Article

Estimating Ideal Points from Roll-Call Data: Explore Principal Components Analysis, Especially for More Than One Dimension?

Richard F. Potthoff

Department of Political Science and Social Science Research Institute, Duke University, Box 90420, Durham, NC 27708, USA; potthoff@duke.edu

Abstract: For two or more dimensions, the two main approaches to estimating legislators' ideal points from roll-call data entail arbitrary, yet consequential, identification and modeling assumptions that bring about both indeterminateness and undue constraints for the ideal points. This paper presents a simple and fast approach to estimating ideal points in multiple dimensions that is not marred by those issues. The leading approach at present is that of Poole and Rosenthal. Also prominent currently is one that uses Bayesian techniques. However, in more than one dimension, they both have several problems, of which nonidentifiability of ideal points is the most precarious. The approach that we offer uses a particular mode of principal components analysis to estimate ideal points. It applies logistic regression to estimate roll-call parameters. It has a special feature that provides some guidance for deciding how many dimensions to use. Although its relative simplicity makes it useful even in just one dimension, its main advantages are for more than one.

Keywords: ideal points; nonidentifiability; Poole-Rosenthal scores; principal components analysis; spatial voting model; US Congress

1. Introduction

The use of roll-call data to place the positions of legislators within a political or ideological spectrum or space is now common, not only in academic research but also, to some degree, in mainstream media and in relation to political campaigns. One may alternatively refer to these spatial positions as locations, or scores, or *ideal points*. In the bulk of applications, the space is one-dimensional, with the single dimension interpreted in terms of a spectrum of "left"-"right" or "liberal"-"conservative" political ideology. However, space with two or more dimensions is of potential importance beyond the limited attention it has yet received, as will be covered in Section 2 below.

Ideal-point estimates can have different origins and different applications. In academic work, use of results from the Poole-Rosenthal approach has predominated, though certain Bayesian methodology has become available more recently. On the nonacademic side, election forecasting by Nate Silver (2014) has used (unidimensional) Poole-Rosenthal scores as part of its input. Other nonacademic pursuits will be cited in Sections 2 and 5.2.

The two main existing avenues for estimation of ideal points from roll-call data are the Poole-Rosenthal approach and a Bayesian approach. We examine both of them critically, particularly for more than one dimension, before turning to detailed study of principal components analysis, a technique that has rarely seen use for ideal-point estimation but offers much promise.

As a tool for appraisal of legislators' spatial locations, Poole-Rosenthal NOMINATE scores (McCarty et al. 1997; Poole 2005; Poole and Rosenthal 1985, 1991, 1997, 2001, 2007) have received wide attention and use. We will confine our analysis mainly to their W-NOMINATE (hereafter P-R), though DW-NOMINATE, a version that is much the same except that it can span across more than

one legislative term, is also often applied. We also will not deal with somewhat related approaches that are less often used such as the Optimal Classification (OC) method of Poole (2000), albeit in some circumstances (see, e.g., Rosenthal and Voeten 2004) OC may produce better results than P-R.

The Bayesian approach that we will consider (hereafter CJR) is the one associated with Jackman (2000, 2001) and Clinton et al. (2004a). A very new Bayesian approach is that of Imai et al. (2016). It is computationally fast but also not simple. In addition, its applications are shown only for ideal points in one dimension, whereas our prime focus here is on two or more dimensions.

This paper embraces two main themes, both with primary relevance for an issue space that is beyond one dimension. First, we highlight defects of P-R and CJR. The most serious problem for both of them is nonidentifiability of ideal points. Also raising questions are other issues concerning indeterminateness and arbitrary model assumptions. We note a greater number of difficulties for P-R than for CJR.

Second, we present a particular way of using principal components analysis (hereafter PCA) to estimate ideal points. Although our main emphasis is on contending that PCA has a pronounced edge over P-R or CJR for two or more dimensions, PCA has some advantage even in one dimension and may be a sound choice even for that case.

In regard to PCA, our main goal is to enable ideal-point estimates that have evident validity and avoid indeterminateness and arbitrary modeling and identification assumptions. Such estimates from PCA will then compare favorably with those from P-R and CJR. Estimates of ideal points can be (and have been) used in many types of applications, including those related to prediction of legislators' votes, measurement of ideology, dimensionality, historical evolvement of the US Congress, and polarization. Poole and Rosenthal (2007, chp. 11 and elsewhere) cite or describe numerous specific applications. See also Section 2 below for relevant information. Our PCA ideal-point estimates could be used for any and all types of applications, with the intent of obtaining more justifiable analyses especially with more than one dimension.

Section 2 presents numerous examples of issue spaces that may have more than one dimension. It thus demonstrates the meaningfulness of such issue spaces and of our work here.

Section 3 covers the shortcomings of P-R and CJR. Mathematical support comes from Appendix A. The shortcomings and their impacts have been largely unrecognized.

Section 4 presents our PCA methodology. Among other things, it covers extraction of the ideal-point estimates; estimation of the roll-call parameters, which is via logistic regression; a special example, for two dimensions, small enough to show the main features of our PCA approach in a single table; and treatment of missing votes. Estimates of roll-call parameters, sometimes used in analysis of individual roll calls, are produced jointly with the ideal-point estimates in both P-R and CJR—but not under PCA, where we produce them separately through logistic regression.

After discussion in Section 5 of certain techniques or applications that bear some relation to our PCA approach or provide alternatives to it, Section 6 presents four empirical examples of PCA applications from the 105th and 106th US Senates (1997–1998 and 1999–2000), with varying numbers of roll calls and with both one- and two-dimensional issue spaces. A benefit of using examples from those two Senates is the availability of comparable published data, especially on model fit, for P-R and CJR. Comparisons are satisfactory, thus attesting to face validity for PCA.

How does one decide how many issue dimensions are best for given data? For judging the number of dimensions to use, PCA provides guidance of a type not available with P-R or CJR. Section 7 gives details.

Why has principal component analysis been used only rarely for estimation of ideal points? We explore this matter in Section 8, before summarizing further in Section 9.

2. Application of Ideal Points in More Than One Dimension

Largely unidimensional voting spaces in recent US Congresses may have curtailed recent attention to issue spaces with more than one dimension. Historically, though, even the US has seen dimensions

beyond the traditional "left"-"right" or "liberal"-"conservative" one: for example, slavery before the Civil War, bimetallism in the late 19th century, and civil rights in the mid-20th century (Poole and Rosenthal 2007).

Moreover, a new second dimension, such as one pertaining to undue police force, foreign military involvement, surveillance by government, drug laws, or some combination thereof (cf. Hook 2014), could become prominent at some point. In a somewhat similar vein, the political preferences of US citizens (as opposed to elites or legislators) were found by Carmines et al. (2012) to fall onto not just one dimension, but rather two—one involving economic and social-welfare issues and the other related to social and cultural matters.

Varied dimensions may arise in different times and places. Social or cultural issues may be largely separable from economic ones in some situations. Other issues that may supply extra dimensions include transnational integration, as in some European countries; secession, as in Scotland and Catalonia; language, as in Belgium and Canada; Peronism (Argentina); ethnicity; religion; immigration; corruption and reform; and nationalism. Dimensions in Israel, in addition to economic ones, have included religious-secular and hawk-dove.

Bornschier (2010) examined manifestations of two dimensions of political space in six countries in Western Europe. Hix et al. (2006) found two dimensions in the European Parliament.

Some recent work has challenged the notion that the US political space is largely one-dimensional. Perhaps paradoxically, extra dimensions for the US may become more evident with fewer roll calls rather than more, and with fewer legislators rather than more. Upon examining separate subsets of votes in different subject areas, Crespin and Rohde (2010) found more multidimensionality than in the totality of roll calls. Aldrich et al. (2014) concluded that greater dimensionality is evinced within each party than in the two parties taken together. Thus, analysis of smaller data sets may uncover new and extra dimensions that are muffled in a comprehensive data set. Along a different line, Dougherty et al. (2014) found that agenda setting that prevents certain votes from taking place can suppress revelation of multidimensionality.

Perhaps the most extensive representation of two-dimensional political space is in a video shown on Voteview Blog (2014). It presents graphs of two-dimensional Poole-Rosenthal ideal points, separately for the US Senate and House, for each two-year period from 1789 to 2013.

A chart with 10 two-dimensional graphs of Poole-Rosenthal scores from Voteview, for the US Senate and House for each of the terms of 1889–90, 1963–64, 1983–84, 1997–98, and 2011–12, appeared in an article in *The Economist* (Anonymous 2014b) to illustrate increasing polarization in American politics in recent years. The first dimension was labeled as "Liberal" versus "Conservative", reflecting "differences in political ideology, based on votes on government intervention in the economy" (an oversimplification, it would seem). The second dimension, labeled as "Southern values" versus "Northern values", was represented as pertaining to "values traditionally identified as Northern and Southern, based on votes on race-related issues" (a description that would seem to be less fitting for the later years).

All in all, either for the US or elsewhere there is fertile ground for studying more than one dimension in roll-call analysis, a pursuit for which PCA can offer some unique advantages.

3. Flaws and Drawbacks of P-R and CJR

In this section, we scrutinize both P-R and CJR. In many respects they do not compare favorably with PCA.

In two or more dimensions (there is no problem in one dimension), the gravest flaw of both P-R and CJR is their nonidentifiability of ideal points—nonidentifiability that goes beyond the simple indefiniteness of location, scale, and orientation. Unlimited numbers of transformations (ones that do not just change location or scale) produce shifts to new sets of ideal points that are substantively different but leave the maximized log likelihood unchanged. Troubles escalate as the number of dimensions increases. Identifiability of ideal points in more than one dimension should

be viewed as an essential property for any approach to roll-call analysis. Nonidentifiability and associated nonestimability (cf., e.g., Basu 1983; Schmidt 1983) mean essentially that the ideal points in more than one dimension are not even defined—a critical pitfall. Nonidentifiability is hardly a trifling transgression.

For CJR, the nonidentifiability is acknowledged explicitly (e.g. see (Jackman 2001), especially pp. 233, 235). Recognizing the nonidentifiability of CJR in two dimensions, Jackman (2001) sought to circumvent it by assigning certain priors to two roll calls chosen to serve as second-dimension anchors. However, this patch seems arbitrary and subjective. Any singling out of a pair of roll calls is open to question, since different anchor pairs yield different results. With more dimensions, the arbitrary identification assumptions create worsening troubles. Two-dimensional CJR applications other than Jackman (2001) seem to be isolated. We could find only one, noted briefly in Jessee (2009, footnote 11) with no indication of what anchors he chose or how otherwise he dealt with the nonidentifiability. The nonidentifiability issue seems to cripple the use of CJR for more than one dimension.

For P-R, one may not suspect nonidentifiability, as one does find P-R applications in more than one dimension. However, Part A.6 of Appendix A below provides formal mathematical proof of the P-R nonidentifiability. This result for P-R may seem almost obvious, though: For two or more dimensions, P-R has far more parameters than CJR (thus entailing worse identifiability troubles) yet even CJR itself lacks identifiability.

Also pertaining to indeterminateness of ideal-point estimates are two other characteristics of both P-R and CJR. First, both of them lack the orthogonality property, under which the ideal-point vectors for any pair of dimensions are orthogonal. Any flexibility gained from forgoing orthogonality might be deemed an advantage until one realizes that it comes at the price of nonidentifiability (or, at least, no suitable way to get identifiability without orthogonality is evident). Second, neither P-R nor CJR preserves lower-dimension ideal-point estimates when calculating those for a higher dimension. Although not very lucid about the matter, the mathematical details for both CJR (Clinton et al. 2004a, p. 367) and P-R (Poole 2005, pp. 107–10) convey that the respective procedures yield first coordinates of their two-dimensional ideal points that will differ from the single coordinates of their one-dimensional ideal points. Thus, where computations are done for both one and two dimensions, there are two different sets of ideal points for the first dimension. With more dimensions, multiplicity increases. At least some users will object to such indefiniteness.

As will be seen shortly, PCA does not have any of the characteristics that have just been described. It does not suffer from nonidentifiability, it enjoys the orthogonality property, and its ideal-point estimates for a given dimension do not change when the number of dimensions changes.

Several further modeling issues, in the form of arbitrary parameter constraints or identification assumptions, cause trouble for P-R. None of them apply to either CJR or PCA.

