Appendix

A Statistical Details

Here we describe the latent variable specification in which deference to some non-ideological factors can also affect justice behavior. Generalizing the model to include multiple non-ideological factors is straightforward. Let i = 1,...,N index individuals and v = 1,...,V index votes. The utility of actor i of voting for the conservative alternative is

\[ u_i(\lambda^C) = -(\theta_{it} - \lambda^C_v)^2 + \delta_i \hat{D}^C_v + \eta^C_{iv} \]  

where \( \lambda^C_v \) is the spatial location of the conservative alternative, \( \theta_{it} \) is the ideal point of the actor at the time of proposal t, \( \hat{D}^C_v \) is the non-policy value of voting for the conservative alternative, \( \delta_i \) is the weight placed by i on non-policy values and \( \eta^C_{iv} \) is a random shock. The utility of voting for the liberal alternative with spatial location of \( \lambda^L_v \) is analogous.

Let \( y^*_itv \) be the utility difference between the conservative and liberal alternatives. It is

\[ y^*_itv = -(\theta_{it} - \lambda^C_v)^2 + \delta_i \hat{D}^C_v + \eta^C_{iv} + (\theta_{it} - \lambda^L_v)^2 + \delta_i \hat{D}^L_v + \eta^L_{iv} \]

\[ = 2\theta_{it}(\lambda^C_v - \lambda^L_v) + \lambda^L_v^2 - \lambda^C_v^2 + \delta_i(\hat{D}^C_v - \hat{D}^L_v) + \eta^C_{iv} - \eta^L_{iv} \]

\[ = (\lambda^C_v - \lambda^L_v)(2\theta_{it} - (\lambda^L_v + \lambda^C_v)) + \delta_i(\hat{D}^C_v - \hat{D}^L_v) + \eta^C_{iv} - \eta^L_{iv} \]  

(2)
Let $\kappa_v = \frac{\lambda^L_v + \lambda^C_v}{2}$ be the vote “cutpoint,” $\alpha_v = 2(\lambda^C_v - \lambda^L_v)$ be the vote “discrimination parameter,” $D_v = (\hat{D}^C_v - \hat{D}^L_v)$ be an observed non-ideological deference variable and $\epsilon_{iv} = \eta_{iv}^C - \eta_{iv}^L$ be a $\mathcal{N}(0, 1)$ random variable; then

$$y_{itv}^* = \alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v + \epsilon_{iv} \quad (3)$$

Observed votes (as opposed to unobserved latent values above) are denoted by $y_{itv}$. To address rotational identification (e.g. liberals can have high values or low values) conservative votes are coded as $y_{itv} = 1$. The location and scale of ideal points is identified by assuming they have mean 0 and variance 1; this is equivalent to fixing two individuals at arbitrary points (see, e.g., Bafumi, Gelman, Park and Kaplan 2005).

The estimation process uses a Gibbs sampler algorithm. This algorithm allows us to draw samples from the posterior distribution of the parameters (Gelman, Carlin, Stern and Rubin 1995, 326; see also Johnson and Albert 1999, 194-197). After a “burn in” period, the following iterative procedure will produce random samples from the underlying posterior distribution.

1. Equation 3 implies that $y_{itv}^*$ (where $i$ indicates individual, $t$ indicates term and $v$ indicates vote) will be distributed according to one of the two truncated distributions (see

\footnote{This parameter is standard in ideal point estimation theory and its precursor, item response theory (Baker 1992). Votes for which the alternatives are relatively close (meaning $(\lambda^C_v - \lambda^L_v)$ is relatively small) will have a low discrimination parameter as the non-spatial error term will be more likely to induce actors with preferences higher than the cutpoint to vote liberally and vice versa.}
e.g. Jackman 2000, 311)

\[ y^*_{itv} | y_{itv} = 1 \sim N(\alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v, 1) I(y^*_{itv} > 0) \] (4)

\[ y^*_{itv} | y_{itv} = 0 \sim N(\alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v, 1) I(y^*_{itv} \leq 0) \] (5)

where I is an indicator function that serves to truncate distributions above or below zero.

