Correction for "Does Legal Doctrine Matter? Unpacking Law and Policy Preferences on the U.S. Supreme Court"

Michael A. Bailey, Georgetown University
Forrest Maltzman, George Washington University

A coding error affected the estimates of law. We have corrected the error and conducted extensive diagnostic tests to ensure that the new code is correct. Details on the problem, the solution and the diagnostic testing are available at http://www9.georgetown.edu/faculty/baileyma/Data/APSR_Correction_Feb2010.pdf.

Table 1 presents the revised results. Figure 1 presents the simulated effects of the variables based on the parameter estimates in Table 1. These correspond to Table 1 and Figure 5 in the original paper. For the precedent parameter, five justices have 90% or more of the posterior distribution of the parameter greater than 0. For the deference to Congress parameter, eight justices have 90% or more of the posterior distribution of the parameter greater than 0 (which corresponds to a p-value less than 0.10). For the speech parameter, nine justices have 90% or more of the posterior distribution of the parameter greater than 0.[1]

We also test the joint significance of the legal parameters. In this test, the null hypothesis is that all legal parameters equal zero ($\pi_i = \delta_i = \sigma_i = 0$). We can reject this null hypothesis at the 0.05 level for all 16 justices and at the 0.01 level for 15 of 16 justices. These results are reported in the column on the right.

[1] The precedent variable for Stewart is estimated very imprecisely because there are only four Stewart votes for which precedent is not equal to zero (all other justices have 40 or more observations for which the precedent variable is not equal to zero).

Table 1: Raw Estimates of Effect of Legal Variables

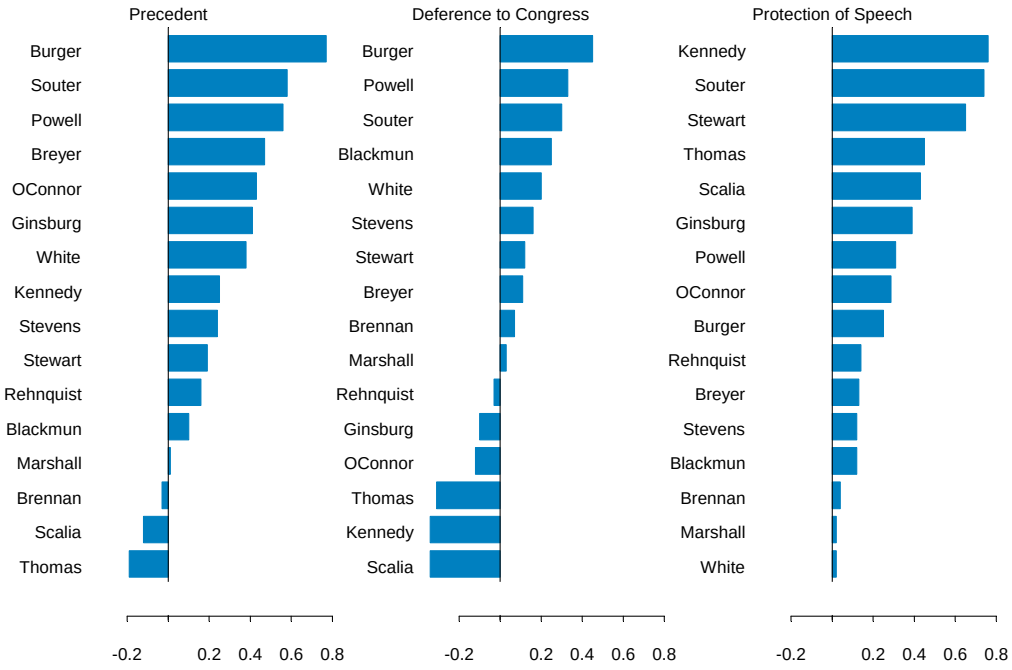
Justice	Legal measure			Joint significance
	Precedent (π)	Congress (δ)	Speech (σ)	
Blackmun	0.17	0.41**	0.19	***
	p = 0.23	*p = 0.03*	*p = 0.12*	*p = -0.002*
Brennan	-0.20	0.49*	0.27	**
	p = 0.65	*p = 0.07*	*p = 0.18*	*p = 0.01*
Breyer	0.84***	0.19	0.22	***
	p = 0.00	*p = 0.29*	*p = 0.24*	*p = -0.005*
Burger	1.51***	0.77***	0.41**	***
	p = 0.00	*p = 0.01*	*p = 0.05*	*p = -0.005*
Ginsburg	0.85***	-0.20	0.8***	***
	p = 0.00	*p = 0.73*	*p = 0.00*	*p = -0.005*
Kennedy	0.39**	-0.54	1.44***	***
	p = 0.02	*p = 0.98*	*p = 0.00*	*p = -0.005*
Marshall	0.10	0.45*	0.27	**
	p = 0.40	*p = 0.10*	*p = 0.19*	*p = 0.008*
O'Connor	0.64***	-0.17	0.41***	***
	p = 0.00	*p = 0.78*	*p = 0.00*	*p = -0.005*
Powell	0.86***	0.47**	0.43**	***
	p = 0.00	*p = 0.05*	*p = 0.02*	*p = -0.005*
Rehnquist	0.43**	-0.09	0.37**	***
	p = 0.02	*p = 0.65*	*p = 0.03*	*p = -0.004*
Scalia	-0.27	-0.78	0.98***	***
	p = 0.90	*p = 0.99*	*p = 0.00*	*p = -0.005*
Souter	0.99***	0.48*	1.38***	***
	p = 0.00	*p = 0.06*	*p = 0.00*	*p = -0.005*
Stevens	0.51***	0.34**	0.26**	***
	p = 0.00	*p = 0.04*	*p = 0.05*	*p = -0.005*
Stewart	0.26	0.16	1.01***	***
	p = 0.34	*p = 0.32*	*p = 0.00*	*p = -0.005*
Thomas	-0.53	-0.85	1.2***	***
	p = 0.99	*p = 0.98*	*p = 0.00*	*p = -0.005*
White	0.54***	0.28*	0.03	***
	p = 0.00	*p = 0.10*	*p = 0.43*	*p = -0.004*