P-R requires legislators' ideal points to lie within a circle or sphere of unit radius for two or three dimensions, respectively, and similarly for higher dimensions. Some ideal points may thus be forced unnaturally to lie exactly on the boundary (surface) of the unit circle (sphere) in two (three) dimensions. In two dimensions under P-R, a legislator with a location of -0.8 or $+0.8$ on the first dimension could not have a second-dimension position below -0.6 or above $+0.6$ (because $0.8^2 + 0.6^2 = 1$). A legislator who is at an extreme (-1 or $+1$) on one dimension cannot be at an extreme (in fact, can be nowhere except at 0) on the other.[1] This limitation of P-R is of more than theoretical concern. For instance, out of the 102 members who served in the 106th US Senate, 23 have two-dimensional ideal points that lie exactly on the boundary of the unit circle.[2] The most extreme case is Barbara Boxer of California, who has a position of -0.988 on the first dimension and -0.156 on the second. Given the value of -0.988,

---

[1]   Note that using a larger circle would not help, because a legislator would still be unable to have extreme scores on both dimensions. Shape matters: A square, a circle, and (e.g.) a four-pointed star will all work differently.

[2]   Source of data is given in Appendix C.

the second-dimension P-R location has to lie in the short interval from -0.156 to $+0.156$. For more on the ideal-point constraints, see Part A.4 of Appendix A.

Roll-call parameters bear similar restrictions. P-R imposes constraints on its estimates of its roll-call midpoints like those on its ideal-point estimates (limitation to the unit sphere in two dimensions, e.g.). Again, some (midpoint) estimates may be unnaturally confined. See Part A.5 of Appendix A for more discussion.

In two or more dimensions, P-R is severely overparameterized. This trait relates to, and aggravates, nonidentifiability. For any dimensionality, CJR and PCA use exactly the same number of ideal-point and roll-call parameters as each other. Although P-R uses the same number of ideal-point parameters as CJR and PCA, with two or more dimensions it uses far more roll-call parameters than the other two. The excess P-R roll-call parameters may have an appealing rationale but exact a steep price in terms of overparameterization and nonidentifiability. Part A.3 of Appendix A explains further.

P-R abides an arcane inflexibility for a (hypothetical) roll call whose votes have no relation at all to the legislator locations. Full details are in Part A.2 of Appendix A.

One finds indications that the use of P-R may be questionable if the number of roll calls and/or legislators is small (e.g., Crespin and Rohde 2010, pp. 980–81; Lewis 2001; Peress 2009; Poole et al. 2011, p. 5). It may not be clear how well CJR behaves with small data sets. As for PCA, Example 1 below applies it to data with only eight roll calls and 14 legislators, and textbooks abound with examples of principal components analyses for small data sets.

PCA yields a unique solution for ideal points.[3] P-R and CJR, by contrast, both involve iterative procedures, and may thus produce ideal-point estimates that vary depending on different choices of such elements as starting values for the ideal points, number of iterations, and stopping rules. With CJR, choice of Bayesian priors can also affect the solution.[4,5]

4. Specifics for the Use of PCA, with an Example

As will be seen shortly, our PCA technique avoids the difficulties of P-R and CJR just described; it is simple, fast, and powerful; and its results are acceptable. We now devote detailed attention to it.

The statistical theory and application of principal components has a lengthy history and is thoroughly covered in various textbooks (e.g., Morrison 1976, chp. 8; Jackson 1991; Jolliffe 2004) that provide details of the technique. For estimating ideal points from roll-call data, however, the use of the methodology has been scant. This section covers the general method and related matters, deals with estimation of roll-call parameters, and provides a detailed small example.

4.1. How PCA Obtains Ideal Points

With I legislators and J roll calls, we suppose that we have a vote matrix $\mathbf{Y}_0$ ($I \times J$) with general element y_{0ij} ($i = 1, ..., I; j = 1, ..., J$) equal to 1 or 0 if legislator i votes, respectively, yea or nay on roll call j. For the moment we assume no missing data (no missed votes), but later, in Section 4.5, we show how to skirt this assumption. Unanimous votes provide no usable information and are excluded from $\mathbf{Y}_0$. All of our calculations use SAS®.

Our approach treats the elements of $\mathbf{Y}_0$ as I observations from a J-variate distribution and then calculates principal components and their associated scores in the standard fashion. Let $\mathbf{u}_I$ denote

---

[3] It is unique except for possible different choices for how to impute for any missing votes.
[4] The same elements whose choices can cause differing CJR results can also lead to differences under the approach of Imai et al. (2016).
[5] For an extensive and informative comparison of CJR and P-R, though only for the case of one dimension, see Carroll et al. (2009); see also Clinton and Jackman (2009). For other works that provide differing views about P-R or how it compares with CJR, see (e.g.) Krehbiel and Peskowitz (2015); Caughey and Schickler (2016); Bateman and Lapinski (2016) and McCarty (2016).

an ($I \times 1$) vector with all 1's. Then the vote matrix adjusted for the roll-call means is $\mathbf{Y}$ ($I \times J$) = $\mathbf{Y}_0 - \mathbf{u}_I \mathbf{u}_I' \mathbf{Y}_0 / I$, from which the sample covariance matrix $\mathbf{S}$ ($J \times J$) = $\mathbf{Y}'\mathbf{Y}/(I-1)$ is obtained.

PCA is based on the eigenvalues and eigenvectors of $\mathbf{S}$. Let L_k denote the k-th largest eigenvalue ($k = 1, 2, ...$) and $\mathbf{g}_k$ ($J \times 1$) the corresponding eigenvector. (Here we consider only nonzero eigenvalues, and also disregard the highly unlikely case where two nonzero sample eigenvalues are equal.) Then, for any $\mathbf{g}$ ($J \times 1$) such that $\mathbf{g}'\mathbf{g} = 1$, the maximum value of $\mathbf{g}'\mathbf{S}\mathbf{g}$ (which is the maximum variance of any linear function of the J votes using such a $\mathbf{g}$) is L_1 and is attained at $\mathbf{g} = \mathbf{g}_1$. Furthermore, for any $\mathbf{g}$ subject to $\mathbf{g}'\mathbf{g}_1 = 0$ (or $\mathbf{g}'\mathbf{S}\mathbf{g}_1 = 0$) and $\mathbf{g}'\mathbf{g} = 1$, the maximum value of $\mathbf{g}'\mathbf{S}\mathbf{g}$ is L_2 and is attained at g = $\mathbf{g}_2$. Analogous results hold for $k > 2$.

We use the formula $\mathbf{x}_1$ ($I \times 1$) = $(3L_1)^{-0.5}\mathbf{Y}\mathbf{g}_1$ for the legislators' first-dimension scores, or ideal points. The mean score, $\mathbf{u}_I'\mathbf{x}_1 / I$, is 0. The variance of the scores is $1/3$, equal to that of the rectangular (uniform) distribution from -1 to $+1$. Similarly, the formula $\mathbf{x}_2$ ($I \times 1$) = $(3L_2)^{-0.5}\mathbf{Y}\mathbf{g}_2$ gives the second-dimension ideal points. Both $\mathbf{x}_1'\mathbf{x}_2$ and $\mathbf{g}_1'\mathbf{g}_2$ are 0; that is, the two ideal-point vectors are orthogonal, as are the two eigenvectors. The scores for general k for the I legislators are found from $\mathbf{x}_k$ ($I \times 1$) = $(3L_k)^{-0.5}\mathbf{Y}\mathbf{g}_k$. For all dimensions k, $\mathbf{x}_k$ has mean 0 and variance $1/3$. For all $k \neq k^*$, $\mathbf{x}_k'\mathbf{x}_{k*}$ and $\mathbf{g}_k'\mathbf{g}_{k*}$ are both 0.

Although in most applications of principal components analysis the number of observations exceeds the number of variables, the reverse condition can also occur (e.g., Jackson 1991, pp. 32, 73, 190). Some of our examples have more legislators (observations) than roll calls (variables), that is, $I > J$, whereas others have $J > I$.

The simplicity of the main element of the PCA approach becomes evident upon noting that the single line of SAS® code

<div align="center">

proc princomp data=d1 cov out=d2 outstat=d3 noprint;

</div>

produces (in the output data set d2, for all dimensions) the legislators' scores before multiplication by $(3L_k)^{-0.5}$ (i.e., the values $\mathbf{Y}\mathbf{g}_k$) as well as (in the output data set d3) the eigenvalues, L_k, and the eigenvectors, $\mathbf{g}_k$. The input data set d1 is the vote matrix $\mathbf{Y}_0$ (after resolving any missing votes). For some limited timing results for PCA, see the end of Section 6.1 below.

4.2. Roll-Call Parameters

Though generally less important than the legislator scores, estimates of roll-call parameters are often desired. However, with our routine (unlike others), they need not be calculated at all if one is interested only in the scores. On the other hand, they are required if one wants to evaluate model fit (see Sections 4.3 and 6.2 below). They are also needed if one wants to study individual roll calls by (e.g.) examining their (two-dimensional) cutting lines as in numerous examples in Chapters 5–7 of Poole and Rosenthal (2007). In one dimension, they are used to calculate, for a given roll call, the cut point that separates predicted nay-voters from predicted yea-voters based on the first-dimension scores.

Once the scores $\mathbf{x}_k$ have been obtained, our PCA approach estimates the roll-call parameters through logistic regression. Our technique for these parameter estimates bears limited resemblance to the method of joint maximum-likelihood estimation used in item-response theory in educational testing. Let x_{ik} denote the i-th element of $\mathbf{x}_k$, that is, the estimated ideal point of legislator i on dimension k. With p_{ij} denoting the probability that legislator i votes yea on roll call j, our logistic-regression model takes the form

$$\log \frac{p_{ij}}{1 - p_{ij}} = a_j + b_j x_{i1} \tag{1}$$

with one dimension and

$$\log \frac{p_{ij}}{1 - p_{ij}} = b_{j0} + b_{j1} x_{i1} + b_{j2} x_{i2} \tag{2}$$

for two dimensions, with obvious extensions for more than two dimensions. Each roll call j has two parameters, denoted by (a_j, b_j), in (1), and three parameters, (b_{j0}, b_{j1}, b_{j2}), in (2).

Because $p_{ij} = \frac{1}{2}$ if the right side of (1) is equal to 0, a cut point (or midpoint) for roll call j under (1) may be defined by $m_j = -a_j/b_j$. Thus, any legislator with ideal point x_{i1} to the right (left) of m_j votes yea (nay) with probability greater than 1/2 if $b_j > 0$—and vice versa if $b_j < 0$. Under (2), the cutting line $b_{j0} + b_{j1}x_{i1} + b_{j2}x_{i2} = 0$ separates the legislators according to whether their probabilities of voting yea on roll call j are greater or less than 1/2.

For each roll call (j) separately, for either (1) or (2), we estimate the roll-call parameters through logistic regression with the y_{0ij}'s (equal to 1 for yea, 0 for nay) as the response variable and the x_{i1}'s [in (1)] or x_{i1}'s and x_{i2}'s [in (2)] as the independent variable(s). Although the principal-components computation of Section 4.1 requires imputed values for missing votes (obtained as described in Section 4.5), the logistic-regression calculations are run for each roll call individually and thus do not need the imputed values. If we define n_{ij} to be 1 if legislator i provides a vote on roll call j and 0 if the vote is missing, then the logistic-regression computation for a roll call j is based only on those legislators i for whom $n_{ij} = 1$.

Because logistic regression can encounter complete separation of points (see, e.g., Albert and Anderson 1984), we can employ special steps to detect this condition and bypass the logistic-regression calculation on any roll call where it occurs. With either (1) or (2), complete separation of points on roll call j entails perfect fit for that roll call.

With (1), the separation prevents convergence of the logistic-regression procedure that estimates (a_j, b_j). It occurs if the x_{i1}'s of the legislators who vote yea on roll call j are either all above or all below the x_{i1}'s of all legislators who vote nay. If that condition exists on roll call j, one proceeds as follows. Let m_j be the point halfway between the highest x_{i1} of a nay (yea) voter and the lowest x_{i1} of a yea (nay) voter if all yea voters have x_{i1}'s above (below) those of all nay voters. Then set $a_j = -Hm_j$ and $b_j = H$ ($a_j = Hm_j$ and $b_j = -H$) if the yea voters have the higher (lower) x_{i1}'s. H is a large positive number (we use $H = 100$). With the (a_j, b_j) pair thus specified, the line represented by the right side of (1) is almost vertical and cuts the horizontal axis at $m_j = -a_j/b_j$.