2. Generate individual-specific preference parameters on an individual-by-individual basis.

Let \( \theta_{it} = T_{it}^\rho_i \) and substitute the equation for the time-path of policy preferences into

Equation 3 yields

\[ y^*_{itv} + \alpha_v \kappa_v = \sum_{p=0}^{3} \alpha_v \rho_p T_{it}^p + \sum_{m=1}^{M} \delta_{im} D_{mv} + \epsilon_{itv} \]

\[ = X_i^\prime \gamma_i + \epsilon_{itv} \]

where \( X_i \) is a \( V_i \times M \) matrix of covariates for individual i (based on M non-policy variables) and \( V_i \) the number of observations for individual i. The first column of \( X_i \) is a column of \( \alpha_v \) for the votes for individual i. The second column of \( X_i \) is \( \alpha_v \) multiplied by the time variable for individual i for each vote and so on for the third through fifth columns. The last three columns of \( X \) are the deference variables for each of the votes for individual i. For Chapter 3 these are deference to precedent, Congress and speech. For Chapter 4 these are precedent and separation of powers. For Chapter 5 these are the three Solicitor General variables. This is only relevant for justices as the \( \delta \) parameters for non-justices are constrained to zero.
The distribution of $\gamma$ is therefore

$$\gamma_i \sim N((X_i'X_i)^{-1}X_i'y_i, (X_i'X_i)^{-1})$$  \hspace{1cm} (6)

where $\tilde{y} = y_{i,v}^* + \alpha_v \kappa_v$. A $N(0, \Omega)$ prior on $\gamma$ identifies the preferences of individuals who vote conservatively or liberally all the time. Without this prior, their estimated ideal points could become unbounded. The implementation of the prior follows Gelman, Carlin, Stern and Rubin (1995, 260). The coefficients on the higher order elements of time (e.g. the coefficient on $T^3$) are restricted to 0 for individuals who served relatively short periods of time. Specifically, $\rho_3 = \rho_4 = 0$ for all individuals who served 24 or fewer years, $\rho_2 = 0$ for all individuals who served 16 or fewer years and $\rho_1 = 0$ for all individuals who served eight or fewer years.

3. Generate $\alpha, \alpha \kappa$ on a vote-by-vote basis. If we let $\beta_v = [\alpha_v, \alpha_v \kappa_v]'$ and $\Theta_{it} = [\theta_{it}, -1]$ (indicating the preference parameter of individual $i$ for vote $v$ which occurred during term $t$) we can re-write Equation 3 as

$$y_{i,v}^* - \delta_i D_v = \Theta_v \beta_v + \epsilon_{iv}.$$  \hspace{1cm} (7)

By standard GLS results,

$$\beta_v \sim N((\Theta_v'\Theta_v)^{-1}\Theta_v'y_v^{**}, (\Theta_v'\Theta_v)^{-1})$$

where $y_v^{**} = y_{i,v}^* - \delta_i D_v$ for all individuals who voted on vote $v$, $\Theta_v$ is a $N_v \times 2$ matrix of $\Theta_{j,it}$ and $N_v$ is the number of votes cast on vote $v$. 
The discrimination parameter is, in part, a measure of vote-specific variance and, as a variance parameter is subject to becoming unbounded as discussed above (see also Baker 1992, 97-98; Mislevy and Bock 1990, 8). Therefore there are normal priors and maximum values for \( \alpha \) and \( \kappa \); the priors follow Gelman, Carlin, Stern and Rubin (1995, 254, 260); see also Johnson and Albert (1999, 192).

A model is unidentified “if the same likelihood function is obtained for more than one choice of the model parameters” (Gelman, Carlin, Stern and Rubin 1995, 422). For fixed-preference one-dimensional models, identification can be achieved by fixing the polarity (meaning, for example, conservative preferences are high values and liberal preferences are low values) and two observations (which is equivalent to setting the mean \( \theta = 0 \) and variance of \( \theta = 1 \)) (see discussions in Clinton, Jackman and Rivers (2004, 356) and Bafumi, Gelman, Park and Kaplan (2005)).

B Data

B.1 Sources for bridge observations

The data for 1950 through 2002 are from Bailey (2007). We have updated the data through 2008 following the same strategies.

- Amicus briefs. Filings by the Solicitor General are from Gibson (1997) for the period 1953 through 1987 and from Lexis-Nexis Academic Universe and the Solicitor General’s website thereafter. Amicus filings for members of Congress are identified in Epstein,
Segal, Spaeth and Walker (2007) and from Lexis-Nexis Academic Universe. Only amicus filings on merit are included.

