The Geometry of Multidimensional Quadratic Utility in Models of Parliamentary Roll Call Voting

by

Keith T. Poole
Graduate School of Industrial Administration
Carnegie Mellon University

24 April 2000

Abstract

The purpose of this paper is to show how the geometry of the quadratic utility function in the standard spatial model of choice can be exploited to estimate a model of Parliamentary roll call voting. In a standard spatial model of Parliamentary roll call voting, the legislator votes for the policy outcome corresponding to Yea if her utility for Yea is greater than her utility for Nay. The voting decision of the legislator is modeled as a function of the difference between these two utilities. With quadratic utility, this difference has a simple geometric interpretation that can be exploited to estimate legislator ideal points and roll call parameters in a standard framework where the stochastic portion of the utility function is normally distributed. The geometry is almost identical to that used in Poole (2000) to develop a non-parametric unfolding of binary choice data and the algorithms developed by Poole (2000) can be easily modified to implement the standard maximum likelihood model.

1. Introduction

The purpose of this paper is to show how the geometry of the quadratic utility function in the standard spatial model of choice can be exploited to estimate a model of parliamentary roll call voting. The quadratic utility function has a long history. Beginning with the earliest papers of Davis and his colleagues (Davis and Hinich, 1966, 1967; Davis, Hinich, and Ordeshook, 1970), it has played an important role in the spatial theory of voting and elections. The quadratic utility function is analytically simple and has a number of mathematical properties that make it easy to work with for modeling purposes. For example, it is symmetric around, and has a unique maximum at, the individual's ideal point.

In a standard spatial model of Parliamentary roll call voting, the legislator votes for the policy outcome corresponding to Yea if her utility for Yea is greater than her utility for Nay. The voting decision of the legislator is modeled as a function of the *difference* between these two utilities. The difference between two quadratic utilities has a simple geometric interpretation that can be exploited to estimate legislator ideal points and roll call parameters in a standard framework where the stochastic portion of the utility function is normally distributed. In particular, the geometry is almost identical to that used in Poole (2000) to develop a non-parametric unfolding of binary choice data. The algorithms developed by Poole (2000) can be easily modified to implement a standard maximum likelihood model where the deterministic portion of the utility function of the legislators is quadratic and the stochastic portion is normally distributed.

Section 2 defines the problem and explains the notation used in the paper. Section 3 shows the geometry of the quadratic utility model and discusses the plausibility of the three major error distributions – normal, logit, and uniform – that have been used by researchers to estimate the parameters of spatial models. Section 4 briefly discusses the scaling method developed by Poole (2000). The geometry of this scaling method is essentially the same as that shown in section 3. Section 5 shows how the algorithms developed in Poole (2000) can be used to estimate the parameters of a standard maximum likelihood model where the deterministic portion of the utility function of the legislators is quadratic and the stochastic portion is normally distributed. A quadratic utility scaling of the 90th House of Representatives shows that the algorithm is stable and produces sensible results. Section 6 concludes.

2. Notation[1] and Definitions

Assume that legislators have Euclidean preferences defined over some multidimensional ideological/policy space and that they vote sincerely for the alternative closest to their ideal point. Let p be the number of legislators (i=1,…,p) and s be the number of dimensions (k=1,…,s). The ith legislator's ideal point on the kth dimension is denoted by x_{ik} and let **X** be the p by s matrix of legislator ideal points. Each roll call vote has two policy points in the space corresponding to the policy consequences of a Yea or Nay vote on the roll call.[2] Let q be the number of roll calls (j=1,…,q) and coordinates for the Yea and Nay outcomes are denoted by z_{jky} and z_{jkn} respectively. Let "c" indicate the outcome (Yea or Nay) chosen by legislator i, and let "b" indicate the outcome not chosen by legislator i. This notation will considerably simplify the exposition below.

If there were no voting error, a plane can be placed in the space such that it separates all the legislators voting Yea from all the legislators voting Nay. Geometrically, this *cutting plane* is both perpendicular to the line joining the Yea and

Nay policy points and passes through the midpoint of the Yea and Nay policy points. Because the normal vector to a plane is perpendicular to the plane, the normal vector to this cutting plane, by definition, is parallel to the line joining the Yea and Nay policy points. Specifically, let $\underline{\mathbf{n}}_{\mathbf{j}}$ be the s by 1 normal vector for the jth roll call and let **N** be the q by s matrix of normal vectors for the q cutting planes. A plane is defined as the vector equation, $\underline{\mathbf{z}}'\underline{\mathbf{n}}_{\mathbf{j}} = \underline{\mathbf{v}}'\underline{\mathbf{n}}_{\mathbf{j}}$, where $\underline{\mathbf{z}}$, $\underline{\mathbf{n}}_{\mathbf{j}}$, and $\underline{\mathbf{v}}$ are s by 1 vectors and the plane consists of all points $\underline{\mathbf{z}}$ such that ($\underline{\mathbf{z}}$ - $\underline{\mathbf{v}}$) is perpendicular to the normal vector, $\underline{\mathbf{n}}_{\mathbf{j}}$, and $\underline{\mathbf{v}}$ is a *specific* point in the plane. Note that if $\underline{\mathbf{v}}_{\mathbf{1}}$, and $\underline{\mathbf{v}}_2$ are both points in the plane then, $\underline{\mathbf{v}}_{\mathbf{1}}'\underline{\mathbf{n}}_{\mathbf{j}} = \underline{\mathbf{v}}_2'\underline{\mathbf{n}}_{\mathbf{j}} = m_{\mathbf{j}}$**,** where $m_{\mathbf{j}}$ is a scalar constant. Geometrically, every point in the plane projects onto the same point on the line defined by the normal vector, $\underline{\mathbf{n}}_{\mathbf{j}}$ and its reflection $-\underline{\mathbf{n}}_{\mathbf{j}}$. Because the midpoint of the Yea and Nay policy points is on the cutting plane, it too projects to the point $m_{\mathbf{j}}$.

