

LexisNexis Appeals Data Review (Data Received March 13, 2019)

Response to Stuart's Questions of June 9, 2019

Duke University, University of Michigan Appeals Project

1 Responses to Points in Stuart's 6-9-2019 E-mail

Points are in black, responses are in red

- I I want to put two more things on the table items 1 and 3 from my email of 5-1-2019 5:49pm to LN. In brief, they involve (item 1) missing cases and (item 3) missing Shepards treatments. I will forward you an email showing those items and a bit of followup from me on them.

- 1) ByungKoo and Tom, can you review item 1 from the 5-1-2019 email and see if you can add any detail that might be helpful to LN (and to us)?

Text from item 1 of 5-1-2019 e-mail:

The total number of cases in the data you sent for 1996 was 44,659. Between 1996 and 2004, the data you sent have a fairly steadily decreasing number of cases, to a low of 27,379 in 2004. The LN client on the Web has similar numbers. The decrease seems to be coming from the cases in the unidentified column (cases that dont have a case type, i.e., neither civil nor criminal). When we looked at the caseload data for those years that on the Administrative Office of the US Courts website, there was no decrease in case load at either the district court or the appellate court level during this time period. We also checked in the Westlaw database to see if their case numbers reflected a similar decrease in number over this time period, but they did not. Accordingly, it doesnt seem like this is something that happened on the part of the courts, but instead is something that changed on Lexis end. We are wondering if Lexis started separating out some of the specialty courts around this time into separate databases that were previously included in the overall number (i.e., bankruptcy appellate panels or veterans claims appeals?). So the questions are:

- a. what cases are included in the unidentified column;*
- b. was there a change in the type of cases included in the United States Courts of Appeals database that started to take effect around 1996?; and*
- c. are there separate databases (for example, Bankruptcy Courts or Court of Appeals for Veterans Claims) that were created around this time that include cases that were formerly included within the United States Courts of Appeals database?*

Figures 1 and 2 show the distribution of cases by case type, court, and year. These were previously distributed in the LexisNexisData-2019-03-13-Review document.

I do not have access to LN on-line, so that I cannot detect missing Shepard's treatments. Perhaps the cases from rows 158, 367, 433, and 1432 of the LNCASESAMPLETreatment-10-25-SMB-4-26-2019.xlsx sheet (5-1-2019 e-mail) can be sent again.

- II Tom and ByungKoo, can you put together the tables and graphs that illustrate the key things we are missing? I think there are four key things in addition to the two items noted above:

1) Missing/duplicate panels

Figures 3 and 4 show the distribution of cases by panel size in aggregate and by court. The issue of empty panel (0 judges) and two judge panels is clearly visible. These graphs were previously distributed in the LexisNexisData-2019-03-13-Review document.

Example cases with missing panel are:

3XNJ-6V30-0038-X1MH-00000-00 Wilder v. United States, 182 F.3d 934 1999-04-29
3S4X-D540-003B-5011-00000-00 Carter v. Lack, 869 F.2d 1489 1989-02-13
3S4X-3S60-003B-G10W-00000-00 Silverman v. Hannah Inland Waterways, 671 F.2d 1376 1982-03-04
3S4X-2NB0-001B-K4DV-00000-00 U.S. v. Bridger, 845 F.2d 1027 1988-02-22
3S4X-BT70-003B-P36H-00000-00 Saephanh v. Shalala, 9 F.3d 1553 1993-10-28
3S4X-38D0-003B-P171-00000-00 Cooley v. Whitley, 35 F.3d 560 1994-08-25
4G0T-GYC0-0038-X427-00000-00 Drain v. Rumsfeld, 126 Fed. Appx. 463 2004-12-14
3S4X-4C00-008H-V500-00000-00 UNITED STATES v. RHYMES, 960 F.2d 153 1992-04-16
3S4X-FY40-001T-D4P0-00000-00 Levy v. Lerner, 52 F.3d 312 1995-03-31
3S4X-5XK0-008H-V329-00000-00 United States v. Brown, 953 F.2d 640 1992-02-03
4J3X-M570-TVSH-535T-00000-00 Mogyorossy v. Dep't of the Air Force, 161 Fed. Appx. 956 2006-01-12

An attempt was made to identify, for each case record with missing panel, a separate case record for the same case, but one with a panel. To accomplish this, the reporter was removed from case titles, all title text was converted to lower case, and abbreviations were standardized (“u.s.” to “united states”, “co.” to “company”, etc.). 38% of cases with missing panel were found to be duplicated by a case record with panel.

Following are examples, from the 9th Circuit on 5-13-1988, of apparent duplicate cases, each pair having one record with and one without a panel:

Friedman v. United States, 1988 U.S. App. LEXIS 22018, panel
Friedman v. United States, 1988 U.S. App. LEXIS 6749, no panel

United States v. Crabtree, 1988 U.S. App. LEXIS 22076, panel
United States v. Crabtree, 846 F.2d 1383, no panel

Figure 5 plots the distribution (count) of cases with no panel in the LN data, by court and year. The solid line represents all cases that have no panel. The dashed line is limited to cases with no panel, but with a matching case (court, date, title) that does have a panel. In addition to the appearance of spikes by court and year, many non-panel cases for court IDs 1, 4, 8, and 10 appear to have matching cases with a panel. Court IDs and names appear in table 2.

Figure 6 plots cases with missing panel as a proportion of total cases by court and year. The similarity of total missing panel and duplicated with panel lines for court IDs 1, 4, 8, and 10 is seen here as in 5. For all courts, the proportion of cases with no panel, but duplicated with panel, is near zero prior to approximately 1990.

2) Cases with two judges listed, even though there were three

21,556 cases have two judge-panels. Byungkoo’s remarks appear in table 3.

3) Definitions of legal topics

LN is conducting their process of review and approval of this request.

4) Missing legal topics

Over 600,000 cases have no legal topics. Table 4 is a random sample of 100 cases from those that are missing legal topics.

III Perhaps should include, though less important

1) Obsolete treatment letters

Several treatment letters that appear on citations are labeled as “obsolete” in the treatment definition table (affected treatment codes are listed in the LexisNexisData-2019-03-13-Review document). We tested whether “qr” and “qv” had corresponding “q” treatments (a possible replacement or substitute), but that is not the case. LN has been asked whether or not citations with “obsolete” treatments should be ignored.

2) IDs are for authoring (opinion, concurring, and dissenting) judges

LN has supplied IDs for panels and should be able to provide them for authors also.

IV Things I’m not sure are worth pursuing right now

1) Cases with missing court identifier

Forty cases with missing court identifier (listed in the LexisNexisData-2019-03-13-Review document) had their court identified using LN on-line or Westlaw. If LN on-line was used then there is an error in LN’s extraction process and, ideally, this would be corrected, as opposed to our backfilling from data they possess. Often, missing data are merely a symptom of flawed query logic and correction of the logic may resolve not only the apparent error, but also others that remain, lurking, to be identified later.

2) Judges missing a name

Twenty judges are missing a name (listed in the LexisNexisData-2019-03-13-Review document). Although none appear in panels at present, they may appear once the missing panel issue is resolved. Also, as stated, missing data can be caused by flaws in query logic. Ideally, the root cause of apparent data errors should be corrected.