*** p < 0.01, ** p < 0.05, * p < 0.10

Bayesian analog to p-value in italics below point estimate indicates proportion of posterior distrtibution less than zero.

The test of joint significance tests the null hypothesis for each individiual justice that all three legal parameters are zero; a value close to zero indicates that one or more parameters was above zero in almost all draws from the posterior distribution.

Figure 5: Effect of Legal Variables on Individual Justices



The numbers on the x-axes are the average percentage point change in the probability of voting conservatively associated with the various legal concepts. The number for each justice and each legal concept is calculated via simulations based on parameter estimates in Table 1. For each simulation, we calculated the average difference in the probability of a conservative vote when the legal concept (e.g. precedent) implies a conservative and a liberal vote. For each concept, we conducted simulations based on cutpoint values of κ=-0.5; κ=0.0; and κ=0.5.

Erratum On-line Appendix

1. Nature of the problem

The model described in the paper was estimated using Matlab. In some of the steps in which the individual-level legal parameters were to be sampled from the posterior distribution (see Equation 6 in the on-line appendix) what was supposed to be a draw from a normal distribution ("randn" in Matlab syntax) was inadvertently set to be a draw from a uniform distribution ("rand" in Matlab syntax). Using a uniform instead of a normal distribution here creates an upward bias for coefficient estimates.

2. Correction

We have corrected the code and conducted a number of tests in order to ensure that the Matlab code follows the procedures described in the paper and related literature.

