

Federal Judicial Center
Integrated Data Base
Civil Codebook

August 26, 1988

Introduction

In 1982 the Research Division of the Federal Judicial Center initiated the Integrated Data Base (IDB) project. The goals of the project were (1) to document the information recorded on data tapes compiled by the Administrative Office of the United States Courts (AO) on cases filed in the federal court system, (2) to gather several years of data together into a single data base to facilitate multiyear analyses, (3) to provide a common format for the data so that those analyses could be done without having to contend with record formats that changed from year to year, and (4) to provide a vehicle for linking cases from the district to the appellate level. The time period covered by the initial project was statistical year 1970 through statistical year 1983 (July 1, 1969 through June 30, 1983). These data have since been updated to include cases through statistical year 1987 (June 30, 1987). The data base includes all federal appellate and district court cases active during this time period. The data base currently is comprised of 56 data sets. There is one data set of terminated cases for each statistical year (SY) for each of the three data categories: district civil (18 data sets, SY70 - SY87), district criminal (18 data sets, SY70 - SY87), and appellate (17 data sets, SY71 - SY87). In addition, in each category there is a data set of cases still pending at the end of the time period (3 data sets, SY87). This codebook describes the format and fields of the district civil data sets. Separate codebooks are available for the district criminal and appellate data sets.

Creating the Data Base

We obtained all the raw data on computer tapes directly from the Administrative Office. The problem of the varying physical structure of the data records was relatively easy to solve. A core set of data fields was identified from among all the fields occurring on the records during any time period. Although this core set is basically a union of all possible fields, some fields that were determined to be of minor importance to our research goals were excluded.

A standard format accommodating all the fields was then designed. Whenever possible, mixed alphanumeric fields were recoded as numeric-only fields and field widths were enlarged to allow for greater value ranges in the future. Each raw data file was first passed through a reformatting program to convert the data to the standard structure; fields that existed in the standard record but not in the original record were left blank.

Next, the restructured data were processed by a conversion program that checked for errors and performed all value recoding. Because the new standard record included some data items that were captured only during a limited time span¹, many fields were recoded to a "not collected" missing value. In addition, some recoding was done in an attempt to standardize values for a particular field across all years². More often, however, the evolution of values did not follow a clear pattern and longitudinal consistency recoding was not possible. This means that for some fields the same coded value may have different interpretations depending on when the information was recorded³.

The lack of a consistent correspondence between a coded value and a single interpretation is the most serious challenge to multiyear analysis presented by the data and the researcher must be extremely careful to account for value inconsistency when doing analyses of this type. The field-by-field descriptions that are presented below are designed to help the researcher identify value changes. Unless limitations are noted, the listed correspondence between a coded value and its interpretation is consistent throughout the data base time period.⁴

¹ For example, information regarding the level of magistrate involvement in the case (field 28) was collected only during statistical years 1981 through 1985.

² For example, the Middle District of Florida was represented by the district code 30 for SY70 and SY71, but the code changed to 3A for subsequent years. In the IDB, the Middle District of Florida is always represented by the code 3A; the older code was detected and recoded.

³ For example, for the field Disposition (#21) the coded value 1 denotes "Default judgment" terminations for SY71 data but "Remanded" terminations for SY81.

⁴ Value consistency should not be confused with value validity. Circuit values of 11 always refer to the Eleventh Circuit but are not valid prior to SY82. Procedural Progress values of 2, on the other hand, are

Original Documents

The records of the district civil data base are derived from information routinely collected by the Administrative Office on standard case status forms filed monthly by the Clerk of Court in each of the federal district courts. These forms, the JS-5 Filing Report and the JS-6 Termination Report, are part of the standard Civil Docket Set. The JS-44 Civil Cover Sheet is also an important raw data form from which JS-5 information is drawn.

valid throughout but assume three different interpretations.