For (2), we first set $b_{j0} = a_j$, $b_{j1} = b_j$, and $b_{j2} = 0$ for any roll call j that has complete separation of points under model (1) [that is, we duplicate the two parameters from (1) and set $b_{j2} = 0$]. Beyond that, complete separation of points is harder to identify in two dimensions than in one, but can be detected through linear programming. Appendix B gives the details.

4.3. Measuring Model Fit

Geometric mean probability or GMP (e.g., Poole and Rosenthal 2007, pp. 37–38) provides a useful means for evaluating model fit. Under our PCA approach, the log-likelihood function for legislator i and roll call j can be taken as

$$V_{ij1} = y_{0ij}\left(a_j + b_j x_{i1}\right) - \log\left(1 + e^{a_j + b_j x_{i1}}\right)$$

under the one-dimensional model (1) and as

$$V_{ij2} = y_{0ij}\left(b_{j0} + b_{j1} x_{i1} + b_{j2} x_{i2}\right) - \log\left(1 + e^{b_{j0} + b_{j1} x_{i1} + b_{j2} x_{i2}}\right)$$

under the two-dimensional model (2), with obvious generalization to k dimensions for a function V_{ijk}. If roll call j has perfect model fit (complete separation of points) in k dimensions ($k \geq 1$), though, then V_{ijk} is set to 0 for all legislators i for that roll call. Excluding those (i, j) combinations for which $n_{ij} = 0$, let $V_{i.k}$, $V_{.jk}$, and $V_{..k}$ denote the sum of V_{ijk} over j, over i, and over both i and j, respectively (where the V_{ijk}'s are evaluated using the estimates of the ideal points and roll-call parameters). Similarly, let $n_{i.}$, $n_{.j}$, and $n_{..}$ denote the sum of n_{ij} over j, over i, and over both i and j, respectively. Then for dimension k the GMP can be written as $G_{i.k} = e^{V_{i.k}/n_{i.}}$ for legislator i, $G_{.jk} = e^{V_{.jk}/n_{.j}}$ for roll call j, and $G_{..k} = e^{V_{..k}/n_{..}}$ overall. As a V-value becomes less negative, the corresponding G-value increases, indicating better fit; G approaches 1 as V approaches 0.

4.4. A "Toy" Example

Perhaps the best way to illustrate the workings of our PCA methodology described in Sections 4.1–4.3 is through a small example. Example 1, all of whose calculated values are shown in Table 1, has $I = 14$ legislators, who are US House members, and $J = 8$ roll calls, taken from the 2006 (second) session of the 109th US Congress. The eight roll calls were chosen from the 12 "key votes" selected by Congressional Quarterly (2007) for that session. Although the choices of the 14 representatives and of the eight key votes from the 12 were made with an eye toward trying (successfully) to obtain an example with a rather strong second dimension, the choices were done without any trial-and-error explorations. In any case, the basic aim of the example is to illustrate the PCA features so as to provide understanding. Any substantive results, even if reasonable, are secondary and incidental. For much larger examples, see Section 6 below.

Table 1. Data and results for Example 1 (14 US House members, eight of the 12 *CQ* key votes for 2006).

| Member | Party | State | Roll-Call Votes on 2006 Congressional Quarterly Vote No. * | | | | | | | | Legislator Results ** | | | | |
			135	239	288	372	388	479	502	511	x_{i1}	x_{i2}	x_{i2} rank	$G_{i.1}$	$G_{i.2}$
Levin	D	Mich.	N	Y	N	N	Y	N	N	N	−0.895	−0.166	6	0.83	0.91
Thompson	D	Miss.	N	Y	Y	N	Y	N	N	Y	−0.714	−0.126	7	0.77	0.92
Lynch	D	Mass.	N	Y	Y	N	Y	Y	N	N	−0.600	0.472	10	0.67	0.85
Baca	D	Calif.	N	N	N	N	Y	N	N	Y	−0.542	−0.431	4	0.66	0.81
Kaptur	D	Ohio	N	Y	N	N	N	N	N	Y	−0.471	−0.782	2	0.56	0.97
Spratt	D	S. C.	N	N	Y	N	Y	Y	Y	N	−0.159	0.919	14	0.58	0.94
Kirk	R	Ill.	Y	N	Y	N	Y	N	Y	N	−0.048	0.697	13	0.57	0.73
Matheson	D	Utah	Y	Y	Y	N	Y	Y	Y	Y	0.091	0.479	11	0.60	0.73
Paul	R	Texas	Y	N	N	Y	N	N	N	Y	0.361	−0.907	1	0.51	0.90
Jones	R	N. C.	Y	Y	Present	Y	N	Y	N	+	0.401	−0.565	3	0.56	0.77
Flake	R	Ariz.	Y	N	Y	N	N	Y	N	Y	0.402	−0.033	8	0.70	0.95
Walsh	R	N. Y.	Y	N	Y	N	N	Y	Y	N	0.490	0.679	12	0.63	0.90
Duncan	R	Tenn.	Y	N	N	Y	N	Y	Y	Y	0.814	−0.318	5	0.79	0.98
Cantor	R	Va.	Y	N	Y	Y	N	Y	Y	Y	0.870	0.082	9	0.90	0.99
Roll-call results *															
	g_{j1}		0.523	−0.339	0.082	0.379	−0.444	0.357	0.319	0.188					
	g_{j2}		0.020	−0.123	0.516	−0.305	0.330	0.307	0.454	−0.465					
	a_j		10.32	−0.42	0.50	−3.17	0.06	0.44	−0.42	0.66					
	b_j		100.00	−2.61	0.60	7.56	−4.43	2.85	2.41	1.31					
	m_j		−0.10	−0.16	−0.84	0.42	0.01	−0.15	0.17	−0.51					
	b_{j0}		10.32	−0.39	21.76	−85.31	−13.78	0.45	−49.41	4.91					
	b_{j1}		100.00	−2.65	8.90	138.05	−151.94	3.52	110.10	8.11					
	b_{j2}		0.00	−0.79	100.00	−100.00	100.00	2.52	100.00	−13.10					
	$G_{.j1}$		1.00	0.60	0.52	0.78	0.70	0.62	0.59	0.55				0.658	
	$G_{.j2}$		1.00	0.62	1.00	1.00	1.00	0.71	1.00	0.82					0.879

* Roll-call descriptions: #135, tax cuts, 5/10/06; #239, network neutrality, 6/8/06; #288, Iraq war resolution, 6/16/06; #372, end bilingual voting help, 7/13/06; #388, stem cell research, 7/19/07; #479, abortion notification, 9/26/06; #502, warrantless surveillance, 9/28/06; #511, easier challenges to eminent domain, 9/29/06. ** For legislator *i* (in row *i*): x_{i1}, x_{i2} = Score (estimated ideal point) for dimension 1, 2; $G_{i.1}$, $G_{i.2}$ = GMP (geometric mean probability) for 1, 2 dimensions. *** For roll call *j* (in column *j*): g_{j1}, g_{j2} = Element of eigenvector for dimension 1, 2; a_j, b_j = Roll-call parameters for 1 dimension [see Equation (1)]; $m_j = -a_j/b_j$ = Midpoint (cut point) for roll call (1 dimension); b_{j0}, b_{j1}, b_{j2} = Roll-call parameters for 2 dimensions [see Equation (2)]; $G_{.j1}$, $G_{.j2}$ = GMP for 1, 2 dimensions.

The table shows the vote of each House member on each roll call. To obtain the 14×8 matrix $\mathbf{Y}_0$, in Table 1 one changes each Y (yea) to 1 and each N (nay) to 0. For purposes of this example, since Congressman Jones voted "Present" on roll call #288 and was "Announced for" (shown as "+" in Table 1) on #511, his value was set to $1/2$ for #288 and to 1 for #511.

Under "Legislator results" Table 1 shows, for each House member, the scores for the first and second dimensions, x_{i1} and x_{i2}, along with the rank of x_{i2}. The x_{i1}'s and x_{i2}'s each have mean 0 and variance $1/3$. The members are listed in the order of their first-dimension scores (low to high, or traditional "left" to "right"), with Levin and Cantor at the two ends. Not surprisingly, there is almost complete separation between Democrats and Republicans on the first dimension, the only exception being that Matheson (D) has a higher x_{i1} than Kirk (R).

The second-dimension scores, x_{i2}, bear little or no relation to party or to x_{i1}, as would be expected in view of the orthogonality of x_{i1} and x_{i2}. Paul and Spratt are at opposite ends on x_{i2}.

In the lower part of Table 1, the first two lines under "Roll-call results" show the first two eigenvectors, $\mathbf{g}_1'$ and $\mathbf{g}_2'$, whose elements are denoted by g_{j1} and g_{j2}. Both $\mathbf{g}_1'\mathbf{g}_1$ and $\mathbf{g}_2'\mathbf{g}_2$ are equal to 1. Note that $|g_{j1}|$ is highest for vote #135 (tax cuts) and lowest for #288 (Iraq war) whereas $|g_{j2}|$ is highest for #288 and lowest for #135, thus suggesting that the first dimension is heavily influenced by the tax-cut vote and the second dimension by positions on the Iraq war. Other votes that contribute strongly to the second dimension are #511 (easier challenges to eminent domain) and #502 (warrantless surveillance). The second largest $|g_{j1}|$ is for #388 (stem cell research).

The first two eigenvalues (not shown in Table 1) are $L_1 = 0.741$ and $L_2 = 0.555$. Since the sum of the eight eigenvalues is 2.029, the first and second dimensions account, respectively, for $0.741/2.029 = 36.5\%$ and $0.555/2.029 = 27.4\%$ of total variability. The ratio of L_2 to L_1, $0.555/0.741 = 0.75$, is unusually high (as is evident upon comparison with examples in Section 6.1) and thus suggests a strong second dimension. As a result, Example 1 provides a good illustration of how PCA works when more than one dimension is important, though the prominence of the second dimension seems attributable in large part to the particular choices of House members and roll calls (e.g. these members include most of the few Republicans who voted nay or present on vote #288).

For each roll call Table 1 shows the estimates of the first-dimension roll-call parameters, a_j and b_j, and the associated cut point, $m_j = -a_j/b_j$. The sign of a b_j is generally positive (negative) for a roll call that attracts its yea votes largely from the Republicans (Democrats). Cut points are close to 0, the mean of x_{i1}, for five of the eight roll calls.

Model fit is perfect in one dimension for roll call #135 (as indicated by the values $G_{.j1} = 1$ for roll-call GMP and 100 for $|b_j|$) and in two dimensions for four additional roll calls [for which $G_{.j2} = 1$ (and $|b_{j2}| = 100$; see Appendix B)]. Roll call #288 has by far the highest value of the ratio $|b_{j2}/b_{j1}|$. Not surprisingly, the three roll calls with the highest $|b_{j2}/b_{j1}|$ ratios are the same three whose $|g_{j2}|$ values are greatest, thus suggesting further that these three contribute heavily to the second dimension.

Improvement in legislator GMP upon adding the second dimension (as measured by $G_{i.2}/G_{i.1}$ or its logarithm) is greatest for Paul, second greatest for Kaptur, and least for Cantor and Levin. Cantor has the largest GMP in both one and two dimensions; Paul has the lowest $G_{i.1}$, but Kirk and Matheson have the lowest $G_{i.2}$ values. Overall GMP's, shown in the lower right corner of Table 1, are $G_{..1} = 0.658$ for one dimension and $G_{..2} = 0.879$ for two.

4.5. Handling of Missing Votes

Before the calculations described in Section 4.1 can run, there must be values for all *IJ* elements of $\mathbf{Y}_0$; Section 4.1 assumed no missing data. In practice, though, one must generally deal with missing votes (although how they are handled may be unimportant if their percentage is small). To get PCA started, one has to assign values for the holes in the data.

Our basic concept is that, if y_{0ij} is missing, we set it equal to a value between 0 and 1 that represents the estimated probability that the (i, j) vote is yea. One might get these values through different means, but our routine is as follows:

1　　Obtain a preliminary $\mathbf{Y}_0$ by setting y_{0ij} equal to the party mean on roll call j if the (i, j) vote is missing. This party mean is the proportion of yea votes to total (yea plus nay) votes on roll call j among those legislators in the same party as legislator i. (In the absence of party data, one could use the proportion of yea votes to total votes on roll call j among all legislators who voted.)

2　　Use the preliminary $\mathbf{Y}_0$ to run a principal-components computation in the same manner as indicated in Section 4.1. The first-dimension scores that result will constitute a preliminary $\mathbf{x}_1$.

