigration and naturalization)
- 20120 employability of aliens (cf. immigration and naturalization)
- 20130 sex discrimination (excluding sex discrimination in employment)
- 20140 sex discrimination in employment (cf. sex discrimination)
- 20150 Indians (other than pertains to state jurisdiction over)
- 20160 Indians, state jurisdiction over
- 20170 juveniles (cf. rights of illegitimates)
- 20180 poverty law, constitutional
- 20190 poverty law, statutory: welfare benefits, typically under some Social Security Act provision.
- 20200 illegitimates, rights of (cf. juveniles): typically inheritance and survivor's benefits, and paternity suits
- 20210 handicapped, rights of: under Rehabilitation, Americans with Disabilities Act, and related statutes
- 20220 residency requirements: durational, plus discrimination against nonresidents
- 20230 military: draftee, or person subject to induction
- 20240 military: active duty
- 20250 military: veteran
- 20260 immigration and naturalization: permanent residence
- 20270 immigration and naturalization: citizenship
- 20280 immigration and naturalization: loss of citizenship, denaturalization
- 20290 immigration and naturalization: access to public education
- 20300 immigration and naturalization: welfare benefits
- 20310 immigration and naturalization: miscellaneous
- 20320 indigents: appointment of counsel (cf. right to counsel)
- 20330 indigents: inadequate representation by counsel (cf. right to counsel)
- 20340 indigents: payment of fine
- 20350 indigents: costs or filing fees
- 20360 indigents: U.S. Supreme Court docketing fee
- 20370 indigents: transcript
- 20380 indigents: assistance of psychiatrist
- 20390 indigents: miscellaneous

- 20400 liability, civil rights acts (cf. liability, governmental and liability, nongovernmental; cruel and unusual punishment, non-death penalty)
- 20410 miscellaneous civil rights (cf. comity: civil rights)

C.3.3 Issue area 3: First Amendment

- 30010 First Amendment, miscellaneous (cf. comity: First Amendment)
- 30020 commercial speech, excluding attorneys
- 30030 libel, defamation: defamation of public officials and public and private persons
- 30040 libel, privacy: true and false light invasions of privacy
- 30050 legislative investigations: concerning internal security only
- 30060 federal or state internal security legislation: Smith, Internal Security, and related federal statutes
- 30070 loyalty oath or non-Communist affidavit (other than bar applicants, government employees, political party, or teacher)
- 30080 loyalty oath: bar applicants (cf. admission to bar, state or federal or U.S. Supreme Court)
- 30090 loyalty oath: government employees
- 30100 loyalty oath: political party
- 30110 loyalty oath: teachers
- 30120 security risks: denial of benefits or dismissal of employees for reasons other than failure to meet loyalty oath requirements
- 30130 conscientious objectors (cf. military draftee or military active duty) to military service
- 30140 campaign spending (cf. governmental corruption):
- 30150 protest demonstrations (other than as pertains to sit-in demonstrations): demonstrations and other forms of protest based on First Amendment guarantees
- 30160 free exercise of religion
- 30170 establishment of religion (other than as pertains to parochial aid)
- 30180 parochial aid: government aid to religious schools, or religious requirements in public schools
- 30190 obscenity, state (cf. comity: privacy): including the regulation of sexually explicit material under the 21st Amendment
- 30200 obscenity, federal

C.3.4 Issue area 4: Due Process

- 40010 due process: miscellaneous (cf. loyalty oath), the residual code
- 40020 due process: hearing or notice (other than as pertains to government employees or prisoners' rights)
- 40030 due process: hearing, government employees
- 40040 due process: prisoners' rights and defendants' rights
- 40050 due process: impartial decision maker
- 40060 due process: jurisdiction (jurisdiction over non-resident litigants)
- 40070 due process: takings clause, or other non-constitutional governmental taking of

C.3.5 Issue area 5: Privacy

- 50010 privacy (cf. libel, comity: privacy)
- 50020 abortion: including contraceptives

- 50030 right to die
- 50040 Freedom of Information Act and related federal or state statutes or regulations

C.3.6 Issue area 6: Attorneys

- 60010 attorneys' and governmental employees' or officials' fees or compensation or licenses
- 60020 commercial speech, attorneys (cf. commercial speech)
- 60030 admission to a state or federal bar, disbarment, and attorney discipline (cf. loyalty oath: bar applicants)
- 60040 admission to, or disbarment from, Bar of the U.S. Supreme Court

C.3.7 Issue area 7: Unions

- 70010 arbitration (in the context of labor-management or employer-employee relations) (cf. arbitration)
- 70020 union antitrust: legality of anticompetitive union activity
- 70030 union or closed shop: includes agency shop litigation
- 70040 Fair Labor Standards Act
- 70050 Occupational Safety and Health Act
- 70060 union-union member dispute (except as pertains to union or closed shop)
- 70070 labor-management disputes: bargaining
- 70080 labor-management disputes: employee discharge
- 70090 labor-management disputes: distribution of union literature
- 70100 labor-management disputes: representative election
- 70110 labor-management disputes: antistrike injunction
- 70120 labor-management disputes: jurisdictional dispute
- 70130 labor-management disputes: right to organize
- 70140 labor-management disputes: picketing
- 70150 labor-management disputes: secondary activity
- 70160 labor-management disputes: no-strike clause
- 70170 labor-management disputes: union representatives
- 70180 labor-management disputes: union trust funds (cf. ERISA)
- 70190 labor-management disputes: working conditions
- 70200 labor-management disputes: miscellaneous dispute
- 70210 miscellaneous union

C.3.8 Issue area 8: Economic Activity

- 80010 antitrust (except in the context of mergers and union antitrust)
- 80020 mergers
- 80030 bankruptcy (except in the context of priority of federal fiscal claims)
- 80040 sufficiency of evidence: typically in the context of a jury's determination of compensation for injury or death
- 80050 election of remedies: legal remedies available to injured persons or things
- 80060 liability, governmental: tort or contract actions by or against government or governmental officials other than defense of criminal actions brought under a civil rights action.
- 80070 liability, other than as in sufficiency of evidence, election of remedies, punitive damages
- 80080 liability, punitive damages

- 80090 Employee Retirement Income Security Act (cf. union trust funds)
- 80100 state or local government tax
- 80105 state and territorial land claims
- 80110 state or local government regulation, especially of business (cf. federal pre-emption of state court jurisdiction, federal pre-emption of state legislation or regulation)
- 80120 federal or state regulation of securities
- 80130 natural resources - environmental protection (cf. national supremacy: natural resources, national supremacy: pollution)
- 80140 corruption, governmental or governmental regulation of other than as in campaign spending
- 80150 zoning: constitutionality of such ordinances, or restrictions on owners' or lessors' use of real property
- 80160 arbitration (other than as pertains to labor-management or employer-employee relations (cf. union arbitration)
- 80170 federal or state consumer protection: typically under the Truth in Lending; Food, Drug and Cosmetic; and Consumer Protection Credit Acts
- 80180 patents and copyrights: patent
- 80190 patents and copyrights: copyright
- 80200 patents and copyrights: trademark
- 80210 patents and copyrights: patentability of computer processes
- 80220 federal or state regulation of transportation regulation: railroad
- 80230 federal and some few state regulations of transportation regulation: boat
- 80240 federal and some few state regulation of transportation regulation: truck, or motor carrier
- 80250 federal and some few state regulation of transportation regulation: pipeline (cf. federal public utilities regulation: gas pipeline)
- 80260 federal and some few state regulation of transportation regulation: airline
- 80270 federal and some few state regulation of public utilities regulation: electric power
- 80280 federal and some few state regulation of public utilities regulation: nuclear power
- 80290 federal and some few state regulation of public utilities regulation: oil producer
- 80300 federal and some few state regulation of public utilities regulation: gas producer
- 80310 federal and some few state regulation of public utilities regulation: gas pipeline (cf.
- 80320 federal transportation regulation: pipeline) federal and some few state regulation of public utilities regulation: radio and television (cf. cable television)
- 80330 federal and some few state regulation of public utilities regulation: cable television (cf. radio and television)
- 80340 federal and some few state regulations of public utilities regulation: telephone or telegraph company
- 80350 miscellaneous economic regulation