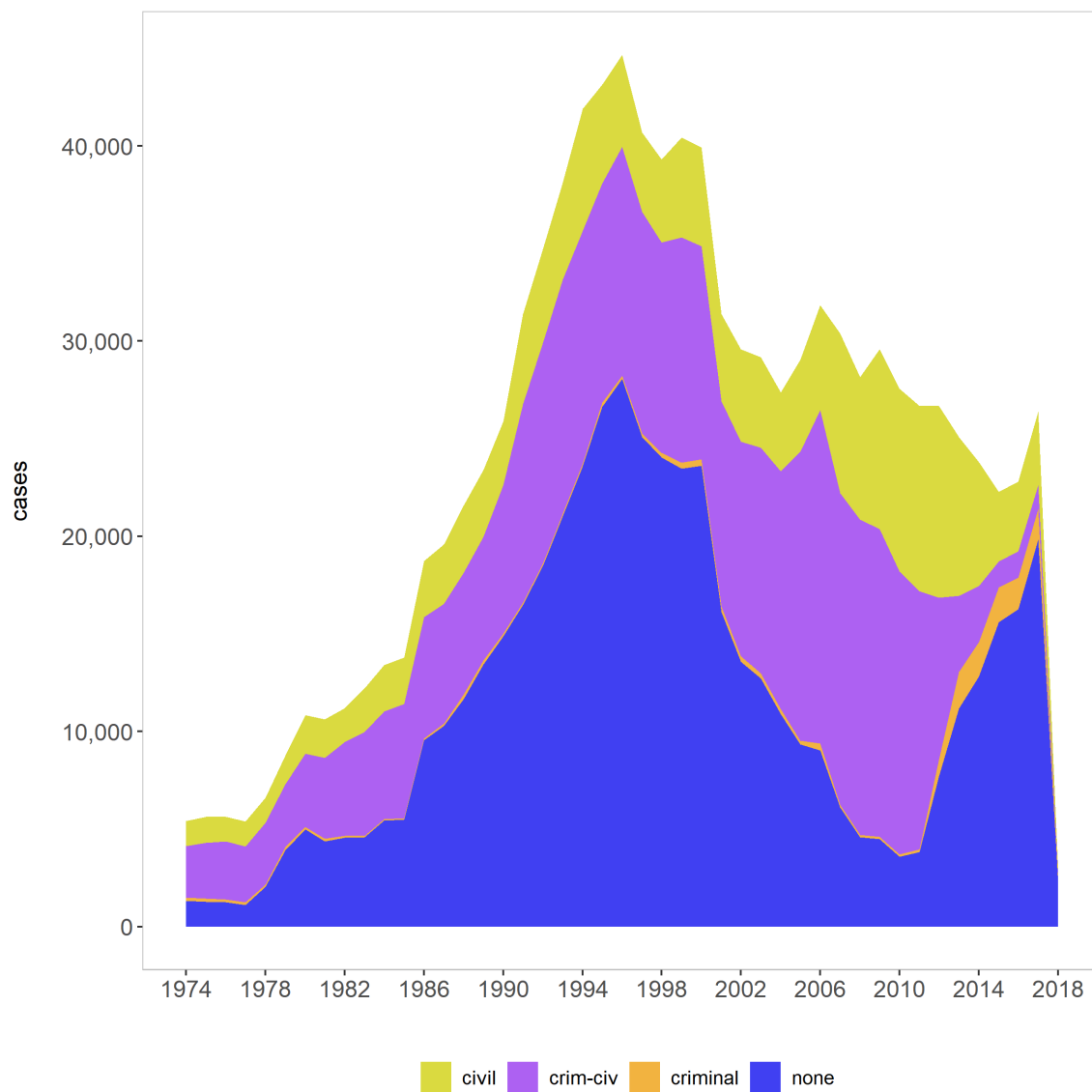


Figure 1: Stacked distribution of cases by year and type

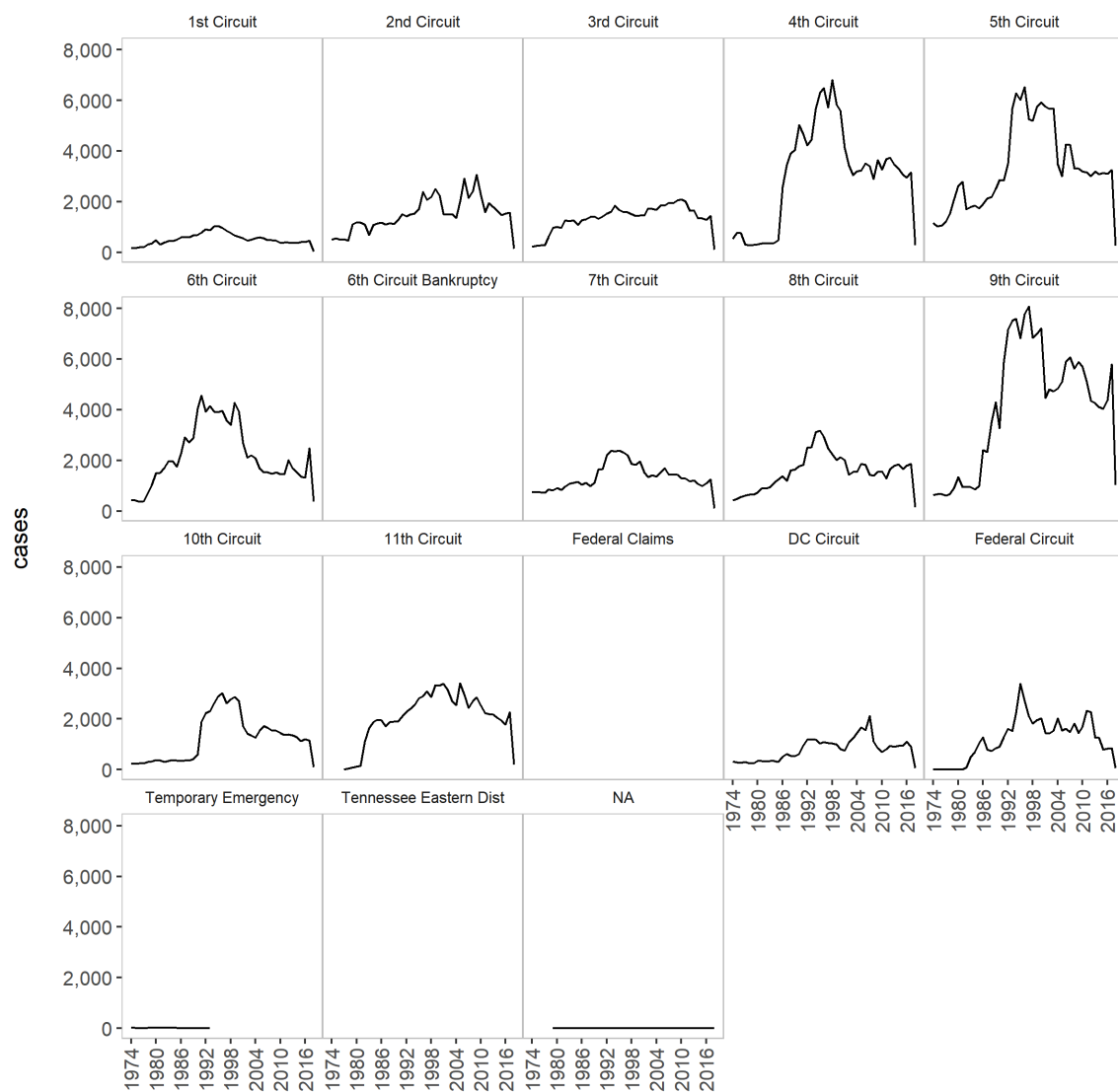


Figure 2: Cases by court and year

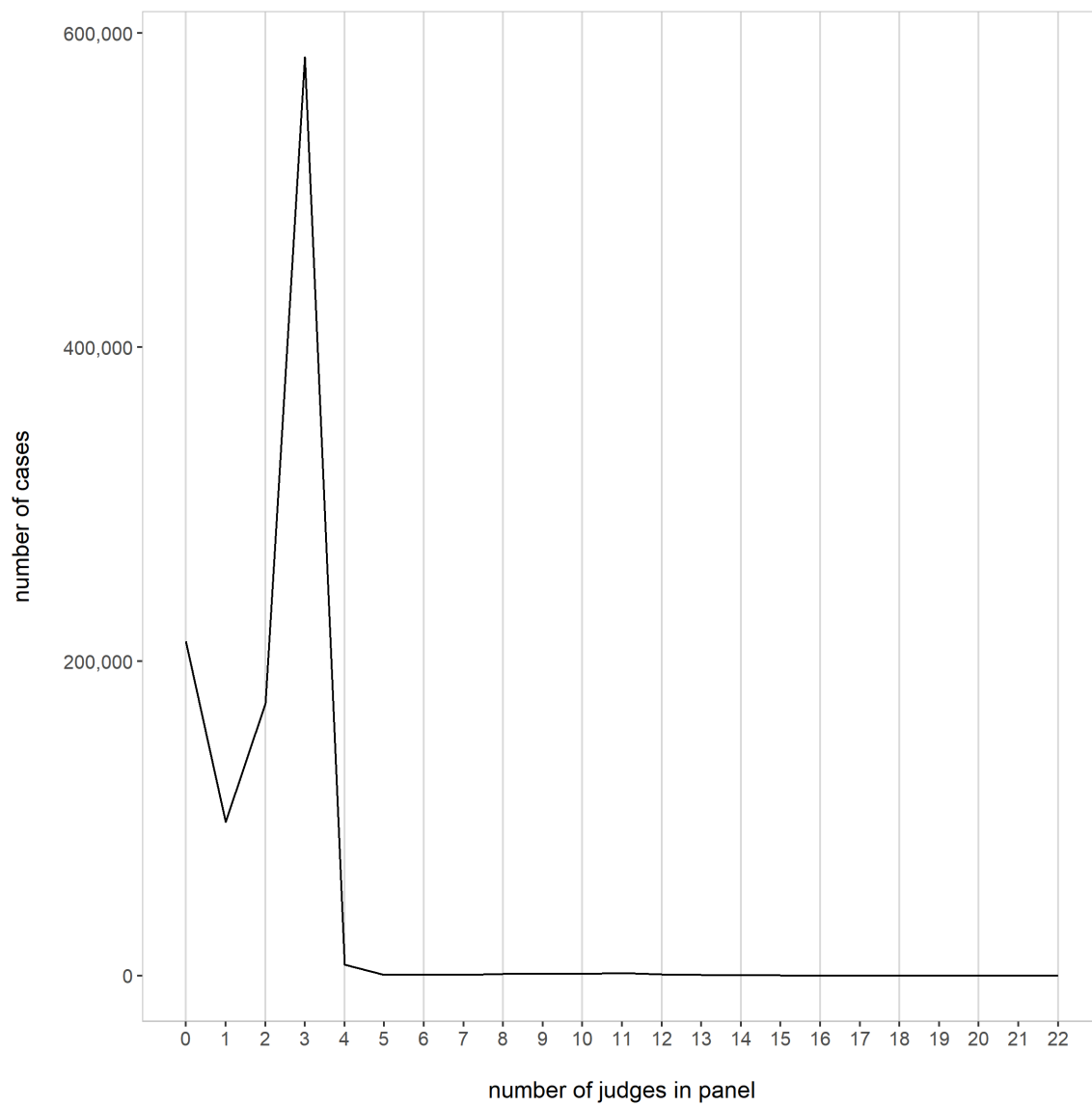


Figure 3: Distribution of cases by panel size

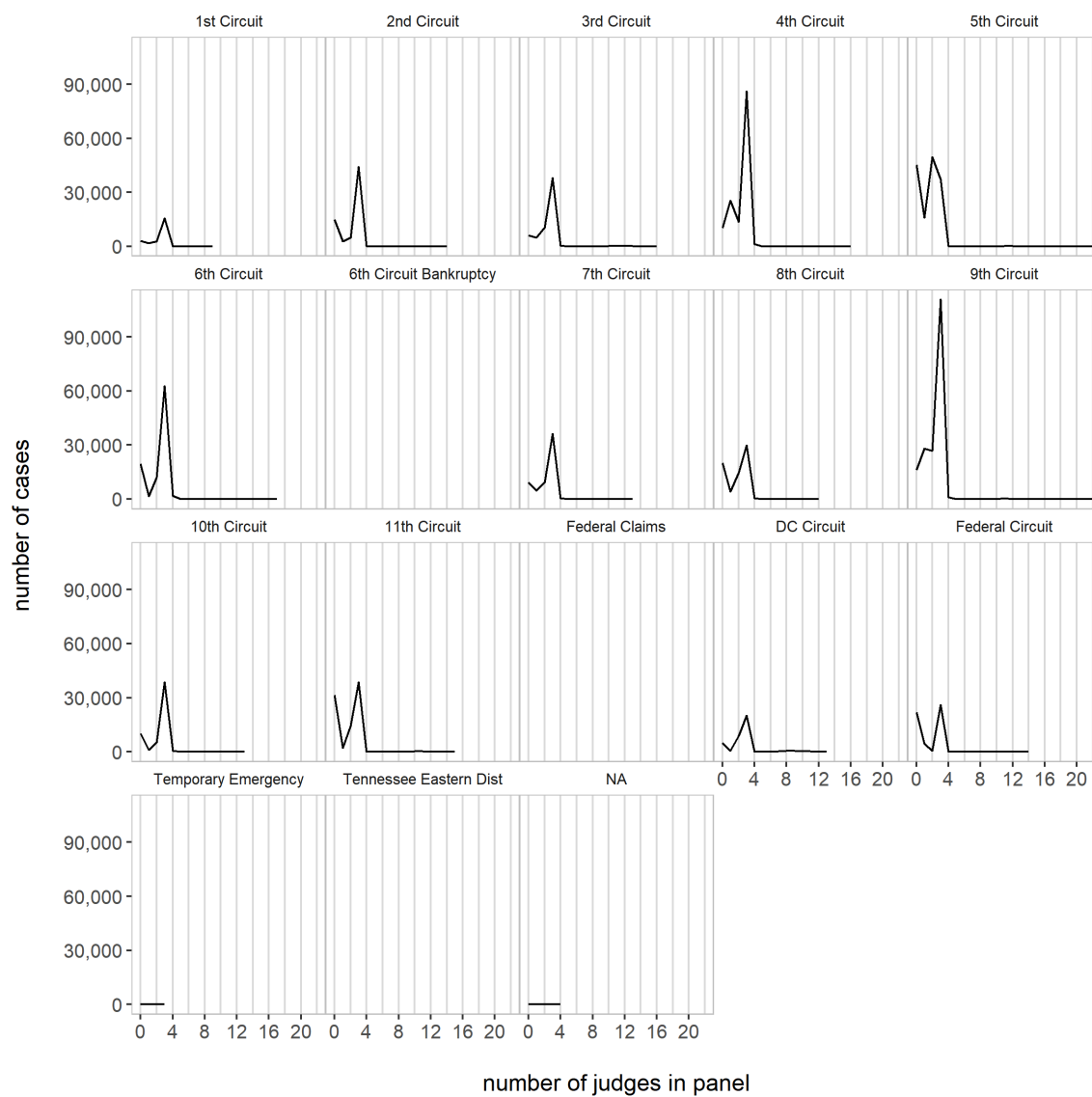


Figure 4: Distribution of panel size by court

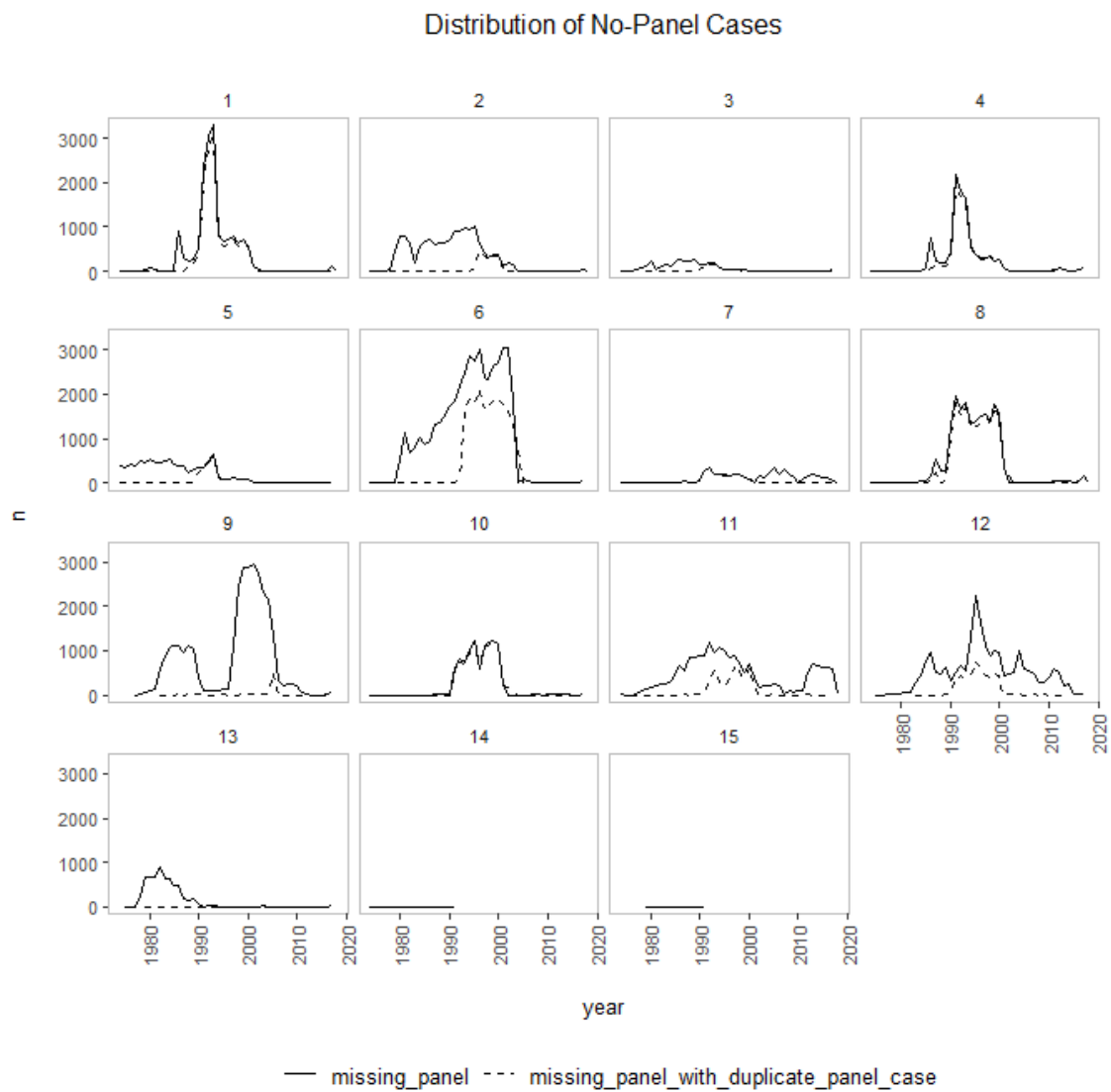


Figure 5: Frequency distribution of cases with missing panel, by court and year

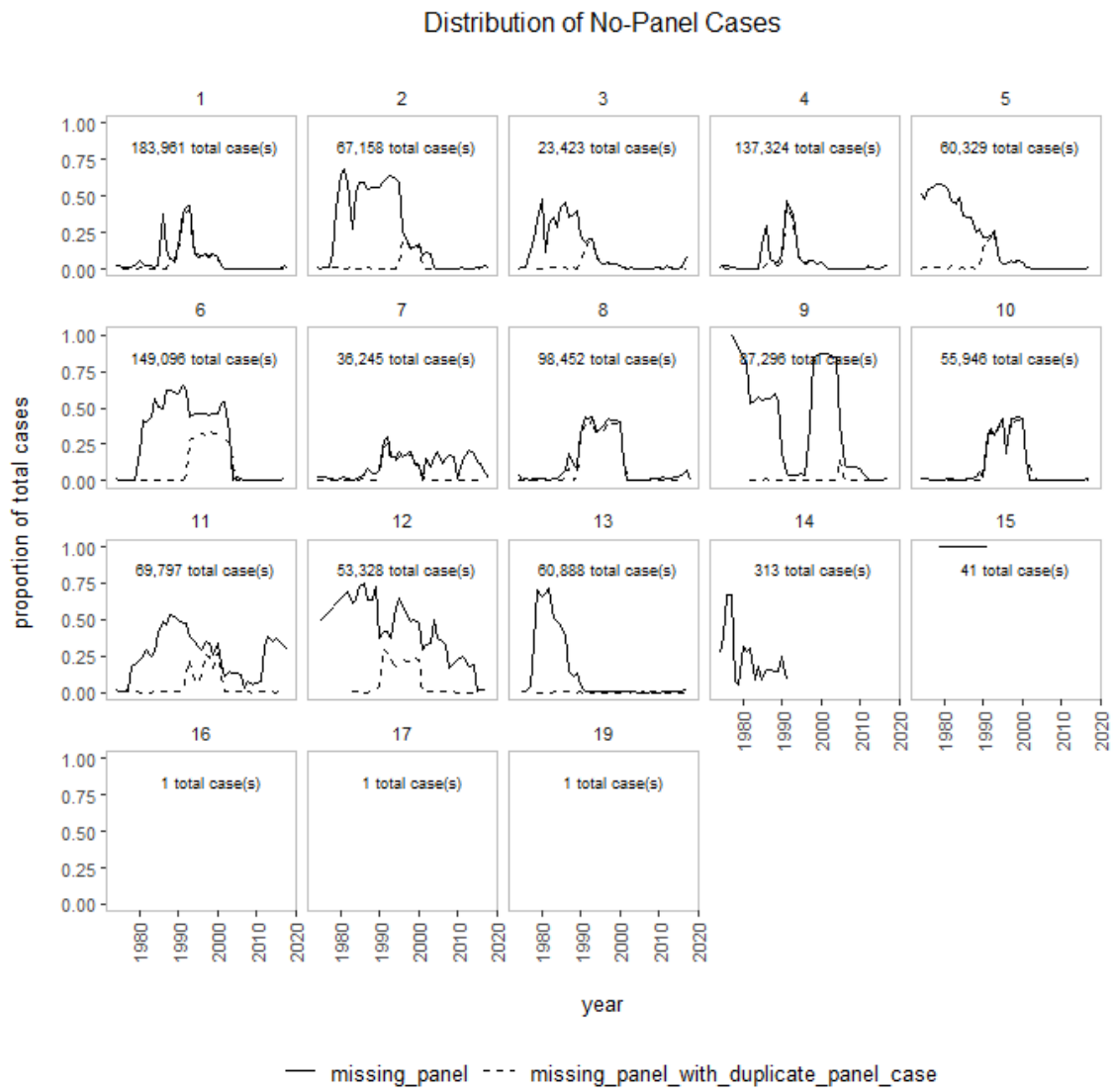


Figure 6: Proportion distribution of cases with missing panel, by court and year

Table 2: Court IDs for figures 5 and 6

| ID | Court |
|----|--|
| 1 | 9th Circuit Court of Appeals |
| 2 | 2nd Circuit Court of Appeals |
| 3 | 1st Circuit Court of Appeals |
| 4 | 4th Circuit Court of Appeals |
| 5 | 7th Circuit Court of Appeals |
| 6 | 5th Circuit Court of Appeals |
| 7 | DC Circuit Court of Appeals |
| 8 | 6th Circuit Court of Appeals |
| 9 | 11th Circuit Court of Appeals |
| 10 | 10th Circuit Court of Appeals |
| 11 | 8th Circuit Court of Appeals |
| 12 | Federal Circuit Court of Appeals |
| 13 | 3rd Circuit Court of Appeals |
| 14 | Temporary Emergency Court of Appeals |
| 15 | |
| 16 | 6th Circuit Bankruptcy Appellate Panel |
| 17 | Tennessee Eastern District Court |
