SUPPLEMENTAL APPENDIX FOR
Measuring Court Preferences, 1950 - 2011: Agendas, Polarity and Heterogeneity

December, 2012

Statistical Details

In the general form of the latent variable model, justices' votes are a function of unobserved ideology and non-ideological factors related to observed characteristics of the votes. The utility of actor i at time t of voting for the conservative alternative on vote v is

$$u_i(\lambda_v^C) = -(\theta_{it} - \lambda_v^C)^2 + \delta_i \hat{D}_v^C + \eta_{iv}^C \quad (1)$$

where λ_v^C is the spatial location of the conservative alternative, θ_{it} is the ideal point of the actor at the time of proposal t, $\hat{D}_v^C$ is a vector of non-ideological variables reflecting legal and political factors that may influence judicial voting, δ_i is the weight placed by i on non-ideological factors and η_{iv}^C is a random shock.

The utility difference between the conservative and liberal alternatives is

$$\begin{aligned} y_{itv}^* &= -(\theta_{it} - \lambda_v^C)^2 + \delta_i \hat{D}_v^C + \eta_{iv}^C + (\theta_{it} - \lambda_v^L)^2 + \delta_i \hat{D}_v^L + \eta_{iv}^L \\ &= 2\theta_{it}(\lambda_v^C - \lambda_v^L) + \lambda_v^{L2} - \lambda_v^{C2} + \delta_i(\hat{D}_v^C - \hat{D}_v^L) + \eta_{iv}^C - \eta_{iv}^L \\ &= (\lambda_v^C - \lambda_v^L)(2\theta_{it} - (\lambda_v^L + \lambda_v^C)) + \delta_i(\hat{D}_v^C - \hat{D}_v^L) + \eta_{iv}^C - \eta_{iv}^L \quad (2) \end{aligned}$$

The vote "cutpoint," κ_v, is $\frac{\lambda_v^L + \lambda_v^C}{2}$ and the vote "discrimination parameter," α_v, is $2(\lambda_v^C - \lambda_v^L)$.[1] The observed non-ideological variable D_v is $(\hat{D}_v^C - \hat{D}_v^L)$. The error term ϵ_{iv} is a N(0, 1) random

[1] The terminology comes from item response theory (Baker 1992). Votes for which the alternatives are relatively close (meaning $(\lambda_v^C - \lambda_v^L)$ is relatively small) will "discriminate" poorly between liberals and conservatives 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.

variable, $\eta_{iv}^C - \eta_{iv}^L$.

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

This implies that

$$\text{Prob}(y_{itv} = 1) = \Phi(\alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v)$$

Assuming independence across individuals and votes, the joint posterior probability of the observed data is

$$g(\theta, \alpha, \kappa|Y) \propto L(\theta, \alpha, \kappa, \delta|Y, D)g(\theta, \alpha, \kappa, \delta)$$

where

$$L(\theta, \alpha, \kappa|Y) = \Pi_{i=1}^N \Pi_{v=1}^V \Phi(\alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v)^{y_{iv}} \times (1 - \Phi(\alpha_v(\theta_{it} - \kappa_v) + \delta_i D_v))^{1-y_{iv}}$$

and $g(\theta, \alpha, \kappa, \delta)$ is the prior distribution over the parameters to be estimated.

The model is estimated with a modified Gibbs sampler algorithm that draws 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

e.g. Jackman 2000, 311)

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

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

where I is an indicator function that truncates distributions above or below zero.

2. The separate θ_{it} are estimated for each justice-year. That is, a separate ideal point is estimated for Justice Black in 1950 and in 1951 and so forth. Ideal points for an individual justice are linked over time by means of a Bayesian prior that is the estimated ideal point of that justice in the previous year. For a justice's first year, we use the estimated ideal point in the second year as the prior. The strength of the prior depends on a smoothing parameter that is set to 0.01. Lower values of this parameter mean more smoothing; higher values mean less smoothing and hence less dependence of ideal points across years for each justice. This is similar to, but not identical to, the way Martin and Quinn (2002) deal with preference dynamics within given justices.