3　　Separately for each roll call j, feed this preliminary $\mathbf{x}_1$ into a logistic regression based on model (1) to obtain preliminary (a_j, b_j) values, using the same procedure as in Section 4.2.

4　　Separately for each legislator i, feed these preliminary (a_j, b_j) pairs into another logistic regression, also based on model (1) but with x_{i1} to be solved for and the (a_j, b_j) values supplied rather than the reverse. The resulting values of x_{i1} form a second (and more refined) preliminary $\mathbf{x}_1$. The calculation for a legislator i is based only on those roll calls j for which $n_{ij} = 1$, that is, for which y_{0ij} is not missing. Because a_j is given, there is no intercept term to be solved for, and so a_j is treated as an offset variable (for which PROC LOGISTIC of SAS® makes provision). Any legislator i who, except for missed roll calls, votes yea (nay) on every roll call with $b_j > 0$ and nay (yea) on every one with $b_j < 0$ is given the value $x_{i1} = H_0$ ($x_{i1} = -H_0$) and is excluded from the logistic-regression calculation upon being detected beforehand. (We use $H_0 = 3$.) The exclusion for such an "extreme" legislator is necessary because otherwise x_{i1} would be unbounded.

5　　For any (i, j) for which $n_{ij} = 0$, use the preliminary (a_j, b_j) values together with x_{i1} from the second preliminary $\mathbf{x}_1$ to obtain $y_{0ij} = 1 / \left[1 + e^{-(a_j + b_j x_{i1})} \right]$, which estimates the p_{ij} of (1). Use these results to produce a full $\mathbf{Y}_0$, with no empty cells. Starting with this new $\mathbf{Y}_0$, one can run the PCA calculations of Section 4.1 (and then the ones in Sections 4.2 and 4.3).

The imputation procedure just described is obviously not the only way to deal with missing votes. However, it seems relatively sophisticated in that it allows missing votes to have values on a continuum from 0 to 1 (not just 0 and 1 themselves), in addition to using the full non-missing data in determining those values. P-R (whose iterations calculate for one legislator at a time, as well as for one roll call at a time) and CJR [which imputes in each iteration (Clinton et al. 2004a, p. 367)] each have their own ways of taking care of missing votes. Rosas and Shomer (2008) express general concern (but with specific mention of P-R and CJR) that techniques for handling missing votes are open to question because the missingness may not be ignorable. Although any method for attacking missingness will be imperfect, ours appears to be relatively satisfactory. Generally, one would expect that adverse impact of missing votes will be less overall if they are fewer, and will be less for a legislator who has fewer of them. In addition, our scheme for imputation of missing votes may work better for legislatures with high intra-party vote cohesion (as is typical for a parliamentary system) than for those where such cohesion is less.

4.6. Bridging Across Sessions

P-R, or, more exactly, D- or DW-NOMINATE, can span multiple legislative terms that entail changing memberships, and can thus aim to estimate legislators' ideal points over time in a common space (Poole 2005; Poole and Rosenthal 1991, 2001, 2007). With PCA (or for any approach, including P-R), much the same objective can be achieved through estimation based on a standard unbalanced two-way design. Thus, under single-dimension PCA (e.g.) let x_{i1t} be the ideal point that was found for legislator i for time or term t (for those t's in which legislator i did serve). Then the estimate of the "treatment" effect w_i calculated under the linear model

$$\mathrm{E}(x_{i1t}) = w_i + v_t, \tag{3}$$

using the restriction that the "block" (term) effects v_t sum to 0 ($\sum_t v_t = 0$), will serve as an estimate of the ideal point of legislator i over time in a common space. The model (3) is "constant" in that it

allows for no time variation beyond that provided by the v_t's. To allow for linear trends for individual legislators (perhaps a questionable move, since a more complex model applies), just augment (3) by replacing w_i with $(w_{i0} + w_{i1}t)$, and add a second restriction.

Questions can be raised, though, as to how meaningful it is to try to place legislators in a common space over time. Concerns may be minor if the time covers just a few terms. However, if it covers a number of decades, to say nothing of two centuries (e.g., Poole and Rosenthal 1991), then serious doubts may be expressed. For further comments on this issue, see (e.g.) Bailey (2007), Bateman and Lapinski (2016), Cillizza (2014), and Sides (2011).

5. Developments Related to Our PCA Approach

Sections 5.1–5.5 deal, respectively, with other work involving principal components in estimating ideal points; ratings by *National Journal*; factor analysis; a method of Heckman and Snyder (1997); and miscellaneous pursuits, including Bayesian approaches other than CJR.

5.1. Other Use of Principal Components in Ideal-Point Estimation

The possibility of making use of principal components in the estimation of ideal points has not been altogether ignored in the past. One finds some applications whose descriptions are less than totally clear as to what was done. In footnote 4 of Clinton et al. (2004a, p. 359), however, is a detailed description of one technique. The CJR method uses the principal-components estimator only for start values of the ideal points, though (Clinton et al. 2004a, p. 368).

Aside from this confinement to the initialization, there are several differences between the CJR approach of footnote 4 and our PCA approach. First, the vote matrix is double-centered under CJR, whereas with **Y** our PCA follows the usual practice in most applied work with principal components, by adjusting only for the variable (roll-call) means and not for the observation (legislator) means. Second, CJR uses pairwise deletion of missing data in computing its correlation matrix (which could result in negative eigenvalues), whereas we impute for missing votes (with values between 0 and 1, as described above in Section 4.5). Third, rather than a correlation matrix as used by CJR, our PCA uses the covariance matrix, **S**. That is largely because inference results are mostly unavailable with a correlation matrix (e.g., Jackson 1991, sct. 4.7) but can be obtained (see Section 7 below), though with some assumptions, with a covariance matrix. Although a correlation matrix rather than a covariance matrix generally has to be used if variables have different units of measurement, such a consideration hardly applies to a vote matrix with all its values in the set [0, 1]. Fourth, CJR uses roll calls as observations and legislators as variables, rather than the reverse as in our PCA. Thus, CJR obtains its ideal points from the $(I \times 1)$ eigenvectors of its $(I \times I)$ correlation matrix, whereas our ideal points are the scores $\mathbf{x}_k$ $(I \times 1)$ that are derived using **Y** and the eigenvectors of our $(J \times J)$ covariance matrix.

One might think that there should be no relation at all between the eigenvectors of the $(I \times I)$ matrix and the score vectors that are based on the eigenvectors of the $(J \times J)$ matrix. By virtue of singular value decomposition (Jolliffe 2004, pp. 44–45), though, the two can be the same except for a multiplicative constant, but generally just under certain restrictive conditions. Specifically, both matrices have to be covariance (not correlation) matrices, and, in essence, both matrices have to originate from a double-centered vote matrix. Of course, the two $(I \times 1)$ vectors could be similar in some cases even if these conditions do not both hold.[6]

---

[6] With the roll calls as variables and the legislators as observations as in our PCA approach, one could ask whether the list of variables might be augmented to include, besides the roll calls, some covariates that would be legislator attributes (e.g., party). No attempt has been made to study that possibility or how it might be used, but it could lead to some interesting applications. For a covariance (rather than correlation) matrix to be used, though, an attribute might need to meet certain conditions, such as confinement to the interval [0, 1].

5.2. The Ratings from National Journal

From selected US Congressional roll-call votes from each year starting with 1981 (and continuing at least through votes from 2013), *National Journal* calculated ideological ratings for all members of the Senate and House. The methodology was explained each year (e.g., Anonymous 2014a) and ever since the beginning was described (briefly) as using "principal-components analysis". However, mathematical details of that analysis are absent; because one can do principal components analyses in differing ways, the ratings methodology is hard to appraise.

The *National Journal* ratings have found their way into political campaigns. They placed presidential candidate John Kerry (during the 2004 campaign) as the most liberal of all senators and Barack Obama (during the 2008 campaign) also as the most liberal senator, thereby sparking attacks from political opponents (see, e.g., Harris 2004; Montopoli 2008). Clinton et al. (2004b) and Clinton and Jackman (2009, p. 603) found the claims based on the ratings to be overstepping but said nothing about the mathematical basis for the principal components analyses and did not (or could not) evaluate it. More recently, North Carolina Senator Kay Hagan in her 2014 losing reelection campaign advertised herself as "the most moderate senator" based on the *National Journal* ratings (Christensen 2014).

5.3. Relation to Factor Analysis

Unfortunately, in the broad literature there has been much disagreement as to what constitutes factor analysis as well as great confusion between it and principal components analysis, with (e.g.) many authors saying that they are using the former when they are really using the latter (Jackson 1991, scts. 17.1, 17.10; Jolliffe 2004, p. 150). In addition, there are myriad varieties of (true) factor analysis, with different techniques for both parameter estimation and factor rotation, as well as for determination of scores. Published reports may give unclear or inadequate details of methods, or use ambiguous language.

For ideal-point estimation as well as more generally, the question may arise as to how much similarity there is between results from (true) factor analysis and those from (true) principal components analysis. A proposition proved by Bentler and Kano (1990) states that, under mild conditions, if a J-variate vector follows the factor-analysis model with just a single factor, then the squared correlation coefficient between that factor and the first-dimension score from principal components analysis will approach 1 as J increases. This result suggests that, for a large number (J) of roll calls, first-dimension ideal points estimated by principal components analysis and factor analysis may be quite close if the issue space is strongly one-dimensional. However, the result is limited since it just applies when there is only one factor.

Thus, in more than one dimension, principal components analysis and factor analysis will generally give results that are different. That is not to say, of course, that the former produces better outcomes than the latter (or vice versa).

5.4. The Heckman-Snyder Method

Heckman and Snyder (1997) estimate ideal points through a factor-analytic approach. The method is often mentioned but has not seen extensive use. It resembles PCA more closely than either P-R or CJR does. Its properties include the following:

1   It uses an $I \times I$ matrix (rather than $J \times J$, as in PCA).
2   It apparently uses some sort of randomization to handle missing votes (Heckman and Snyder 1997, p. S160, footnote 13), a practice that seems questionable.
3   It pays little regard to estimation of roll-call parameters (which are sometimes desired).
4   It uses an unusual distributional assumption in relation to its utility function.
5   It provides no standard errors.

For further comments on this method, especially some critical ones regarding the fourth property above, see Poole and Rosenthal (2001) and Clinton et al. (2004a). P-R, CJR, and Heckman-Snyder each use a particular utility function of their own (and have never considered any approach that refrains from using one). By contrast, PCA makes no use of a utility function and does not need to do so.

5.5. Further Endeavors, Bayesian and Other

Estimation of ideal points through Bayesian approaches other than CJR has been more specialized than CJR. Bailey (2001) focused on situations with a tiny number of votes, utilized covariates, and provided an example based on five US Senate roll calls dealing with international trade. Martin and Quinn (2002) and Bafumi et al. (2005) each examined US Supreme Court decisions over a span of more than 45 years and estimated ideal points for 29 justices, with allowance for temporal change in the case of the former paper. All three of these Bayesian works dealt only with a one-dimensional issue space.

For various possible extensions of principal components analysis, see Jolliffe (2004).

6. Large Empirical Examples Using PCA

Because Example 1 in Section 4.4 was small, it could illustrate various PCA details. Examples 2–5 for PCA, which we now present, are large but less comprehensive. They illustrate various aspects of application of our PCA technique. There are comparisons with P-R, and also some with CJR. Section 6.1 deals with general results, and Section 6.2 with model fit. The comparisons mostly show that PCA differs little from P-R or CJR, thus suggesting close equivalence of PCA with the other two from the standpoint of their results.

For details about data, see Appendix C.[7]

6.1. General Results for Examples 2–5

All four examples use US Senate roll calls, from the 105th Congress (1997–1998) for Examples 2 and 4 and from the 106th Congress (1999–2000) for Examples 3 and 5. To be consistent with previous P-R and CJR published results and thus allow proper comparisons, Examples 2 and 3 are based on all roll calls except those with a vote more extreme than 97.5% to 2.5%: 486 such roll calls for the former example, 540 for the latter. Examples 4 and 5 are not nearly as large, and use, respectively, the 23 and 25 "key votes" selected by Congressional Quarterly (1998, 1999 and 2000, 2001) from among the 486 and 540.

In Example 2, the PCA x_{i1}'s of all Democrats are less than (to the left of) those of all Republicans. The same is true for the P-R x_{i1}'s in Example 2, as well as for the PCA x_{i1}'s in Examples 3 and 5. It is true also for the P-R x_{i1}'s in Example 3 except that Lincoln Chafee (R) of Rhode Island is to the left of Miller (D) of Georgia, and for the PCA x_{i1}'s in Example 4 except that John Chafee (R) of Rhode Island is left of Hollings (D) of South Carolina and Breaux (D) of Louisiana.