Technically, the general equation for a line is:

$$\mathbf{Y(t) = A + t(B - A)}$$

Where **A** and **B** are points in the space and **t** is a scalar. In this instance, **A** is placed at the origin of the space so that the equation for the line defined by the normal vector, $\underline{\mathbf{n}}_{\mathbf{j}}$ and its reflection $-\underline{\mathbf{n}}_{\mathbf{j}}$ is simply

$$\mathbf{Y(t) = t}\underline{\mathbf{n}}_{\mathbf{j}} \qquad \textbf{(1)}$$

where $-1 \le \mathbf{t} \le +1$.

To set the scale of the voting space, let the legislator coordinates lie within the s dimensional unit hypersphere and let the origin of the space be placed at the centroid of the legislator coordinates; that is, let

$$\sum_{k=1}^{s} x_{ik}^2 \le 1 \quad , i=1,...,p \quad \text{and} \quad \sum_{i=1}^{p} x_{ik} = 0 \quad , k=1,...,s$$

In addition, without loss of generality, the normal vector, $\underline{\mathbf{n}}_{\mathbf{j}}$, can be constrained to be of unit length; i. e., $\underline{\mathbf{n}}_{\mathbf{j}}'\underline{\mathbf{n}}_{\mathbf{j}} = 1$.

Let the projections of the legislator points onto the line defined by equation (1) be:

$$\mathbf{X}\underline{\mathbf{n}}_{\mathbf{j}} = \underline{\mathbf{w}} \qquad \textbf{(2)}$$

Note that the elements in the p-length vector, $\underline{\mathbf{w}}$, range from -1 to +1. The elements in $\underline{\mathbf{w}}$ all lie on the line defined by equation (1) that passes through the origin of the s-dimensional unit hypersphere in the direction of the normal vector with exit points $-\underline{\mathbf{n}}_{\mathbf{j}}$ and $+\underline{\mathbf{n}}_{\mathbf{j}}$ respectively. This line will hereafter be referred to as the *projection line*.

3. The Quadratic Utility Model

Given these definitions, legislator i's utility for her chosen outcome, c, on roll call j is:

$$\mathbf{U_{ijc} = u_{ijc} + \boldsymbol{\varepsilon}_{ijc} = \sum_{k=1}^{s}(x_{ik} - z_{jkc})^2 + \boldsymbol{\varepsilon}_{ijc}} \qquad \textbf{(3)}$$

Where $\mathbf{u_{ijc}}$ is the deterministic portion of the utility function and $\boldsymbol{\varepsilon}_{\mathbf{ijc}}$ is the stochastic portion.

The probability that legislator i votes for her chosen outcome, c, is

$$\mathbf{P(U_{ijc} > U_{ijb}) = P(\boldsymbol{\varepsilon}_{ijb} - \boldsymbol{\varepsilon}_{ijc} < u_{ijc} - u_{ijb})} \qquad \textbf{(4)}$$

The Deterministic Portion of the Utility Function

The difference between the deterministic utilities can be simplified as follows:

$$\mathbf{u_{ijc} - u_{ijb} = \sum_{k=1}^{s}(x_{ik} - z_{jkc})^2 - \sum_{k=1}^{s}(x_{ik} - z_{jkb})^2}$$

$$= \sum_{k=1}^{s} x_{ik}^2 - 2\sum_{k=1}^{s} x_{ik} z_{jkc} + \sum_{k=1}^{s} z_{jkc}^2 - \sum_{k=1}^{s} x_{ik}^2 + 2\sum_{k=1}^{s} x_{ik} z_{jkb} - \sum_{k=1}^{s} z_{jkb}^2$$

$$= 2\sum_{k=1}^{s} x_{ik}(z_{jkb} - z_{jkc}) - \sum_{k=1}^{s} (z_{jkb} - z_{jkc})(z_{jkb} + z_{jkc}) \qquad \textbf{(5)}$$

Now, note that the s by 1 vector:

$$\underline{\mathbf{z}}_{jb} - \underline{\mathbf{z}}_{jc} = \begin{bmatrix} z_{j1b} - z_{j1c} \\ z_{j2b} - z_{j2c} \\ \vdots \\ z_{jsb} - z_{jsc} \end{bmatrix}$$

is equal to a constant times the normal vector, $\underline{\mathbf{n}}_j$ (see Figure 1)**.** Namely,

$$\gamma_j \underline{\mathbf{n}}_j = \underline{\mathbf{z}}_{jb} - \underline{\mathbf{z}}_{jc} \qquad \textbf{(6)}$$

where

$$\gamma_j = +\left[\sum_{k=1}^{s} (z_{jkb} - z_{jkc})^2\right]^{\frac{1}{2}} \text{ if } \underline{\mathbf{z}}_{jb}'\underline{\mathbf{n}}_j > \underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j \text{ or}$$

$$\gamma_j = -\left[\sum_{k=1}^{s} (z_{jkb} - z_{jkc})^2\right]^{\frac{1}{2}} \text{ if } \underline{\mathbf{z}}_{jb}'\underline{\mathbf{n}}_j < \underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j$$

γ_j is the *directional distance* between the Yea and Nay outcomes in the space.

Figure 1 about Here

The s by 1 vector

$$\underline{\mathbf{z}}_{jb} + \underline{\mathbf{z}}_{jc} = \begin{bmatrix} z_{j1b} + z_{j1c} \\ z_{j2b} + z_{j2c} \\ \vdots \\ z_{jsb} + z_{jsc} \end{bmatrix}$$

divided by 2 is simply the s by 1 vector of midpoints for the Yea and Nay outcomes for roll call j. That is:

$$\underline{\mathbf{z}}_{\mathbf{mj}} = \frac{\underline{\mathbf{z}}_{\mathbf{jb}} + \underline{\mathbf{z}}_{\mathbf{jc}}}{\mathbf{2}}$$