- Comments by members of Congress. These were primarily taken from the *Congressional Record*. For 1989 to present, the Thomas.gov database was searched for entries with “Supreme Court.” For years before that every entry under “Supreme Court” in the annual indices was researched. Some observations were found in other sources such as Eskridge (1991) and the *Congressional Almanac*.

- Roll call votes. One must be careful when using roll call votes to ascertain members’ of Congress positions on Supreme Court cases. First, provisions that address court cases are often embedded in broader legislation. This makes it impossible to know if the vote indicates an opinion on the court case or some other matter. An example is *Denver Area Educational Telecommunications Consortium v. Federal Communications Commission* (518 U.S. 727) (1996) which struck some elements and upheld other elements of the Cable Television Consumer Protection and Competition Act of 1992. This act was passed over the veto of President Bush with nearly universal support of Democrats and substantial support of Republicans (although 85 of the 114 votes against it in the House on October 5, 1992 came from Republicans). The court ruled only on one small part of the bill, a part that put various restraints on cable operators in the interest of controlling “indecent” programming. Using a vote on the overall bill as an indicator
of congressional positions on the issue addressed by the Supreme Court would not be reasonable. However, it turns out that the Court explicitly addressed Sections 10 (a) and (b) of the law (upholding the first and striking the second) and that these were added in an amendment by Sen. Helms (R, NC) that passed 95-0. We use the vote on the amendment, but not a vote on passage. Section 10(c) of the law was also explicitly addressed by the court. There was no roll call vote on this, but the legislative history reveals that Sen. Fowler (D, GA) and Sen. Wirth (D, CO) sponsored this language, meaning that the position of these two on this section is clear.

- Cosponsorship. We used Thomas.gov to search for bills that had Supreme Court in their text and assessed whether these bills directly related to a Supreme Court case. Data is less comprehensive for periods predating Thomas.gov coverage.

B.2 Selection of Supreme Court cases

We use the Spaeth database and limit cases to those VALUE < 6 (criminal procedure, civil rights, First Amendment, due process and privacy). Citations are the unit of analysis (ANALU =0 in Spaeth’s data set) and add split-vote decisions (ANALU = 4) when there are bridging observations. Bakke is a prominent example of a case with a split votes and many members of Congress taking positions on one or the other (or both) of the main holdings. We do not include memorandum cases and decrees (DEC TYPE = 3 or 4).

Selected cases are those for which at least one of the following is true: discussed directly
in the *Harvard Law Review’s* annual court review, included as a landmark case in the Legal Information Institute’s database of cases (see supct.law.cornell.edu/supct/cases/name.htm), coded as a salient case in Epstein and Segal (2000), included in the CQ’s key cases list, a president or member of Congress or non-contemporaneous justice took a position on the case, the case has clear cutpoint relation to another case, the case implicates precedent, deference or speech as coded.

There are some instances in which we do not use Spaeth’s coding of the liberal/conservative directionality of a decision because it does not comport with the underlying politics of the case. This occurs in a number of campaign finance cases. For example, Spaeth codes *Buckley*’s decision to strike a limit on campaign expenditures as a conservative decision, when it is clear by the coalition on the court and in Congress that expenditure limits were a liberal reform targeting wealthy contributors.

### B.3 Coding of precedent

To identify those cases where precedent was in play, we relied upon Segal and Howard (2001) for the 1984-1995 period. For the other time periods, we relied upon a three stage process. First, we identified phrases or words associated with overturning precedent based on reading the cases identified in Segal and Howard. Second, we searched for all such phrases in petitioner and respondent briefs in the appropriate times. Third, we read and manually coded each identified case. See the section on alternative specifications for results based on
other coding approaches for precedent.