The PCA eigenvalue ratios L_2/L_1 are 0.07, 0.05, 0.19, and 0.09 for Examples 2–5, respectively. All are far below the 0.75 ratio for Example 1 (Section 4.4) and thus show smaller roles for the second dimension. In Examples 2–5 the respective percentages of PCA total variability accounted for by the first dimension are 54.8, 63.6, 46.3, and 66.5, and by the second dimension are 4.0, 3.1, 8.9, and 5.7.

For 65 of the 100 senators in Example 2, the PCA two-dimension GMP, $G_{i.2}$, exceeds the corresponding value for P-R. In Example 3, these GMP's are higher for PCA than for P-R for 74 of the 102 senators. Although the 65 senators in Example 2 are disproportionately Republicans, the 74 in Example 3 are disproportionately Democrats.

For PCA in Example 2, the highest $G_{i.2}$ value, 0.869, is for Sessions of Alabama, and the lowest, 0.613, is for Byrd of West Virginia. P-R GMP's are likewise largest for Sessions and least for Byrd.

---

[7] Besides Examples 2–5, we also have one more example, Example 6. It is for the 90th U.S. Senate (1967–1968) and is used only in Section 7 below.

Table 2 shows correlation coefficients of location scores for both the first dimension (x_{i1}'s, top half of table) and second dimension (x_{i2}'s, bottom half) and for both the 105th Senate (left half of table) and the106th (right half). For the 105th Senate, the correlations are among the senators' scores from PCA using just the 23 key votes (PCA/23); from PCA using all 486 non-lopsided votes (PCA/486); and from P-R (P-R/486, also using these 486 votes). For the 106th Senate, PCA/25, PCA/540, and P-R/540 are analogous, respectively, to PCA/23, PCA/486, and P-R/486. In each 3×3 square in the table, Pearson (product-moment) and Spearman (rank-order) correlation coefficients are, respectively, below and above the main diagonal.

Table 2. Pearson and Spearman correlation coefficients (lower left and upper right triangle, respectively, in each 3×3 square) for senator location estimates, among "key vote" PCA, full PCA, and Poole-Rosenthal W-NOMINATE, for one and two dimensions and 105th and 106th Senates.

	105th Senate				106th Senate		
First dimension (x_{i1}'s)							
	PCA/23	PCA/486	P-R/486		PCA/25	PCA/540	P-R/540
PCA/23	—	0.959	0.957	PCA/25	—	0.936	0.919
PCA/486	0.981	—	0.995	PCA/540	0.990	—	0.990
P-R/486	0.983	0.995	—	P-R/540	0.989	0.996	—
Second dimension (x_{i2}'s)							
	PCA/23	PCA/486	P-R/486		PCA/25	PCA/540	P-R/540
PCA/23	—	0.702	0.729	PCA/25	—	0.643	0.638
PCA/486	0.679	—	0.718	PCA/540	0.685	—	0.930
P-R/486	0.719	0.670	—	P-R/540	0.687	0.906	—

The x_{i1}'s, x_{i2}'s are the legislator scores for dimension 1, 2; PCA/23, PCA/25 refer to the 23, 25 "key votes" in the 105th, 106th Senates; PCA/486 and P-R/486, PCA/540 and P-R/540 refer to the 486, 540 non-lopsided votes in the 105th, 106th Senates.

In the comparisons of P-R x_{i1} scores versus those of PCA based on all (486 or 540) votes, all (four) of the correlation coefficients, both Pearson and Spearman and for both Senates, are greater than 0.99. Table 2 also shows that even the correlations involving the PCA x_{i1} scores derived from the key votes, though lower, are still high (all greater than 0.9), despite the small numbers of key votes that provide the basis for the corresponding scores.

The second-dimension correlations of PCA/540 versus P-R/540 (106th Senate) are high. However, otherwise the correlations for x_{i2} that appear in Table 2, though well above zero, are not very large, perhaps an indirect result of weakness of the second dimension.

The first dimension in each of Examples 2–5 is obviously related to party and to traditional "left"-"right" factors. The second dimension, though, is rather elusive and not easy to pin down. However, the PCA $\mathbf{g}_2$ eigenvectors can shed at least some amount of light in Examples 2 and 3.

For Example 2 (105th Senate), 19 roll calls have $|g_{j2}| > 0.1$. (The value 0.1 was picked arbitrarily.) All but two of those roll calls are budget or appropriations votes. On all 19, the Democrats are at least 80% united and the Republicans are less than 80% united; in fact, on all but three the GOP senators are no more than 2/3 united. Broadly speaking, Republicans who tend to vote with the Democrats on these roll calls are at one end of the x_{i2} scale, whereas those who tend not to do so are at the other end. The senators with x_{i2} ranks of 78 through 100 are, except for Lieberman (D) of Connecticut, all Republicans in the former group, whereas those with ranks 1 through 12 are, except for Feingold (D) of Wisconsin, all GOP senators in the latter group. Senators with extreme x_{i2} scores tend also to have high $G_{i.2}/G_{i.1}$ ratios.

For Example 3 (106th Senate), the picture for the second dimension is far different. All 29 of the roll calls with $|g_{j2}| > 0.1$ deal with some aspect of foreign trade (CQ votes #54, 178, 213, 344, 346, 348–350, and 352–353 in 1999; and #97–98, 231, 234–236, 238–246, and 248–251 in 2000). On the x_{i2} scale, senators on one end favor free trade whereas those on the other end are protectionist. The difference in the complexion of the second dimension in the two Senates stems from the fact that the 106th but not the 105th has a large number of votes related to foreign trade. (Generally, of course, the results of any method of roll-call analysis will be fundamentally affected by the nature of the votes in the data set.)

Section 7 below will provide some suggestion that a third dimension could play a role in the 106th Senate (though not in the 105th). In line with this, in the 16 roll calls in Example 3 for which $|g_{j3}| > 0.1$, Democrats are at least 93% united in all but one, whereas Republicans are less than 2/3 united in all but three. The pattern is much the same as for the second dimension in Example 2, except that the 16 votes cover a medley of issues rather than mainly budgets and appropriations.

For one dimension, results for cut points (analogous to P-R midpoints) may be of interest. Of the 486 roll calls in Example 2, 106 have cut points (values of m_j) that are outside the range of senators' x_{i1}'s (below -0.764 or above $+0.673$ for this case). The corresponding figures for Examples 3–5 are, respectively, 130 out of 540 roll calls, six out of 23, and four out of 25. For such roll calls with cut points outside the range of the x_{i1}'s, the senators' one-dimension p_{ij}'s are either all greater than 1/2 or all less than 1/2. These roll calls generally have poor single-dimension model fit and/or a vote that is not at all close.

Time to run PCA is minimal. We did no analyses with US House roll-call data, which involves a much larger matrix than US Senate data. However, for a dummy vote matrix with $I = 450$ legislators and $J = 1250$ roll calls (comparable in size to a roll-call matrix for one two-year period for the US House), computer time (clock time) to run the single line of code at the end of Section 4.1 above was less than 14 seconds. With the code changed to extract just one dimension rather than all 449, the time was under five seconds. These times are for an Intel dual-core processor running at 2.4 gigahertz with 2.00 gigabytes of random-access memory.

6.2. Measures of Model Fit

Section 4.3 above defined GMP. Other possible measures of model fit are also available. They include percentage of correct classifications or %CC (Poole and Rosenthal 2007, p. 33) and aggregate proportional reduction in error or APRE (Poole and Rosenthal 2007, pp. 36–37). All the measures are applicable to each of the techniques of ideal-point estimation that we consider here and can be used to compare them.[8]

GMP can be seen as more sensitive than the other two measures. That is, the other two differentiate less than GMP. For example, consider a roll call whose yeas and nays in the order of the spatial locations of the legislators (in a one-dimensional issue space) are

$$\text{N N Y Y N Y N Y N N N N N N.}$$

No matter what the estimation technique, APRE (calculated for this roll call by itself) could not have a value greater than zero, because no placement of the cut point separating predicted yeas from predicted nays could do better than predicting nays for all 13 legislators. The measure %CC is similarly constrained. However, GMP can yield varied results for different estimation techniques.

Table 3 presents measures of model fit (GMP, %CC, and APRE) that compare PCA, CJR, and P-R, for both one and two dimensions and for both the 105th and 106th Senates (see end of Appendix C for

---

[8]   The last two measures can be illustrated for Example 1 (Section 4.4). Their values are, for one and two dimensions respectively, $1 - 27/111 = 0.757$ and $1 - 6/111 = 0.946$ for % CC and $1 - 27/45 = 0.400$ and $1 - 6/45 = 0.867$ for APRE. Here, 111 is the total number of votes cast, of which 45 were on the losing side; and 27 ($k = 1$) and 6 ($k = 2$) are the numbers of incorrect predictions (classifications), based on the sign of the estimate of (1) or (2) disagreeing with the vote.

some details). Differences among the results from the three methods are small. On its four available comparisons, CJR is better than both PCA and P-R on the two for GMP, and better than PCA but about the same as P-R on the two for %CC. PCA is better than P-R on all four comparisons for GMP, and worse than P-R or about the same on the eight comparisons for %CC and APRE. The various differences are so narrow, though, that they seem to be inconsequential.

Table 3. Measures of model fit for PCA, CJR, and P-R, for results in one and two dimensions, 105th and 106th Senates.

	105th Senate 486 Roll Calls, 100 Senators			106th Senate 540 Roll Calls, 102 Senators	
	PCA	CJR	P-R	PCA	P-R
One dimension					
GMP	0.758	0.762	0.751	0.791	0.781
% CC	87.5%	87.8%	87.9%	89.8%	90.1%
APRE	0.628	—	0.642	0.711	0.720
Two dimensions					
GMP	0.775	0.780	0.771	0.813	0.808
% CC	88.5%	88.6%	88.6%	91.1%	91.1%
APRE	0.659	—	0.662	0.749	0.748

GMP = Geometric mean probability (see Section 4.3); % CC = Percentage of correct classifications; APRE = Aggregate proportional reduction in error.

7. Number of Dimensions

Our PCA treatment throughout Sections 4 and 6 provides only descriptive results and thereby avoids any need for distributional assumptions. In this section, though, our PCA methodology entails a model based on continuous variables, whereas votes (if not missing) are, of course, binary. Specifically, the model assumes that each legislator's set of votes (i.e., each row of the matrix $\mathbf{Y}_0$) is drawn independently from a multivariate normal distribution. In addition, the theoretic results rely on large-sample (asymptotic) distributions of the relevant statistics. Our methods here in Section 7 are suitable to the extent that assumption violations (e.g., votes being binary) do not have serious effects. Our results below, though, suggest that the methodology does work well. (One can speculate that the binariness is less of an issue with larger sample sizes.)

There can be controversy over what number of dimensions to use in estimating ideal points from roll-call data. In particular, for the US Congress Heckman and Snyder (1997, pp. S165, S184) contended that the number should be much higher than the one or two favored by Poole and Rosenthal (1991) and others. Questions involving number of dimensions are not easy to judge. PCA, however, can at least furnish some clues.

A result for principal components (e.g., Anderson 1963, pp. 130–33; Morrison 1976, p. 294; Jackson 1991, pp. 86–87) provides a test of the null hypothesis that the k_0-th through k_1-th population eigenvalues (inclusive) are equal to one another. The test uses only the sample eigenvalues, L_k. One refers

$$(I-1)\left[(k_1-k_0+1)\log\left(\frac{\sum_{k=k_0}^{k_1} L_k}{k_1-k_0+1}\right)-\sum_{k=k_0}^{k_1}\log L_k\right] \tag{4}$$

to the chi-square distribution with $(k_1 - k_0 + 3)(k_1 - k_0)/2$ degrees of freedom. The statistic (4) will suggest that the k_0-th through k_1-th population eigenvalues are not all alike if (4) is significantly large,

or are all about the same if (4) is nonsignificant. Because the rough equality of eigenvalues for some dimensions $k \geq k_0$ would generally signal that dimensions $k \geq k_0$ are of little benefit and should not be kept, the use of (4) may help to decide how many dimensions are appropriate to use.