This allows equation (5) to be rewritten as the vector equation:

$$\mathbf{u_{ijc}} - \mathbf{u_{ijb}} = 2\boldsymbol{\gamma}_{\mathbf{j}}(\underline{\mathbf{x}}_{\mathbf{i}}'\underline{\boldsymbol{\nu}}_{\mathbf{j}} - \underline{\mathbf{z}}_{\mathbf{mj}}'\underline{\boldsymbol{\nu}}_{\mathbf{j}}) = 2\boldsymbol{\gamma}_{\mathbf{j}}(w_{\mathbf{i}} - m_{\mathbf{j}}) \qquad \mathbf{(7)}$$

where $w_{\mathbf{i}}$ is the projection of the ith legislator's ideal point onto the projection line as defined by equation (1), and $m_{\mathbf{j}}$ is the projection of the midpoint of the roll call outcomes onto the projection line. Equation (7) shows that:

if $\boldsymbol{\gamma}_{\mathbf{j}} > \mathbf{0}$ and $w_{\mathbf{i}} > m_{\mathbf{j}}$, or

if $\boldsymbol{\gamma}_{\mathbf{j}} < \mathbf{0}$ and $w_{\mathbf{i}} < m_{\mathbf{j}}$, then $\mathbf{u_{ijc}} > \mathbf{u_{ijb}}$

In one dimension, $\underline{\boldsymbol{\nu}}_{\mathbf{j}}$ is equal to 1 and $\boldsymbol{\gamma}_{\mathbf{j}} = z_{jc} - z_{jb}$. Hence, equation (1) becomes simply $2(z_{jc} - z_{jb})(x_i - z_{mj}) = 2\boldsymbol{\gamma}_{\mathbf{j}}(x_{\mathbf{i}} - m_{\mathbf{j}})$. Except for an added "valence" dimension, this is identical to the one dimensional model developed by Londregan (2000, p. 40-41). Specifically, in Londregan's notation, $g = (z_{jc} - z_{jb}.)$, $x_v = x_i$, and $m = z_{mj}$.[3] Now, note that

if $z_{jc} > z_{mj}$ and $x_{\mathbf{i}} > z_{mj}$ or

if $z_{jc} < z_{mj}$ and $x_{\mathbf{i}} < z_{mj}$, then $\mathbf{u_{ijc}} > \mathbf{u_{ijb}}$

If voting is sincere and without error then $2\boldsymbol{\gamma}_{\mathbf{j}}(w_{\mathbf{i}} - m_{\mathbf{j}}) > 0$ for all i and j and, in one dimension, the legislator ideal points and the roll call midpoints are only identified up to a joint rank ordering. With "perfect" voting in more than one dimension, if a variety of voting coalitions form amongst the legislators, then the cutting planes will intersect one another in a myriad of directions creating a maximum of $\sum_{k=0}^{s}\binom{q}{k}$ regions in the policy space (Coombs, 1964, p. 262). Each region corresponds to a unique voting pattern on the q roll calls – e.g., YYYNNYNYNYYY…. Hence, a legislator's ideal point is identified up to a region in the space. Note however, that as q gets large the number of regions explodes so that the volume of these regions is extremely small. For example, with 500 roll calls, there are a maximum 125,251 regions in two dimensions and a maximum of 20,833,751 in three dimensions.[4] Most of these regions are so small that a typical legislator's point is very precisely pinned down (Poole, 2000).[5] Similarly, with a large number of legislators, the cutting plane – defined by the normal vector, $\underline{\mathbf{n}}_{\mathbf{j}}$, and the midpoint of the roll call, $m_{\mathbf{j}}$ -- is also precisely pinned down (Poole, 2000).

To reiterate, if voting is perfect, that is, sincere and without error, then in one dimension the legislator ideal points and the roll call midpoints are only identified up to a joint rank ordering. In more than one dimension, legislators are identified up to regions in the space (polytopes) and roll calls are identified up to cone shaped regions containing the normal vectors and line segments on the normal vectors for the midpoints. These limits on identification arise because the data are simply Yea and Nay. If legislators could report "thermometer scores" for the alternatives then the perfect case would have an exact solution.

The Stochastic Portion of the Utility Function

Superficially, it is the stochastic portion of the utility function that allows for more precise solutions for the legislator ideal points and roll call parameters. However, as Londregan (2000) proves, this precision is, to an extent, an illusion. Londregan shows that consistency in its usual statistical sense does not hold in the roll call voting problem outlined above. With nominal choices standard maximum likelihood estimators that attempt to *simultaneously* recover legislators' ideal points and roll call parameters inherit the "granularity" of the choice data and so cannot recapture the underlying *continuous* parameter space. However, when the number of roll calls and legislators is large, the bias in the estimated parameters is not severe (Londregan, 2000).[6]

Turning to the stochastic portion of the utility function stated in equation (3) above, three probability distributions have been used to model the error; the normal (Ladha, 1991; Londregan, 2000), uniform (Heckman and Snyder, 1997), and logit (Poole and Rosenthal, 1997). The normal is clearly the best from both a theoretical *and* a behavioral standpoint.

From a statistical standpoint, given the difference between the two random errors, $\varepsilon_{ijb} - \varepsilon_{ijc}$, the standard assumptions are that ε_{ijb} and ε_{ijc} are a *random sample* (independent and identically distributed random variables) from a known distribution. Hence, it is therefore easy to write down the probability distribution of the difference --$\varepsilon_{ijb} - \varepsilon_{ijc}$. From a behavioral standpoint, it seems sensible to assume that the distributions of ε_{ijb} and ε_{ijc} are *symmetric and unimodal* and that ε_{ijb} and ε_{ijc} are uncorrelated. The normal distribution is the only one of the three distributions to satisfy all these criteria. To illustrate, assume that ε_{ijb} and ε_{ijc} are drawn (a random sample of size two) from a normal distribution with mean zero and variance one-half. The difference between the two errors has a standard normal distribution; that is

$$\varepsilon_{ijb} - \varepsilon_{ijc} \sim N(0, 1)$$

Hence, the probability that legislator i votes for her chosen outcome, c, can be rewritten as:

$$P_{ijc} = P(U_{ijc} > U_{ijb}) = P(\varepsilon_{ijb} - \varepsilon_{ijc} < u_{ijc} - u_{ijb}) =$$

$$\Phi[2\beta_j(\underline{x}_i'\underline{n}_j - \underline{z}_{mj}'\underline{n}_j)] = \Phi[2\beta_j(\omega_i - \mu_j)] \qquad \textbf{(8)}$$

Heckman and Snyder (1997) assume that $\varepsilon_{ijb} - \varepsilon_{ijc}$ has a uniform distribution. This is an extremely problematic assumption because ε_{ijb} and ε_{ijc} *cannot be a random sample*![7] Assuming that $\varepsilon_{ijb} - \varepsilon_{ijc}$ has a uniform distribution enables Heckman and Snyder to develop a linear probability model but the price for this simplicity is that they have no intuitive basis for a behavioral model.