| 18 | Court of Federal Claims |

Table 3: Cases with more or less than three judges in panel

-
- Around 200 random samples of cases were taken for close examination. We discovered that many cases with less or more than 3 judges in panel are actually 3 judge panel cases. They have the following patterns.
 - Cases with some of the judges missing
 - 1975; 3S4X-3RD0-0039-M292-00000-00 Wolf v. Jane Phillips Episcopal-Memorial Medical Center (Only Seth appears in the data, missing Judge Lewis and McWilliams)
 - 1991; 3S4X-BPJ0-008H-V351-00000-00 Love v. Love (Only Brorby appears in the data, missing Judge Anderson and Tacha)
 - 1979; 3S4W-WJ90-0039-M01F-00000-00 St. Regis Paper Co. v. Marshall (Only Seth appears in the data, missing Judge Lewis and McWilliams)
 - 1998; 3S3D-VVD0-0038-X18P-00000-00 Brown v. Gome (Only Leavy and Canby appear, missing Pregerson)
 - 1990; 3S4X-1NY0-003B-51X8-00000-00 Pledger v. Frank (Only Russell and Butzner appear, missing Phillips)
 - The list goes on. For the majority of these cases, it seems that missing judges do not have hyperlinks.
 - Cases with redundant judges
 - Judge of different court level included in the panel
 - * 1987; 3S4X-BKC0-001B-K27G-00000-00 Richard A. Plehn, Inc. v. Gould (Hilton and Merhige, district court judges prior to appeal, included in the panel in addition to the three judges)