3. Generate $\alpha, \alpha\kappa$ on a vote-by-vote basis. 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) and re-write Equation 3 as

$$y^*_{itv} - \delta_i D_v \quad = \quad \Theta_v\beta_v + \epsilon_{iv}. \tag{6}$$

By standard GLS results,

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

where $y_v^{**} = y_{itv}^* - \delta_i D_v$ for all individuals who voted on vote v, Θ_v is a $N_v \times 2$ matrix of Θ_{it} and N_v is the number of votes cast on vote v.

The linkage information about the relative position of certain case cutpoints is implemented via rejection sampling. For each vote on which we have information about the position of its cutpoint relative to another case we draw a cutpoint according to standard MCMC sampling, but reject it and repeat the process if the sampled cutpoint does not satisfy the cutpoint constraint information we have. This means that when the cutpoint of Case A is identified as being to the left of Case B, the program samples from the distribution for the cutpoint of Case A until it produces one that is to the left of the cutpoint estimated for Case B in that iteration. Similarly, the program samples from the distribution for the cutpoint of Case B until it produces one to the right of the cutpoint estimated for Case A in that iteration.

The discrimination parameter is, in part, a measure of vote-specific variance and, as a variance parameter is subject to becoming unbounded (see also Baker 1992, 97-98; Mislevy and Bock 1990, 8). Therefore we use normal priors and maximum values for α (mean is one, variance is 0.6 and maximum is 4.0) and κ (mean is 0, variance is 1 and maximum absolute value is 10). The priors are implemented as in 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)).

Data

Justices with partial year service I use calendar years (in order to line up with congressional voting). A number of justices voted on very few cases in either their first or last year of service. While it would be possible to estimate a separate ideal point even in years with small number of votes (and let the smoothing parameter pull the estimate toward the next year's estimated ideal point), experience shows that these estimates fluctuate significantly and then become priors for the following years, causing instability. Therefore, for justice-years with less than 14 votes I coded those votes as being from the year before/after depending on whether it was at the start or end of a justice's career. In other words, for years in which a justice voted less than 14 times I'm assuming his or her ideal point was the same as the following or previous year. For example, Justice Frankfurter voted on only 13 cases in 1962, his last year on the Court. We assume Frankfurter ideal point's for 1962 is the same as in 1961 and therefore report only his ideal point in 1961. For applications needing a Frankfurter ideal point in 1962 (for example in appointment studies that use departing

justice ideal points), one would simply use his ideal point from 1961.

Coding of ideological direction Instances in which I 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 ten campaign finance cases. For campaign finance related cases, Spaeth codes votes to uphold the limits on contributions or expenditures as conservative votes (as these are limiting speech and Spaeth coding rules take protecting speech as necessarily liberal). The political valence of campaign finance regulations is the opposite: overwhelmingly the political supporters of limits on contributions or expenditures come from the left and opponents of such laws come from the right. We therefore re-code these cases accordingly (as do Epstein and Segal 2006). 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. This also occurs in four agricultural advertising cases where the Spaeth coding rules take votes to favor mandatory government marketing programs (and the fees inherent in them) as conservative votes (as these are interpreted as limiting the speech rights by forcing farmers to support certain kinds of speech and the Spaeth rules take limits on speech as being inherently conservative). Here, the political coalition is again the opposite, with opposition to these programs coming from the right, as conservatives have opposed the government intervention in the economy inherent in the programs. Coding changes on an additional eight cases. For example, in *Hurley v. Irish American Gay, Lesbian and Bisexual Group of Boston* (1995) Spaeth coded as liberal the Court's decision to side with conservatives who argued that parade organizers did not have to allow a gay group in

a St. Patrick's Day parade.

Congress and the President Votes and positions by members of Congress are used to position Supreme Court justices in political ideology space. Bridge observations for 1950 through 2008 are from Bailey and Maltzman (2011). This paper updates the data through 2011.

