Analysis of the Supreme Court Database

AUTHOR
Allan Hegedus, Johns Hopkins University

Introduction

Created and maintained by Harold Spaeth (Michigan State University), Lee Epstein (Washington University of St. Louis), Ted Ruger (University of Pennsylvania), Jeffrey Segal (Stony Brook University), Andrew D. Martin (University of Michigan), and Sara Benesh (University of Wisconsin), the Supreme Court Database (SCDB) is a resource that details every case that has come before the Supreme Court. It has numerous pieces of information that would be useful to any law firm, including the number of votes in the affirmative/negative for each case and issue (depending on the version of the database used), the Court that voted on each issue, the issue area the vote was based on, the reason that a writ of certiorari was granted, and many other important variables to understand Supreme Court operations.[1]

Dataset Description

The SCDB allows for various breakdowns of the data, including by case, issue and Justice votes. For this analysis, I will utilize the dataset at the issue-level, organized by Supreme Court citation. This means that cases can and will appear multiple times if there were multiple legal provisions are ruled on by the Supreme Court. The dataset has 53 variables and 13,854 observations/issues ruled on.

Variables of Interest

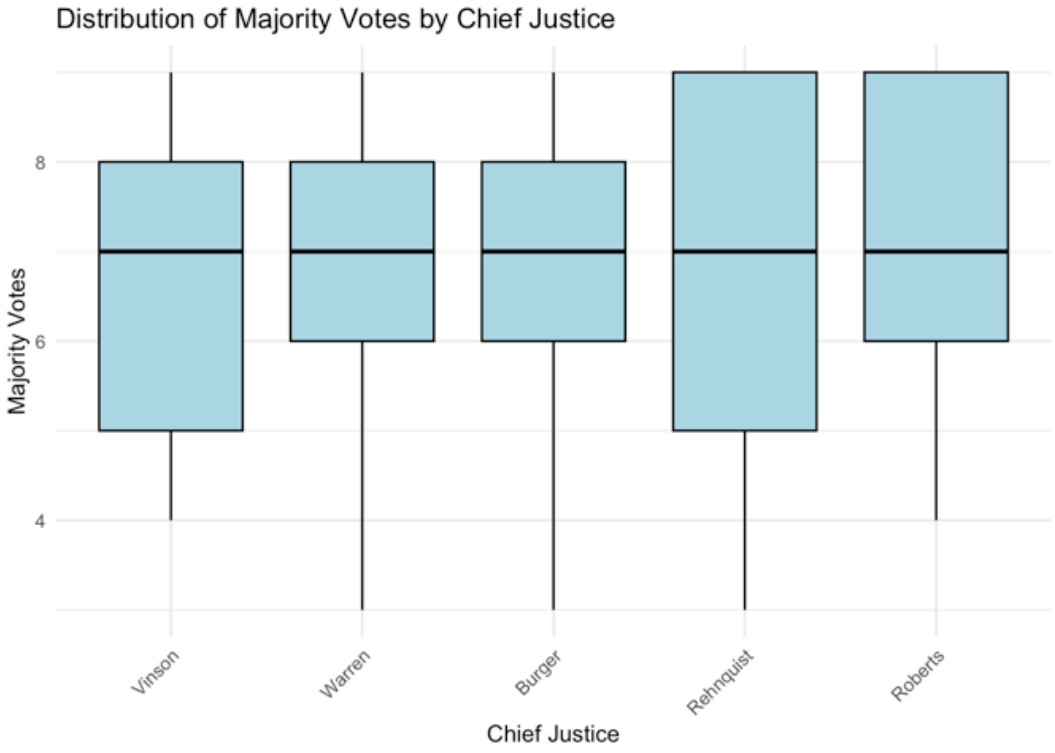
For this analysis, I will utilize the following key variables:

- Chief Justice (chief): The Chief Justice at the time of the case. This will serve as a natural divider between different eras of the court.