 - * 1988; 3S4W-YDM0-001B-K27K-00000-00 MARTIN v. UNITED STATES (Scalia, the Supreme court judge, is included in the panel when Scalia appears in the opinion text as reference)
 - One judge counted multiple times
 - * 1989; 3S4X-C640-003B-555C-00000-00 Vaccaro v. Stephens (Herbert Y.C. Choy appears twice as Herbert Y.C. Choy and Herbert Young Cho Choy)
 - One judge leads to multiple judges with the same last name
 - * 1985; 3S4X-JKR0-0039-P50D-00000-00 Weems v. Ball Metal & Chemical Div., Inc. (Bailey Brown and Jones Brown both included in the panel. Only Brown shows up in LN search result. The hyperlink to Brown leads to Bailey Brown, suggesting that Jones Brown may be added by mistake)
 - Note that the above are selected example of cases. The issue is pervasive such that for 115 sampled cases with 2 judges in panel (according to LN SQL database), we found that 110 of them are actually 3 judge panel cases.
 - More information can be found in the random samples we collected (csv files will be provided)

- Theory

- In LexisNexis SQL data, we see the full name of judges while we mostly see last names of judges with hyperlinks in LexisNexis online searches.
 - The mismatches I found had several patterns:
 - * judges with no hyperlinks are missing
 - * district court judges appear in panel
 - * one judge counted multiple times
 - * unrelated judge is added
 - My theory is that LexisNexis created the SQL dataset by going through the webpage, getting last names of judges, and matching them with their judge database for more information -e.g. their full names. This explains why judges without hyperlinks are missing in general because they are likely to have not been linked to their judge database. This also explains why judges of different court level (either in prior history section or in opinion section) show up in the panel. It does not exactly explain how a judge's last name with hyperlink leads to two different judges (e.g. In one of the cases, 'Brown' in search result led to 2 different judges in LN SQL: Bailey Brown and Jones Brown. The hyperlink of Brown is linked to Bailey Brown.) But it suggests that it also has something to do with the process of getting the last name of the judge and finding them in their judge database (this might have resulted in multiple matches all of which are added to the SQL database, hence showing 4 judges in panel instead of 3.)
 - This is still a wild guess since there are cases that missing judge had a hyperlink. But if some of it is correct, LN can reflect on those and improve their method to get us better data.
-