C.3.9 Issue area 9: Judicial Power

- 90010 comity: civil rights
- 90020 comity: criminal procedure
- 90030 comity: First Amendment
- 90040 comity: habeas corpus
- 90050 comity: military
- 90060 comity: obscenity
- 90070 comity: privacy

90080 comity: miscellaneous
 90090 comity primarily removal cases, civil procedure (cf. comity, criminal and First Amendment); deference to foreign judicial tribunals
 90100 assessment of costs or damages: as part of a court order
 90110 Federal Rules of Civil Procedure including Supreme Court Rules, application of the Federal Rules of Evidence, Federal Rules of Appellate Procedure in civil litigation, Circuit Court Rules, and state rules and admiralty rules
 90120 judicial review of administrative agency's or administrative official's actions and procedures
 90130 mootness (cf. standing to sue: live dispute)
 90140 venue
 90150 no merits: writ improvidently granted
 90160 no merits: dismissed or affirmed for want of a substantial or properly presented federal question, or a nonsuit
 90170 no merits: dismissed or affirmed for want of jurisdiction (cf. judicial administration: Supreme Court jurisdiction or authority on appeal from federal district courts or courts of appeals)
 90180 no merits: adequate non-federal grounds for decision
 90190 no merits: remand to determine basis of state or federal court decision (cf. judicial administration: state law)
 90200 no merits: miscellaneous
 90210 standing to sue: adversary parties
 90220 standing to sue: direct injury
 90230 standing to sue: legal injury
 90240 standing to sue: personal injury
 90250 standing to sue: justiciable question
 90260 standing to sue: live dispute
 90270 standing to sue: parens patriae standing
 90280 standing to sue: statutory standing
 90290 standing to sue: private or implied cause of action
 90300 standing to sue: taxpayer's suit
 90310 standing to sue: miscellaneous
 90320 judicial administration: jurisdiction or authority of federal district courts or territorial courts
 90330 judicial administration: jurisdiction or authority of federal courts of appeals
 90340 judicial administration: Supreme Court jurisdiction or authority on appeal or writ of error, from federal district courts or courts of appeals (cf. 753)
 90350 judicial administration: Supreme Court jurisdiction or authority on appeal or writ of error, from highest state court
 90360 judicial administration: jurisdiction or authority of the Court of Claims
 90370 judicial administration: Supreme Court's original jurisdiction
 90380 judicial administration: review of non-final order
 90390 judicial administration: change in state law (cf. no merits: remand to determine basis of state court decision)
 90400 judicial administration: federal question (cf. no merits: dismissed for want of a substantial or properly presented federal question)
 90410 judicial administration: ancillary or pendent jurisdiction
 90420 judicial administration: extraordinary relief (e.g., mandamus, injunction)

- 90430 judicial administration: certification (cf. objection to reason for denial of certiorari or appeal)
- 90440 judicial administration: resolution of circuit conflict, or conflict between or among other courts
- 90450 judicial administration: objection to reason for denial of certiorari or appeal
- 90460 judicial administration: collateral estoppel or res judicata
- 90470 judicial administration: interpleader
- 90480 judicial administration: untimely filing
- 90490 judicial administration: Act of State doctrine
- 90500 judicial administration: miscellaneous
- 90510 Supreme Court's certiorari, writ of error, or appeals jurisdiction
- 90520 miscellaneous judicial power, especially diversity jurisdiction

C.3.10 Issue area 10: Federalism

- 100010 federal-state ownership dispute (cf. Submerged Lands Act)
- 100020 federal pre-emption of state court jurisdiction
- 100030 federal pre-emption of state legislation or regulation. cf. state regulation of business. rarely involves union activity. Does not involve constitutional interpretation unless the Court says it does.
- 100040 Submerged Lands Act (cf. federal-state ownership dispute)
- 100050 national supremacy: commodities
- 100060 national supremacy: intergovernmental tax immunity
- 100070 national supremacy: marital and family relationships and property, including obligation of child support
- 100080 national supremacy: natural resources (cf. natural resources - environmental protection)
- 100090 national supremacy: pollution, air or water (cf. natural resources - environmental protection)
- 100100 national supremacy: public utilities (cf. federal public utilities regulation)
- 100110 national supremacy: state tax (cf. state tax)
- 100120 national supremacy: miscellaneous
- 100130 miscellaneous federalism

C.3.11 Issue area 11: Interstate Relations

- 110010 boundary dispute between states
- 110020 non-real property dispute between states
- 110030 miscellaneous interstate relations conflict
- 110033 incorporation of foreign territories

C.3.12 Issue area 12: Federal Taxation

- 120010 federal taxation, typically under provisions of the Internal Revenue Code
- 120020 federal taxation of gifts, personal, business, or professional expenses
- 120030 priority of federal fiscal claims: over those of the states or private entities
- 120040 miscellaneous federal taxation (cf. national supremacy: state tax)

C.3.13 Issue area 13: Miscellaneous

130010	legislative veto
130015	executive authority vis-a-vis congress or the states
130020	miscellaneous

C.3.14 Issue area 14: Private Action

140010	real property
140020	personal property
140030	contracts
140040	evidence
140050	civil procedure
140060	torts
140070	wills and trusts
140080	commercial transactions

C.4 Disposition and winning party

Disposition of case `caseDisposition` *numeric* [Source: SCDB codebook]

How the Supreme Court disposed of the case under review. Note that the project's LLM coder (`scdb_first_disposition`; see §4.2) uses a Songer-style 0–11 numbering rather than this SCDB native 1–11 scheme; remap when merging.

- 1 = stay, petition, or motion granted
- 2 = affirmed (includes modified)
- 3 = reversed
- 4 = reversed and remanded
- 5 = vacated and remanded
- 6 = affirmed and reversed (or vacated) in part
- 7 = affirmed and reversed (or vacated) in part and remanded
- 8 = vacated
- 9 = petition denied or appeal dismissed
- 10 = certification to or from a lower court
- 11 = no disposition

Winning party `partyWinning` *numeric* [Source: SCDB codebook]

Whether the petitioning party (the appellant in CoA terms) emerged victorious. Coded identically to the project's `scdb_first_appellant_win` (see §4.2).

- 0 = no favorable disposition for petitioning party apparent
- 1 = petitioning party received a favorable disposition
- 2 = favorable disposition for petitioning party unclear

C.5 Direction of decision

Decision direction `decisionDirection` *numeric* [Source: SCDB codebook]

Ideological “direction” of the decision relative to the issue. The SCDB native 3-code scheme is shown below; the project’s LLM coder (and the rest of this codebook) instead uses the Songer 4-code convention 0/1/2/3 (not ascertained / conservative / mixed / liberal); see the direction-convention callout in §4.4.

- 1 = conservative
- 2 = liberal
- 3 = unspecifiable

C.6 Decision type

Decision type `decisionType` *numeric* [Source: SCDB codebook]

How the case was decided procedurally (signed opinion, per curiam, decree, etc.). Note that the project’s LLM coder (`scdb_first_method`; see §4.5) uses an adapted 4-code scheme oriented toward Courts of Appeals (panel / panel rehearing / en banc / per curiam), not the SCDB native scheme below.

- 1 = opinion of the court (orally argued)
- 2 = per curiam (no oral argument)
- 4 = decrees
- 5 = equally divided vote
- 6 = per curiam (orally argued)
- 7 = judgment of the Court (orally argued)
- 8 = seriatim

C.7 Declaration of unconstitutionality

Declaration of unconstitutionality `declarationUncon` *numeric* [Source: SCDB codebook]

Whether the Court declared a law unconstitutional, and if so what type of law. The project’s LLM coder (`scdb_first_unconst`; see §4.5) collapses this into a 3-code scheme (0 = no strike-down, 1 = facially invalid, 2 = unconstitutional as-applied) rather than distinguishing among types of law.

- 1 = no declaration of unconstitutionality
- 2 = act of congress declared unconstitutional
- 3 = state or territorial law, regulation, or constitutional provision unconstitutional
- 4 = municipal or other local ordinance unconstitutional

C.8 Legal provisions

Type of legal provision `lawType` *numeric* [Source: SCDB codebook]

Broad type of the constitutional provision, statute, or court rule the Court considered. The project's LLM coder (`scdb_first_legal_provisions`; see §4.5) instead emits a free-text JSON list of cited provisions, not an SCDB `lawType` code; the values below are provided for reference when cross-referencing raw SCDB.