1. Bridge observations of members of Congress taking positions on Supreme Court cases come from the following sources.

   - 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*.
   - Cosponsorship. We used www.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. Cosponsorship data is less comprehensive for periods predating Thomas.gov coverage.
   - A small number of roll call votes were coded as bridge observations because they

were directly on a case at hand. Given the complexities of voting in Congress it is generally the case that a potentially relevant vote in Congress involved non-trivial considerations unrelated to the case at hand and cannot therefore be directly linked as bridge observations.

I exclude repeat bridge observations (e.g. a Senator taking the same position on the same case) that are within five years of an observation already in the data set.

Coding of precedent To identify those cases where respect for a stare decisis doctrine was particularly likely to shape outcomes, we relied upon Segal and Howard's (2001) identification of these cases 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. The precedent variable is +1 if the parties or justices on the liberal side advocated overturning precedent; it is -1 if the parties or justices on the conservative side advocated overturning precedent.

Deference to Congress Spaeth's (2011) *authDecision* variables are the starting point in identifying those cases where a doctrine of deference to Congress is particularly applicable. Each case for which *authDecision1* or *authDecision2* was equal to 1 was read to ensure that it involved the constitutionality of a law enacted by Congress and the President and then coded as described in the paper. The congressional deference variable is +1 if the parties or justices on the liberal side advocated overturning a congressional statute; it is -1 if the

parties or justices on the conservative side advocated overturning a congressional statute.

Free speech The free speech variable is +1 if the parties or justices on the liberal side advocated limiting speech rights; it is -1 if the parties or justices on the conservative side advocated limiting speech rights.

Deference to Executive These were coded based on Spaeth coding of parties to the case. If the petitioner or respondent was from the U.S. government (with values of 1, 27 or between 300 and 500 in Speath's coding) this was coded as +1 or -1 depending on whether the government was advocating a liberal or conservative position (which is determined by Speath's *partyWinning* and *decisionDirection* variables). In one case both parties were from the U.S. government, but the Solicitor General argued for petitioners so that was treated that as the government position. The executive deference variable is +1 if the parties or justices on the liberal side advocated against a federal party to a case; it is -1 if the parties or justices on the conservative side advocated against a federal party to a case.

Sixth Amendment Sixth amendment cases are identified primarily based on *LawSupp* codes in the Supreme Court database. Cases for which this variable was equal to 213 (right to confront and cross-examine, compulsory process) or 215 (right to trial by jury) were coded, as were seven additional cases that feature prominently in those areas of Sixth Amendment case law. (They were *In Re Winship, Apprendi v. New Jersey, Almendarez-Torres v. U.S., Jones v. U.S., Richardson v. U.S., Harris v. U.S., U.S. v. Booker.*) This yielded 85 cases, all of which were coded -1 for this variable as voting consistently with Sixth Amendment in

these cases implied a liberal vote. Few cases before 1987 are coded as dealing with Sixth Amendment, rendering coefficients for this variable for justices who served in that era very imprecise. The Sixth Amendment variable is -1 if the parties or justices on the liberal side advocated on behalf of an accused person's rights under the confrontation clause or right to a jury; there are no cases in which invocation of these rights implied a conservative outcome.[2]

Software The models are estimated in Matlab code. First, the data is pre-processed in R in order to produce the core data for the estimation: vectors that indicate for each vote the identity of the individual, the case or roll call number, the year of the observation and the observation itself (1 for conservative, 0 for liberal). There are also vectors that indicate the values of the precedent, congressional deference, free speech, executive deference and Sixth Amendment variables as well. The program is made more computationally efficient by providing for each individual a list of roll calls/case on which he or she took positions and for each roll call/case a list of individuals who took positions on the case/roll call.

Results

Figures 1 - 3 show the ideal points estimated by the method described in the paper for each justice across their terms on the Court.