Reference Documents

The primary reference for information collected on the JS-5, JS-6, and JS-44 is Title II "Instructions for Completing District Court Report Forms DC-111 (including JS-5 and JS-6), JS-5A, JS-6M, JS-9, JS-9C, JS-44, and JS-44B" of the Statistical Analysis Manual, which is volume XI of the Guide to Judiciary Policies and Procedures published by the Administrative Office of the United States Courts. This manual and the periodic transmittals used to update it are the official instructions to the Clerks of Court for completing the original case forms. Information regarding the values coded on the computer record from the original documents is contained in a booklet issued periodically by the Administrative Office entitled "Statistical Codes for Civil Reports Submitted by Clerks of Court, United States District Courts." The several versions of these documents issued over the course of our data period as well as computer tape layouts and occasional special memoranda were the sources of the field and value documentation presented below.

Documentation materials available to the project for the period prior to 1979 are very sparse; therefore it is frequently impossible to accurately identify when a change in values or interpretation occurred. The practice used in this document is to list the statistical year in which the value is known to have been valid (primarily because documentation from the same statistical year lists the value). A value may have been valid prior to the listed date but we cannot confirm it. Conversely, the assumption is made that a value continues to have the same interpretation until a change is documented. Values that are supposedly no longer valid are sometimes found in the data as a result of coding habit or the use of an outdated form.

Published Tables

Tables based on the original data compiled by the Administrative Office are published in the Annual Report of the Director of the Administrative Office of the United States Courts. The "C" tables (C1 - C6) in the appendix of "Detailed Statistical Tables" are based on the civil data. Case counts achieved with the IDB data should match the published tables. If there are discrepancies, contributing factors may be: (1) population restrictions (for example, excluding land condemnation cases); (2) case groupings employed to simplify presentation (for example, "All Other" categories); (3) statistical year skews effecting inclusion in a cohort based on calendar dates (see "Statistical Year Designation" below); and

(4) any IDB processing that may have altered values (for example, out-of-range values falling into a separate "missing" category rather than part of "other").

Statistical Year Designation

The statistical year used in the federal courts runs from July through June (for example, SY84 covers the period July 1, 1983 through June 30, 1984). Due to reporting time lags, there is often a slight skew to the correspondence between calendar date and statistical year at the borders of the statistical year. Transactions which by date would logically fall into one statistical year may actually be counted in a different statistical year. For example, June 1985 filings and terminations should be counted in SY85. However, if the reports are received by the AO after the statistical year data have been processed, the late-arriving transactions are counted in SY86 instead. The tables published by the AO for a particular statistical year are based on all the cases designated as belonging to that statistical year regardless of the actual filing or termination dates.

The IDB files are organized according to statistical year of termination based on the termination date used by the AO to identify the statistical year in which the case was counted⁵. This organization was chosen to facilitate matching published terminations data. To match published data based on the statistical year of filing, the "Filing Date Used by AO" field (#20), can be used to identify cases⁶. Cohorts based on actual filing and termination dates, rather than on these Used-by-AO dates, are unlikely to provide counts that can be matched with published tables.

Special Notes

During SY86 the Administrative Office conducted a pilot program in four district courts: Eastern District of Pennsylvania, Northern District of Texas, Middle District of Florida, and Western District of Missouri. The program tested the online data entry of the JS-5 and JS-6 information and used a slightly more detailed version of the JS-6 form. This new form contained different interpretations for some of the values for the Procedural Progress, Disposition, Nature of Judgment, and

⁵ Starting in SY81, this date is recorded in field 30, "Termination Date Used by AO".

⁶ Continuing the example given above, a case that was docketed in June 1985 but was processed late and thus was counted among SY86 cases instead of SY85 cases will have a value of 8507 for Filing Date Used by AO which properly places it in the statistical year in which it was counted.

Judgment For fields than are listed on the standard JS-6 which was used by all the other courts in SY86. This more detailed coding was adopted by all courts in SY87 for cases filed on or after July 1, 1986. Termination information for cases filed prior to July 1, 1986 will be coded according to the older (standard SY86) scheme.

Some civil data files include a final blank record. The researcher may want to detect the existence of such records and exclude them from any analysis.

