Appendix: Supporting Information for "Racial Diversity and Judicial Influence on Appellate Courts"

Jonathan P. Kastellec

American Journal of Political Science

Case selection and coding procedures

Sunstein et al. assembled their sample of affirmative action cases by doing a Lexis keyword search of "affirmative action and constitution or constitutional," and through a Westlaw Key Cite search of *United Steelworkers of America, AFL-CIO-CLC v. Weber*, 443 U.S. 193 (1979), and *Regents of University of California v. Bakke*, 438 U.S 265 (1978). This resulted in 162 cases decided between 1980 and 2003. For each of these cases, I rechecked both that they were indeed affirmative action cases, and the coding of the judges and votes. To backdate and update the dataset, I performed the same Lexis keyword search, resulting in 27 cases decided between 1971 and 1980, and 15 cases decided after 2003. In addition, some of the original cases involved non-race-based affirmative action; I excluded all such cases.

For each case, I identified the three judges serving on the panel. I then read the case and double-checked the coding of each judge's vote, following the coding procedures set forth in Sunstein et al. (2006). I coded a decision as "conservative" if a judge voted to strike down any part of an affirmative action plan as unconstitutional, and "liberal" if a judge upheld the program in its entirety. In a few cases, the coding of votes did not correspond to the coding procedures outlined in Sunstein et al. (2006), and I corrected these votes. A small number of cases involved only gender-based affirmative action; these were not included in my analysis. I also dropped any cases that were directly on remand from the Supreme Court, since these are qualitatively different from initial decisions by three-judge panels. For each case, I also coded the direction of the lower court or agency's decision, using the same coding protocols.

Note that in this calculation and the discussion of dissents, for the purpose of reliability,

dissents were coded with respect to the overall coding scheme, not simply based on whether a judge dissented in a case. For example, if a majority of the panel ruled unconstitutional part of an affirmative action program, and a judge dissented because she would have overturned the entire program, all three judges were coded as voting conservatively, and thus no judge was coded as dissenting.

Information on the race and gender of each judge was gathered from various sources. To conduct the actual analyses at both the case level and judge-level, I merged the affirmative action datasets with the appeals court judges attribute database (Gryski and Zuk 2008) and the district court judges attribute database (Gryski, Zuk and Goldman 2008) (for district court judges sitting by designation), both of which include judges appointed through 2004. In a few cases where non-Article III judges sat by designation, or for Article III judges appointed after 2004, I obtained their biographical information, including race and gender, from *History of the Federal Judiciary* (2011). Information on each judge's appointing president, party of the appointing president, home state and year of appointment was also gathered from these sources.

The measure of judicial ideology used in the regression analyses—*judge's conservatism*—are the scores created by Giles, Hettinger and Peppers (2001). They involve using the common space scores of the appointing president and/or a nominee's home state senators (Poole 1998). The procedure is the same for all appeals court judges and district court judges. The first step is to determine whether senatorial courtesy is in effect. Following Giles, Hettinger and Peppers (2001), I assume that senatorial courtesy exists whenever one senator from a nominee's home state is of the same party as the president. If one (and only one) senator is of the same party, then the GHP score takes on that senator's Common Space score. If both senators are of the home state party, the GHP score is average of their common space scores. If neither senator is of the president's party, the GHP score takes on the president's common space score. I assume that senatorial courtesy is not in effect for

judges appointed to the D.C. Circuit, judges who come from U.S. territories, all non-Article III judges. Thus, for these judges, their GHP scores is the common space score of their appointing president.

For each judge I coded their appointing president's common space score, the common space scores of the judge's home state senator, and whether senatorial courtesy was in effect during the judge's nomination. In some cases, more than two senators served during the Congress in which a nominee was appointed. Using the "Biographical Directory of the United States Congress," I determined which two senators were in office at the time of the judge's nomination.[1] I then created GHP scores using the above criteria.

Robustness check: sensitivity analysis with Rosenbaum bounds

While matching provides the twin benefits of increasing model transparency and reducing model dependence in drawing inferences, it does nothing to mitigate the potential problem of omitted variable bias. Fortunately, Rosenbaum (2002 2002, ch.4) has devised a method for assessing how sensitive estimated average causal effects are to potential "hidden bias." Such bias can take two forms: 1) if two observations with the same observed covariates nevertheless differed in their probability of receiving the treatment (various assignment probabilities); 2) an unobserved covariate that might predict the outcome but that was not controlled for (omitted variable bias).