Table 4: Sample of cases with no legal topics

| LNI | Case Title |
|------------------------------|---|
| 3RR4-V7T0-00B1-D45F-00000-00 | Lutin v. New Jersey Steel Corp., 122 F.3d 1056 |
| 3RS7-P8W0-006F-M398-00000-00 | USA v. Acklen, 1996 U.S. App. LEXIS 23537 |
| 3RV5-HWT0-0038-X2TH-00000-00 | United States v. Kimbrough, 1998 U.S. App. LEXIS 698 |
| 3S3V-G7T0-0038-X1TD-00000-00 | United States v. Randall, 1998 U.S. App. LEXIS 2902 |
| 3S4V-JP00-0039-V3CP-00000-00 | A.N. Deringer, Inc. v. United States, 832 F.2d 592 |
| 3S4V-JVN0-0039-V4V2-00000-00 | Connelly v. DOT, 802 F.2d 468 |
| 3S4V-JW40-0039-V518-00000-00 | Yamamoto v. U.S., 795 F.2d 1018 |
| 3S4W-TVW0-003B-G19W-00000-00 | Miller v. Bestline Corp., 749 F.2d 731 |
| 3S4W-XNG0-003B-G49F-00000-00 | UNITED STATES v. MATOOK, 729 F.2d 1464 |
| 3S4W-Y700-003B-G2ND-00000-00 | Myles v. Commissioner, IRS, 719 F.2d 373 |
| 3S4W-YP00-001B-K4CY-00000-00 | Bajoa v. U.S. Dep't of Immigration & Naturalization, 855 F.2d 860 |
| 3S4X-1720-0039-M53W-00000-00 | Burleigh House Condominium, Inc. v. Buchwald, 546 F.2d 57 |
| 3S4X-1JK0-001B-K1BF-00000-00 | United States v. Jefferson, 843 F.2d 1388 |
| 3S4X-1NK0-003B-P49D-00000-00 | United States v. Hicks, 37 F.3d 1496 |
| 3S4X-1RF0-003B-P4W3-00000-00 | Harvey v. Gramley, 37 F.3d 1501 |
| 3S4X-1X10-006F-M154-00000-00 | Wooden v. Oliver, 91 F.3d 136 |
| 3S4X-1XN0-006F-M1D0-00000-00 | Cawley v. Mony Life Ins. Co., 95 F.3d 47 |
| 3S4X-2C20-006F-M1PY-00000-00 | United States v. Whittlesey, 85 F.3d 639 |
| 3S4X-2CF0-003B-G0B3-00000-00 | Lewis v. U.S.A., 691 F.2d 503 |
| 3S4X-2FK0-008H-V4Y3-00000-00 | UNITED STATES v. BERBESSI-ACOSTA, 972 F.2d 1332 |
| 3S4X-2N70-006F-M0ST-00000-00 | United States v. Mickle, 1996 U.S. App. LEXIS 7001 |
| 3S4X-2V10-003B-51X6-00000-00 | Collins v. Allied Chemical Corp., 912 F.2d 465 |
| 3S4X-3190-006F-M3YH-00000-00 | Bivins v. Hudson, 81 F.3d 163 |
| 3S4X-3780-008H-V377-00000-00 | Enemy Boy v. McCormick, 967 F.2d 585 |
| 3S4X-3C40-008H-V48F-00000-00 | UNITED STATES v. CHACON, 966 F.2d 680 |
| 3S4X-3FM0-001B-K147-00000-00 | Gradford v. AIG Consultants, Inc., 838 F.2d 473 |
| 3S4X-3K00-008H-V049-00000-00 | Carter v. Demmy, 968 F.2d 12 |
| 3S4X-54S0-008H-V3RK-00000-00 | UNITED STATES v. BRANHAM, 958 F.2d 369 |
| 3S4X-60P0-003B-P3W4-00000-00 | Davidson v. Barratt, 24 F.3d 239 |
| 3S4X-8VB0-003B-P230-00000-00 | United States v. Blackburn, 1994 U.S. App. LEXIS 2149 |
| 3S4X-9NP0-003B-P169-00000-00 | DeTreville v. Department of the Army, 1993 U.S. App. LEXIS 37038 |
| 3S4X-BFG0-001T-D11W-00000-00 | Hayden v. Parker, 68 F.3d 474 |
| 3S4X-C2X0-003B-54NP-00000-00 | Coffin v. Murray, 875 F.2d 314 |
| 3S4X-CGR0-003B-51VD-00000-00 | Hoang v. U.S. Office of Personnel Management, 873 F.2d 1443 |
| 3S4X-D140-003B-54B9-00000-00 | NLRB v. USPS, 872 F.2d 1023 |
| 3S4X-DKM0-008H-V0W2-00000-00 | UNITED STATES v. WERTS, 931 F.2d 57 |
| 3S4X-G2K0-001T-D0CJ-00000-00 | Carter v. St. Louis Airport Officials, 50 F.3d 12 |
| 3S4X-G610-008H-V36B-00000-00 | Williams v. Wells, 925 F.2d 1467 |
| 3S4X-GCB0-001T-D2F3-00000-00 | United States v. Leal, 1995 U.S. App. LEXIS 4978 |
| 3S4X-HC60-003B-P4D6-00000-00 | Hassan v. FBI, 1993 U.S. App. LEXIS 12813 |
| 3S4X-JDV0-00B1-D3W4-00000-00 | Rodriguez v. Quinn, 107 F.3d 873 |
| 3S4X-JRG0-00B1-D2X5-00000-00 | United States v. Attar, 1997 U.S. App. LEXIS 79 |
| 3S65-BK00-0038-X4XV-00000-00 | Smith v. Westlake PVC Corp., 132 F.3d 34 |
| 3T2G-V5X0-0038-X175-00000-00 | United States v. Cornejo, 141 F.3d 1180 |
| 3TVN-0C70-0038-X3S8-00000-00 | Meester v. Runyon, 1998 U.S. App. LEXIS 25459 |
| 3W24-DMG0-0038-X54B-00000-00 | Milton v. Scott, 172 F.3d 868 |
| 3WJN-7W70-0038-X3JY-00000-00 | United States v. Khachatryan, 172 F.3d 60 |
| 3WR2-1FB0-0038-X121-00000-00 | Camastro v. City of Wheeling, 173 F.3d 849 |
| 3WVH-BFD0-0038-X45T-00000-00 | Johnson v. Cain, 182 F.3d 914 |
| 3X6J-06G0-0038-X4F5-00000-00 | United States v. Durham, 178 F.3d 1286 |
| 3X70-S4G0-003B-91K0-00000-00 | Deggins v. United States, 178 F.3d 1308 |
| 3YK4-RK00-0038-X0XW-00000-00 | COMSAT Corp. v. FCC, 2000 U.S. App. LEXIS 2002 |
| 4019-X3X0-0038-X440-00000-00 | U.S. v. Jackson, 207 F.3d 663 |
| 405X-NKB0-0038-X1TH-00000-00 | Foote v. Apfel, 2000 U.S. App. LEXIS 8637 |
| 42G6-HP00-0038-X430-00000-00 | Revelles v. Stout, 238 F.3d 430 |