Purpose and Limitations of the Documentation

The purpose of this document is to briefly describe the data contained in the IDB files: the physical structure of the records, the fields that are included, and the range of valid values that can be expected for each field (particularly in those instances where IDB processing programs have altered the value recorded in the original data). The document also points out known problems associated with using the IDB data, especially with respect to multiyear analyses. This document does not provide detailed explanations of fields or values, coding instructions for special cases, definitions of legal terms, information regarding court management or procedures, or changes in the federal rules or the law. Although information on these latter topics may be important when conducting analyses using these data, it is impossible to present that information here. The researcher may want to consult the reference materials listed above ("Reference Documents"), the federal rules of procedure, and the U.S. code for more detailed information.

Record Format

<u>Number</u>	<u>Field Description</u>	<u>Format</u>	<u>Length</u>
<u>Position</u>			
1	Record Quality Indicator	Integer	2 1-2
2	Circuit	Integer	2 3-4
3	District	Alpha	2 5-6
4	Office	Alpha	1 7
5	Docket Number	Alpha	7 8-14
6	Filing Date	YYMMDD	6 15-20
7	Jurisdiction	Integer	2 21-22
8	Nature of Suit	Integer	3 23-25
9	Origin	Integer	2 26-27
10	Residence	Integer	2 28-29
11	Class Action	Integer	2 30-31
12	Termination Judge	Alpha	4 32-35
13	Filing Judge	Alpha	4 36-39
14	Trial Date	YYMM	4 40-43
15	Demand	Integer	4 44-47
16	Filing Magistrate	Alpha	4 48-51
17	County	Integer	5 52-56
18	Style (Parties Involved)	Alpha	40 57-96
19	Termination Date	YYMMDD	6 97-102
20	Filing Date Used by AO	YYMM	4 03-106
21	Disposition	Integer	2 07-108
22	Termination Magistrate	Alpha	4 109-112
23	Procedural Progress	Integer	2 113-114
24	Nature of Judgment	Integer	2 15-116
25	Amount Received	Integer	4 17-120
26	Unknown Date	YYMM	4 121-124
27	Judgment For	Integer	2 125-126
28	Magistrate Involvement	Integer	2 127-128
29	Other Involvement	Integer	2 129-130
30	Termination Date Used by AO	YYMM	4 131-134

Field Descriptions

1. RECORD QUALITY INDICATOR i2 1-2
- This field is used to quantify the accuracy of the record. The evaluation programming has not yet been implemented; therefore this field is always blank.
2. CIRCUIT i2 3-4
- The federal circuit where the district court is located.
- | | |
|------------------------------|---------------------------------------|
| 00 - District of Columbia | 06 - Sixth Circuit |
| 01 - First Circuit | 07 - Seventh Circuit |
| 02 - Second Circuit | 08 - Eighth Circuit |
| 03 - Third Circuit | 09 - Ninth Circuit |
| 04 - Fourth Circuit | 10 - Tenth Circuit |
| 05 - Fifth Circuit | 11 - Eleventh Circuit
(since SY82) |
| -8 - Missing or out-of-range | |
3. DISTRICT a2 5-6
- The U.S. District Court where the case was filed.
- See District/Office Code Appendix.
- 8 - Missing or out-of-range.
- Note: Prior to SY74, the district code 70 referred to a district that included both Arizona and Alaska, since SY74 the district code for Arizona has been 70, while the code for Alaska is 7- (listed as 7X in the appendix).
4. OFFICE a1 7
- The office of the district court where the case was filed. See District/Office Code Appendix.
- 0 substituted for missing data.
5. DOCKET NUMBER a7 8-14
- The case number as assigned by the Clerk of Court at the time of the original filing of the case.

The format is YYNNNNN, where YY is the last two digits of the calendar year in which the case was filed, and NNNNN is the five digit sequence number from 00001 to 99999. Docket numbers run on a calendar year basis; the serially assigned numbers are restarted on January 1.

6. FILING DATE (YYMMDD) i6 15-20

The date on which the case was filed in district court.

00 substituted for missing or out-of-range segments.