Poole and Rosenthal (1985, 1991, 1997) assume that ε_{ijb} and ε_{ijc} are a random sample from the log of the inverse exponential distribution. Consequently, $\varepsilon_{ijb} - \varepsilon_{ijc}$ has the logit distribution. The log of the inverse exponential and the logit distribution which is derived from it, are unimodal *but not symmetric*.[8] However, they are not too skewed and the distribution function of the logit distribution is reasonably close to the normal distribution function.[9]

A further difficulty with the approaches taken by both Heckman and Snyder and Poole and Rosenthal is that they assume that the error variance is homoskedastic. A more realistic assumption is that the error variance varies across the roll call votes and across the legislators. For the roll calls, it is impossible to distinguish between the underlying unknown error variance and the distance between the Yea and Nay alternatives (Ladha, 1991; Londregan, 2000). The intuition behind this is straightforward. As the distance between the Yea and Nay alternatives increases, the easier it is for legislators to distinguish between the two policy outcomes and the less likely it is that they make an error. Conversely, if the Yea and Nay alternatives are very close together, then the utility difference is small and it is more likely that voting errors occur. Increasing/decreasing the distance is equivalent to decreasing/increasing the variance of the underlying error.

Because γ_j is picking up the roll call specific variance, the difference between the two errors for legislator i on roll call j can be modeled as:

$$\varepsilon_{ijb} - \varepsilon_{ijc} \sim N(0, \sigma_i^2)$$

With heteroskedastic error equation (8) becomes:

$$P_{ijc} = P\left(\frac{\varepsilon_{ijb} - \varepsilon_{ijc}}{\sigma_i} < \frac{u_{ijb} - u_{ijc}}{\sigma_i}\right) = \Phi\left[\frac{2\gamma_j}{\sigma_i}(w_i - m_j)\right] \qquad (9)$$

The corresponding likelihood function is therefore:

$$L = \prod_{i=1}^{p}\prod_{j=1}^{q} P_{ijc} \qquad (10)$$

The approach developed in section 5 below allows the error to be heteroskedastic. The algorithms developed by Poole (2000) can be used to obtain excellent estimates for the legislator points, the $\underline{\mathbf{x}}_i$'s, the roll call normal vectors, the $\underline{\mathbf{n}}_j$'s, and the cutpoints, the m_j's. With these fixed, the γ_j and σ_i can be estimated.

In sum, the normal distribution is the most sensible model of error both from a mathematical standpoint as well as a behavioral standpoint. Consequently, it will be the focus of the model developed in this paper.

4. The Classification Algorithm

Poole (2000) develops a new scaling method for analyzing parliamentary roll call data. The scaling method uses almost exactly the same geometry as that shown for the difference between two quadratic utilities shown above. Given the legislator coordinates, the scaling method estimates cutting planes for each roll call; and given the cutting planes, the method finds the region in the space that best matches the legislator's roll call choices. The scaling method is *non-parametric* because no assumptions are made about the probability distribution of the legislators' errors in making choices. The only assumptions made are that the choice space is Euclidean and that individuals making choices behave as if they utilize symmetric, single-peaked preferences.

Strictly speaking, the scaling method developed in Poole (2000) is *not* a statistical model. However, as shown in the next section, the algorithms developed by Poole (2000) can be easily modified to implement a standard maximum likelihood model where the deterministic portion of the utility function of the legislators is quadratic and the stochastic portion is normally distributed.

The classification algorithm uses the geometry outlined above along with the assumption that preferences are symmetric and single peaked to find estimates of **X** and **N.** The rule for correct classification is:

If legislator i votes c: $\boldsymbol{\delta_{ij} = 1}$ if $w_i \geq m_j$ and $\underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j > m_j$, or $w_i < m_j$ and $\underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j < m_j$

$\boldsymbol{\delta_{ij} = 0}$ if $w_i < m_j$ and $\underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j > m_j$, or $w_i > m_j$ and $\underline{\mathbf{z}}_{jc}'\underline{\mathbf{n}}_j < m_j$

In other words if the legislator votes "Yea"/"Nay" and her ideal point is on the Yea/Nay side of the plane, the legislator's vote is correctly classified. Note that the assumption of *symmetric* single-peaked preferences means that if a legislator votes "Yea" and her ideal point is *anywhere on the Yea side of the plane* then that counts as a correct classification. If preferences are *not symmetric* then this might not be true.

The total correct classification is therefore:

$$\boldsymbol{\delta(X, N) = \sum_{i=1}^{p}\sum_{j=1}^{q}\delta_{ij}} \qquad \textbf{(11)}$$

In sum, given the number of dimensions, s, the classification problem consists of finding estimates of **X** and **N** that maximize equation (11).

5. An Algorithm to Estimate the Multidimensional Quadratic Utility Model

Note that if $\boldsymbol{\delta}_{\mathbf{ij}} = \mathbf{1}$, then $\frac{2\gamma_j}{\sigma_i}(w_i - m_j) > 0$ and $\Phi\left[\frac{2\gamma_j}{\sigma_i}(w_i - m_j)\right] > .5$; and

$$\text{if } \boldsymbol{\delta}_{\mathbf{ij}} = \mathbf{0}, \text{ then } \frac{2\gamma_j}{\sigma_i}(w_i - m_j) < 0 \text{ and } \Phi\left[\frac{2\gamma_j}{\sigma_i}(w_i - m_j)\right] < .5$$

In other words, the classification algorithm, which is intended to maximize equation (11), will also tend to maximize equation (10). Given estimates of **X, N**, and the m_j**'s** from the classification algorithm -- denoted as **X*, N*,** and m_j***** respectively -- it is a simple matter to estimate the γ_j and σ_i terms because the likelihood function is convex if the roll call *cannot be classified without error*. With a finite number of legislators, there will be roll calls where $|\gamma_j|$ will be very large because the roll call is so important or the information is so complete that no legislator makes an error. However, if the roll call can be classified without error then $|\gamma_j| \rightarrow +\infty$ -- that is, the probabilities assigned to the choices of the p legislators will go to 1 on a "perfect" roll call. This does not present a problem since $\mathbf{P_{ijc}}$ can be set equal to 1 and its corresponding log-likelihood can be set equal to 0.

