Phase I and Phase II Supreme Court Database (compiled by Vanessa A. Baird and archived at the University of Colorado). Individuals can access these datasets through the Ulmer Project by clicking on the appropriate web links.

FUTURE DIRECTION OF THE ULMER PROJECT AND ITS CONTRIBUTION TO LAW AND COURTS

It is my hope that the Ulmer Project contributes positively to the Law and Courts community. By archiving judicial datasets in a single location, and distributing the most recent information across multiple formats, I believe a valuable environment is established; one in which individuals can access updated versions of data and pose questions to individuals knowledgeable in judicial politics and empirical methodology. As more individuals either archive their data with the Ulmer Project directly, or allow links from the Ulmer Project to their locations, then this benefit will continue to expand. The datasets do not have to be large, NSF-sponsored files. Smaller databases are equally welcome and I will work with any and all individuals to make their data publicly available. Additionally, I welcome working papers from all individuals with research interests in law and judicial politics. Finally, suggestions from the Law and Courts community on future directions are always appreciated. Without input from the community of judicial scholars, it is difficult to determine how best to evolve and develop the Ulmer Project.

Notes
[1] The Ulmer Project is accessed at http://www.as.uky.edu/polisci/ulmerproject
[2] This biography was provided by Sidney Ulmer for inclusion on the website
[3] These data were archived initially at the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan, and later at the Michigan State University Program for Law and Judicial Politics.

References

Barrow, Deborah J., Gary Zuk, and Gerard S. Gryski. 1996. *The Federal Judiciary and Institutional Change*. Ann Arbor: University of Michigan Press.

Bradley, Robert C. 2003. "S. Sidney Ulmer: The Multidimensionality of Judicial Decision Making." In *The Pioneers of Judicial Behavior*. Nancy Maveety, editor. Ann Arbor: University of Michigan Press.

Maveety, Nancy. 2003. *The Pioneers of Judicial Behavior*. Ann Arbor: University of Michigan Press.

Songer, Donald R., Reginald S. Sheehan, and Susan B. Haire. 2000. *Change and Continuity in the U.S. Courts of Appeals*. Ann Arbor: University of Michigan Press.

BECOMING AN INTELLIGENT USER OF THE SPAETH SUPREME COURT DATABASES
Sara C. Benesh
University of Wisconsin - Milwaukee
sbenesh@uwm.edu

Scarcely an article on Supreme Court behavior fails to avail itself of one of the Spaeth Supreme Court Databases. However, that they are widely used does not necessarily mean they are widely understood. Here, I attempt to highlight the contents of the various databases, provide instructions on how and where to obtain them, and give some insight on how to intelligently use them. Questions will, no doubt, remain. Please feel free to direct them to me via email (sbenesh@uwm.edu). Harold Spaeth himself will be happy to help as well (spaeth@msu.edu).

There are currently six (soon to be more) separate databases. One covers final voting behavior from the Vinson Court through the entirety of the Rehnquist Court. This is called "The Original Supreme Court Database" on the Ulmer Project Site (http://www.as.uky.edu/polisci/ulmerproject/). The second covers all stages of decision making (cert, merits, and final vote) for the Vinson and Warren Courts and is labeled "The Vinson-Warren Court Database." The third adds the Burger conference data to its final vote data and is named "The Burger Court Database." The last three change the unit of analysis from the case to the individual justice vote: "The Justice-Centered Rehnquist Court Database," "The Justice-Centered

Burger Court Database," and "The Justice-Centered Warren Court Database." There are a plethora of variables in all of these databases, providing an expansive array of information from which to draw. Spaeth puts these variables into several categories: background variables (including jurisdiction, administrative agency action, the source and origin of the case, the reason given for hearing the case, the parties to the case, and the disposition of the case in the court prior to the Supreme Court); chronological variables (including the date of oral argument, of reargument, if any, and the date of decision); substantive variables (including legal provision, authority for decision, issue and issue area, and the ideological direction of the decision); outcome variables (including the type of decision, the disposition of the case, the winning party, whether there was a formal alteration of precedent, and whether there was a declaration of unconstitutionality); and voting and opinion variables (including the vote in the case, and the votes, opinions, and interagreements of the justices) (Spaeth and Segal 2000). All of these are available for free downloading at the Ulmer Project's website (referenced above) along with all relevant documentation. However, this is an overwhelming and sometimes intimidating set of information. I hope to make it a bit more accessible here.