7. JURISDICTION i2 21-22

The basis of jurisdiction for filing the case in federal court.

- 1 - U.S. Plaintiff
- 2 - U.S. Defendant
- 3 - Federal Question
- 4 - Diversity of Citizenship
- 5 - Local Question (territorial districts only)
- 8 - Missing or out-of-range

8. NATURE OF SUIT i3 23-25

The nature of the action filed.

See Nature of Suit Appendix. -8 Missing or out-of-range.

9. ORIGIN i2 26-27

The source of the case filing.

- 1 - Original proceeding
- 2 - Removed from state court
- 3 - Remanded from appellate court
- 4 - Reinstated or reopened
- 5 - Transferred from another district
- 6 - Multi-district litigation
(Transfer by order of Panel for
Multi-district Litigation)

- 7 - Appeal of magistrate judgment
to a district judge (since SY81)
- 8 - 2nd reopening (since SY83)
- 9 - 3rd reopening (since SY83)
- 10 - 4th reopening (since SY83)
- 11 - 5th reopening (since SY83)
- 12 - 6th reopening (since SY83)

- 0 - Missing
- 8 - Out-of-range

10. RESIDENCE

i2 28-29

Describes the citizenship of the principal parties in cases in which diversity of citizenship is the basis of federal jurisdiction (JURISDICTION value of 4). The two-digit code is a composite of a single-digit plaintiff's code and a single-digit defendant's code (that is, a 1 for plaintiff and a 3 for defendant yields a residence code of 13).

SY70 - SY76

- 1 - Resident of your state
- 2 - Non-resident corporation doing business in state
- 3 - Non-resident corporation not doing business in state
- 4 - Other non-resident of your state

SY77 - SY85 (two variations active)

- 1 - Citizen of this state
- 2 - Incorporated this state
- 3 - Foreign corporation--principal place of business in _____
- 4 - Other non-citizen of this state

- 1 - Citizen or resident business of this state
- 2 - Non-resident business incorporated in this state
- 3 - Foreign business corporation, incorporated in another state
- 4 - Other non-citizen of this state

SY85 - SY87

- 1 - Citizen of this state
- 2 - Citizen of another state
- 3 - Citizen or subject of a foreign country
- 4 - Incorporated or principal place of business in this state
- 5 - Incorporated and principal place of business in another state
- 6 - Foreign nation

SY70 - SY87

- 8 - Missing, out-of-range, or not applicable (that is, non-diversity cases).

11. CLASS ACTION i2 30-31
 Class Action allegation under Rule 23, Federal Rules of Civil Procedure.
- 0 - No allegation of class action made
 - 1 - Allegation of class action
 - 2 - Allegation dismissed
 - 3 - Allegation affirmed
- 8 - Out-of-range
 - 9 - Data not collected (SY70 - SY72)
12. TERMINATION JUDGE a4 32-35
 Blank on public use files.
13. FILING JUDGE a4 36-39
 Blank on public use files.
14. TRIAL DATE (YYMM) i4 40-43
 Year and month that trial commenced.
- 00 substituted for missing or out-of-range segments.
 9912 inserted for SY74 - SY87 because the data were no longer being collected.
- Note: This should be regarded as a highly suspicious variable because there is limited documentation that describes its use and the observed values do not consistently reflect the expected patterns for date information.
15. DEMAND i4 44-47
 Amount demanded in thousands of dollars.
- 9999 - coded for amounts greater than \$9,999,000
 - 0 - missing (blank)

SY70 - SY76 values less than \$1000 were left blank.
SY77 - SY81 coding of values less than \$1000 is unknown.
SY82 - SY87 values less than \$1000 were coded as 1.

16. FILING MAGISTRATE a4 48-51

Blank on public use files.

17. COUNTY i5 52-56

Code for the county of residence of the first-listed plaintiff, except where the United States is the plaintiff; in that case it is the code of the county of residence for the first-listed defendant. In land condemnation cases, the county code of the land in question is used.