The multidimensional quadratic utility model can be efficiently estimated in four steps. First, using **X*, N*,** and the m_j***** from the classification algorithm, set all the σ_i equal to 1 and estimate the γ_j's using a simple grid search. Given these γ_j's estimate the σ_i's using a simple grid search. Repeat this process until there is no significant improvement in the log-likelihood. In practice, this takes no more than three repetitions. Second, given **N*** and the m_j***** from the classification algorithm and the estimated γ_j's and σ_i's from step 1, estimate new legislator coordinates, **X**. This is easily accomplished using standard gradient techniques. Third, given **N*** from the classification algorithm, the estimated $\mathbf{g}_j$'s and $\mathbf{s}_i$'s from step 1, and the estimated legislator coordinates, **X,** from step 2, estimate new projected midpoints, the m_j, using a simple grid search. Fourth, given the estimated $\mathbf{g}_j$'s and $\mathbf{s}_i$'s from step 1, the estimated legislator coordinates, **X,** from step 2, and the estimated projected midpoints, the m_j, from step 3, estimate new normal vectors, the $\underline{\mathbf{n}}_j$. This is easily accomplished using standard gradient techniques with the constraints that $\underline{\mathbf{n}}_j'\underline{\mathbf{n}}_j = 1$ and that $\underline{\mathbf{z}}_{mj} = m_j\underline{\mathbf{n}}_j$. In other words, the point defined by the end of the normal vector is moved along the surface of the unit hypersphere with the position of the projected midpoint held fixed on the normal vector as it is moved. Geometrically, this is equivalent to moving the cutting plane rigidly through the space as its normal vector is moved.

In summary, the algorithm is:

Step 1a: Estimate the $\mathbf{g}_j$

Step 1b: Estimate the $\mathbf{s}_i$

Step 1c: Repeat a and b Until convergence

Step 2: Estimate the $\underline{\mathbf{x}}_i$

Step 3: Estimate the m_j

Step 4: Estimate the $\underline{\mathbf{n}}_j$

Go to 1a

In one dimension, given a joint rank ordering of the legislators and roll call midpoints from the classification algorithm, step 4 is not necessary and the legislator coordinates in step 2 can be found through a simple grid search. In practice only 3 overall passes through steps 1 to 4 are required for the program to converge.

Table 1 and Figures 2 to 6 show an application of the quadratic utility algorithm to the 90th House of Representatives. Table 1 shows the scaling results for the quadratic algorithm for 1 to 10 dimensions. The corresponding correct classifications from the optimal classification algorithm are also shown for comparative purposes. Figure 2 graphs the increase in fit from adding dimensions. The increase in geometric mean probability (GMP) from adding a dimension was multiplied by 100 so it could be graphed on the same scale as the increase in classification. After the 2nd dimension the incremental increase from adding a dimension is quite small. This is a classic "elbow" indicating the correct dimensionality is *at most* three and almost certainly two.[10]

Table 1 and Figure 2 about Here

Figure 3 shows a plot of the legislator ideal points for the 90th House. The "d", "s", and "r" tokens indicate Northern (Non-Southern) Democrats, Southern Democrats, and Republicans respectively.[11] The two dimensions are liberal-conservative (government intervention in the economy) and Race (North vs. South). The configuration is the same as the recovered by NOMINATE, Heckman-Snyder, and the optimal classification algorithm.[12] An analysis of the structure of voting during this period of American history can be found in Poole and Rosenthal (1997) and McCarty, Poole, and Rosenthal (1997).

Figure 3 about Here

Figure 4 shows a histogram of the estimated normal vectors in terms of their angles from –90 to +90 degrees. A normal vector at an angle of –45 degrees produces a cutting plane at +45 degrees parallel to the "channel" between the two political parties. A normal vector at an angle between +45 degrees and about +20 degrees produces cutting planes that are potential "conservative coalition" votes. That is, cutting planes that run between the two wings of the Democratic Party with a majority of Republicans on the side of the Southern Democrats.

Figure 4 about Here

Figures 5 and 6 show histograms of the $\mathbf{s_i}$'s and $\mathbf{g_j}$'s respectively. The mean and standard deviation of the $\mathbf{s_i}$'s is .995 and .40 respectively. The mean and standard deviation of the $\mathbf{g_j}$'s is 4.46 and 2.55 respectively. Both are weakly related to their corresponding correct classifications. The Pearson r between the $\mathbf{s_i}$'s and the correct classification percentages for the legislators is -.59. The Pearson r between the $\mathbf{g_j}$'s and the correct classification percentages for the roll calls is .66. Although the distributions of the $\mathbf{s_i}$'s and $\mathbf{g_j}$'s appear to be quite reasonable, without a comprehensive Monte-Carlo study of the quadratic procedure it is not possible to make definitive statements about them at this time.

Figures 5 and 6 about Here

6. Conclusion

The purpose of this paper was to show how the geometry of the multidimensional quadratic utility function could be exploited to estimate legislator ideal points and roll call normal vectors and cutpoints in a standard framework where the stochastic portion of the utility function is normally distributed. The algorithm shown in section 5 appears to be quite stable and produces sensible results. However, a comprehensive Monte-Carlo study is needed to pin down all the properties of the scaling algorithm.

References

Best, Alvin M., Forrest W. Young, and Robert G. Hall. 1979. "On the Precision of a Euclidean Structure." *Psychometrika*, 44:395-408

Coombs, Clyde. 1964. *A Theory of Data.* New York: Wiley.