The general concept is that, if the list of successively lower sample eigenvalues reaches a point where no further eigenvalues differ much, then the dimensions starting at that point probably have little meaning and can be dropped. Using (4) is thus similar to using the classical scree graph (e.g., Jolliffe 2004, p. 115 ff.). However, the latter may involve extra subjective judgment whereas the former provides a more formal statistical criterion for judging how many dimensions to retain. For PCA or any other approach, one could use model fit to try to assess dimensionality as in (e.g.) Poole and Rosenthal (2007, pp. 63–64), but that is also more subjective than using (4).

For selected (k_0, k_1) and the associated degrees of freedom, Table 4 shows the value of (4), along with its chi-square probability, for Examples 2–6 from the US Senate. As indicated before, Examples 2 and 3 involve all the non-lopsided votes, and Examples 4 and 5 each use a small set of "key" votes, from the 1997–1998 and 1999–2000 terms. That period of time, unlike that of Example 6, is generally thought to have strongly unidimensional voting patterns.

Table 4. Chi-square values from (4) along with their upper-tail probabilities, for Examples 2-6, for certain (k_0, k_1) and corresponding degrees of freedom.

			Example 2 105th Senate 1997–1998 486 Votes		Example 3 106th Senate 1999–2000 540 Votes		Example 4 105th Senate 1997–1998 23 Key Votes		Example 5 106th Senate 1999–2000 25 Key Votes		Example 6 90th Senate 1967–1968 518 Votes	
k_0	k_1	d.f.	χ^2	Prob.	χ^2	Prob.	χ^2	Prob.	χ^2	Prob.	χ^2	Prob.
1	2	2	136.88	0.0000	174.98	0.0000	61.25	0.0000	125.04	0.0000	35.69	0.0000
1	3	5	275.41	0.0000	319.92	0.0000	119.39	0.0000	227.36	0.0000	99.16	0.0000
1	4	9	416.50	0.0000	480.21	0.0000	171.35	0.0000	336.01	0.0000	171.64	0.0000
1	5	14	546.18	0.0000	639.28	0.0000	220.57	0.0000	443.54	0.0000	239.67	0.0000
2	3	2	5.84	0.0539	0.62	0.7334	2.90	0.2341	0.92	0.6304	13.89	0.0010
2	4	5	17.13	0.0043	10.35	0.0658	6.67	0.2465	7.71	0.1730	34.10	0.0000
2	5	9	27.01	0.0014	23.48	0.0052	11.71	0.2302	17.51	0.0413	52.66	0.0000
3	4	2	2.76	0.2519	5.66	0.0590	0.62	0.7329	3.29	0.1930	3.52	0.1717
3	5	5	5.24	0.3870	13.10	0.0224	2.19	0.8225	8.42	0.1346	7.28	0.2004
3	6	9	8.68	0.4677	22.32	0.0079	5.20	0.8168	14.02	0.1215	12.69	0.1770
4	5	2	0.24	0.8883	1.21	0.5469	0.47	0.7892	1.04	0.5954	0.47	0.7910
4	6	5	1.15	0.9494	3.84	0.5733	1.91	0.8618	2.64	0.7549	2.08	0.8379

The value of χ^2 tests for the equality of the k_0-th through k_1-th population eigenvalues (inclusive); d.f. = Degrees of freedom = $(k_1 - k_0 + 3)(k_1 - k_0)/2$; Prob. = Probability of finding a value of χ^2 greater than the one shown.

Although some caution is needed because of multiple testing, one can draw several conclusions of varying tenability. First, consider Examples 2–5. Not only for the full-vote data (Examples 2 and 3) but also for the key votes (Examples 4 and 5), the probabilities for the basic statistic (4) with $k_0 = 1$ show decisively that the first population eigenvalue differs from the others, thus (not surprisingly) indicating a strong first dimension. For Example 2, the low probabilities for $k_0 = 2$ and the nonsignificant ones for $k_0 = 3$ suggest acknowledging a second (albeit weak) dimension but not a third. For Example 3, though, the high probability for $(k_0, k_1) = (2, 3)$ combined with the low remaining probabilities for $k_0 = 2$ and $k_0 = 3$ suggests that the second and third population eigenvalues differ from those for higher dimensions but perhaps not greatly from each other—a condition consistent with recognizing both a second and a third dimension (see Section 6.1 above for related discussion). The results for $k_0 = 2$ and 3 for Examples 4 and 5 are largely nonsignificant, though their patterns slightly resemble the ones in

Examples 2 and 3, respectively. The results for $k_0 = 4$ in the different examples provide no evidence for a fourth dimension. All of the results for Examples 2–5 are credible.

Example 6 pertains to the 90th Senate (1967–1968). This Senate was chosen for inclusion in Table 4 because it was cited by Lewis and Poole (2004, p. 106) as one with "two dominant dimensions". In conformity with that description, the probabilities for $k_0 = 2$ are far lower than in Examples 2–5. Thus, the findings for the 90th Senate reinforce other conclusions from Table 4 regarding usefulness of (4) for judging dimensionality.[9]

8. Discussion

In comparison with alternatives, our PCA approach to estimating ideal points from roll-call data is simple in both concept and implementation, and its computation is fast. Together with its simplicity come less programming and easier understanding. It evidently has face validity, based on the results in Sections 6 and 7. It also avoids the difficulties (noted in Section 3) that, especially for more than one dimension, affect both P-R and CJR, with the former facing more issues than the latter. Why, then, is principal components analysis seldom used for ideal points—either in one dimension, where applications are more frequent, or in more than one, where its advantages are greater? Outside of plain inertia and adherence to the status quo, two considerations may be playing a role. We do not find either one to be generally compelling.

First, PCA provides no means to assess uncertainty. CJR uses Bayesian methodology to estimate uncertainty for estimates of ideal points and other parameters. Lewis and Poole (2004) proposed parametric-bootstrap standard errors to handle uncertainty assessment for P-R. Both the P-R and CJR techniques, however, rely on an ungrounded assumption of (mutual) conditional independence of a legislator's votes given the ideal points and roll-call parameters. Although for PCA one may try to find complex standard-error formulas that steer clear of that assumption, it appears that, whether for P-R, CJR, or PCA, any effort to assess uncertainty is fraught with impediments. In addition, standard errors may see infrequent use in applications anyway.

Second, unlike P-R and CJR, PCA does not use a utility function in deriving its model. However, in its theory it is still based on a spatial voting model, just as P-R and CJR are. Parts A.1 and A.3 of Appendix A show a close mathematical similarity between P-R, which uses a utility function, and the model (1)–(2) above, which does not. Table 2 indicates that PCA and P-R yield first-dimension ideal points that are barely distinguishable. This close parallel between PCA and P-R, both theoretic and pragmatic, suggests that the lack of a utility function for PCA need not be a general concern. Principal-components methods have, of course, seen use in various fields of application.

9. Summary

For analysis of roll-call data, this paper builds a case for considering our PCA approach as an alternative to P-R and CJR, two well-established methods. P-R has been used for years and is deeply entrenched. CJR has made recent inroads.

For unidimensional applications anyway, many users may thus hesitate to lightly eschew P-R (or CJR) and embrace PCA instead, despite the strong points of PCA even in one dimension. However, for two or more dimensions, whose study may be fruitful for varied situations (e.g., for certain locales and time periods or certain subsets of votes or legislators), the relative benefits of PCA are especially striking and should suffice to make PCA a preeminent contender.

---

[9]　Also consonant with a relatively strong second dimension in the 90th Senate are its eigenvalue ratio L_2/L_1, equal to 0.29, and its GMP gain from the second dimension, $G_{.2} - G_{.1} = 0.711 - 0.665 = 0.046$, both much larger than the respective values, for the 105th and 106th Senates (Examples 2 and 3), of 0.07 and 0.05 for L_2/L_1, and of $0.775 - 0.758 = 0.017$ and $0.813 - 0.791 = 0.022$ for $G_{.2} - G_{.1}$.

Appendix A. Some Mathematical Details Concerning P-R

This appendix deals mainly with the mathematical relationship of P-R to PCA and CJR. In so doing, it concentrates on comparing P-R with the model of (1) and (2) above. In Parts A.1–A.5 we make many comparisons but also need to cover some details of the P-R approach in order to do so. Part A.6 deals with nonidentifiability.

Appendix A.1. P-R for One Dimension

In an early P-R model with one dimension (Poole and Rosenthal 1985, p. 361, Equation (3)), p_{ij} (the probability that legislator i votes yea on roll call j) takes the form

$$p_{ij} = \frac{e^{\beta e^{-1/2\omega^2(x_{i1}-z_{j1})^2}}}{e^{\beta e^{-1/2\omega^2(x_{i1}-z_{j1})^2}} + e^{\beta e^{-1/2\omega^2(x_{i1}-z_{j0})^2}}},$$ (A1)

where β and ω are parameters to be estimated, x_{i1} is (as before) the spatial position of legislator i, and z_{j1} and z_{j0} are the respective positions of the yea and nay outcomes for roll-call vote j. A later one-dimensional model (McCarty et al. 1997, Equation (A-3)); (Poole and Rosenthal 2001, Equation (A3)); (Poole 2005, Equation (4.11)) assumes the form

$$p_{ij} = \Phi(\beta[e^{-1/2\omega^2(x_{i1}-z_{j1})^2} - e^{-1/2\omega^2(x_{i1}-z_{j0})^2}]),$$ (A2)

where $\Phi(\bullet)$ is the cumulative distribution function of the normal distribution with zero mean and unit variance. Note, though, that experimentation led to the conclusion that the parameter ω in a p_{ij} formula could effectively be dropped (McCarty et al. 1997, p. 53; Poole and Rosenthal 1997, pp. 235, 249).

Though it may not seem so at first glance, the model Equation (1) above does bear mathematical resemblance to (A1) and (A2). Suppose that the exponential utility function in Equation (2) of Poole and Rosenthal (1985, p. 361) is replaced by a quadratic one by dropping "exp", and also dropping β and ω^2, so that the right side of that equation takes the form $[-2(x_i - z_{jv})^2 + \varepsilon_{ijv}]$, where $v = 0$ or 1. Then, in place of (A1) above, Equation (3) of Poole and Rosenthal (1985) changes to the form

$$p_{ij} = \frac{e^{-1/2(x_{i1}-z_{j1})^2}}{e^{-1/2(x_{i1}-z_{j1})^2} + e^{-1/2(x_{i1}-z_{j0})^2}},$$ (A3)

from which it follows that

$$\log \frac{p_{ij}}{1-p_{ij}} = -1/2[(x_{i1}-z_{j1})^2 - (x_{i1}-z_{j0})^2] = (z_{j1}-z_{j0})\left(x_{i1} - \frac{z_{j1}+z_{j0}}{2}\right).$$ (A4)

If one defines the midpoint of the roll-call outcome positions as

$$m_j = \frac{z_{j1}+z_{j0}}{2}$$ (A5)

and their difference as

$$b_j = z_{j1} - z_{j0},$$ (A6)

then (A4) becomes

$$\log \frac{p_{ij}}{1 - p_{ij}} = b_j(x_{i1} - m_j),$$ (A7)

which (assuming $b_j \neq 0$) is identical with (1) if one writes

$$a_j = -b_j m_j \qquad \text{or} \qquad m_j = -\frac{a_j}{b_j}.$$ (A8)

The merits of exponential versus quadratic utility functions can be debated. Poole and Rosenthal (2001, p. 9) did note that others (Ladha 1991; Heckman and Snyder 1997; Londregan 2000) have used quadratic functions. Earlier, Poole and Rosenthal (1985, p. 363; 1991, p. 237) had briefly explained their choice of exponential utility and mentioned the possibility of quadratic utility. Carroll et al. (2013) considered a mixture of exponential and quadratic utility. Poole (2001) used a quadratic function, as did CJR (Clinton et al. 2004a). The argument that invokes (A3) in order to get from (A1) to (1) thus appears appropriate.

The close relation between (A1) and (1) can be seen in a second way. From (A1) one obtains

$$\log \frac{p_{ij}}{1 - p_{ij}} = \beta \left[e^{-\frac{1}{2}\omega^2 (x_{i1} - z_{j1})^2} - e^{-\frac{1}{2}\omega^2 (x_{i1} - z_{j0})^2} \right].$$ (A9)

If the first-order Taylor expansion $e^w \approx 1 + w$ is applied twice on the right side of (A9), it then becomes the same as the right side of (A4) after dropping of the terms β and ω^2. Thus, (1) is shown through a different route to approximate (A1) (cf. Carroll et al. 2009, p. 564, also).