- Majority Votes (majVotes): The number of the majority and minority votes in the case/issue.
- Issue Area (issueArea): a descriptor of the subject area that the case/issue was detailing. This is important to group out the issues that each court was dealing with.
- Case Disposition (caseDisposition): The ultimate outcome of the case, including, but not limited to, being affirmed, reversed, and/or remanded.
- Decision direction (decisionDirection): Whether the decision goes in a more liberal or conservative direction (or is unknown).[2]

Analysis

We start by analyzing the distribution of majority votes over time for each Chief Justice. We find that while the median has always sat at 7 votes, the variance decreases and then increases some over time, as indicated by the interquartile range over the Chief Justices:

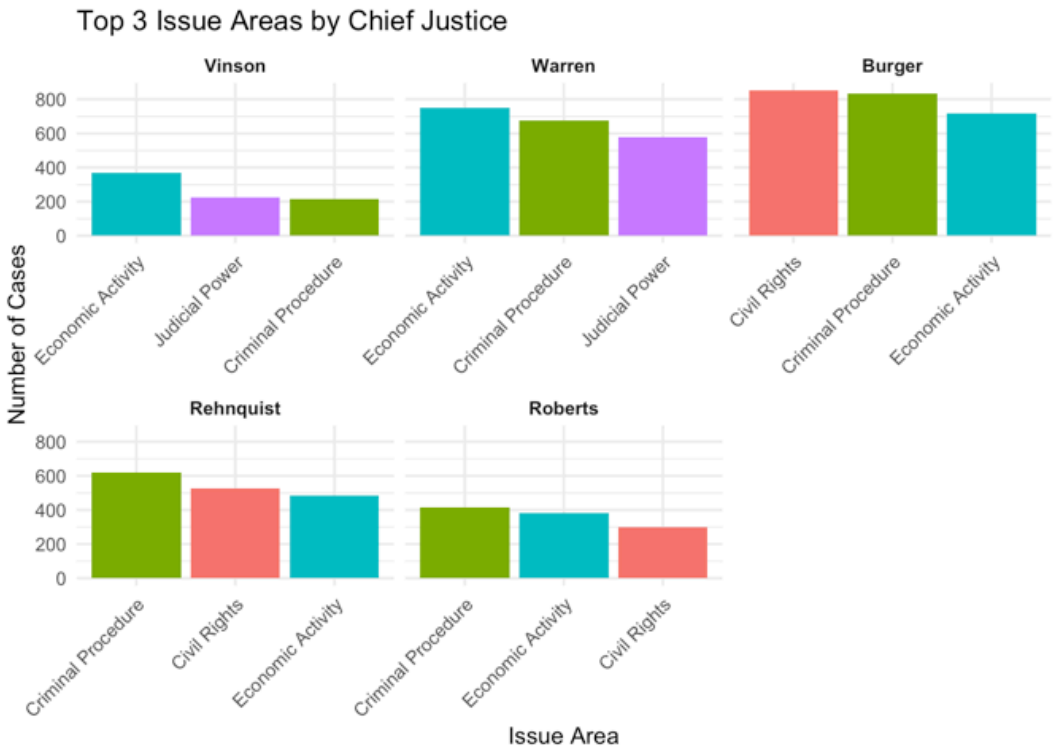


- Vinson: 3 votes

- Warren: 2 votes
- Burger: 2 votes
- Rehnquist: 4 votes
- Roberts: 3 votes

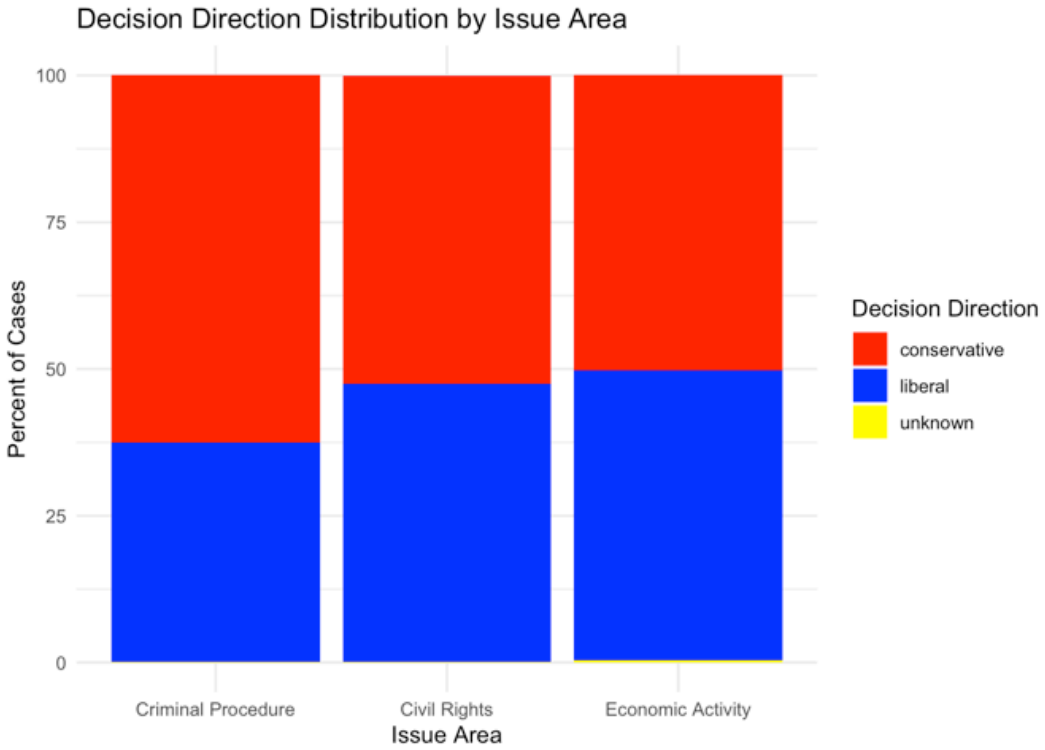
Since a majority would require 5 votes for the entirety of the Court's time since 1946, when this data started, it makes sense that the 25th percentile is no lower than 5, though unknown votes are likely the reason that the minimums do go below 5. While the lower 25th percentile would indicate closer votes, the Rehnquist and Roberts courts also have more 9-0 majority votes, as illustrated by the 75th percentile being at 9 votes.

In addition, we can see what each Chief Justice dealt with most while presiding over the Supreme Court. All 5 Chief Justices' courts feature Criminal Procedure and Economic Activity in their top 3 issues, while judicial power has ceased to be an area where cert is granted as often.



Given that the issues that appear most for the most recent Chief Justices are related to Criminal Procedure, Economic Activity, and Civil Rights, it is worth considering how those issues shake out on average. From the chart below, we see that Criminal Procedure matters tend to fall more on the conservative side, with 62.6% of Criminal Procedure cases since the Burger courts. The other issues have 52.5% (Civil Rights) and 50.3%