- 1 = Constitution
- 2 = Constitutional Amendment
- 3 = Federal Statute
- 4 = Court Rules
- 5 = Other
- 6 = Infrequently litigated statutes
- 8 = State or local law or regulation
- 9 = No Legal Provision

D. Alphabetical variable index

An alphabetical listing of every column in the dataset, with a cross-reference to the section where its definition appears. Repeated patterns (e.g., `j1_id-j3_id`, `opinion_1_type-opinion_11_type`, `songer_gt_j2maj1-songer_gt_j16maj1`) are listed individually so that ctrl-F finds each one. The 10 SCDB-native names (e.g., `petitioner`, `decisionDirection`) point at Appendix C; they are reproduced from the SCDB codebook for joining convenience and are not columns in this dataset. *This index is generated automatically by `data/code/process/build_codebook_index.py` from the parquet schema and the codebook’s own `\var{}` entries. Rebuild after schema changes.*

<code>age_max</code> — Panel age (max) (§3.4)	(grouped) (§2.2)
<code>age_max_slot</code> — Panel age (max slot) (§3.4)	<code>author_name</code> — Majority author: judge name (§2.2)
<code>age_mean</code> — Panel age (mean) (§3.4)	<code>author_nonwhite</code> — Majority author: nonwhite indicator (§2.2)
<code>age_med</code> — Panel age (median) (§3.4)	<code>author_race</code> — Majority author: race or ethnicity (§2.2)
<code>age_med_slot</code> — Panel age (median slot) (§3.4)	<code>author_role_dime</code> — Author’s role on the ideology distribution (§2.2)
<code>age_min</code> — Panel age (min) (§3.4)	<code>author_role_ghp</code> — Author’s role on the ideology distribution (§2.2)
<code>age_min_slot</code> — Panel age (min slot) (§3.4)	<code>author_role_judjis</code> — Author’s role on the ideology distribution (§2.2)
<code>all_authors_matched</code> — All opinion authors matched (§1.5)	<code>author_senior</code> — Majority author: senior-status indicator (§2.2)
<code>all_female</code> — All-female panel (§3)	<code>author_senior_date</code> — Majority author: senior status date (§2.2)
<code>all_judges_matched</code> — All judges matched (§1.5)	<code>author_tenure</code> — Majority author: tenure at decision (years) (§2.2)
<code>all_male</code> — All-male panel (§3)	<code>caseDisposition</code> — Disposition of case (§C)
<code>all_nonwhite</code> — All-nonwhite panel (§3)	<code>caseid</code> — Case ID (§1)
<code>all_same_law</code> — Same law school (§3)	<code>circuit</code> — Circuit (§1)
<code>all_white</code> — All-white panel (§3)	<code>citation</code> — Citation (§1)
<code>author</code> — Majority opinion author (§1.4)	<code>cite_count</code> — Citation count (placeholder) (§1)
<code>author_aba</code> — Majority author: ABA rating at appointment (§2.2)	<code>conc1</code> — First concurrence author (§1.4)
<code>author_age</code> — Majority author: age at decision (§2.2)	<code>conc1_aba</code> — First concurrence: ABA rating at appointment (§2.3)
<code>author_appointing_president</code> — Majority author: appointing president (§2.2)	<code>conc1_age</code> — First concurrence: age at decision (§2.3)
<code>author_birth_city</code> — Majority author: birth city (§2.2)	<code>conc1_appointing_president</code> — First concurrence: appointing president (§2.3)
<code>author_birth_date</code> — Majority author: birth date (§2.2)	<code>conc1_birth_city</code> — First concurrence: birth city (§2.3)
<code>author_birth_state</code> — Majority author: birth state (§2.2)	<code>conc1_birth_date</code> — First concurrence: birth date (§2.3)
<code>author_commission_date</code> — Majority author: commission date (§2.2)	<code>conc1_birth_state</code> — First concurrence: birth state (§2.3)
<code>author_court_name</code> — Majority author: court name (§2.2)	<code>conc1_commission_date</code> — First concurrence: commission date (§2.3)
<code>author_dem</code> — Author is on a Democratic-appointee panel slot (§2.2)	<code>conc1_court_name</code> — First concurrence: court name (§2.3)
<code>author_dime</code> — Majority author: DIME ideology score (§2.2)	<code>conc1_dem</code> — First concurrence: democratic appointee (§2.3)
<code>author_duty_state</code> — Majority author: duty state (§2.2)	<code>conc1_dime</code> — First concurrence: DIME ideology score (§2.3)
<code>author_fem</code> — Majority author: female indicator (§2.2)	<code>conc1_duty_state</code> — First concurrence: duty state (§2.3)
<code>author_ghp</code> — Majority author: GHP ideology score (§2.2)	<code>conc1_fem</code> — First concurrence: female indicator (§2.3)
<code>author_id</code> — Majority author: FJC ID (§2.2)	<code>conc1_ghp</code> — First concurrence: GHP ideology score (§2.3)
<code>author_is_appellate</code> — Majority author: court-of-appeals appointee indicator (§2.2)	<code>conc1_id</code> — First concurrence: FJC ID (§2.3)
<code>author_judjis</code> — Majority author: JudJIS ideology score (§2.2)	
<code>author_law</code> — Majority author: law school (normalized) (§2.2)	
<code>author_law_grouped</code> — Majority author: law school	

conc1_is_appellate — First concurrence: court-of-appeals appointee indicator (§2.3)
conc1_judjis — First concurrence: JudJIS ideology score (§2.3)
conc1_law — First concurrence: law school (normalized) (§2.3)
conc1_law_grouped — First concurrence: law school (grouped) (§2.3)
conc1_name — First concurrence: judge name (§2.3)
conc1_nonwhite — First concurrence: nonwhite indicator (§2.3)
conc1_race — First concurrence: race or ethnicity (§2.3)
conc1_role_dime — First concurrence: role on DIME ideology distribution (§2.3)
conc1_role_ghp — First concurrence: role on GHP ideology distribution (§2.3)
conc1_role_judjis — First concurrence: role on JudJIS ideology distribution (§2.3)
conc1_senior — First concurrence: senior-status indicator (§2.3)
conc1_senior_date — First concurrence: senior status date (§2.3)
conc1_tenure — First concurrence: tenure at decision (years) (§2.3)
decisionDirection — Decision direction (§C)
decisionType — Decision type (§C)
decision_date — Decision date (§1)
declarationUncon — Declaration of unconstitutionality (§C)
dime_max — Panel DIME score (max) (§3.4)
dime_max_slot — Panel DIME score (max slot) (§3.4)
dime_mean — Panel DIME score (mean) (§3.4)
dime_med — Panel DIME score (median) (§3.4)
dime_med_slot — Panel DIME score (median slot) (§3.4)
dime_min — Panel DIME score (min) (§3.4)
dime_min_slot — Panel DIME score (min slot) (§3.4)
dis1 — First-dissent author (§1.4)
dis1_aba — First dissent: ABA rating at appointment (§2.3)
dis1_age — First dissent: age at decision (§2.3)
dis1_appointing_president — First dissent: appointing president (§2.3)
dis1_birth_city — First dissent: birth city (§2.3)
dis1_birth_date — First dissent: birth date (§2.3)
dis1_birth_state — First dissent: birth state (§2.3)
dis1_commission_date — First dissent: commission date (§2.3)
dis1_court_name — First dissent: court name (§2.3)
dis1_dem — First dissent: democratic appointee (§2.3)
dis1_dime — First dissent: DIME ideology score (§2.3)
dis1_duty_state — First dissent: duty state (§2.3)
dis1_fem — First dissent: female indicator (§2.3)
dis1_ghp — First dissent: GHP ideology score (§2.3)
dis1_id — First dissent: FJC ID (§2.3)
dis1_is_appellate — First dissent: court-of-appeals appointee indicator (§2.3)
dis1_judjis — First dissent: JudJIS ideology score (§2.3)
dis1_law — First dissent: law school (normalized) (§2.3)
dis1_law_grouped — First dissent: law school (grouped) (§2.3)
dis1_name — First dissent: judge name (§2.3)
dis1_nonwhite — First dissent: nonwhite indicator (§2.3)
dis1_race — First dissent: race or ethnicity (§2.3)
dis1_role_dime — First dissent: role on DIME ideology distribution (§2.3)
dis1_role_ghp — First dissent: role on GHP ideology distribution (§2.3)
dis1_role_judjis — First dissent: role on JudJIS ideology distribution (§2.3)
dis1_senior — First dissent: senior-status indicator (§2.3)
dis1_senior_date — First dissent: senior status date (§2.3)
dis1_tenure — First dissent: tenure at decision (years) (§2.3)
docket_number — Docket number (§1)
enbanc_aba — En banc: ABA rating at appointment (§2.4)
enbanc_age — En banc: age at decision (§2.4)
enbanc_appointing_president — En banc: appointing president (§2.4)
enbanc_birth_city — En banc: birth city (§2.4)
enbanc_birth_date — En banc: birth date (§2.4)
enbanc_birth_state — En banc: birth state (§2.4)
enbanc_commission_date — En banc: commission date (§2.4)
enbanc_court_name — En banc: court name (§2.4)
enbanc_dem — En banc: democratic appointee (§2.4)
enbanc_dime — En banc: DIME ideology score (§2.4)
enbanc_duty_state — En banc: duty state (§2.4)
enbanc_fem — En banc: female indicator (§2.4)
enbanc_ghp — En banc: GHP ideology score (§2.4)
enbanc_id — En banc: FJC ID (§2.4)
enbanc_is_appellate — En banc: court-of-appeals appointee indicator (§2.4)
enbanc_judjis — En banc: JudJIS ideology score (§2.4)
enbanc_law — En banc: law school (normalized) (§2.4)
enbanc_law_grouped — En banc: law school (grouped) (§2.4)
enbanc_name — En banc: judge name (§2.4)
enbanc_nonwhite — En banc: nonwhite indicator (§2.4)
enbanc_race — En banc: race or ethnicity (§2.4)
enbanc_senior — En banc: senior-status indicator (§2.4)
enbanc_senior_date — En banc: senior status date (§2.4)
enbanc_tenure — En banc: tenure at decision (years) (§2.4)
ghp_max — Panel GHP score (max) (§3.4)
ghp_max_slot — Panel GHP score (max slot) (§3.4)
ghp_mean — Panel GHP score (mean) (§3.4)
ghp_med — Panel GHP score (median) (§3.4)
ghp_med_slot — Panel GHP score (median slot) (§3.4)
ghp_min — Panel GHP score (min) (§3.4)
ghp_min_slot — Panel GHP score (min slot) (§3.4)
idb_agency — Reviewing agency (§B)