Because of the well-known similarity between the logistic and normal distribution functions, it is clear that (A1) and (A2) are close approximations of each other. It thus follows that there is a strong relation of (1) not only with (A1) but also with (A2).

Appendix A.2. Roll Call with No Relation to Ideal Points

From a theoretical standpoint at least, (1) possesses an advantage over (A1) and (A2) for the hypothetical situation where, on roll call j, p_{ij} is the same for all legislators and does not vary with x_i ($p_{ij} = p_{*j}$, say). The model of (1) readily handles this situation by setting $b_j = 0$ and $a_j = \log[p_{*j}/(1 - p_{*j})]$ (cf. also Jackman 2001, p. 229). However, with (A1), (A2), (A3), (A4), and (A7), no choice of the parameters with j subscripts will yield $p_{ij} = p_{*j}$ except in the special case where $p_{*j} = \frac{1}{2}$; in addition, (A5), (A6), and (A8) are uninterpretable if $p_{ij} = p_{*j}$ is required.

The difficulty regarding (A1) and (A2) would remain even without any constraints on the roll-call parameters, and persists even with generalization to more than one dimension. Moreover, the concern may be more than just theoretical, in the sense that, if the p_{ij}'s for a given roll call j are almost but not exactly the same for all legislators i (all x_i), then the yea and nay outcome positions z_{j1} and z_{j0} in (A1) or (A2) can assume values boundlessly far from 0. That condition might lead to added computational complications.

Curiously, recent works differ with respect to whether their model parameterizations resemble the right side of (1) (arguably better) or the right side of (A7). For instance, Jackman (2000, 2001) and Martin and Quinn (2002), as well as CJR (Clinton et al. 2004a), are in the former camp, whereas Bailey (2001) and Bafumi et al. (2005), like P-R, are in the latter.

Appendix A.3. P-R for Two and More Dimensions

If the issue space has two dimensions rather than one, then (A1) and (A2) stay the same except that $\omega^2(x_{i1} - z_{j1})^2$ and $\omega^2(x_{i1} - z_{j0})^2$ are replaced with

$$\sum_{k=1}^{2} \omega_k^2 \left(x_{ik} - z_{jk1} \right)^2 \text{ and } \sum_{k=1}^{2} \omega_k^2 \left(x_{ik} - z_{jk0} \right)^2,$$ (A10)

respectively, where the index k refers to the dimension (cf. McCarty et al. 1997, Appendix A, e.g.). The parameter ω_1 may be set to 1 (McCarty et al. 1997, p. 53) and thus effectively dropped.

For either (A1) or (A2), the total number of parameters to be estimated (excluding ω_1) is $(2I + 4J + 2)$ after expansion to two dimensions, compared with $(I + 2J + 1)$ for one dimension. The number of parameters for PCA (and also for CJR) is $(2I + 3J)$ for two dimensions and $(I + 2J)$ for one. Thus, not counting ω_1, (A1) and (A2) after expansion to two dimensions each have $(J + 2)$ more parameters than (2), but for one dimension they have only one more parameter than (1).

If (A1) is replaced by substituting quadratic for exponential utility *and* by allowing for two dimensions instead of one through substitution of (A10), then (A3) is unchanged except for the substitution of (A10), given that β and both ω_k^2 terms are dropped from the two-dimensional version of (A1). Thus, (A4) now becomes

$$\begin{aligned} \log \tfrac{p_{ij}}{1-p_{ij}} &= -\tfrac{1}{2}\big[(x_{i1} - z_{j11})^2 + (x_{i2} - z_{j21})^2 - (x_{i1} - z_{j10})^2 - (x_{i2} - z_{j20})^2\big] \\ &= (z_{j11} - z_{j10})\Big(x_{i1} - \tfrac{z_{j11}+z_{j10}}{2}\Big) + (z_{j21} - z_{j20})\Big(x_{i2} - \tfrac{z_{j21}+z_{j20}}{2}\Big) \\ &= b_{j1}(x_{i1} - m_{j1}) + b_{j2}(x_{i2} - m_{j2}), \end{aligned} \qquad \text{(A11)}$$

where the m_{jk}'s and b_{jk}'s are defined analogously to (A5) and (A6), respectively. In order for (A11) to be the same as (2), it is necessary to substitute

$$b_{j0} = -b_{j1}m_{j1} - b_{j2}m_{j2} \qquad \text{(A12)}$$

in (A11), thereby reducing the number of parameters for each roll call j from 4 to 3. In fact, if the substitution (A12) is not made, then the m_{jk} parameters in (A11) encounter exacerbated difficulties with identifiability. The necessary decrease in number of parameters is just what one would expect given the results stated in the preceding paragraph.

For the general case of K dimensions, (A1) and (A2) generalize in obvious fashion through substitution of (A10) again but with 2 replaced by K as the upper limit of the summations in (A10). For general K, the number of parameters (with ω_1 excluded) is then $K(I + 2J + 1)$ for the expanded version of either (A1) or (A2). The model (2) as generalized has $(KI + KJ + J)$ parameters for general K (as does CJR—see (Clinton et al. 2004a, p. 357), formula for what they call p). Thus, P-R has $(KJ - J + K)$ more parameters than (2). The use by P-R of the yea and nay roll-call outcome parameters z_{jk1} and z_{jk0} may appear reasonable until one becomes aware of the resulting overparameterization and its consequences.

Appendix A.4. P-R Ideal-Point Constraints

P-R constrains legislators' spatial locations to lie within a circle, sphere, or hypersphere of unit radius for (respectively) two, three, or more dimensions, and from -1 to $+1$ in one dimension. This is covered (e.g.) by Poole and Rosenthal (1997, p. 250); see also Poole (2005, p. 107). CJR and PCA have no such restraints.

The P-R restriction clearly places undue limitations on the locations of the ideal points in two or more dimensions. However, there can be troubles even in one dimension: Some of the ideal points may be amassed at -1 and $+1$. Thus, Clinton and Jackman (2009, pp. 610–11) show a problematic P-R example with 99 roll calls where 15 out of 101 senators are bunched at -1 and $+1$.

In Section 3 we noted a case where 23 out of 102 senators have two-dimensional ideal points that are on the edge of the unit circle. The problem of "rimming", which this example illustrates, has been noted before (e.g., Rosenthal and Voeten 2004).

Suppose, further, that a third dimension were to be tried for this case. Then none of the 23 senators could have a third-dimension coordinate other than 0 if their first two coordinates still put them on the rim of the unit circle. Matters worsen with more added dimensions.

Appendix A.5. P-R Roll-Call Constraints

A P-R roll-call midpoint is the midpoint of the line segment joining the yea and nay outcome positions of the roll call. As with ideal points, P-R constrains estimates of roll-call midpoints to lie within a circle, sphere, or hypersphere of unit radius in more than one dimension and between -1 and $+1$ in one dimension (Poole and Rosenthal 1997, p. 250). Even in one dimension, this can be seen as a peculiar restriction that forces unnatural results. The PCA unidimensional results for examples in Section 6.1 above find many roll calls whose cut points lie below the lowest x_{i1} or above the highest (roughly akin to P-R midpoints lying below -1 or above $+1$). That suggests that many P-R midpoints would fall outside the range of -1 to $+1$ if allowed to do so.

The roles of the P-R yea and nay locations themselves may be even more idiosyncratic. For instance, consider Example 1 (Section 4.4 above). Note that, if the first-dimension PCA roll-call parameters were reported as $z_{j1} = m_j + b_j/2$ and $z_{j0} = m_j - b_j/2$ [which correspond, respectively, to the positions of the yea and nay outcomes as used in (A1) and later equations], then one would find both $|z_{j1}| > 1$ and $|z_{j0}| > 1$ for every roll call except votes 288 and 511 in Table 1. Moreover, even for those two roll calls, only one of the two positions (z_{j1}, z_{j0}) lies between -1 and $+1$, the approximate range of the x_{i1}'s. This result raises a question as to the meaningfulness of the yea and nay positions (z_{j1}, z_{j0}) vis-à-vis (a_j, b_j) or (b_j, m_j). Instability of the estimates of the yea and nay positions under P-R has already been recognized, however (Poole and Rosenthal 1997, p. 245).

Appendix A.6. Nonidentifiability of P-R for Two-Dimensional Version of (A1) or (A2)

By directly using (A1) or (A2) as modified by the replacement (A10) with $\omega_1 = 1$, one can verify the nonidentifiability of P-R two-dimensional ideal points, under transformations involving both ideal-point coordinates. We show that a set of parameters in two dimensions can be transformed to an infinite number of other sets whose likelihood is the same, thus establishing nonestimability and nonidentifiability.[10] For the version of either (A1) or (A2) as modified by the replacement, note first that p_{ij} (and thus also the log likelihood) depends on the expression inside the square brackets of the modified (A2). Then, for any r with $|r| < 1$ and any nonzero q, that expression is unchanged under the transformation from $(x_{i1}, x_{i2}, \omega_2, z_{j10}, z_{j11}, z_{j20}, z_{j21})$ to $(x_{i1\#}, x_{i2\#}, \omega_{2\#}, z_{j10\#}, z_{j11\#}, z_{j20\#}, z_{j21\#})$, where $x_{i1\#} = rx_{i1} - \omega_2(1 - r^2)^{0.5}x_{i2}$, $x_{i2\#} = qx_{i1} + \omega_2 qr(1 - r^2)^{-0.5}x_{i2}$, $\omega_{2\#} = (1 - r^2)^{0.5}/q$, $z_{j1v\#} = rz_{j1v} - \omega_2(1 - r^2)^{0.5}z_{j2v}$ ($v = 0, 1$; i.e., $v = 0$ for $z_{j10\#}$ and $v = 1$ for $z_{j11\#}$), and $z_{j2v\#} = qz_{j1v} + \omega_2 qr(1 - r^2)^{-0.5}z_{j2v}$ ($v = 0, 1$ again). It is easily seen that the *ranking* of legislators based on either set of transformed ideal points ($x_{i1\#}$'s or $x_{i2\#}$'s) can differ from that based on the corresponding original set (x_{i1}'s or x_{i2}'s), thus showing that either transformed set can be substantively different from the original one.

Appendix B. Linear Programming to Detect Complete Separation

Here we continue from Section 4.2 above and show, for model (2), how one can identify complete separation of points. The separation is harder to identify in two dimensions than in one, because it involves a line, rather than a point, that completely separates the positions of the yea and nay voters. To find whether such a line exists for a given roll call j, one solves two linear-programming problems. In the first problem, one maximizes C (the objective function) subject to the constraints

$$(1 - 2y_{0ij})A + (1 - 2y_{0ij})x_{i1}B - C \geq (1 - 2y_{0ij})x_{i2} \text{ for all } i \text{ such that } n_{ij} = 1,$$

where A, B, and C, the variables to be solved for, are not bounded either above or below (a condition that SAS® provides for). In the second problem, one maximizes C subject to

$$(1 - 2y_{0ij})A + (1 - 2y_{0ij})x_{i1}B + C \leq (1 - 2y_{0ij})x_{i2} \text{ for all } i \text{ such that } n_{ij} = 1$$

---

[10] The transformations are other than those that only involve changes to location, scale, or orientation.

If the solution for neither problem yields a positive value of *C*, then for roll call *j* there exists no line $X_2 = A + BX_1$ that completely separates the two-dimensional locations of the yea voters from those of the nay voters, and so the standard logistic-regression calculation for the roll-call parameters can proceed. (X_2 and X_1 relate to x_{i2} and x_{i1}, respectively.) If the first problem yields a positive value for *C*, then the line $X_2 = A + BX_1$ has all yea voters above it and all nay voters below it. One then uses the solution values of *A* and *B* and sets $b_{j0} = -HA$, $b_{j1} = -HB$, and $b_{j2} = H$ (as in Section 4.2, we use *H* = 100). If the second problem yields a positive value for *C*, then the line $X_2 = A + BX_1$ has all nay voters above it and all yea voters below it. One then sets $b_{j0} = HA$, $b_{j1} = HB$, and $b_{j2} = -H$. The routine of this paragraph easily generalizes for more than two dimensions.

If desired, one could do all logistic regressions at the start and then, just for those roll calls for which convergence fails, do the linear programming. Or, for those roll calls, one could even accept the non-convergent (b_{j0}, b_{j1}, b_{j2}) estimates (forgoing any linear programming) and declare perfect model fit; such values differ from the ones above but may entail no other consequences.