[2]Sixth amendment cases are primarily drawn from cases for which the "LawSupp" variable in the Supreme Court database is 213 or 215.

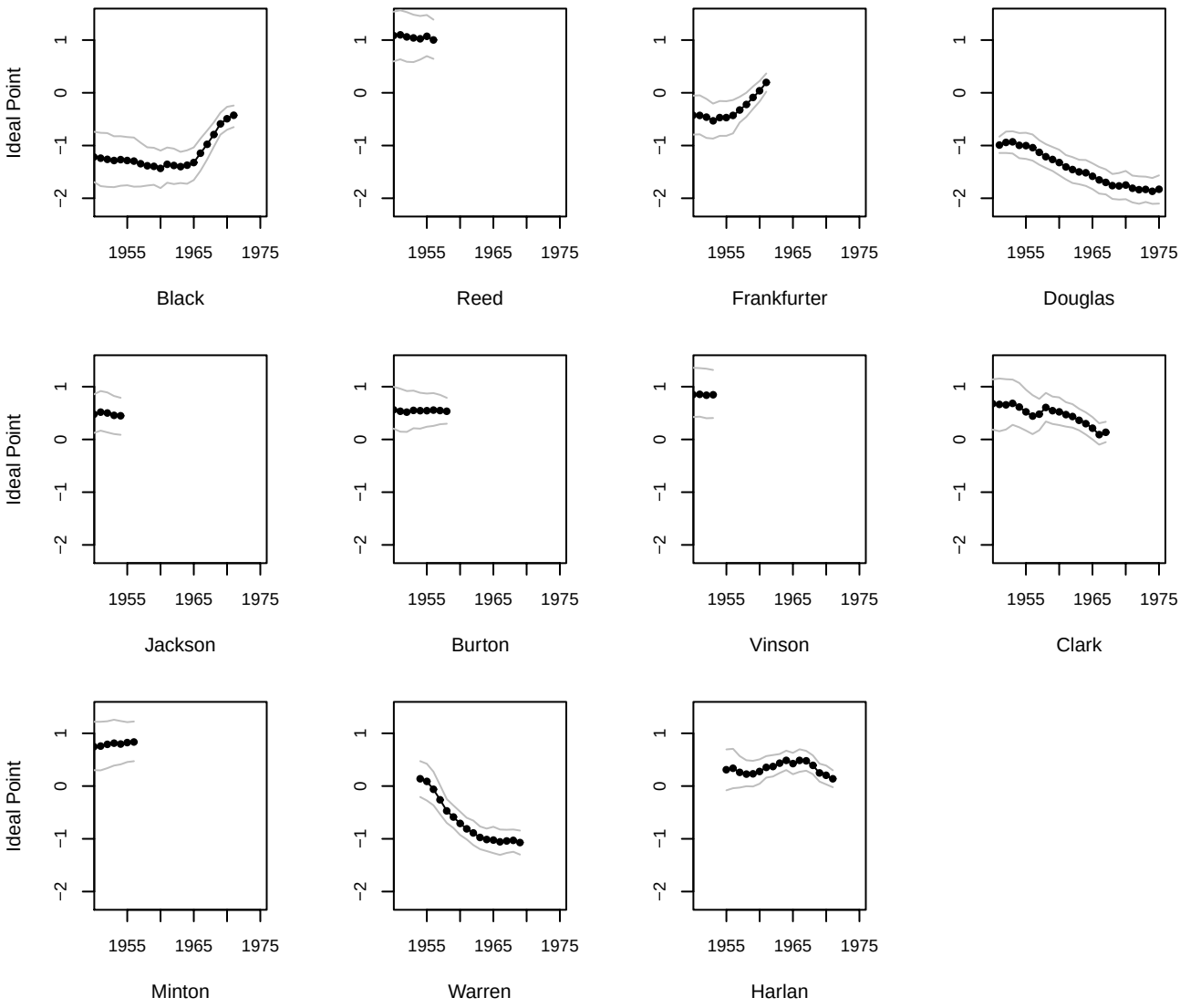

Figure 1: Estimated ideal points of Supreme Court Justices, 1950-1975

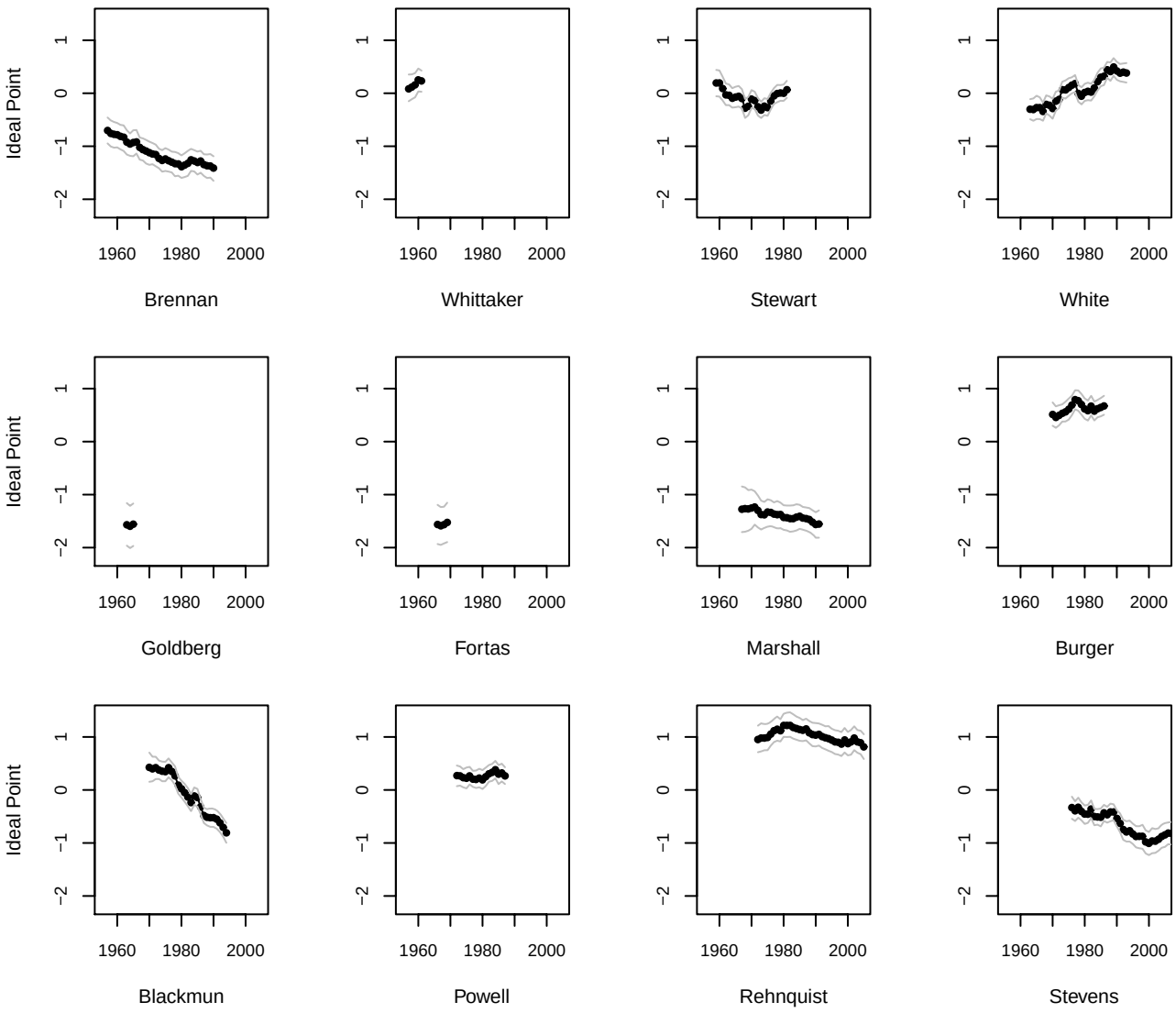

Figure 2: Estimated ideal points of Supreme Court Justices, 1955-2005

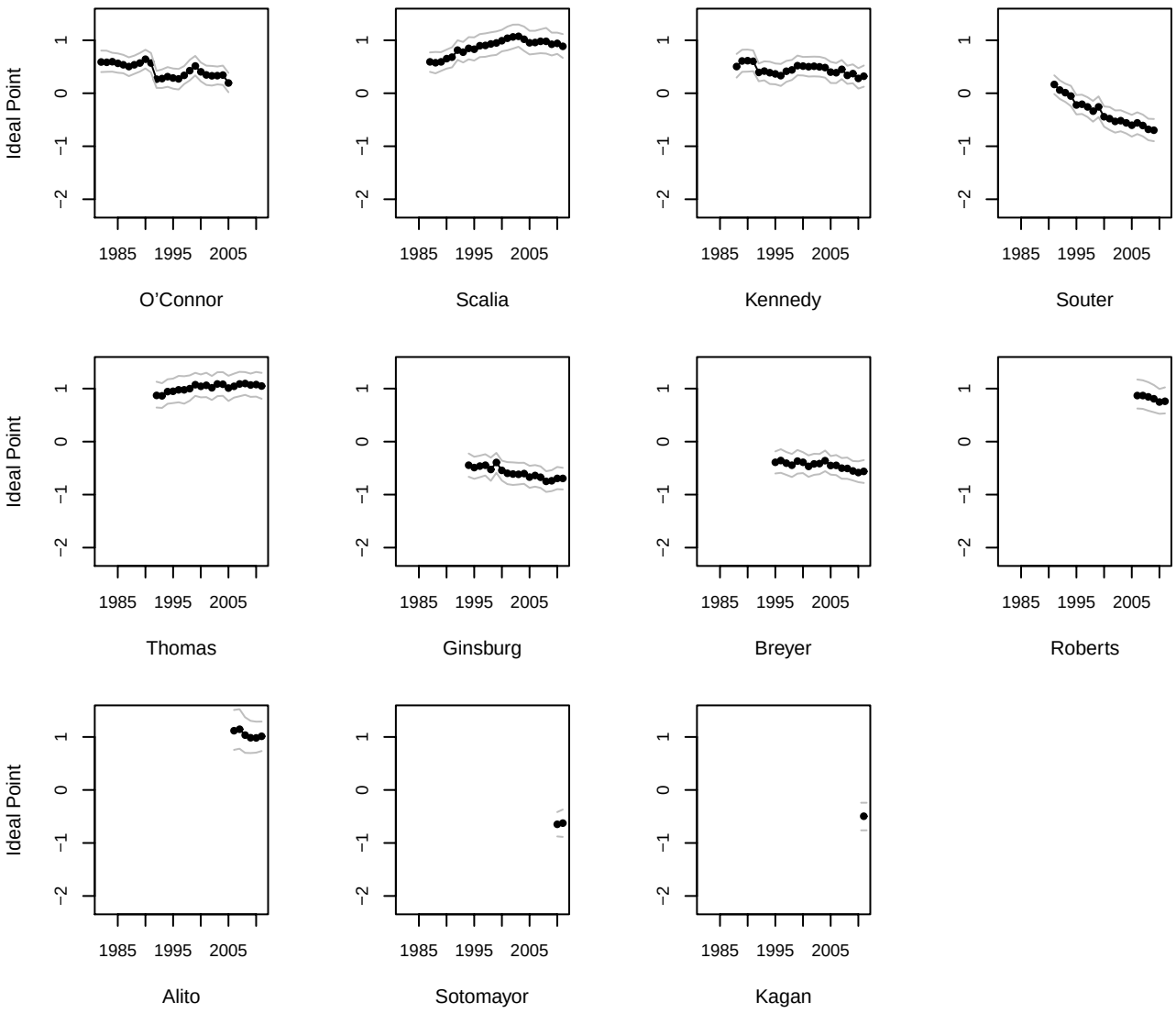

Figure 3: Estimated ideal points of Supreme Court Justices, 1982-2011

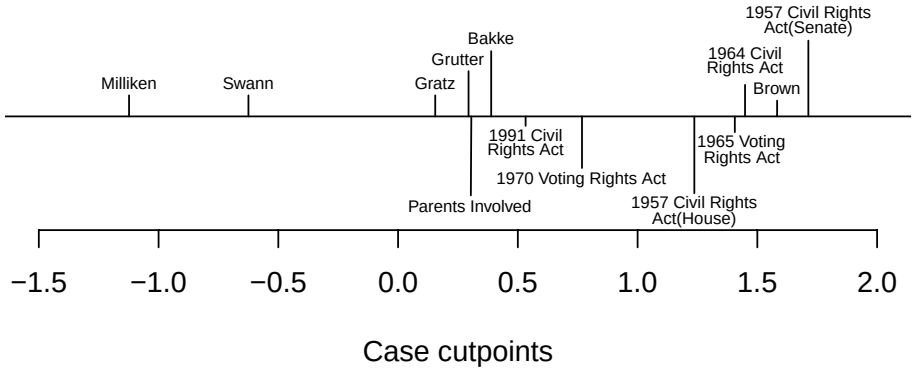

Figure 4: Case Cutpoints on Important Civil Rights Cases and Votes

Case cutpoints

One way to get a sense of the evolution of the agenda facing the Court and Congress is to look at the estimated case cutpoints on important cases over time. Figure 4 shows cutpoints for important civil rights