OBTAINING THE DATA

First, to download the databases, navigate to http://www.as.uky.edu/polisci/ulmerproject/ and click on *Databases*. Several different databases appear on this page (including the Songer Courts of Appeals database, the Auburn Attributes of Courts of Appeals judges database, and the State Supreme Court Database, among others). Click on *United States Supreme Court*. The next page offers free software used for viewing and downloading information contained (PKZip, Stuff-It, and Acrobat Reader) and a listing of the various databases discussed above, each with several download options. The user can download the data as an SPSS file, a STATA file, a SAS file, or an ASCII file, and can download the documentation as a PDF file. For users familiar with SPSS, STATA, or SAS, it is easiest to download the data in those formats. If you use a different statistical software program, the ASCII file should allow for relatively painless conversions. Simply click on the relevant format and save the file to your hard drive. Then double click the file and, if you have an unzip program, the file will be unzipped. Then either double click the unzipped file (if your computer recognizes the extension and opens SPSS, STATA or SAS) or click "Extract" and choose a file location to save the database, then open it in the relevant program.

I advise anyone using the databases for the first time to also download the documentation at this time. To do so, click on *Documentation*. If Acrobat Reader is installed on your computer, the file should come up directly. You can either save or print this document. If you plan to use the databases extensively, I would advocate printing the documentation. If not, the user may wish to read the documentation on screen as it is a very long document (101 pages for the original databases). Either way, it is nearly impossible to use the databases without reading the documentation (although it is more feasible now with the addition of value labels), so I urge every user to do so.

SELECTING A UNIT OF ANALYSIS AND DECISION TYPE

If researchers were to open the database and start running models, they would suffer an unfortunate fate, for they will have included non-orally argued cases and per curiams without determining whether they should, given their research question. They will have double- and triple-counted certain cases because they will have included multiple dockets, multiple issues, and multiple legal provisions and probably counted each as a separate case. The first step in intelligently using the databases, therefore, is to make the primary case selection decisions about the relevant unit of analysis and decision type.

ANALU

As you may know, case citations may have several docket numbers as the Court oftentimes combines several cases under a single opinion. Spaeth deals with this by having an additional record for each docket number. In addition, many cases have more than one issue or more than one legal provision at bar. Because issue and legal provision are single variables, Spaeth deals with multiple issues and multiple legal provisions by duplicating the record for a given case with the new issue or legal provision noted. (All other entries in that new record will be the same as the first record for that case citation.) If there are multiple dockets and multiple issues, each docket/issue will receive a record. Finally, cases in which there was a split vote – e.g., a justice dissents on one issue and is in the majority on another, or five justices voted for one part of the opinion and a different five for another – also have two entries, one for the first part of the decision and one for the second. This oftentimes leads to some confusion for researchers.

The following are the options available to the researcher:

ANALU	UNIT OF ANALYSIS
0	Case Citation
1	Docket Number
2	Multiple Issue Case
3	Multiple Legal Provisions Case
4	Split Vote Case
5	Multiple Issues and Multiple Legal Provisions

In the Burger Court database, there is an additional unit of analysis: analu=6, which indicates a relisted case. For information on relisting, see Spaeth 2004.