Davis, Otto A. and Melvin J. Hinich. 1966. "A Mathematical Model of Policy Formation in a Democratic Society." In *Mathematical Applications in Political Science II*, edited by J. Bernd. Dallas Texas: Southern Methodist University Press.

Davis, Otto A. and Melvin J. Hinich. 1967. "Some Results Related to a Mathematical Model of Policy Formation in a Democratic Society." In *Mathematical Applications in Political Science III*, edited by J. Bernd. Charlottesville, VA: University of Virginia Press.

Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *American Political Science Review*, **64:**426-448.

Dhrymes, Phoebus J. 1978. *Introductory Econometrics*. New York: Springer-Verlag.

Heckman, James J. and James M. Snyder. 1997. "Linear Probability Models of the Demand for Attributes With an Empirical Application to Estimating the Preferences of Legislators." *Rand Journal of Economics*, 28:142-189.

Ladha, Krishna K. 1991. "A Spatial Model of Legislative Voting With Perceptual Error." *Public Choice*, 68:151-174.

Londregan, John B. 2000. "Estimating Legislators' Preferred Points. *Political Analysis,* 8(1), 35-56.

Lord, F. M. 1983. "Unbiased Estimates of Ability Parameters, of their Variance, and of their Parallel Forms Reliability." *Psychometrika*, 48:477-482.

MacRae, Duncan, Jr. 1958. *Dimensions of Congressional Voting*. Berkeley: University of California Press.

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

Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science*, 29:357-384.

Poole, Keith T. and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science*, 35:228-278.

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

Poole, Keith T. 2000. "Non-Parametric Unfolding of Binary Choice Data." *Political Analysis* (forthcoming).

Rasch, G. 1961. "On General Laws and the Meaning of Measurement in Psychology." *Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability*, 4:321-333.

Ross, John and Norman Cliff. 1964. "A Generalization of the Interpoint Distance Model."" *Psychometrika*, 29:167-176.

Schonemann, Peter H. 1966. "A Generalized Solution of the Orthogonal Procrustes Problem." *Psychometrika*, 31:1-10.

Young, Gale and A. S. Householder. 1938. "Discussion of a Set of Points in Terms of their Mutual Distances." *Psychometrika*, 3:19-22.