cases and votes. The early votes (such as the 1957 Civil Rights bill and *Brown*) had cutpoints far to the right, implying that only the most conservative justices and/or members of Congress opposed them. As the civil rights agenda progressed, the cutpoints typically moved to the left with the Civil Rights bill of 1991 reaching around 0.5 by 1991. The busing decisions (*Swann* and *Milliken*) had cutpoints to the left of the scale, implying these were supported by some of the more liberal justices and members of Congress in the time period. The more current cases, such as *Gratz, Grutter* and *Parents Involved* hover in the middle of the space.

Note that the movement of case cutpoints is not the same as the movement of Court jurisprudence. The Court held against execution of people for crime committed when they were under 16 *Thompson v. Oklahoma* (1988). A year later, in *Stanford v. Kentucky* (1989), the Court allowed execution of those convicted of crimes committed between the ages of 16

and 18. The Court did not move left in *Stanford*, but the cutpoint did as justices voting liberally on *Stanford* (advocating against execution of minors over 16) is logically consistent with voting liberally on *Thompson* (opposing execution of individuals under 16). (The Court eventually did move left in *Roper v. Simmons* (2005) holding that it is unconstitutional to execute people for crimes committed under 18.)

References for Supplemental Appendix

Bafumi, Joseph, Andrew Gelman, David Park and Noah Kaplan. 2005. Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation. *Political Analysis* 13, 2: 171-187.

Baker, Frank. 1992. *Item Response Theory.* New York: Marcel Dekker.

Clinton, Josh, Simon Jackman and Doug Rivers. 2004. The Statistical Analysis of Legislative Roll Call Data. *American Political Science Review* 98, 2: 355-370.

Gelman, Andrew, John Carlin, Hal Stern and Donald Rubin. 1995. *Bayesian Data Analysis.* Boca Raton: Chapman Hall.

Jackman, Simon. 2000. Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo. *American Journal of Political Science* 44, 2: 375-404.

Johnson, Valen and James Albert. 1999. *Ordinal Data Modeling.* New York: Springer.

Mislevy, Robert and R. Darrell Bock. 1990. BILOG 3 [computer program]. Mooresville, Indiana: Scientific Software, Inc.

Poole, Keith and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* Oxford: Oxford University Press.