The Case Salience Index (CSI)
Documentation for Phase 2

July 10, 2014

Co-Principal Investigators:

Todd Collins
Associate Professor
Department of Political Science & Public Affairs
tcollins@email.wcu.edu

Christopher Cooper
Associate Professor
Department of Political Science & Public Affairs
ecoooper@email.wcu.edu

General Overview:

This brief codebook provides details on the data files (CSI Phase 2) in Excel, STATA, SPSS, R, and SAS formats. The data represent newspaper media coverage of Supreme Court decisions from all orally argued cases for four papers, Los Angeles Times, Chicago Tribune, Washington Post, and New York Times. Coverage dates include cases from the 1953 to the 2004 terms. Phase 2 data includes all cases from Phase 1 so that only Phase 2 needs to be downloaded. Each observation represents newspaper coverage for one case. This dataset includes 6258 observations (all orally argued cases with decisions over the time period).

Data Collection Process:

For each orally-argued case, multiple search terms were used (the names of the parties, the keywords of the case) on and around the day each case’s written decision was released. The search would begin with the parties’ names. If no articles appeared using the names, researchers would search for key terms appearing in the decision. For example, in a case such as Roe v. Wade, if articles did not appear based on the names, we would search for keywords such as “abortion,” “privacy,” “Texas,” “Dallas County,” and other terms that appeared in the syllabus of the United States Reports. Most cases were searched using ProQuest, a commercial search engine that provides PDF files of various newspapers across time. ProQuest searches were supplemented with searches on the individual paper web sites or other full-text databases, such as Lexis-Nexis when needed.

Variables:

Each case includes six fields:

A. “caseId”: This variable serves as the unique identifier for each decision and is taken from the Supreme Court Database (available at http://scdb.wustl.edu/data.php). This comprehensive dataset and updated version of the original “Spaeth Database” includes a wealth of information and case characteristics. Our data should be easily merged with the Supreme Court Database using caseId.
B. “CSI”: This is the Case Salience Index. It is the addition of all four of our paper scores and ranges from zero to eight. A value of zero (0) indicates that the decision is not reported anywhere in any of our four papers. A value of eight (8) indicates that the decision was covered on the front page of all four of the four papers. Values in between indicate various combinations of coverage.

C. Individual Paper Scores (“laScore,” “chScore,” “washScore,” “nyScore”): These variables indicate individual paper coverage of the decision in each of our four newspapers (the Los Angeles Times, Chicago Tribune, Washington Post, and New York Times, respectively):

- **Zero (0):** Indicates no coverage anywhere in the paper
- **One (1):** Indicates that the decision was covered somewhere in the paper, but not on the front page.
- **Two (2):** Indicates that the decision was covered on the front page of the paper.

Other Formatting or Problems:

If other problems arise with the files or further information is needed, please do not hesitate to contact us.

Acknowledgements:

We gratefully acknowledge the assistance of several entities at Western Carolina University, including the Hunter Library, the Graduate School, the Public Policy Institute, and the Department of Political Science and Public Affairs. Without the commitment from all of these groups, we would not have been able to complete this project. We also are indebted to the several graduate students that assisted with the data collection process.