What follows are some ideas of usages of the databases and the appropriate unit of analysis:

RESEARCH QUESTION	UNIT OF ANALYSIS
% Liberal Decisions Per Term	Case Citation
All Issues Addressed Each Term	Case Citation and Multiple Issue
All Legal Provisions Used Each Term	Case Citation and Multiple Legal Provision
Number of Opinions Written by Rehnquist	Case Citation and Split Vote
Assignments Made by Associate Justices	Case Citation and Split Vote
Number of Cases Declaring Laws Unconstitutional	Case Citation
Number of Cases Altering Precedent	Case Citation
Which Courts Supreme Court Reviews	Docket Number
Disposition of Cases at Supreme Court	Docket Number
Parties to Supreme Court Cases	Docket Number

In general, any time one wants to study decisions in the aggregate and wants to count each decision only once, one should use the case citation. That record is the first record for any case and contains the "most important" issue and legal provision dealt with by the Court (according to Spaeth). If one wants an idea of the range of issues or legal provisions dealt with in a single term, one may wish to use the multiple issue unit of analysis and then combine several issues into one case for any percentages computed. If one needs to know about the treatment of certain lower courts or the success of certain parties, one should use the docket as the unit as different dockets have different parties and points of origin. Do note that if one wanted a picture of all issues, one would need to choose both case citation and multiple issue case in order to obtain all issues in a given case. Finally, note also that certain variables locate only in a specific record in a given citation; e.g., uncon and alt_prec. Hence, the focus in such research should be on the specific variable and analu should, for the most part, be ignored.

So what do these choices mean in terms of outcome? Below are some examples of cases with multiple dockets, issues, legal provisions, and split votes and how they appear in the database. This again demonstrates the need to sometimes select more than one unit of analysis and certainly to always use the research question to guide which analu to select.

Case #1: One record, one docket, one law, one issue

us	docket	analu	rec	law	issue
519/0316	95-789	0	.	ERIS	911

Case #2: Three records, three dockets, one law, one issue

us	docket	analu	rec	law	issue
517/0343	95-7587	0	.	SCTR	385
517/0343	95-7588	1	2	SCTR	385
517/0343	95-7589	1	0	SCTR	385

Case #3: Five records, one docket, five laws, one issue

us	docket	analu	rec	law	issue
517/0748	94-1966	0	.	EO	362
517/0748	94-1966	3	4	11	362
517/0748	94-1966	3	0	1814	362
517/0748	94-1966	3	0	172	362
517/0748	94-1966	3	0	8A	362

Case #4: Six records, three dockets, two laws, one issue

us	docket	analu	rec	law	issue
517/0952	94-805	0	.	14A=	250
517/0952	94-805	3	1	321	250
517/0952	94-806	1	2	14A=	250
517/0952	94-806	3	1	321	250
517/0952	94-988	1	0	14A=	250
517/0952	94-988	3	1	321	250

Case #5: Ten records, three dockets, two laws, one issue

us	docket	analu	rec	law	issue
354/0298	6	0	.	SMIT	422
354/0298	7	1	2	SMIT	422
354/0298	8	1	0	SMIT	422
354/0298	7	3	1	STOP	422
354/0298	6	4	2	SMIT	422
354/0298	7	4	2	SMIT	422
354/0298	8	4	2	SMIT	422
354/0298	6	4	0	SMIT	422
354/0298	7	4	0	SMIT	422
354/0298	8	4	0	SMIT	422

Case #6: Two records, one docket, one law, two issues

us	docket	analu	rec	law	issue
355/0083	20	0	.	54S899	673
355/0083	20	2	1	54S899	721

Note that rec is basically a counting variable. When analu = 0, it is empty (either . or 0). When analu = 1, rec gives the number of additional docket numbers in the case. When analu = 2, rec gives the number of additional issues in the case. When analu = 3, rec gives the number of additional legal provisions. When analu = 4, rec gives the number of additional split vote records. When analu = 5, rec gives the number of additional legal provisions and issues. All of these entries for rec occur only in the first record of the new analu type.