idb_alts_json — Alternate IDB matches (§B)
idb_aofilemo — AOUSC filing month (§B)
idb_aofileyr — AOUSC filing year (§B)
idb_appdate — Notice of appeal filed (§B)
idb_appellan — Appellant name (§B)
idb_appellee — Appellee name (§B)
idb_apptype — Type of appeal (§B)
idb_brfiled — Last brief filed (§B)
idb_circuit — Circuit (§B)
idb_cnslapp — Counsel appointed (§B)
idb_cnslcir — Counsel source — circuit (§B)
idb_cnslcont — Counsel continued (§B)
idb_cnsldis — Counsel source — district (§B)
idb_consdkt — Consolidated/cross-appeal docket (§B)
idb_cpc — Certificate of probable cause decision (§B)
idb_cpcdisc — Certificate of probable cause decision-maker (§B)
idb_crecdate — Briefing schedule issued (§B)
idb_dcirc — Originating district circuit (§B)
idb_ddefnum — District-court defendant number (§B)
idb_ddist — Originating district (§B)
idb_ddktdate — Docketed in lower court (§B)
idb_ddocket — District-court docket number (§B)
idb_disp — Type of disposition action (§B)
idb_divoff — Divisional office (legacy) (§B)
idb_djudge — District judge code (§B)
idb_dktdate — Docketed in court of appeals (§B)
idb_docket — Appellate docket number (§B)
idb_doffice — Divisional office (§B)
idb_enbanc — En banc indicator (§B)
idb_fedcap — Federal capital case (§B)
idb_filefee — In forma pauperis status at filing (§B)
idb_heardate — Oral argument date (§B)
idb_ifp — In forma pauperis (pre-SY81) (§B)
idb_jdg1inv — Judge 1 involvement (§B)
idb_jdg2inv — Judge 2 involvement (§B)
idb_jdg3inv — Judge 3 involvement (§B)
idb_jdgcode1 — Panel judge codes (§B)
idb_jdgcode2 — Panel judge codes (§B)
idb_jdgcode3 — Panel judge codes (§B)
idb_jdgpanel — Full judge panel indicator (§B)
idb_joinapp — Joined-appeal type (§B)
idb_judgdate — Final judgment date (§B)
idb_juris — Jurisdictional basis (§B)
idb_match_stage — Match stage (§B)
idb_method — Clerk-only termination method (§B)
idb_newsytrm — Statistical year of termination (legacy) (§B)
idb_nos — Nature of suit (§B)
idb_offense — Statutory offense (§B)
idb_opdisp — Original-proceeding determination (§B)
idb_opinion — Type of separate opinions (§B)
idb_orgproc — Nature of original proceeding (§B)
idb_othtype — Other type (bankruptcy) (§B)
idb_outcome — Merits disposition outcome (§B)
idb_present — IDB record available (§1.5)
idb_procterm — Procedural-termination type (§B)
idb_prosefle — Pro-se status at filing (§B)
idb_prosetrm — Pro-se status at termination (§B)
idb_pubstat — Publication status (§B)
idb_reopen — Reopened or remanded indicator (§B)
idb_statuscd — Record status (§B)
idb_subdate — Submission date (§B)
idb_tapeyear — Statistical / fiscal year of record (§B)
idb_termdate — Termination date (legacy) (§B)
idb_torprosub — Original-proceeding subtype (§B)
idb_transcode — Last transaction code (§B)
idb_transdate — Last transaction date (§B)
idb_trmfee — In forma pauperis status at termination (§B)
idb_trmtype — Termination-time type of appeal (§B)
idb_trusedmo — AOUSC termination month (§B)
idb_trusedyr — AOUSC termination year (§B)
idb_usape — U.S. as appellee (§B)
idb_usapt — U.S. as appellant (§B)
is_enbanc — En banc indicator (§1.5)
is_percuriam — Per curiam indicator (§1.5)
issue — Specific issue (§C)
issueArea — Issue area (§C)
j1_aba — ABA rating at appointment (§2.1)
j1_age — Judge’s age at decision (§2.1)
j1_appointing_president — Appointing president (§2.1)
j1_birth_city — Birth city (§2.1)
j1_birth_date — Birth date (§2.1)
j1_birth_state — Birth state (§2.1)
j1_commission_date — Commission date (§2.1)
j1_court_name — Court name (§2.1)
j1_dem — Democratic-appointee indicator (§2.1)
j1_dime — DIME ideology score (§2.1)
j1_duty_state — Duty state (§2.1)
j1_fem — Female indicator (§2.1)
j1_ghp — Giles-Hettinger-Peppers ideology score (§2.1)
j1_id — Judge identifier (§2.1)
j1_is_appellate — Court-of-Appeals appointee indicator (§2.1)
j1_judjis — JudJIS ideology score (§2.1)
j1_law — Law school (normalized) (§2.1)
j1_law_grouped — Law school (grouped) (§2.1)
j1_name — Judge name (§2.1)
j1_nonwhite — Nonwhite indicator (§2.1)
j1_race — Race or ethnicity (§2.1)
j1_senior — Senior status indicator (§2.1)
j1_senior_date — Senior status date (§2.1)
j1_tenure — Years on the bench at decision (§2.1)
j2_aba — ABA rating at appointment (§2.1)
j2_age — Judge’s age at decision (§2.1)
j2_appointing_president — Appointing president (§2.1)
j2_birth_city — Birth city (§2.1)
j2_birth_date — Birth date (§2.1)
j2_birth_state — Birth state (§2.1)
j2_commission_date — Commission date (§2.1)
j2_court_name — Court name (§2.1)
j2_dem — Democratic-appointee indicator (§2.1)
j2_dime — DIME ideology score (§2.1)