While Rosenbaum demonstrates that the two types of bias are mathematically equivalent, because of the quasi-random assignment to panels on the Courts of Appeals, it makes more sense to conceptualize the potential for hidden bias as arising from an unobserved covariate. The most likely candidate might be the true differences in judicial ideology between non-black and black judges, as the existing measures might not sufficiently capture these differences in ideology. If black judges are sufficiently more liberal than non-black judges, then the race-based effects uncovered in the analysis could actually be due to ideological differences.

---

[1]This can be accessed at `http://bioguide.congress.gov/biosearch/biosearch.asp`.

Thus, we can think of the hidden bias as the extent to which black judges would have to be more liberal than non-black judges (beyond what is measured) to confound the inferences.

The analysis proceeds by defining a sensitivity parameter Γ, which one can think of the size of the effect on an unobserved covariate u, which is highly predictive of the outcome (Keele 2011, 8-9). If $\Gamma = 1$, then no hidden bias exists. As Γ increases, so does the level of hidden bias. Of course, the true Γ is unknown; the sensitivity analysis proceeds by asking how high Γ would have to be before our conclusions about the significance of our estimated average treatment effects would change. The maximum Γ at which our conclusions would hold provide bounds on the estimate (and are hence known as "Rosenbaum bounds"). (See Rosenbaum (2002, ch. 4) for technical details.)

I calculated Rosenbaum bounds for the estimated counter-judge effects based on matched data (models (5) and (6)).[2] The analysis reveals that for the nearest-neighbor and optimal matching, Γ would have to be as large as 1.9 and 2.0, respectively, before we would conclude that the estimated counter-judge effects were not statistically different from zero (based on a one-tailed test). For this to occur, an unobserved variable would have to be both more than double the probability of being assigned to a panel with a black judge, and nearly perfectly predict liberal voting in affirmative action cases. Both the quasi-random assignment to panels on the Courts of Appeals and the theoretical implausibility of a covariate (besides race) predicting voting so well would seem to make such a scenario highly unlikely. Alternatively, black Democratic judges would have to be more than twice as liberal as the average non-black Democratic judge, which also is implausible. Thus, we can conclude that the estimated counter-judge effects are *not* sensitive to hidden bias.

---

[2]To estimate the bounds, I used the `binarysens` function in the `Rbounds` package in `R` (Keele 2011). The function implements McNemar's test statistic, which requires matched pairs—i.e. the same number of observations in the treatment and control groups after matching. Thus, for the nearest neighbor matching, I matched *without* replacement. I confirmed that the estimated average counter-judge effect was statistically and substantively similar to the results based on nearest-neighbor matching with replacement.

References

Giles, Micheal W., Virginia A. Hettinger and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54:623–41.

Gryski, Gerald S. and Gary Zuk. 2008. "A Multi-User Data Base on the Attributes of U.S. Appeals Court Judges, 1801-2000." available at `http://www.cas.sc.edu/poli/juri/auburndata.htm`, accessed 15 August 2008.

Gryski, Gerald S., Gary Zuk and Sheldon Goldman. 2008. "A Multi-User Data Base on the Attributes of U.S. District Court Judges, 1789-2000." available at `http://www.cas.sc.edu/poli/juri/auburndata.htm`, accessed 15 August 2008.

History of the Federal Judiciary. 2011. Website of the Federal Judicial Center. Washington, DC. available at `http://www.uscourts.gov/JudgesAndJudgeships/BiographicalDirectoryOfJudges.aspx`.

Keele, Luke J. 2011. "`Rbounds`: An `R` package for Rosenbaum bounds sensitiviy analysis with matched data." Ohio State University typescript.

Poole, Keith T. 1998. "Recovering a Basic Space from a Set of Issue Scales." *American Journal of Political Science* 42:954–93.

Rosenbaum, Paul R. 2002. *Observational Studies.* New York: Springer.

Sunstein, Cass R., David Schkade, Lisa M. Ellman and Andres Sawicki. 2006. *Are Judges Political? An Empirical Analysis of the Federal Judiciary.* Washington, D.C.: Brookings Institution Press.