| | |
|------------------------------|---|
| 44WB-FP70-0038-X0H1-00000-00 | United States v. Fuell, 23 Fed. Appx. 173 |
| 462C-XH70-0038-X103-00000-00 | USA v. Brown, 37 Fed. Appx. 713 |
| 47JD-5P70-0038-X3H8-00000-00 | United States v. Hargrove, 54 Fed. Appx. 796 |
| 49WN-82S0-0038-X0M5-00000-00 | Hall v. Sullivan, 67 Fed. Appx. 587 |
| 4CN6-8PT0-0038-X3C5-00000-00 | Parker v. True, 100 Fed. Appx. 930 |
| 4D1M-W7T0-0038-X329-00000-00 | United States v. Thorn, 2004 U.S. App. LEXIS 16253 |
| 4DN8-MH30-0038-X4GT-00000-00 | Benson v. Thompson, 112 Fed. Appx. 283 |
| 4FPP-KVC0-0038-X1PB-00000-00 | Hunt v. Johnson, 123 Fed. Appx. 122 |
| 4G0T-GYD0-0038-X439-00000-00 | Leger v. Sec'y, DOC, 126 Fed. Appx. 463 |
| 4GG7-NHK0-0038-X3H9-00000-00 | United States v. Cortes-Melendez, 135 Fed. Appx. 781 |
| 4GWB-4YN0-0038-X21F-00000-00 | Blackwood v. Fairways & Greens, 138 Fed. Appx. 300 |
| 4PSD-3820-TX4N-G1B1-00000-00 | United States v. Pina, 243 Fed. Appx. 574 |
| 4SJ9-SWV0-TXFX-82J3-00000-00 | Sanchez v. City of Franklin, 278 Fed. Appx. 490 |
| 4TKB-9R40-TXFX-D39G-00000-00 | D.A.D.S. Denim, Inc. v. Ramirez, 295 Fed. Appx. 206 |
| 4VNC-3S70-TXFX-D2TD-00000-00 | United States v. Ramos, 312 Fed. Appx. 852 |
| 4VTK-1CY0-TXFX-D1VW-00000-00 | Rimac v. Duncan, 2009 U.S. App. LEXIS 5143 |
| 4VXK-B4R0-TXFX-D31D-00000-00 | Espinoza v. Holder, 320 Fed. Appx. 588 |
| 4W7S-J1C0-TXFX-D2VS-00000-00 | Hall Family Props., Ltd. v. Gosnell Dev. Corp. (In re Gosnell Dev. Corp.), 331 Fed. Appx. 440 |
| 4WS8-H010-TXFX-D2F7-00000-00 | United States v. Gregory Island, 336 Fed. Appx. 759 |
| 4XBP-SH50-TXFX-G1VC-00000-00 | Louis v. United States AG, 333 Fed. Appx. 497 |
| 51CC-GV91-652R-1004-00000-00 | Rolle v. United States, 399 Fed. Appx. 695 |
| 5561-CGN1-F04K-V2BY-00000-00 | Johnson v. Chevron Corp., 472 Fed. Appx. 428 |
| 5578-7B61-F04K-P0WG-00000-00 | Houston v. Logan, 674 F.3d 613 |
| 55WK-4P11-F04K-S10M-00000-00 | Neely v. McDaniel, 2012 U.S. App. LEXIS 12159 |
| 59VN-MV31-F04K-M1FS-00000-00 | James v. Chesapeake City Jail, 1986 U.S. App. LEXIS 37694 |
| 59VN-MX01-JCNH-X072-00000-00 | Docter v. Brown (In re Sports Accessories, Inc.), 1986 U.S. App. LEXIS 37735 |
| 5B1M-8BC1-JCNH-X422-00000-00 | Dammerau v. Wright, 1990 U.S. App. LEXIS 27067 |
| 5B5N-NF61-F04K-M4Y1-00000-00 | Williams v. Brinkan, 2003 U.S. App. LEXIS 28742 |
| 5BNB-F5V1-F04K-M239-00000-00 | McPhatter v. Wilson, 557 Fed. Appx. 250 |
| 5CVC-R6R1-F04K-M1C6-00000-00 | Day v. Rogers, 1997 U.S. App. LEXIS 42420 |
| 5D6V-GC31-F04B-M0BJ-00000-00 | Wagstaff v. United States, 2014 U.S. App. LEXIS 18507 |
| 5DX1-B9T1-JCNH-X25M-00000-00 | Mochelle v. J. Walter, Inc., 1994 U.S. App. LEXIS 42121 |
| 5F5W-DJ01-JCNH-X109-00000-00 | Solis v. Collins, 1993 U.S. App. LEXIS 40874 |
| 5FTS-CB91-F04K-S0BX-00000-00 | Irby v. Steele, 2015 U.S. App. LEXIS 6707 |
| 5GJY-61G1-F04K-M0GD-00000-00 | Chernov v. Lynch, 608 Fed. Appx. 180 |
| 5H07-WPW1-F04K-V0BC-00000-00 | Dowdy v. Curry, 617 Fed. Appx. 772 |
| 5M73-GXB1-F04K-M017-00000-00 | Jassie v. Mariner, 670 Fed. Appx. 789 |
| 5MF7-94C1-JCB9-617F-00000-00 | Ladd v. Neuhooff Packing Co., 1978 U.S. App. LEXIS 13241 |
| 5NR9-HH91-F04K-M15X-00000-00 | Flood v. Warden, Lieber Corr. Inst., 691 Fed. Appx. 711 |
| 5PCM-PJ81-F04K-N0VR-00000-00 | United States v. Berny-Estrada, 697 Fed. Appx. 315 |
| 5PGK-SYX1-F04K-J0G6-00000-00 | United States v. Smith, 697 Fed. Appx. 83 |
| 7XGT-BS10-YB0V-P071-00000-00 | Talavera v. Holder, 361 Fed. Appx. 773 |
| 7Y3F-MJ00-YB0V-K02T-00000-00 | Kim v. Parker, 373 Fed. Appx. 606 |
| 7YTJ-J4M1-652R-200H-00000-00 | Larue v. Matheney, 385 Fed. Appx. 328 |
| 801R-38K0-YB0V-F017-00000-00 | Rhines v. Bledsoe, 388 Fed. Appx. 225 |

2 Important Past Observations and Comments

Following is a compilation of observations made by the team regarding apparent missing or discrepant data. Supporting spreadsheets, tables, and examples are available by viewing referenced e-mails. Have each of these been resolved?

| Source | Content |
|--------------------------|--|
| Stuart, 4-25-2019 e-mail | <p>I am going over the first part of the slice of cases we are checking, and Jane and Wick are going over the second part. I am not close to finished my part. I am emailing all of you with an update that surprises me, and makes the process really, really slow. I discovered that there were some discrepancies between what the spreadsheet shows and the results I get on LN (Shepards, then Table of Authorities, then looking at the substantive positive and negative treatments in majority opinions, and thus no cited treatments). I had been treating the LN client as truth. But after finding a bunch of these, I decided to read the citing case to see if I could figure out why the LN spreadsheet had these errors. And I discovered that most of the errors were actually mistakes in the LN client (the Web service for which law firms pay so much money). I erred on the side of giving the benefit of the doubt to LN when there was a discrepancy (that is, deciding close cases against our data from LN on the spreadsheet). And I did find one discrepancy that (if repeated widely in our data) would be a bit concerning two circuit court cases that arent listed in our spreadsheet but should be (see cell A158 of the attached). There are two other discrepancies that dont bother me (a case from 200 years ago and a state case). I put the discrepancies that favor our spreadsheet over the LN client in brackets (scroll down Column A for all discrepancies).</p> <p>I say all this by way of saying that this is taking me much longer than I had been planning on, and Im not happy about the one discrepancy that matters, though Im not overly concerned at this point. But Im also highlighting that, sadly, to evaluate the discrepancies between the spreadsheet and the LN client we need to read the underlying cases and see how they treat the cited cases.</p> |
| Stuart, 4-26-2019 e-mail | <p>I have now gone over the first part of the spreadsheet with cases that have 10-25 substantive treatments. My notes are in Column A immediately after the relevant case, and I put in an extra hard return after my comments, so they are easy to find if you just scroll through the document.</p> <p>Bottom line: of what I checked, the data they sent us is very near complete for what we want, but not complete. I stopped at row 828 of the attached (of 1505 rows). Many of the citations are duplicates (e.g., Supreme Court cases cited 2 or 3 ways). I would estimate that there are about 425 unique treatments in what I reviewed. I found 4 missing cases that should be in the dataset that is, four federal court opinions that are shown in the LN Client (the web service) but are not in the dataset. You can see them in the following rows of the spreadsheet: Row 158 (two cases), Row 367, Row 433</p> <p>If this pattern is representative, that would suggest that we have about 99% of the cases that we want. (And, as I noted in my previous email, this dataset avoids some mistakes that I found in the LN client.) Still, I am a bit disappointed to have found the four omissions I found, as I cant see any reason for them. I know that LN is updating its databases all the time, and its possible that these were recently added despite the fact that the cases are pretty old. Otherwise, I cant figure out why they are missing.</p> |