(Economic Activity) going in a conservative direction.[3] This is essential information for any law firm that is pursuing litigation in one of these areas, and to prepare clients for the outcome in the case (or potentially give them hope).



Finally, we should also consider the potential for remand in each of these topic areas, as that would be essential information to consider when building out a case timeline, budget, and preparing a client for what is to come after a Supreme Court case. For this part of the analysis, we will only consider the Roberts courts, as they are the most recent data we have on this subject, and how they rule will be the most directly relevant. We will consider all scenarios where the Supreme Court elects to remand, which include being ruling both for the plaintiffs and the defendants.

Issue Area	Number of Cases	Percentage of Cases	Standard Deviation
Attorneys	23	78.260870	0.4217412
Miscellaneous	20	75.000000	0.4442617
First Amendment	119	63.025210	0.4847775
Civil Rights	298	60.738255	0.4891543
Judicial Power	227	60.352423	0.4902464

Criminal Procedure	417	59.472422	0.4915351
Economic Activity	380	59.210526	0.4920913
Privacy	35	57.142857	0.5020964
Private Action	7	57.142857	0.5345225
Unions	46	56.521739	0.5012063
Due Process	50	54.000000	0.5034574
Federalism	85	48.235294	0.5026540
Federal Taxation	21	38.095238	0.4976134
NA	7	14.285714	0.3779645
Interstate Relations	17	5.882353	0.2425356

We see that most types of issues go back down on remand on average. Given their high sample sizes, we can be most confident that First Amendment, Civil Rights, Judicial Power, Criminal Procedure, and Economic Activity issues will go back on remand, though they only typically go down about 60% of the time still. Therefore, if the firm deals with these types of issues, they should be sure to prepare for more litigation if the case does make it to the Supreme Court. While this does not mean all legal questions will go back down, it seems that at least one could.