DEC_TYPE

The second decision every researcher needs to make when approaching the database is the decision type s/he wishes to employ. There are seven different decision types coded by Spaeth in his databases. They are as follows:

DEC_TYPE	EXPLANATION
1	Formally decided, full opinion cases. Were granted oral argument and resulted in a signed opinion.
2	Cases decided with an opinion, but without oral argument; e.g., per curiam.
3	Memorandum cases (cert decisions, motions to file as amicus curiae, orders).
4	Decrees (decisions on cases arising under the Court's original jurisdiction).
5	Equally divided vote (due to nonparticipation of one or more justices) with the effect of upholding lower court's decision due to a tie.
6	Per curiam, orally argued cases (unsigned, but orally argued)
7	Judgments of the Court (majority attracts only a plurality of the justices)

Again, it is immediately evident that much violence would be done to one's analysis if one did not decide which of these seven decision types to include in one's analysis. The following research situations would be associated with the following decision types:

RESEARCH QUESTION	DECISION TYPE
Who writes the Court's majority opinions?	1 and possibly 7
Which states win more often in general?	1, 2, 3, 4, 5, 6, and 7
Is the Court more liberal than before?	1, 2, 4, 6, 7
How often does the Chief assign the majority opinion?	1, 7
What role do oral arguments have in decision making?	1, 5, 6, 7
What precedents has the Court set in libel?	1, 2, 6

Of course, people may disagree over whether to include per curiams or judgments in a given research project. The researcher should always report the decision he or she makes in this regard, however, so that others may evaluate the cases on which their analysis is based. Most cases in the database are of type 1, but results could be skewed, depending on the research question, if other types are erroneously included. Note also that for types 2 and 3, not all cases are included. Type 2 cases are only included if a summary is reported in the U.S. Reports or one of the justices issues a separate opinion. Type 3 cases are only included if one of the justices wrote a separate opinion, which does not often occur. In the Vinson-Warren Court Database, all type 3 cases are included except those deadlisted and those in which no justice supported granting the petition. All other types include the universe of cases decided.

Selecting cases intelligently requires that both analu and dec_type (as well as the variables to be analyzed) are carefully taken into account.

CHANGING THE UNIT OF ANALYSIS

As mentioned earlier, the Spaeth databases also include a set of databases with justice vote as the unit of analysis rather than case citation or docket. The Rehnquist, Burger, and Warren Courts are now available at the individual justice level. These databases will further facilitate research on the Supreme Court by allowing the researcher to focus on individual justice behavior. In so doing, one may ask questions like "How often do the justices disagree as to issue in the case? Legal provision? Whether or not to alter precedent? Whether or not to declare a federal or state statute unconstitutional?" This allows for a more nuanced understanding of the decision making proclivities of the justices.

In order to change the unit of analysis, one can employ the routine discussed by Paul Collins in this issue of *Law & Courts*, or the old routine detailed by Benesh and Zorn (1998). The basic idea is this: the database as now constructed has cases as the unit of analysis, as demonstrated above, such that the database looks like this (where the entries are votes in a given case):

TABLE 1

CASE	JUSTICE 1	JUSTICE 2	JUSTICE 3	JUSTICE 4	JUSTICE 5	JUSTICE 6	JUSTICE 7	JUSTICE 8	JUSTICE 9
1	1	1	1	2	1	2	1	1	2
2	2	1	1	1	2	2	1	1	1

In order to study the individual justices, the datafile needs to be changed to look like this:

TABLE 2

Case	Justice	Vote
1	1	1
1	2	1
1	3	1
1	4	2
1	5	1
1	6	2
1	7	1
1	8	1
1	9	2
2	1	2
2	2	1
2	3	1
2	4	1
2	5	2
2	6	2
2	7	1
2	8	1
2	9	1

Suffice it to say here, that the new databases are so transformed. However, the major work in this new database was not in the flipping but rather in the coding. Indeed, we have coded 4500 special opinions (dissents and concurrences) and the 11000 votes represented by these opinions. All variables in the original databases are included, some recoded for individual justice behavior, and others were added. I detail the contents of these databases below.

VARIABLES AT THE INDIVIDUAL LEVEL

In the flipped database, there are several variables aimed at measuring whether or not there were deviations of individual justices from the Court as to legal provision, issue or authority for decision. The first, law_dev measures the number of deviations as to law a given justice exhibited in that particular case. Note that this is measured as a case level (not docket level) variable so it appears only in the first record for the relevant justice. Therefore, the entry does not necessarily appear in the record in which the deviation occurs. In order that the user be able to know not only whether there was a deviation but how many, this entry reflects the number of legal provisions. This was coded as one deviation if they address a different law than the majority or address one less law than the majority or add one more law for consideration.