Appendix C. Data Details for Examples 2–6[11]

Senators' vote data for PCA came from ftp://voteview.com/dtaord/sen105kh.ord for Examples 2 and 4 (105th Senate) and from ftp://voteview.com/dtaord/sen106kh.ord for Examples 3 and 5 (106th Senate). The vote data for Example 6 (90th Senate), used only for Table 4, came from ftp://voteview.com/dtaord/sen90kh.ord. P-R first- and second-dimension scores, used for calculations for Table 2, came from ftp://voteview.com/junkord/massproduction/s105_bs_1000_2.dat for Example 2 and from ftp://voteview.com/junkord/massproduction/s106_bs_1000_2.dat for Example 3. The P-R two-dimension GMP's ($G_{i.2}$), used for some comparisons with PCA in Section 6.1, came from ftp://voteview.com/junkord/massproduction/s105_nom31_1000.dat for Example 2 and from ftp://voteview.com/junkord/massproduction/s106_nom31_1000.dat for Example 3.

In Table 2 the results for the 106th Senate exclude Miller of Georgia, who was in office for just a short time, and are based only on the other 101 senators. If Miller is included, the four values for PCA/540 versus P-R/540 change from (0.996, 0.990, 0.906, 0.930) to (0.994, 0.990, 0.896, 0.923).

In Table 3, the P-R measures came from Poole (2005, p. 165) for the 105th Senate, and from http://www.voteview.com/c106/fits.htm for the 106th Senate. The PCA measures are, of course, those for Examples 2 (105th Senate) and 3 (106th Senate), and are calculated as indicated in Sections 4.3 and 6.2

Measures for CJR, though, are available only for the 105th Senate and only for GMP and %CC. Their source is Table 1 of Jackman (2001). Although that table does not include GMP's, it does show log-likelihood values, which are −12,965.14 for one dimension and −11,844.19 for two. Thus, the respective GMP's can be calculated as $e^{-12,965.14/47,739} = 0.762$ and $e^{-11,844.19/47,739} = 0.780$ (where 47,739 is the total of the yeas and nays across the 486 roll calls, and thus reflects 861 missed votes out of a potential 48,600 for the 100 senators who served in the 105th Senate).

References

Albert, Adelin, and John A. Anderson. 1984. On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* 71: 1–10. [CrossRef]

Aldrich, John H., Jacob M. Montgomery, and David B. Sparks. 2014. Polarization and ideology: Partisan sources of low dimensionality in scaled roll call analyses. *Political Analysis* 22: 435–56. [CrossRef]

Anderson, Theodore Wilbur. 1963. Asymptotic theory for principal component analysis. *Annals of Mathematical Statistics* 34: 122–48. [CrossRef]

---

[11]　The data sources given below are what we actually used but apparently have recently become unavailable. Almost all of them seem to be still available through a different connection, https://legacy.voteview.com/dwnl.htm, under the headings "2. W-NOMINATE" and "Roll Call Data".

Anonymous. 2014a. How the vote rankings are calculated. *National Journal*, 30–31, February 8.

Anonymous. 2014b. Powering down: Voters have chosen change, but America's political system makes that far too hard. *The Economist*, November 8, pp. 25–26, 29.

Bafumi, Joseph, Andrew Gelman, David K. Park, and Noah Kaplan. 2005. Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis* 13: 171–87. [CrossRef]

Bailey, Michael. 2001. Ideal point estimation with a small number of votes: A random-effects approach. *Political Analysis* 9: 192–210. [CrossRef]

Bailey, Michael A. 2007. Comparable preference estimates across time and institutions for the Court, Congress, and Presidency. *American Journal of Political Science* 51: 433–48. [CrossRef]

Basu, Asit P. 1983. Identifiability. In *Encyclopedia of Statistical Sciences*. Edited by Samuel Kotz, Norman L. Johnson and Campbell B. Read. New York: Wiley, vol. 4, pp. 2–6.

Bateman, David A., and John Lapinski. 2016. Ideal points and American political development: Beyond DW-NOMINATE. *Studies in American Political Development* 30: 147–71. [CrossRef]

Bentler, Peter M., and Yutaka Kano. 1990. On the equivalence of factors and components. *Multivariate Behavioral Research* 25: 67–74. [CrossRef] [PubMed]

Bornschier, Simon. 2010. The new cultural divide and the two-dimensional political space in Western Europe. *West European Politics* 33: 419–44. [CrossRef]

Carmines, Edward G., Michael J. Ensley, and Michael W. Wagner. 2012. Who fits the left-right divide? Partisan polarization in the American electorate. *American Behavioral Scientist* 56: 1631–53. [CrossRef]

Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2009. Comparing NOMINATE and IDEAL: Points of difference and Monte Carlo tests. *Legislative Studies Quarterly* 34: 555–91. [CrossRef]

Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2013. The structure of utility in spatial models of voting. *American Journal of Political Science* 57: 1008–28. [CrossRef]

Caughey, Devin, and Eric Schickler. 2016. Substance and change in Congressional ideology: NOMINATE and its alternatives. *Studies in American Political Development* 30: 128–46. [CrossRef]

Christensen, Rob. 2014. Is Hagan an Obama moderate? *News & Observer*, (Raleigh, N. C.), September 10, p. 1B.

Cillizza, Chris. 2014. Is Obama the Most Liberal President Ever? February 4. Available online: http://www.washingtonpost.com/blogs/the-fix/wp/2014/02/04/is-barack-obama-the-most-liberal-president-ever/ (accessed on 29 May 2015).

Clinton, Joshua D., and Simon Jackman. 2009. To simulate or NOMINATE? *Legislative Studies Quarterly* 34: 593–621. [CrossRef]

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004a. The statistical analysis of roll call data. *American Political Science Review* 98: 355–70. [CrossRef]

Clinton, Joshua D., Simon Jackman, and Doug Rivers. 2004b. The most liberal senator? Analyzing and interpreting congressional roll calls. *PS: Political Science and Politics* 37: 805–11.

Congressional Quarterly. 1998. Appendix C: Key votes. In *Congressional Quarterly 1997 Almanac*. Washington: Congressional Quarterly, Inc., vol. 53.

Congressional Quarterly. 1999. Appendix C: Key votes. In *Congressional Quarterly 1998 Almanac*. Washington: Congressional Quarterly, Inc., vol. 54.

Congressional Quarterly. 2000. Appendix C: Key votes. In *Congressional Quarterly 1999 Almanac*. Washington: Congressional Quarterly, Inc., vol. 55.

Congressional Quarterly. 2001. Appendix C: Key votes. In *Congressional Quarterly 2000 Almanac*. Washington: Congressional Quarterly, Inc., vol. 56.

Congressional Quarterly. 2007. Appendix C: Key votes. In *Congressional Quarterly 2006 Almanac*. Washington: Congressional Quarterly, Inc., vol. 62.

Crespin, Michael H., and David W. Rohde. 2010. Dimensions, issues, and bills: Appropriations voting on the House floor. *The Journal of Politics* 72: 976–89. [CrossRef]

Dougherty, Keith L., Michael S. Lynch, and Anthony J. Madonna. 2014. Partisan agenda control and the dimensionality of Congress. *American Politics Research* 42: 600–27. [CrossRef]

Harris, John F. 2004. Truth, consequences of Kerry's 'liberal' label. *Washington Post*, July 19, p. A01.

Heckman, James J., and James M. Snyder Jr. 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *RAND Journal of Economics* 28: S142–89. [CrossRef]

Hix, Simon, Abdul Noury, and Gérard Roland. 2006. Dimensions of politics in the European Parliament. *American Journal of Political Science* 50: 494–511. [CrossRef]

Hook, Janet. 2014. In Ferguson's wake, odd bedfellows. *Wall Street Journal*, August 17, p. A4.

Imai, Kosuke, James Lo, and Jonathan Olmsted. 2016. Fast estimation of ideal points with massive data. *American Political Science Review* 110: 631–56. [CrossRef]

Jackman, Simon. 2000. Estimation and inference are missing data problems: Unifying social science statistics via Bayesian simulation. *Political Analysis* 8: 307–32. [CrossRef]

Jackman, Simon. 2001. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. *Political Analysis* 9: 227–41. [CrossRef]

Jackson, J. Edward. 1991. *A User's Guide to Principal Components*. New York: Wiley.

Jessee, Stephen A. 2009. Spatial voting in the 2004 presidential election. *American Political Science Review* 103: 59–81. [CrossRef]

Jolliffe, Ian T. 2004. *Principal Component Analysis*, 2nd ed. New York: Springer.

Krehbiel, Keith, and Zachary Peskowitz. 2015. Legislative organization and ideal-point bias. *Journal of Theoretical Politics* 27: 673–704. [CrossRef]

Ladha, Krishna K. 1991. A spatial model of legislative voting with perceptual error. *Public Choice* 68: 151–74. [CrossRef]

Lewis, Jeffrey B. 2001. Estimating voter preference distributions from individual-level voting data. *Political Analysis* 9: 275–97. [CrossRef]

Lewis, Jeffrey B., and Keith T. Poole. 2004. Measuring bias and uncertainty in ideal point estimates via the parametric bootstrap. *Political Analysis* 12: 105–27. [CrossRef]

Londregan, John. 2000. Estimating legislators' preferred points. *Political Analysis* 8: 35–56. [CrossRef]

Martin, Andrew D., and Kevin M. Quinn. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10: 134–53. [CrossRef]

McCarty, Nolan. 2016. In defense of W-NOMINATE. *Studies in American Political Development* 30: 172–84. [CrossRef]

McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal. 1997. *Income Redistribution and the Realignment of American Politics*. Washington: AEI Press.

Montopoli, Brian. 2008. National Journal: Obama Most Liberal Senator in 2007. January 31. Available online: http://www.cbsnews.com/news/national-journal-obama-most-liberal-senator-in-2007/ (accessed on 29 May 2015).

Morrison, Donald F. 1976. *Multivariate Statistical Methods*, 2nd ed. New York: McGraw-Hill.

Peress, Michael. 2009. Small chamber ideal point estimation. *Political Analysis* 17: 276–90. [CrossRef]

Poole, Keith T. 2000. Nonparametric unfolding of binary choice data. *Political Analysis* 8: 211–37. [CrossRef]

Poole, Keith T. 2001. The geometry of multidimensional quadratic utility in models of parliamentary roll call voting. *Political Analysis* 9: 211–26. [CrossRef]

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

Poole, Keith T., and Howard Rosenthal. 1985. A spatial model for legislative roll call analysis. *American Journal of Political Science* 29: 357–84. [CrossRef]

Poole, Keith T., and Howard Rosenthal. 1991. Patterns of Congressional voting. *American Journal of Political Science* 35: 228–78. [CrossRef]

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

Poole, Keith T., and Howard Rosenthal. 2001. D-NOMINATE after 10 years: A comparative update to Congress: A Political-Economic History of Roll Call Voting. *Legislative Studies Quarterly* 26: 5–29. [CrossRef]

Poole, Keith T., and Howard Rosenthal. 2007. *Ideology & Congress, Second, Revised Edition of Congress: A Political-Economic History of Roll Call Voting*. New Brunswick: Transaction Publishers.

Poole, Keith, Jeffrey Lewis, James Lo, and Royce Carroll. 2011. Scaling roll call votes with wnominate in R. *Journal of Statistical Software* 42: 1–21. [CrossRef]

Rosas, Guillermo, and Yael Shomer. 2008. Models of nonresponse in legislative politics. *Legislative Studies Quarterly* 33: 573–601. [CrossRef]

Rosenthal, Howard, and Erik Voeten. 2004. Analyzing roll calls with perfect spatial voting: France 1946–1958. *American Journal of Political Science* 48: 620–32. [CrossRef]

Schmidt, Peter. 1983. Identification problems. In *Encyclopedia of Statistical Sciences*. Edited by Samuel Kotz, Norman L. Johnson and Campbell B. Read. New York: Wiley, vol. 4, pp. 10–14.

Sides, John. 2011. The challenge of measuring political ideology. May 4. Available online: http://themonkeycage. org/2011/05/04/the-challenge-of-measuring-political-ideology/ (accessed on 5 June 2015).

Silver, Nate. 2014. How the FiveThirtyEight Senate Forecast Model Works. September 17. Available online: http://fivethirtyeight.com/features/how-the-fivethirtyeight-senate-forecast-model-works/ (accessed on 29 May 2015).

Voteview Blog. 2014. DW-NOMINATE video, 1789–2013. February 10. Available online: https://www.youtube. com/watch?v=_0TE5TWYP-I (accessed on 12 October 2017). (Previously online: http://voteview.com/ blog/?p=1045 (no longer available)).