See County Codes Appendix. -9 Not collected (SY70 - SY80)

18. STYLE (NAMES OF PARTIES INVOLVED) a40 57-96

Format: Plaintiff (et.al.) v. Defendant (et.al.)

19. TERMINATION DATE (YYMMDD) i6 97-102

The date the case was closed in district court. This is the date of the final judgment or of the order disposing of the action. If the case was transferred, the date of transfer is the termination date.

00 substituted for missing or out-of-range segments.

20. FILING DATE USED BY AO (YMMM) i4 103-106

The statistical date when the filing entered the master AO record. The date used by the AO to assign the case to a "statistical year filed" cohort.

00 substituted for missing or out-of-range segments.
9912 - Not collected (SY70 - SY80)

21. DISPOSITION i2 107-108

The method of disposition.

SY70 - SY71

- 0 - Transferred to another district
- 1 - Default judgment
- 2 - Consent judgment
- 3 - Dismissed for want of prosecution

21. DISPOSITION (continued)

- 4 - Dismissed by action of parties
- 5 - Remanded to state court
- 6 - Judgment for plaintiff
- 7 - Judgment for defendant
- 8 - Judgment for both or other party
- 9 - Outcome unknown

SY79 - SY86 (excluding 4 pilot districts in SY86)

- 0 - Transferred to another district
- 1 - Remanded
- 2 - Dismissed for want of prosecution
- 3 - Dismissed, discontinued, settled, withdrawn, etc.
- 4 - Judgment on default
- 5 - Judgment on consent
- 6 - Judgment on motion before trial
- 7 - Judgment on jury verdict
- 8 - Judgment on directed verdict
- 9 - Judgment on court trial
- 10 - Judgment on other
- 11 - Statistical closing (since SY84)

SY86 - SY87 (4 districts in SY86: EDPA, NDTX, MDL, WDMI)

- 0 - Transferred to another district
- 1 - Remanded to state court
- 2 - Dismissed for want of prosecution
- 3 - Dismissed for lack of jurisdiction
- 4 - Judgment on default
- 5 - Judgment on consent
- 6 - Judgment on motion before trial
- 7 - Judgment on jury verdict
- 8 - Judgment on directed verdict
- 9 - Judgment on court trial
- 10 - Multi-district litigation transfer
- 11 - Remanded to U.S. agency
- 12 - Dismissed: voluntarily
- 13 - Dismissed: settled
- 14 - Dismissed: other
- 15 - Judgment on award of arbitrator
- 16 - Judgment on trial de novo after arbitration
- 17 - Judgment on other
- 18 - Statistical closing

SY70 - SY87

- 8 - Missing or out-of-range
- 9 - Data not collected (SY72 - SY78)

22. TERMINATION MAGISTRATE a4 109-112

Blank on public use files.

23. PROCEDURAL PROGRESS i2 113-114

Procedural progress at termination.

SY70 - SY75

- 1 - No issue - no action
- 2 - No issue - action
- 3 - Issue, no pretrial - no action
- 4 - Issue, no pretrial - action
- 5 - Pretrial but no trial
- 6 - Terminated during court trial
- 7 - Terminated during jury trial
- 8 - Judgment after court trial
- 9 - Judgment after jury trial
- 0 - Land condemnation cases

SY76 - SY86 (excluding 4 pilot districts in SY86)

- 1 - Before issue joined
- 2 - After motion decided but before issue joined
- 3 - Issue joined, no other court action
- 4 - Issue joined, and after judgment of court on motion
- 5 - Issue joined, and after pretrial conference but before trial
- 6 - During court trial
- 7 - During jury trial
- 8 - After court trial
- 9 - After jury trial
- 0 - Other

SY86 - SY87 (4 districts in SY86: EDPA, NDTX, MDFL, WDMI)

- 1 - Before issue joined, no court action
- 2 - Before issue joined, order entered
- 3 - After issue joined, no court action

- 4 - After issue joined, judgment on motion
- 5 - After issue joined, pretrial conference held
- 6 - After issue joined, during court trial
- 7 - After issue joined, during jury trial
- 8 - After issue joined, after court trial
- 9 - After issue joined, after jury trial
- 10 - After issue joined, other
- 11 - Before issue joined, hearing held
- 12 - Before issue joined, motion decided

SY70 - SY87

- 8 - Out-of-range (also missing in SY87)

24. NATURE OF JUDGMENT

i2 115-116

The nature of the judgment for those actions disposed of by the entry of a final judgment.