It is coded two if they add two laws, drop two laws, or changed one law and added or dropped another (A change is always counted as one deviation; e.g., the majority considered 14AD but this justice considered 5AD – we don't count a drop and an add here but rather one change). The same coding applies to the issue_dev variable. Auth_dev is largely the same as well, but because there is sometimes a secondary authority for decision this deserves additional comment. In general, when an authority is changed from say, 7 to 2, that is one deviation. If it is changed from 7 to 27, that is also one deviation. We do not count that as dropping 7 and then adding 2 and adding 7. Likewise, if they change authdec = 24 to just 7, that is only two deviations rather than three. Note also that sometimes a justice will not consider a given legal provision (law_dev=1) and in so doing will not address the issue associated with that legal provision. We deem this to be an issue_dev as well since, in ignoring the legal provision, they also ignore the issue. Surely they could address the issue even without the legal provision but do not and so we count this as a deviation as well. So do we count as deviant as to authority for decision a justice who does not address a given issue that is associated with a unique authdec for the majority? That justice is counted as having both an issue and an authdec deviation.

In addition to the mere numbers of deviation, we are also interested in identifying that deviation. Which issues are dropped? Which laws are added? In order to allow the user to explore the qualitative disagreements among the justices, we

code six additional variables. They are addiss1, addiss2, dropiss1, dropiss2, addauth1, addauth2, dropauth1, dropauth2, addlaw1, addlaw2, droplaw1, droplaw2 signifying the additional issues, the dropped issues, the additional authorities for decision, the ignored authorities for decision, the added legal provisions, and the legal provisions not considered, respectively. These variables map the disagreement of any given judge with the majority and therefore give a fuller description of the disagreement that exists on the Court. We could plausibly determine whether certain justices are more likely to add issues or more likely to drop issues; whether certain justices address specific legal provisions more often than others; and whether particular justices use a certain authority for their decisions disproportionately. Again, these entries are made only in the first record for each case and so are case level variables in order to determine valid frequencies of dissensus. Do note that a blank law field means that the justice used a law not coded for in the Spaeth dataset; 0 connotes that the justice did not use a legal provision in reaching his/her decision.

Finally, some of the variables in the existing database are recoded to correspond to individual justice vote. These include: legal provision, issue, and authority for decision (all discussed above), as well as disposition of case (reverse, affirm, remand, etc), directionality of vote (liberal or conservative), declaration of unconstitutionality, and formal alteration of precedent. This allows scholars to efficiently investigate such theoretically important matters as individual voting behavior, models predicting and explaining dissent or concurrence, and opinion assignment, answering some or all of the following questions (as posed in our grant proposal):

> What strategies do individual justices employ in relation to their colleagues? With what degree of individual success? Under what circumstances? Which case-specific variables affect the justices' behavior? Does coalition size affect behavior? Do freshmen justices dissent less often? How does their voting and opinion writing compare with that of the others? How has one individual justice behaved on all counts over the course of his or her tenure (for biographical purposes)?

Potential dependent variables from this database include:

> Justice's final vote
> Whether justice voted with minority or majority
> How justice disposed of the case
> Whether justice voted liberally or conservatively
> Justice(s) with whose opinion(s) the justice agreed
> Whether the justice wrote an opinion
> The legal provision(s) the justice addressed
> The issue(s) the justice addressed
> The basis for the justice's final vote
> Assigning Justice for majority opinion
> Assigned Justice for majority opinion

NSF SUPPORT

REFERENCES

Spaeth, Harold J. 2004. "Relisting: An Unexamined Feature of Supreme Court Decision Making." 25: *Justice System Journal* 143.
Spaeth, Harold J. and Jeffrey A. Segal. 2000. "The U.S. Supreme Court Judicial Data Base: Providing New Insights into the Court." 83: *Judicature* 228.