Limitations of the Data

While the data is remarkably robust, there are a few important notes to make about this data. The level of data selected can make analyzing certain factors would be difficult. For instance, had we chosen the case-level data, we would not have had the full votes on each issue that the issue-level data gives us. In the case[4] of the issue-level data, it does not give us the information to get information on the justice's individual votes, nor more case-level analyses. As such, the level of data we choose for future analyses must be considered. Great care must be taken for future analyses choosing the appropriate information for our clients.

Unfortunately, there are also a fair number of missing values in three ofour variables of interest, and given that these are categorical variables other than the majority/minority votes, there is not a straightforward way to deal with them. The table below shows the number of missing observations in our variables of interest, but there are more missing values in the other

variables of interest.

Percentage of Cases Sent Back to Lower Courts on Remand

Variable	Missing Values
caseDisposition	164
issueArea	70
decisionDirection	43

Finally, there are imputations made in our variables of interest that are worth noting. The majority and minority vote variables have difficulties in determining non-participation in certain instances. Blanks in caseDisposition can be missing in early cases if they are related to certification to or from a lower court or do not have a disposition. It may be necessary to use a more granular variable than issueArea in future analyses, as the generalization of the variable can leave out some nuance that would be helpful. Finally, there are also generalizations made under decisionDirection that must be taken with great care in future analyses.

Conclusion

In this analysis, we found that the Supreme Court tended to have more agreements overall despite the median affirmative vote staying consistent over each chief justice. We found that three main issues the Court consistently deals with in the modern era are Criminal Procedure, Civil Rights, and Economic Activity, and that these issues tend to have conservative findings (though, with the exception of criminal procedure, only marginally). Finally, we also find that many cases will end up going down on remand, which is essential to prepare a client for.

Link to GitHub repo:
https://github.com/AllanHegedus/jhu_dap_470.768_final.git

References

Spaeth, H. J., Epstein, L., et al. (2024). *The supreme court database*. http://scdb.wustl.edu

Footnotes

1. For disclosure, I have worked with the SCDB previously using its case-level data, though I tried to do something different for this project by using issues instead of cases. I do not reuse any code or analysis from that work, though I will come to some similar conclusions at certain points. To see that work, see https://www.hegedusanalytics.com/blog/new-project-using-supreme-court-data-to-predict-case-outcomes and https://www.hegedusanalytics.com/blog/using-logistic-regression-to-predict-supreme-court-cases. ↩
2. "Conservative" and "liberal" do not follow a traditional policy definition, and is dependent on the issue that the court is dealing with. For more details, see Spaeth et al. (2024). ↩
3. For additional details on each: - Liberal on a Criminal Procedure issue includes examples such as "pro-person accused or convicted of crime, or denied a jury trial" and "anti-government in the context of due process, except for takings clause cases where a pro-government, anti-owner vote is considered liberal except in criminal forfeiture cases or those where the taking is pro-business", amongst others, with the reverse being considered conservative. - Liberal on a Civil Rights issue includes the above examples, in addition to "pro-civil liberties or civil rights claimant, especially those exercising less protected civil rights (e.g., homosexuality)," amongst others, with the reverse being considered conservative. - Liberal on an Economic Activity matter includes examples such as being "pro-government," "anti-business," "anti-employer," "pro-competition", amongst others, with the reverse being considered conservative. For more details, see Spaeth et al. (2024). ↩
4. No pun intended. (Okay, maybe slightly intended). - Allan ↩