Figure 1. Vector Example

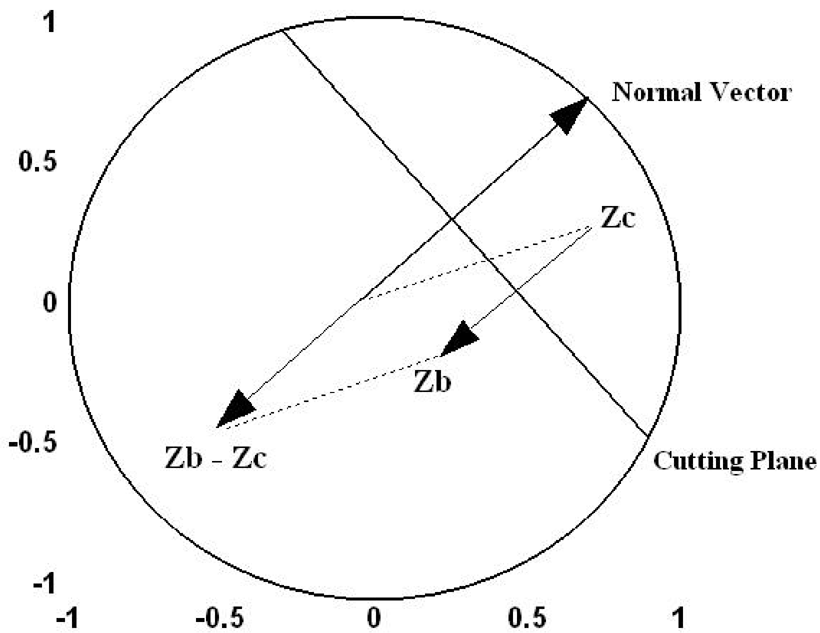

Fig. 2: 90th House
Gain in Fit From Adding Dimensions

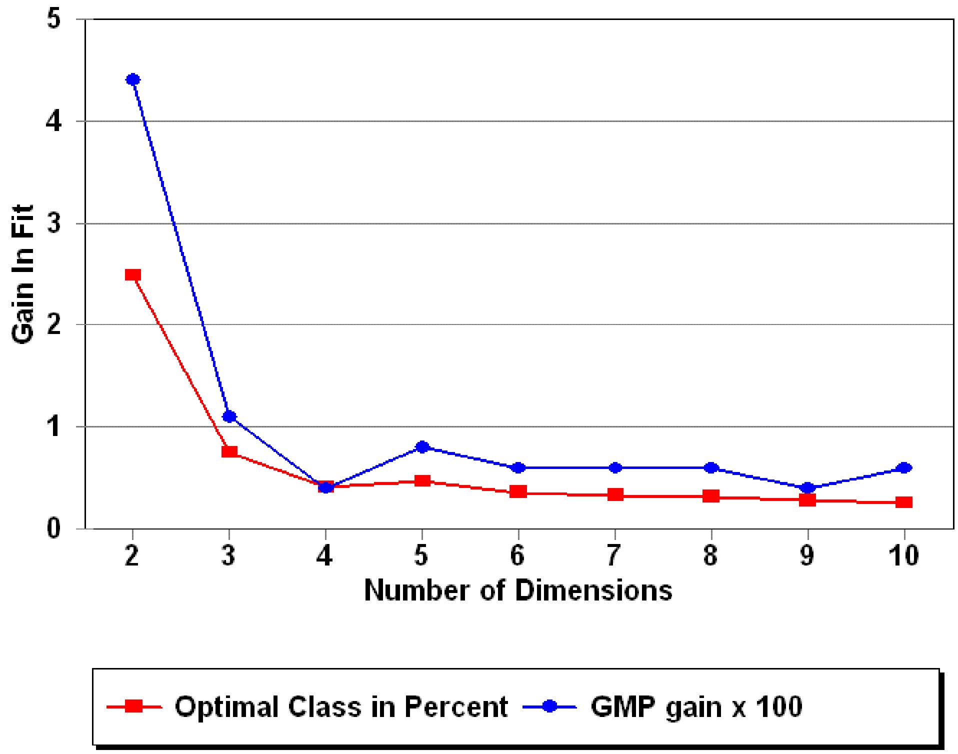

Fig. 3: 90th House
Quadratic Utility Scaling

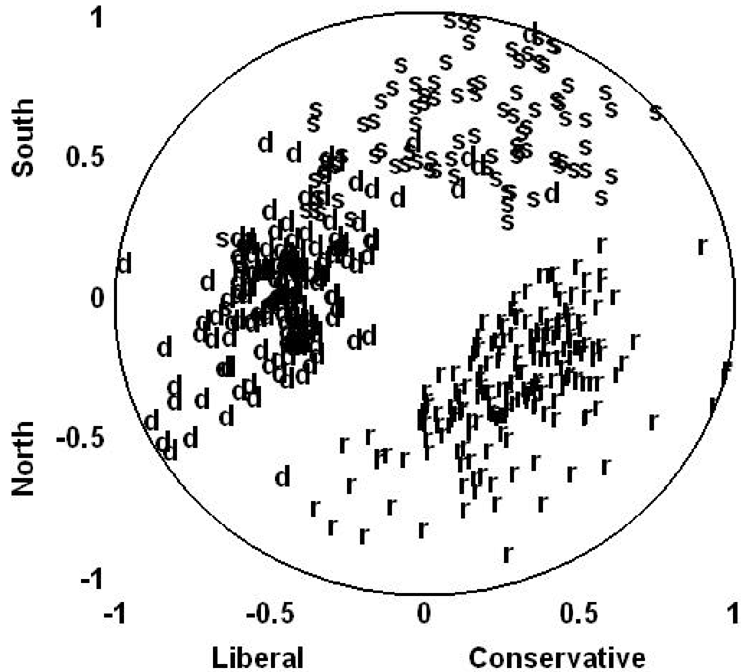

Fig. 4: 90th House
Distribution of Normal Vectors

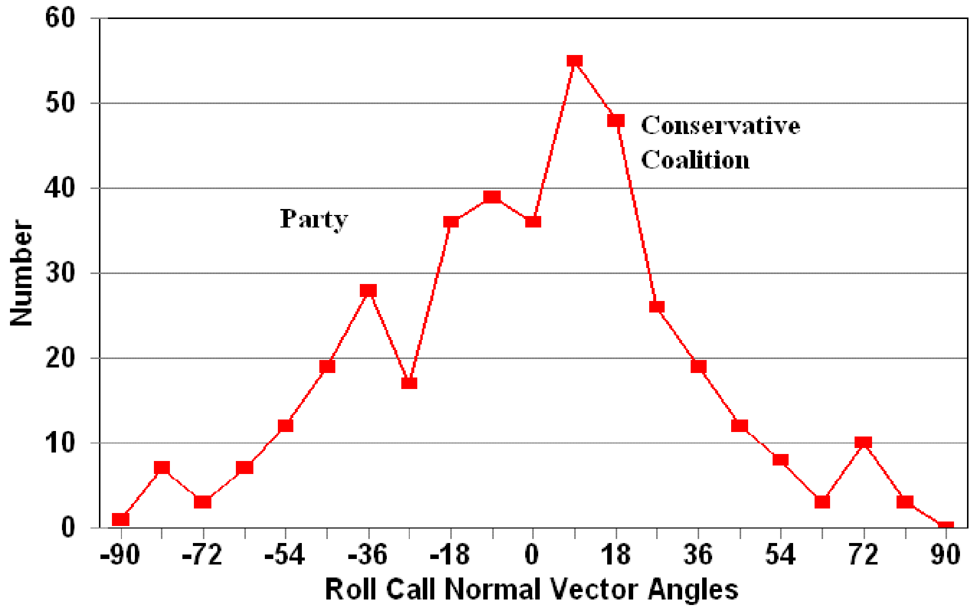

Fig. 5: 90th House
Distribution of Sigma-i

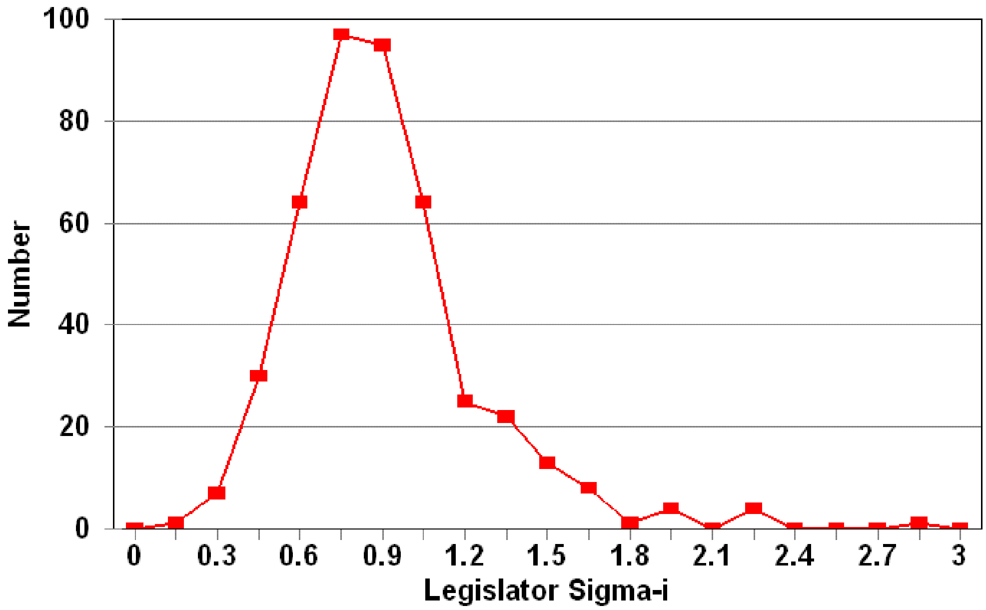

Fig. 6: 90th House
Distribution of Gamma-j

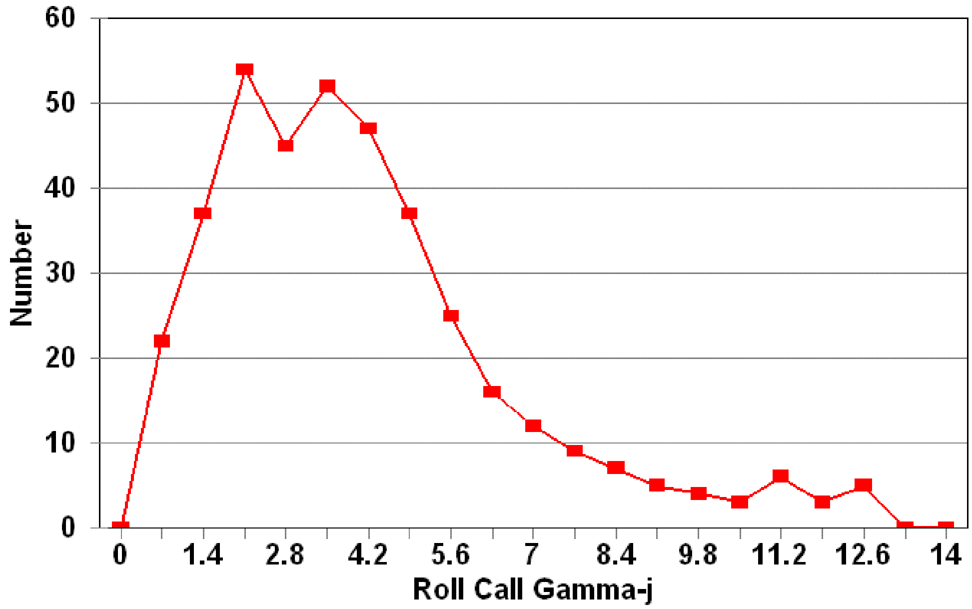

Table 1

Scaling Results for the 90th House:
389 Roll Calls, 438 Legislators, 147,199 Total Choices

	Optimal Classification		Quadratic Utility Scaling		
Dimension	**Percent Correct Class.**	**APRE**	**Percent Correct Class.**	**APRE**	**GMP**
1	**87.85**	**.573[a]**	**85.42**	**.488**	**.728[b]**
2	**90.34**	**.661**	**88.53**	**.597**	**.772**
3	**91.09**	**.687**	**89.31**	**.624**	**.783**
4	**91.50**	**.701**	**89.63**	**.636**	**.787**
5	**91.97**	**.718**	**90.05**	**.651**	**.795**
6	**92.33**	**.731**	**90.42**	**.664**	**.801**
7	**92.66**	**.742**	**90.78**	**.676**	**.807**
8	**92.98**	**.753**	**91.10**	**.687**	**.813**
9	**93.26**	**.763**	**91.34**	**.696**	**.817**
10	**93.52**	**.773**	**91.65**	**.707**	**.823**

[a] $$\text{APRE} = \frac{\sum_{j=1}^{q} \{\text{Minority Vote - Classification Errors}\}_j}{\sum_{j=1}^{q} \{\text{MinorityVote}\}_j}$$

[b] **Geometric Mean Probability: The exponential of the average log –likelihood; that is: GMP = exp[log-likelihood of all observed choices/N].**

Endnotes

[1] The notation in this paper with some minor variations is the same as that used in Poole and Rosenthal (1997), McCarty, Poole, and Rosenthal (1997), and Poole (2000).

[2] This model was first proposed by MacRae (1958) and later developed by Poole and Rosenthal (1997) in their NOMINATE procedure.

[3] Specifically, the utility function used by Londregan is:

$$U(z, q| x_v) = (-1/2)(z - x_v)^2 + \alpha q$$

The αq picks up a "valence" element of policy.

[4] Although the number of regions is very large, there are 2^q possible voting patterns and in practical applications this number will greatly exceed the maximum number of regions in the space. Finding the region that best matches the legislator's observed pattern of roll call votes is extremely difficult. For a solution, see Poole (2000).