B.4 Deferece to Congress

Our statistical approach identifies the effect of legal ideas of deference by looking for differences in Supreme Court behavior relative to congressional behavior. It is possible that members of Congress may also share the concept of legislative deference; that is, it is possible that members who vote against a law would want to the Supreme Court to uphold the law (despite their personal opposition to it) on the grounds that the issue is one for Congress, not the Courts to decide. As discussed in the paper, in this case, our approach is estimating any additional influence the logic of judicial restraint may have on justices. We suspect this issue may not arise very often. It is reasonable to expect, however, that members of Congress who oppose a law would like to see the Supreme Court overturn it. Senator McConnell (R, KY) is a case in point. One could imagine that he would fight for the court to uphold the McCain-Feingold campaign finance bill that he vigorously opposed in the Senate on the grounds that the Court should defer to Congress. In fact, however, he argued for the same substantive outcome before the court as he did in the Senate, opposing the legislation in both venues.
Figure 7: **Supreme Court cases and Congressional roll calls by type**
B.5  Additional data description
B.5.1  Case and roll call data

Figure 7 displays information about the number and type of congressional roll calls and Supreme Cases in the database. There are 1,536 roll calls in the House and Senate and 3,239 cases in the Supreme Court.

Key to the method is use of actors not subject to the legal principle in question to pin down the policy cutpoints of Supreme Court cases. We have 1,039 Supreme Court cases with such observations; 908 of them have direct positions taken by members of Congress or presidents and 288 of them have positions taken by non-contemporaneous justices (there are some cases with both).

There are 399 cases in which precedent was non-zero for at least some observations; 227 had conservative codes and 193 had liberal outcomes. There are 232 cases in which deference to Congress was coded as implying either a conservative (193 times) or liberal (39 times) vote on the court. There are 509 cases in which deference to speech was coded as implying either a conservative (26 times) or liberal (483 times) vote on the court.

While the precedent cases are divided reasonably equally across cases that support liberal and conservative outcomes, the deference to Congress cases tend to imply conservative outcomes and the speech cases tend to imply liberal outcomes. We do not think believe these

---

8 The reason that the number of liberal precedent cases plus the number of conservative precedent cases do not sum to the number of total precedent cases is that there are a couple of cases for which the precedent variable took on different values depending on the time of the observation. For example, when Rust v. Sullivan was considered in 1991, the conservative side was seeking to overturn precedent, meaning the precedent variable was coded as -1. But when Justice Breyer took a position on the Rust v. Sullivan in 1995, the precedent was conservative (since the court decided in the conservative direction in 1991); hence, the value of precedent for the Breyer observation is 1.
distributions exert undue influence on our results. First, note we are not simply looking for
justices to be more or less liberal on these cases, but to be moved in the direction law implies
conditional on ideology (and, therefore, conditional on what co-policy ideologues in Congress
want). Second, our results work across three types of legal variables in a manner that defies
simple categorization in terms of the model coding of the legal variable: some (but clearly
not all) conservatives justices are estimated to have low regard for precedent and deference
even as the precedent cases tend to be liberal and deference cases tend to be conservative.

B.5.2 Bridging observations

There are 17,882 bridge observations on Supreme Court cases; 16,699 of these are by mem-
bers of Congress and presidents and 1,213 are by justices. These are non-voting observations
of individuals taking positions on Supreme Court cases as described in the text and above.

Figure 8 summarizes this data with respect to their support for the majority opinion
on the court case and characteristics of the commenter and case. The top bar shows that
the positions are reasonably evenly distributed across comments that support the court’s
majority opinion and those that oppose it.

Because the comments on cases with legal implications that we have coded are partic-
ularly interesting, we break down these subsets of congressional/presidential comments in
the figure. The comments on cases that we have coded as having clear legal implications
exhibit some more variation with regard to supporting the court majority. Even though the
distribution of comments across our legal coding categories and support and opposition of
Figure 8: Bridge observations on Supreme Court cases by support for Court majority.
the court majority, the fact that we find similar results across these various configurations of observations indicates to us that it is unlikely that the actual configuration of comments is biasing things one way or the other.

Figure 9 summarizes this data with respect to whether the comment supported the liberal or conservative side.

**Distributions of ideal points**  Figure 10 plots the distribution of ideal points for all members of Congress and those who provided either direct or implicit comments. The direct comment data has a skew toward the extremes as liberals and conservatives are more likely to pro-
provide direct comments. The implicit comment has a skew to the right indicating that those providing comment data tended to be more conservative.

While skew one way or the other does not indicate bias, it does affect the efficiency of the estimation. In the limit, a distribution of all conservatives or all liberals would not allow us to make meaningful distinctions about locations of cutpoints, one of the central pieces of this research design. There are two ways to increase efficiency: increase data or increase the
representativeness of the data. The second option is not under our control, as the data is what it is.