3. Diagnostic simulations

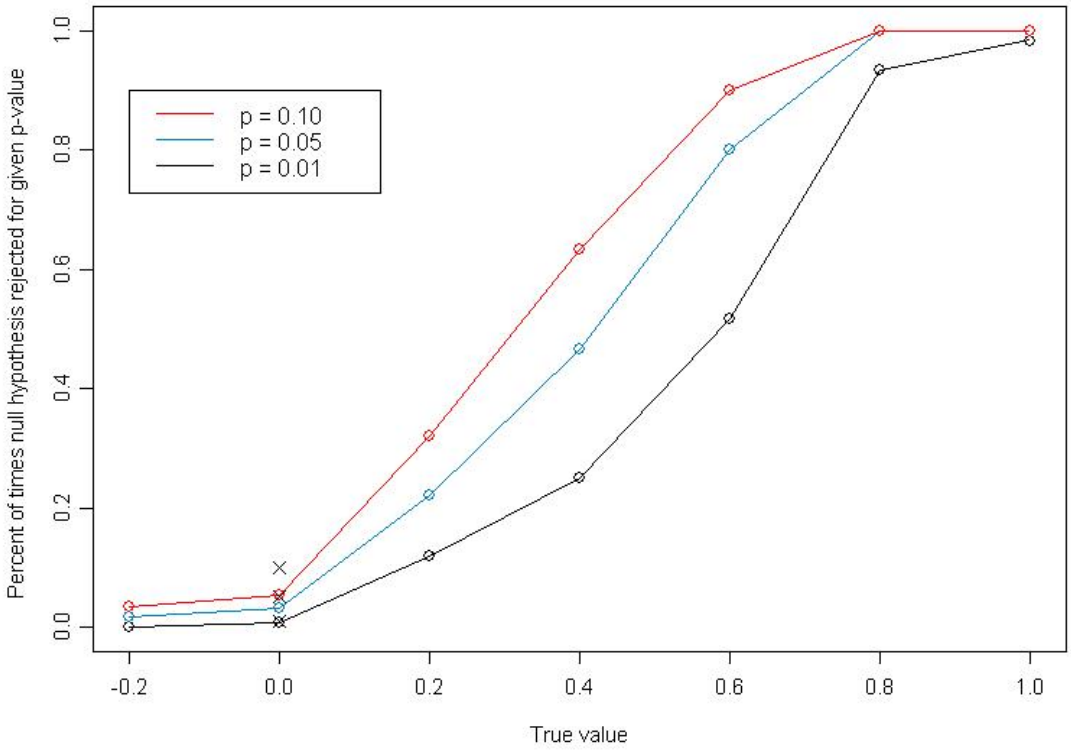
The most important diagnostic is a simulation in which ideal points, legal effects, cutpoints and error terms are simulated for the actors and votes in the data. Based on these known quantities, votes are simulated. Using the revised Matlab code, we then estimated the parameters based on this simulated data and compared them to the known simulated values. We did this eight times, saving the estimated parameter values and comparing them to the "true" values.

The results indicate decent accuracy and no tendency to falsely reject the null of no effects. The correlation between the average estimate and true parameter value is 0.98 for deference, 0.89 for precedent (and 0.98 when excluding Stewart who, as mentioned above has virtually no observations for which precedent is not zero and is therefore extremely imprecisely estimated) and 0.96 for speech.

Figure A-1 plots the statistical power of the simulations. The figure plots the proportion of times the null hypothesis of no effect was rejected across the various values of the known parameter values.[2] It does this for three conventional significance levels: 0.01, 0.05 and 0.10. Most importantly, there is no sign of positive bias. For true parameter values at or below zero, the null hypotheses are seldom rejected. For true parameter values above 1, the nulls of no effect are almost always rejected for p-values of 0.05 and 0.1 and usually rejected for p-values of 0.01. For true parameter values between 0 and 1, the statistical power rises with the true parameter values as it should. We should note, however, that Type II errors (failure to reject the null of no effect when in fact there is an effect) are fairly common at these levels. If anything, this indicates low statistical power that biases *against* finding significant results even when there really are legal effects.

[2] We are using frequentist hypothesis testing language to describe the Bayesian results here.

Figure A-1: Statistical Power in Simulation



The proportion of time we could reject null hypothesis that parameter equals zero for three different p-values. The x-axis is the true value.

4. Additional results

In addition, Chapter Three of our book manuscript, *The Constrained Court: Law, Politics and the Limits of the Attitudinal Model* conducts an analyses of a broader time period using the method described in the paper and the corrected code. We find that legal factors affect the decision-making of Supreme Court justices.