SY70 - SY71

- 0 - No judgment entered
- 1 - Money
- 2 - Money and other
- 3 - Injunction
- 4 - Other, foreclosure, condemnation, HEW remand, divorce

SY79 - SY86 (excluding 4 pilot districts in SY86)

- 0 - No monetary award
- 1 - Monetary award only
- 2 - Monetary award and other
- 3 - Injunction
- 4 - Other, foreclosure, condemnation, remand, etc.
- 5 - Costs only

SY86 - SY87 (4 districts in SY86: EDPA, NDTX, MDFL, WDMI)

- 0 - No monetary award
- 1 - Monetary award only
- 2 - Monetary award and other
- 3 - Injunction
- 4 - Forfeiture, foreclosure, condemnation, etc.
- 5 - Costs only
- 6 - Costs and attorney fees

SY70 - SY87

- 8 - Missing or out-of-range
- 9 - Data not collected (SY72 - SY78)

25. AMOUNT RECEIVED i4 117-120

The monetary judgment amount awarded (excluding costs) in thousands of dollars. (Value associated with nature of judgment codes 1 and 2.)

- 9999 - Coded for amounts greater than \$9,999,000
- 0 - Missing (blank)
- 8 - Out-of-range (contains alphas)

SY70 - SY75 values less than \$1000 were left blank.
SY76 - SY87 coding of values less than \$1000 is unknown.

26. UNKNOWN DATE (Y Y M M) i4 121-124

Although the observed values indicate that this field is a date in the format Y Y M M, it is not clear what date is being represented. It is most likely that it is a statistical year representation of the termination date (therefore very similar, if not equivalent to, the "termination date used by AO", field #30, which started being collected in SY81 when this field stopped being coded). It is unlikely that this field represented the date the judgment award was received which previous codebooks suggested.

00 was substituted for missing or out-of-range segments.
9912 - Data not collected (SY81 - SY87).

27. JUDGMENT FOR i2 125-126

Identifies the party favored by the judgment of the court for actions disposed of by the entry of a final judgment.

- 1 - Plaintiff
- 2 - Defendant
- 3 - Both
- 4 - Unknown (or not applicable)

- 8 - Out-of-range

-9 - Data not collected (SY70 - SY78)

28. MAGISTRATE INVOLVEMENT

i2 127-128

Level of involvement of the U.S. magistrate.

- 1 - Consent of parties to magistrate judgment with appeal directly to U.S. court of appeals
 - 2 - Consent of parties to magistrate judgment with appeal to U.S. district court
 - 3 - Special master reference
 - 4 - All pretrial matters, including dispositive motions
 - 5 - Report and recommendation on dispositive motions
 - 6 - All pretrial matters, except dispositive motions
 - 7 - Full pretrial/settlement conferences
 - 8 - Non-dispositive motions
 - 9 - No involvement of magistrate
- 8 - Out-of-range
- 9 - Data not collected (SY70 - SY80, SY86 - SY87)

29. OTHER INVOLVEMENT i2 129-130

Involvement in the case of persons other than judge or magistrate.

- 1 - Arbitrator
- 2 - Land commissioner
- 3 - Special master (other than magistrate)
- 4 - Other
- 5 - No involvement other than judge or magistrate

- 8 - Out-of-range
- 9 - Data not collected (SY70 - SY80, SY86 - SY87)

30. TERMINATION DATE USED BY AO (YYMM) i4 131-134

The statistical date when the termination entered the AO master record. The date used by the AO to assign the case to a "statistical year terminated" cohort.

00 was substituted for missing or out-of-range segments.
9912 - Data not collected (SY70 - SY80).