j2.duty_state — Duty state (§2.1)
j2.fem — Female indicator (§2.1)
j2.ghp — Giles-Hettinger-Peppers ideology score (§2.1)
j2.id — Judge identifier (§2.1)
j2.is_appellate — Court-of-Appeals appointee indicator (§2.1)
j2.judjis — JudJIS ideology score (§2.1)
j2.law — Law school (normalized) (§2.1)
j2.law_grouped — Law school (grouped) (§2.1)
j2.name — Judge name (§2.1)
j2.nonwhite — Nonwhite indicator (§2.1)
j2.race — Race or ethnicity (§2.1)
j2.senior — Senior status indicator (§2.1)
j2.senior_date — Senior status date (§2.1)
j2.tenure — Years on the bench at decision (§2.1)
j3.aba — ABA rating at appointment (§2.1)
j3.age — Judge's age at decision (§2.1)
j3.appointing_president — Appointing president (§2.1)
j3.birth_city — Birth city (§2.1)
j3.birth_date — Birth date (§2.1)
j3.birth_state — Birth state (§2.1)
j3.commission_date — Commission date (§2.1)
j3.court_name — Court name (§2.1)
j3.dem — Democratic-appointee indicator (§2.1)
j3.dime — DIME ideology score (§2.1)
j3.duty_state — Duty state (§2.1)
j3.fem — Female indicator (§2.1)
j3.ghp — Giles-Hettinger-Peppers ideology score (§2.1)
j3.id — Judge identifier (§2.1)
j3.is_appellate — Court-of-Appeals appointee indicator (§2.1)
j3.judjis — JudJIS ideology score (§2.1)
j3.law — Law school (normalized) (§2.1)
j3.law_grouped — Law school (grouped) (§2.1)
j3.name — Judge name (§2.1)
j3.nonwhite — Nonwhite indicator (§2.1)
j3.race — Race or ethnicity (§2.1)
j3.senior — Senior status indicator (§2.1)
j3.senior_date — Senior status date (§2.1)
j3.tenure — Years on the bench at decision (§2.1)
j.enbanc — En banc judge NIDs (§2.4)
judge_count — Panel size (§1.5)
judjis_max — Panel JudJIS score (max) (§3.4)
judjis_max_slot — Panel JudJIS score (max slot) (§3.4)
judjis_mean — Panel JudJIS score (mean) (§3.4)
judjis_med — Panel JudJIS score (median) (§3.4)
judjis_med_slot — Panel JudJIS score (median slot) (§3.4)
judjis_min — Panel JudJIS score (min) (§3.4)
judjis_min_slot — Panel JudJIS score (min slot) (§3.4)
lawType — Type of legal provision (§C)
llm_present — LLM coding available (§1.5)
lone_man_slot — Lone man (§3)
lone_nonwhite_slot — Lone nonwhite judge (§3)
lone_white_slot — Lone white judge (§3)
lone_woman_slot — Lone woman (§3)
minority_slot — Lone minority-party judge (§3)
mix1 — First-mixed-opinion author (§1.4)
mix1.aba — First mixed opinion: ABA rating at appointment (§2.3)
mix1.age — First mixed opinion: age at decision (§2.3)
mix1.appointing_president — First mixed opinion: appointing president (§2.3)
mix1.birth_city — First mixed opinion: birth city (§2.3)
mix1.birth_date — First mixed opinion: birth date (§2.3)
mix1.birth_state — First mixed opinion: birth state (§2.3)
mix1.commission_date — First mixed opinion: commission date (§2.3)
mix1.court_name — First mixed opinion: court name (§2.3)
mix1.dem — First mixed opinion: democratic appointee (§2.3)
mix1.dime — First mixed opinion: DIME ideology score (§2.3)
mix1.duty_state — First mixed opinion: duty state (§2.3)
mix1.fem — First mixed opinion: female indicator (§2.3)
mix1.ghp — First mixed opinion: GHP ideology score (§2.3)
mix1.id — First mixed opinion: FJC ID (§2.3)
mix1.is_appellate — First mixed opinion: court-of-appeals appointee indicator (§2.3)
mix1.judjis — First mixed opinion: JudJIS ideology score (§2.3)
mix1.law — First mixed opinion: law school (normalized) (§2.3)
mix1.law_grouped — First mixed opinion: law school (grouped) (§2.3)
mix1.name — First mixed opinion: judge name (§2.3)
mix1.nonwhite — First mixed opinion: nonwhite indicator (§2.3)
mix1.race — First mixed opinion: race or ethnicity (§2.3)
mix1.role.dime — First mixed opinion: role on DIME ideology distribution (§2.3)
mix1.role.ghp — First mixed opinion: role on GHP ideology distribution (§2.3)
mix1.role.judjis — First mixed opinion: role on JudJIS ideology distribution (§2.3)
mix1.senior — First mixed opinion: senior-status indicator (§2.3)
mix1.senior_date — First mixed opinion: senior status date (§2.3)
mix1.tenure — First mixed opinion: tenure at decision (years) (§2.3)
mixed_panel — Mixed-party panel (§3)
n.concurrence — Number of concurrences (§1.4)
n.dissent — Number of dissents (§1.4)
n.majority — Number of majority opinions (§1.4)
n.mixed — Number of mixed opinions (§1.4)
n.opinions — Total opinions (§1.4)
n.rehearing — Number of rehearing orders (§1.4)
num.dems — Democratic appointees (§3)
num.law.schools — Distinct law schools (§3)

num_nonwhite — Nonwhite judges on panel (§3)
num_women — Women on panel (§3)
opinion_10.author — Opinion author (§1.4)
opinion_10.type — Opinion type (§1.4)
opinion_10.words — Opinion word count (§1.4)
opinion_11.author — Opinion author (§1.4)
opinion_11.type — Opinion type (§1.4)
opinion_11.words — Opinion word count (§1.4)
opinion_1.author — Opinion author (§1.4)
opinion_1.type — Opinion type (§1.4)
opinion_1.words — Opinion word count (§1.4)
opinion_2.author — Opinion author (§1.4)
opinion_2.type — Opinion type (§1.4)
opinion_2.words — Opinion word count (§1.4)
opinion_3.author — Opinion author (§1.4)
opinion_3.type — Opinion type (§1.4)
opinion_3.words — Opinion word count (§1.4)
opinion_4.author — Opinion author (§1.4)
opinion_4.type — Opinion type (§1.4)
opinion_4.words — Opinion word count (§1.4)
opinion_5.author — Opinion author (§1.4)
opinion_5.type — Opinion type (§1.4)
opinion_5.words — Opinion word count (§1.4)
opinion_6.author — Opinion author (§1.4)
opinion_6.type — Opinion type (§1.4)
opinion_6.words — Opinion word count (§1.4)
opinion_7.author — Opinion author (§1.4)
opinion_7.type — Opinion type (§1.4)
opinion_7.words — Opinion word count (§1.4)
opinion_8.author — Opinion author (§1.4)
opinion_8.type — Opinion type (§1.4)
opinion_8.words — Opinion word count (§1.4)
opinion_9.author — Opinion author (§1.4)
opinion_9.type — Opinion type (§1.4)
opinion_9.words — Opinion word count (§1.4)
parse_count — LLM parse attempts (§4)
partyWinning — Winning party (§C)
petitioner — Petitioner (§C)
regex_disp.concl_found — Conclusion-section header detected (§1.3)
regex_disp.plain.concl — Disposition — plain, conclusion-block (§1.3)
regex_disp.plain.walkback — Disposition — plain, walkback (§1.3)
regex_disp.strict.concl — Disposition — strict, conclusion-block (§1.3)
regex_disp.strict.walkback — Disposition — strict, walkback (§1.3)
respondent — Respondent (§C)
scdb_first_appealed — Who appealed (§4.1)
scdb_first_appellant_category — Appellant category (§4.1)
scdb_first_appellant_description — Appellant description (§4.1)
scdb_first_appellant_name — Appellant name (§4.1)
scdb_first_appellant_win — Winning party (§4.2)
scdb_first_broad_area — Broad issue area (§4.3)
scdb_first_broad_area.alt — Alternate broad issue area (§4.3)
scdb_first_direction_oneshot — Direction (one-shot) (§4.4)
scdb_first_disposition — Disposition (§4.2)
scdb_first_justification — First-pass justification (§4.5)
scdb_first_legal_provisions — Cited legal provisions (§4.5)
scdb_first_method — Method of review (§4.5)
scdb_first_procedural — Procedural holding (§4.3)
scdb_first_respondent_category — Respondent category (§4.1)
scdb_first_respondent_description — Respondent description (§4.1)
scdb_first_respondent_name — Respondent name (§4.1)
scdb_first_unconst — Declaration of unconstitutionality (§4.5)
scdb_second_direction_flag — Direction edge-case flag (§4.4)
scdb_second_direction_mech — Mechanical direction (§4.4)
scdb_second_direction_supermech — Super-mechanical direction (§4.4)
scdb_second_focal — Focal party (§4.4)
scdb_second_focal.alt — Other focal parties considered (§4.4)
scdb_second_focal_appellant — Focal party is appellant (§4.4)
scdb_second_focal_won — Focal party prevailed (§4.4)
scdb_second_justification — Second-pass justification (§4.5)
scdb_second_sub_area — Specific issue area (§4.3)
songer_first_appealed_from — Type of lower-court judgment (§4.10)
songer_first_appellant_category — Appellant category (broad) (§4.6)
songer_first_authority — Basis of decision (authority) (§4.10)
songer_first_broad_area — Broad issue area (§4.8)
songer_first_direction_oneshot — Direction (one-shot LLM) (§4.9)
songer_first_justification — First-pass justification (§4.10)
songer_first_procedural — Procedural holding (§4.8)
songer_first_respondent_category — Respondent category (broad) (§4.6)
songer_first_threshold_failure_to_state_claim — Threshold issue: failure to state a claim (§4.8)
songer_first_threshold_frivolous_appeal — Threshold issue: frivolous appeal (§4.8)
songer_first_threshold_frivolous_case — Threshold issue: frivolous case (§4.8)
songer_first_threshold_immunity — Threshold issue: governmental immunity (§4.8)
songer_first_threshold_jurisdiction — Threshold issue: jurisdiction (§4.8)
songer_first_threshold_late_appeal — Threshold issue: late appeal (§4.8)
songer_first_threshold_mootness — Threshold issue:

mootness (§4.8)
songer_first_threshold.other_appellate_threshold — Threshold issue: other appellate-level threshold (§4.8)
songer_first_threshold.other_trial_threshold — Threshold issue: other trial-level threshold (§4.8)
songer_first_threshold.political_question — Threshold issue: political question (§4.8)
songer_first_threshold.ripeness_exhaustion — Threshold issue: ripeness / exhaustion (§4.8)
songer_first_threshold.standing — Threshold issue: standing (§4.8)
songer_first_threshold.timeliness — Threshold issue: timeliness (§4.8)
songer_first_typeiss — Type of issue (§4.8)
songer_gt_aamicusbriefs — Amicus briefs for appellant (§A.2.1)
songer_gt_aatty — Number of appellant attorneys (§A.2.1)
songer_gt_aattyfirstname — First appellant attorney name (§A.2.1)
songer_gt_aattylastname — First appellant attorney name (§A.2.1)
songer_gt_abusedis — Deference to agency discretion (§A.2.1)
songer_gt_acounsel1 — Appellant counsel (legacy duplicate) (§A.2.1)
songer_gt_adminrev — Administrative review (§A.2.1)
songer_gt_afirmname — First appellant firm name (§A.2.1)
songer_gt_afirms — Number of appellant law firms (§A.2.1)
songer_gt_afirstlistedamicusgroup — First amicus group for appellant (§A.2.1)
songer_gt_agen_acq — Agency acquisition of information (§A.2.1)
songer_gt_agroups — Appellant interest groups (§A.2.1)
songer_gt_alj — ALJ decision (§A.2.1)
songer_gt_altdisp — Alternative dispute resolution (§A.2.1)
songer_gt_amicus — Amicus participation (§A.2.1)
songer_gt_app_stid — State of first state/local appellant (§A.2.1)
songer_gt_appbus — Appellant counts by category (§A.2.1)
songer_gt_appel1 — Appellant detail code (first listed) (§A.2.1)
songer_gt_appel2 — Appellant detail code (second listed) (§A.2.1)
songer_gt_appfed — Appellant counts by category (§A.2.1)
songer_gt_appfiduc — Appellant counts by category (§A.2.1)
songer_gt_applfrom — District-court decision appealed from (§A.2.1)
songer_gt_appnatpr — Appellant counts by category (§A.2.1)
songer_gt_appnonp — Appellant counts by category (§A.2.1)
songer_gt_appstate — Appellant counts by category (§A.2.1)
songer_gt_appsubst — Appellant counts by category (§A.2.1)
songer_gt_astate_code — Appellant state code (legacy alias) (§A.2.1)
songer_gt_attyfee — Attorneys' fees (§A.2.1)
songer_gt.bank_ap1 — Appellant bankruptcy indicators (§A.2.1)
songer_gt.bank_ap2 — Appellant bankruptcy indicators (§A.2.1)
songer_gt.bank_r1 — Respondent bankruptcy indicators (§A.2.1)
songer_gt.bank_r2 — Respondent bankruptcy indicators (§A.2.1)
songer_gt.beginpg — Beginning page (§A)
songer_gt.capric — Arbitrary and capricious (§A.2.1)
songer_gt.casenum — Case number (§A)
songer_gt.casetyp1 — Primary case-type code (§A)
songer_gt.casetyp2 — Second case-type code (§A)
songer_gt.circuit — Circuit (§A)
songer_gt.cite — Citation (§A)
songer_gt.civproc1 — Most-cited Federal Rule of Civil Procedure (§A.2.1)
songer_gt.civproc2 — Second most-cited Federal Rule of Civil Procedure (§A.2.1)
songer_gt.classact — Class action (§A.2.1)
songer_gt.cmbasis — Concur-merits basis (first concurrence) (§A.2.1)
songer_gt.cmbasis2 — Concur-merits basis (second concurrence) (§A.2.1)
songer_gt.codej1 — Panel judge codes (§A.2.1)
songer_gt.codej10 — Panel judge codes (§A.2.1)
songer_gt.codej11 — Panel judge codes (§A.2.1)
songer_gt.codej12 — Panel judge codes (§A.2.1)
songer_gt.codej13 — Panel judge codes (§A.2.1)
songer_gt.codej14 — Panel judge codes (§A.2.1)
songer_gt.codej15 — Panel judge codes (§A.2.1)
songer_gt.codej2 — Panel judge codes (§A.2.1)
songer_gt.codej3 — Panel judge codes (§A.2.1)
songer_gt.codej4 — Panel judge codes (§A.2.1)
songer_gt.codej5 — Panel judge codes (§A.2.1)
songer_gt.codej6 — Panel judge codes (§A.2.1)
songer_gt.codej7 — Panel judge codes (§A.2.1)
songer_gt.codej8 — Panel judge codes (§A.2.1)
songer_gt.codej9 — Panel judge codes (§A.2.1)
songer_gt.comment — Notice-and-comment rulemaking (§A.2.1)
songer_gt.concur — Number of concurrences (§A.2.1)
songer_gt.confess — Confession admissibility (§A.2.1)
songer_gt.const1 — Most-cited constitutional provision (§A.2.1)
songer_gt.const2 — Second most-cited constitutional provision (§A.2.1)
songer_gt.constit — Constitutionality issue (§A.2.1)
songer_gt.counsel — Counsel (project metadata) (§A.2.1)
songer_gt.counsel1 — Nature of appellant counsel (§A.2.1)