Note that the four omissions noted above *dont* include some omissions of materials that are probably not of interest to us, but that I would have expected LN to include in our dataset (given that we asked for all the substantive treatments from a given case). I noted those omissions in the attached document. What our database seems not to include: State cases, Bankruptcy court cases, Orders/opinions from administrative agencies, Really old cases (the one missing case I found is *United States v. Schooner Peggy*, 5 U.S. (1 Cranch) 103, 110 (1801), so I dont have a sense of what really old might mean)

The only of these categories that could potentially be a problem is the last one, in case we are missing more recent (than 1801) cases and thus some of the universe. Tom, can you check to see how far in time the cited cases go back in time? Maybe you could search for 19th century cases?

Jane and Wick, what have you found in your review of the second part of this spreadsheet? (By the way, a LN rep gave me a time-saving suggestion to type shep: [citation], which saves a step. You still have to click on Table of Authorities and then choose the substantive treatments, but it saves a bit of time.)

Kevin, 4-26-2019 e-mail Thanks for the update, Stuart. If we Jane and Wick see similar patterns (both in terms of the percentage of missing cases and the types of missing cases) this actually looks pretty good to me. I dont think many referees are going to complain about this level of incompleteness.

I suppose we could ask LN if they have any ideas about why there are these missing cases, but its probably not necessary to have them track down all such problems unless their explanation for why this is happening suggests that it is actually a bigger problem than we currently think.

Stuart, 4-30-2019 e-mail Perhaps most striking (and perhaps of greatest interest to you), I looked at all the discrepancies between the data you sent and the data on the LN client, and I found that many of the discrepancies were mistakes in the data on the LN client. That is, the data you sent were correct, and what shows on the LN client is not correct. I noted those on the spreadsheet for your benefit (look for brackets in column A). Im obviously not concerned about this in the least. I figured as long as I was going to spend so many hours going over all this, I might as well share what I found with you, in case its helpful to you.

Jane, 5-1-2019 e-mail The total cases in the data for 1996 were 44,659. Between 1996 and 2004, a fairly steadily decreasing number are reported, to a low of 27,379 in 2004. The decrease in number seems to be coming from the cases in the “unidentified column. When we looked at the caseload data for those years that is produced from the Administrative Office of the US Courts website, there was no decrease in case load at either the district court or the appellate court level during this time period. We also checked in the LN client to see if their case numbers reflected a similar decrease in number over this time period, but they did not. So, the question is twofold: (1) what cases are included in the unidentified column, and (2) was there a change in the type of cases included in this database that took effect around 1996?

When I checked the Lexis front end search, there was a similar result. What I looked at was whether their competitor, Westlaw, also documented a decrease in the number of opinions, and it does not.

Tom, 5-29-2019 e-mail Hi Stuart,

A list of cases with missing panel and broader date range is attached. There is nothing in the data supplied by LN that indicates a link to any other case so that, with what we have, there does not appear to be a way to determine if one case record supersedes another. At present, if two records actually represent a single case, then a query would duplicate that case. My opinion is that if we do not want duplicated (linked?) cases then LN should omit them from their query when constructing our data set. We do not have a method to reconstruct your analysis, below, using only the data provided. With over 200,000 cases having no panel and an apparent association between missing panels and links, I think we definitely need to address (get clarification from LN) the scenarios under which the panel is missing and develop a method of omitting any unwanted duplicate cases.

Tom

From: Stuart M. Benjamin ;benjamin@law.duke.edu;

OK, so this is a bit complex, and requires a little more digging.

First thing: all the cases you sent me are from a narrow date range. Can you send me some examples of panels with no judges from much earlier and later years? (Or do they not exist, which itself would be good to know.)

Now, on to the issue: I looked up the 10 cases you list below, and for each of them LN doesn't list any judges. That surprised me, b/c usually even for unpublished opinions that affirm in very few words, members of the panel are usually listed. So I looked them up on Westlaw, and Westlaw shows the names of judges (and opinions ranging from a sentence to a few paragraphs, although all are applying settled law and are very cut-and-dried). So I then went back to LN, and noticed that 7 of the 10 cases on LN has a notation that says Reported in Full-Text Format at and then has a Lexis citation number (see below). And at that link there is the full opinion and a list of judges. So: so long as we have all the *linked to* opinions, we are in good shape. Indeed, if we have the linked to opinion, we wouldn't want to get the judges names for the citations you list in your 3:14pm email, b/c then we would be double-counting a given case (b/c, again, LN includes each case in two places one with no judges listed, and one with judges listed and the opinion). Note: two of the 10 opinions (USA v. McNeese, 127 F.3d 37 and USA v. Garrett, 131 F.3d 155) don't link to another opinion, b/c there appears to be no opinion and no panel. Also, one of the opinions (United States v. Hanson, 1997 U.S. App. LEXIS 30475) is a denial of a petition for rehearing en banc (such denials of rehearing are very common); such rehearings don't list panel members and don't have opinions. So the point in the paragraph above doesn't apply to all 10 opinions you sent me. But we don't need to worry about cases like USA v. McNeese and US v. Hanson, b/c there are no judges listed for those cases *anywhere* – the court never listed the judges.

Bottom line: if we have the linked to opinions, then we are in good shape. So: can you determine if the linked to opinions are in our dataset from LN? Specifically, can you look up to see if the following are in our dataset, separate from the entries you list in the email you sent me at 3:14pm today:

Tom, 5-30-2019 e-mail

I believe that we have two problems:

1. Many cases with a single opinion on the LN site have multiple case header records in the data supplied to us. Some of these have a panel in our data, others do not.

2. Differences, in the LN data supplied, of case titles for a single case prohibit unequivocal consolidation of multiple case records to a single one that represents a single opinion. Example: [United States v. Licea Sandoval] and [United States v. Sandoval].

It would be helpful to have LN explain item 1 so that we can develop a strategy for aligning our data with what you expect it to be, based on an interpretation of on-line case data.

Kevin, e-mail, 6-10-2019

Stuart not sure if you have any law student RAs at the moment. If you do, would they have time to help lookup some of these cases that ByungKoo is looking into? I ask primarily b/c ByungKoo doesn't have Lexis access yet (I'm still in discussions with the law school here about that). With some additional RA help we could also look at more than 100 cases (stratified random samples by circuit and year might be useful).

Alternatively, we probably are getting very close to having enough info about the 2 judge cases to be able to go to LN and ask some questions about what is going on even if we don't fully understand why this is happening.