songer_gt_counsel2 — Nature of respondent counsel (§A.2.1)
songer_gt_cpbasis — Concur-procedural basis (first concurrence) (§A.2.1)
songer_gt_cpbasis2 — Concur-procedural basis (second concurrence) (§A.2.1)
songer_gt_crmproc1 — Most-cited Federal Rule of Criminal Procedure (§A.2.1)
songer_gt_crmproc2 — Second most-cited Federal Rule of Criminal Procedure (§A.2.1)
songer_gt_crossapp — Cross-appeals (§A.2.1)
songer_gt_day — Day of decision (§A)
songer_gt_deathpen — Death penalty (§A.2.1)
songer_gt_decisiondate — Decision date (§A)
songer_gt_decuncon — Type of unconstitutionality declaration (§A.2.1)
songer_gt_denovo — De novo review (§A.2.1)
songer_gt_direct1 — Direction of decision on casetyp1 (§A.2.1)
songer_gt_direct2 — Direction of decision on casetyp2 (§A.2.1)
songer_gt_discover — Discovery (§A.2.1)
songer_gt_dissent — Number of dissents (§A.2.1)
songer_gt_distjudg — District-court judge code (§A.2.1)
songer_gt_district — District within state (§A.2.1)
songer_gt_diverse — Diversity sufficiency (§A.2.1)
songer_gt_dmbasis — Dissent-merits basis (§A.2.1)
songer_gt_docknum — Docket number (legacy alias) (§A)
songer_gt_docnum — Docket number (§A)
songer_gt_dpbasis — Dissent-procedural basis (§A.2.1)
songer_gt_dueproc — Due process (§A.2.1)
songer_gt_endopin — End page of opinion (§A)
songer_gt_endpage — End page of case (§A)
songer_gt_entrap — Entrapment (§A.2.1)
songer_gt_erron — Clearly erroneous (§A.2.1)
songer_gt_execord — Executive order / agency regulation (§A.2.1)
songer_gt_exhaust — Exhaustion / ripeness (threshold) (§A.2.1)
songer_gt_fedlaw — Federal statute interpretation (§A.2.1)
songer_gt_fedlaw2 — Federal-law case-law application (§A.2.1)
songer_gt_fedvst — Federal vs. state law (§A.2.1)
songer_gt_fileid — File identifier (§A)
songer_gt_foreign — Domestic vs. foreign law (§A.2.1)
songer_gt_freeinfo — Freedom of information (§A.2.1)
songer_gt_frivapp — Frivolous appeal (threshold) (§A.2.1)
songer_gt_frivol — Frivolous case (threshold) (§A.2.1)
songer_gt_genapel1 — Appellant category (first listed) (§A.2.1)
songer_gt_genapel2 — Appellant category (second listed) (§A.2.1)
songer_gt_geniss — General issue summary (§A.2.1)
songer_gt_genresp1 — Respondent category (first listed) (§A.2.1)
songer_gt_genresp2 — Respondent category (second listed) (§A.2.1)
songer_gt_genstand — General standard (§A.2.1)
songer_gt_govt — Government-as-party indicator (§A.2.1)
songer_gt_habeas — Habeas corpus (§A.2.1)
songer_gt_immunity — Governmental immunity (threshold) (§A.2.1)
songer_gt_improper — Improper jury influence (§A.2.1)
songer_gt_indict — Indictment (§A.2.1)
songer_gt_indigent — Indigent-defendant rights (§A.2.1)
songer_gt_initiate — Party initiating the appeal (§A.2.1)
songer_gt_injunct — Injunction (§A.2.1)
songer_gt_insane — Insanity defense (§A.2.1)
songer_gt_int_law — International law (§A.2.1)
songer_gt_interven — Intervenor (§A.2.1)
songer_gt_isstyp — Type of issue (legacy alias) (§A.2.1)
songer_gt_j10maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j10maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j10vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j10vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j11maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j11maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j11vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j11vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j12maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j12maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j12vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j12vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j13maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j13maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j13vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j13vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j14maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j14maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j14vote1 — Judge vote direction on casetyp1

(§A.2.1)
songer_gt_j14vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j15maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j15maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j15vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j15vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j16maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j16vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j2maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j2maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j2vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j2vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j3maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j3maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j3vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j3vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j4maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j4maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j4vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j4vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j5maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j5maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j5vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j5vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j6maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j6maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j6vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j6vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j7maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j7maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j7vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j7vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j8maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j8maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j8vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j8vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_j9maj1 — Judge majority indicators on casetyp1 (§A.2.1)
songer_gt_j9maj2 — Judge majority indicators on casetyp2 (§A.2.1)
songer_gt_j9vote1 — Judge vote direction on casetyp1 (§A.2.1)
songer_gt_j9vote2 — Judge vote direction on casetyp2 (§A.2.1)
songer_gt_judgdisc — Judicial discretion (§A.2.1)
songer_gt_judrev — Judicial review of agency action (§A.2.1)
songer_gt_juris — Jurisdiction (threshold) (§A.2.1)
songer_gt_juryinst — Jury instructions (§A.2.1)
songer_gt_keynotes — Total West key numbers (§A.2.1)
songer_gt_late — Late appeal (threshold) (§A.2.1)
songer_gt_majvotes — Majority votes (§A.2.1)
songer_gt_merge_log — Merge audit log (§A)
songer_gt_method — Nature of the proceeding (§A.2.1)
songer_gt_month — Month of decision (§A)
songer_gt_mootness — Mootness (threshold) (§A.2.1)
songer_gt_multdoc — Multiple docket count (§A)
songer_gt_multiple — Multi-citation flag (§A)
songer_gt_notice — Notice (§A.2.1)
songer_gt_numappel — Number of appellants (§A.2.1)
songer_gt_numresp — Number of respondents (§A.2.1)
songer_gt_opinstat — Opinion type (§A.2.1)
songer_gt_origin — Court of origin (§A.2.1)
songer_gt_othadmis — Other evidence admissibility (§A.2.1)
songer_gt_othappth — Other appellate-level threshold (§A.2.1)
songer_gt_othcivil — Other civil issue (§A.2.1)
songer_gt_othcrim — Other criminal grounds (§A.2.1)
songer_gt_othjury — Other jury issues (§A.2.1)
songer_gt_oththres — Other trial-level threshold (§A.2.1)
songer_gt_pagelgth — Page length (§A)
songer_gt_phase — Source phase (§A)
songer_gt_plea — Plea bargaining (§A.2.1)
songer_gt_polques — Political question (threshold) (§A.2.1)
songer_gt_polquest — Political question (threshold) (§A.2.1)
songer_gt_post_tr1 — Post-trial procedure (§A.2.1)
songer_gt_prejud — Prosecutor misconduct (§A.2.1)
songer_gt_pretrial — Pretrial procedure (§A.2.1)
songer_gt_priorpub — Prior published opinion (§A)

songer_gt_procdis — Procedural dismissal (§A.2.1)
songer_gt_procedur — Precedent interpretation (non-statutory) (§A.2.1)
songer_gt_r_bus — Respondent counts by category (§A.2.1)
songer_gt_r_fed — Respondent counts by category (§A.2.1)
songer_gt_r_fiduc — Respondent counts by category (§A.2.1)
songer_gt_r_natpr — Respondent counts by category (§A.2.1)
songer_gt_r_nonp — Respondent counts by category (§A.2.1)
songer_gt_r_state — Respondent counts by category (§A.2.1)
songer_gt_r_stid — State of first state/local respondent (§A.2.1)
songer_gt_r_subst — Respondent counts by category (§A.2.1)
songer_gt_ramicusbrfs — Amicus briefs for respondent (§A.2.1)
songer_gt_ratty — Number of respondent attorneys (§A.2.1)
songer_gt_rattyfirstname — First respondent attorney name (§A.2.1)
songer_gt_rattylastname — First respondent attorney name (§A.2.1)
songer_gt_rcounselcode — Respondent counsel (legacy duplicate) (§A.2.1)
songer_gt_realapp — Real appellant indicator (§A.2.1)
songer_gt_realresp — Real respondent indicator (§A.2.1)
songer_gt_record — Adequacy of agency record (§A.2.1)
songer_gt_respond1 — Respondent detail code (first listed) (§A.2.1)
songer_gt_respond2 — Respondent detail code (second listed) (§A.2.1)
songer_gt_resstate — Respondent state code (legacy alias) (§A.2.1)
songer_gt_rfirmname — First respondent firm name (§A.2.1)
songer_gt_rfirms — Number of respondent law firms (§A.2.1)
songer_gt_rgroups — Respondent interest groups (§A.2.1)
songer_gt_rnameoffirstlistedamicusgroup — First amicus group for respondent (§A.2.1)
songer_gt_rtcouns — Right to counsel (§A.2.1)
songer_gt_sample — Sampling stratum (§A)
songer_gt_sanction — Sanctions (§A.2.1)
songer_gt_search — Search and seizure (§A.2.1)
songer_gt_sentence — Sentencing (§A.2.1)
songer_gt_signed — Signed-opinion indicator (§A.2.1)
songer_gt_source — Source forum (§A)
songer_gt_st_v_st — State vs. state (§A.2.1)
songer_gt_standing — Standing (threshold) (§A.2.1)
songer_gt_state — State of origin (§A.2.1)
songer_gt_statecl — Failure to state a claim (threshold, legacy alias) (§A.2.1)
songer_gt_statecode — State code (legacy alias) (§A.2.1)
songer_gt_stpolicy — State policy (§A.2.1)
songer_gt_subevid — Substantial evidence (§A.2.1)
songer_gt_suffic — Sufficiency of evidence (§A.2.1)
songer_gt_summary — Summary judgment (§A.2.1)
songer_gt_supct — Supreme Court review indicator (§A.2.1)
songer_gt_threshgen — General threshold flag (§A.2.1)
songer_gt_timely — Timeliness (threshold) (§A.2.1)
songer_gt_treat — Court of appeals disposition (§A.2.1)
songer_gt_trialpro — Trial procedure (§A.2.1)
songer_gt_turnthresh — Case turned on threshold issue (§A.2.1)
songer_gt_typeiss — Type of issue (§A.2.1)
songer_gt_unan — Unanimous decision indicator (§A.2.1)
songer_gt_usc1 — Most-cited U.S. Code title (§A.2.1)
songer_gt_usc1sect — Section of most-cited U.S. Code title (§A.2.1)
songer_gt_usc2 — Second most-cited U.S. Code title (§A.2.1)
songer_gt_usc2sect — Section of second most-cited U.S. Code title (§A.2.1)
songer_gt_vol — Federal Reporter volume (§A)
songer_gt_weightev — Weight of evidence (§A.2.1)
songer_gt_whlaws — Which-laws control (§A.2.1)
songer_gt_year — Year of decision (§A)
songer_present — Songer ground truth available (§1.5)
songer_second_appellant_subcategory — Appellant detail code (§4.6)
songer_second_direction_flag — Direction edge-case flag (§4.9)
songer_second_direction_mech — Mechanical direction (§4.9)
songer_second_direction_supermech — Super-mechanical direction (§4.9)
songer_second_focal — Focal party (§4.9)
songer_second_focal_alt — Other focal parties considered (§4.9)
songer_second_focal_appellant — Focal party is appellant (§4.9)
songer_second_focal_won — Focal party prevailed (§4.9)
songer_second_justification — Second-pass justification (§4.10)
songer_second_respondent_subcategory — Respondent detail code (§4.6)
songer_second_sub_area — Specific issue area (§4.8)
songer_second_typeiss_abusedis — Typeiss sub-issue: abuse of discretion (§4.8.2)
songer_second_typeiss_agen_acq — Typeiss sub-issue: agency acquiescence (§4.8.2)
songer_second_typeiss_alj — Typeiss sub-issue: ALJ

findings (§4.8.2)

songer_second.typeiss_altdisp — Typeiss sub-issue: alternative dispute resolution (§4.8.2)

songer_second.typeiss_attyfee — Typeiss sub-issue: attorney’s fees (§4.8.2)

songer_second.typeiss_capric — Typeiss sub-issue: arbitrary/capricious (§4.8.2)

songer_second.typeiss_comment — Typeiss sub-issue: rulemaking comment (§4.8.2)

songer_second.typeiss_confess — Typeiss sub-issue: confession admissibility (§4.8.2)

songer_second.typeiss_counsel — Typeiss sub-issue: counsel issues (§4.8.2)

songer_second.typeiss_deathpen — Typeiss sub-issue: death penalty (§4.8.2)

songer_second.typeiss_denovo — Typeiss sub-issue: de novo review (§4.8.2)

songer_second.typeiss_discover — Typeiss sub-issue: discovery (§4.8.2)

songer_second.typeiss_diverse — Typeiss sub-issue: diversity jurisdiction (§4.8.2)

songer_second.typeiss_dueproc — Typeiss sub-issue: due process (§4.8.2)

songer_second.typeiss_entrap — Typeiss sub-issue: entrapment (§4.8.2)

songer_second.typeiss_erron — Typeiss sub-issue: erroneous finding (§4.8.2)

songer_second.typeiss_execord — Typeiss sub-issue: executive order (§4.8.2)

songer_second.typeiss_fedvst — Typeiss sub-issue: federal vs state law (§4.8.2)

songer_second.typeiss_foreign — Typeiss sub-issue: foreign law (§4.8.2)

songer_second.typeiss_freeinfo — Typeiss sub-issue: FOIA issues (§4.8.2)

songer_second.typeiss_genstand — Typeiss sub-issue: general agency standard (§4.8.2)

songer_second.typeiss_improper — Typeiss sub-issue: improper trial conduct (§4.8.2)

songer_second.typeiss_indict — Typeiss sub-issue: indictment (§4.8.2)

songer_second.typeiss_indigent — Typeiss sub-issue: indigency (§4.8.2)

songer_second.typeiss_injunct — Typeiss sub-issue: injunctive relief (§4.8.2)

songer_second.typeiss_insane — Typeiss sub-issue: insanity defense (§4.8.2)

songer_second.typeiss_int_law — Typeiss sub-issue: international law (§4.8.2)

songer_second.typeiss_judgdisc — Typeiss sub-issue: judicial discretion (§4.8.2)

songer_second.typeiss_judrev — Typeiss sub-issue: judicial review (§4.8.2)

songer_second.typeiss_juryinst — Typeiss sub-issue: jury instructions (§4.8.2)

songer_second.typeiss_notice — Typeiss sub-issue: notice & comment (§4.8.2)

songer_second.typeiss_othadmis — Typeiss sub-issue: other admissibility (§4.8.2)

songer_second.typeiss_othcivil — Typeiss sub-issue: other civil issues (§4.8.2)

songer_second.typeiss_othcrim — Typeiss sub-issue: other criminal issues (§4.8.2)

songer_second.typeiss_othjury — Typeiss sub-issue: other jury issues (§4.8.2)

songer_second.typeiss_plea — Typeiss sub-issue: guilty plea (§4.8.2)

songer_second.typeiss_post_trl — Typeiss sub-issue: post-trial motions (§4.8.2)

songer_second.typeiss_prejud — Typeiss sub-issue: prejudicial conduct (§4.8.2)

songer_second.typeiss_pretrial — Typeiss sub-issue: pretrial motions (§4.8.2)

songer_second.typeiss_procdis — Typeiss sub-issue: procedural disposition (§4.8.2)

songer_second.typeiss_record — Typeiss sub-issue: administrative record (§4.8.2)

songer_second.typeiss_rtcouns — Typeiss sub-issue: right to counsel (§4.8.2)

songer_second.typeiss_search — Typeiss sub-issue: search & seizure (§4.8.2)

songer_second.typeiss_sentence — Typeiss sub-issue: sentencing (§4.8.2)

songer_second.typeiss_st_v_st — Typeiss sub-issue: state vs state (§4.8.2)

songer_second.typeiss_stpolicy — Typeiss sub-issue: state policy (§4.8.2)

songer_second.typeiss_subeid — Typeiss sub-issue: substantial evidence (§4.8.2)

songer_second.typeiss_suffic — Typeiss sub-issue: sufficiency of evidence (§4.8.2)

songer_second.typeiss_summary — Typeiss sub-issue: summary judgment (§4.8.2)

songer_second.typeiss_trialpro — Typeiss sub-issue: trial procedure (§4.8.2)

songer_second.typeiss_weightev — Typeiss sub-issue: weight of evidence (§4.8.2)

songer_second.typeiss_whlaws — Typeiss sub-issue: which-law (choice of law) (§4.8.2)

source — Data source (§1)

tenure_max — Panel tenure (max) (§3.4)

tenure_max_slot — Panel tenure (max slot) (§3.4)

tenure_mean — Panel tenure (mean) (§3.4)

tenure_med — Panel tenure (median) (§3.4)

tenure_med_slot — Panel tenure (median slot) (§3.4)

tenure_min — Panel tenure (min) (§3.4)

tenure_min_slot — Panel tenure (min slot) (§3.4)

year — Year of decision (§1)