[5] Actually measuring the volume of these regions is very difficult and is a problem that has not been satisfactorily solved (Best, Young, and Hall, 1979; Poole, 2000).

[6] Londregan cites the Monte Carlo work of Lord (1983) and Poole and Rosenthal (1991). Lord's Monte Carlo work was on the Rasch (1961) model used in ability tests (which is isomorphic with the one dimensional spatial model with quadratic utility [Ladha, 1991; Londregan, 2000]) and Poole and Rosenthal's work was done on their NOMINATE model. In the NOMINATE model the deterministic utility function is the normal distribution and the stochastic utility is the logit distribution. Both sets of studies indicate that when the questions/legislators are at least 100, the bias is not very large. This is confirmed by Monte Carlo studies shown in Poole (2000, Appendix).

[7] Heckman and Snyder (1997) acknowledge that no distribution exists such that the probability distribution of the difference between two random draws has a uniform distribution. For example, if ε_{ijb} and ε_{ijc} are drawn from a uniform distribution, then the distribution of their difference will be a triangle shaped distribution.

[8] The log of the inverse exponential is: $f(\varepsilon) = e^{-\varepsilon}e^{-e^{-\varepsilon}}$ where $-\infty < \varepsilon < +\infty$. The "logit" distribution is the distribution of the difference between two random draws of the log of the inverse exponential. That is:

$$f(y) = \frac{e^{-y}}{(1 + e^{-y})^2}$$

where $y = \varepsilon_{ijb} - \varepsilon_{ijc}$ and $-\infty < y < +\infty$ (see Dhrymes, 1978, pp. 340-352 for the derivation). Note that integrating **f(y)** from $-\infty$ to **x** yields the distribution function, $F(x) = \frac{1}{1 + e^{-x}}$, which, in the parliamentary roll call voting context is: $P_{ijc} = \frac{1}{1 + e^{-(u_{ijb} - u_{ijc})}}$.

[9] The logit distribution is somewhat of a "Y2K" phenomenon. The normal distribution is a much more sensible model of error but when computing resources were scarce the fact that the distribution function of the logit distribution is a *formula* (what economists call a "closed form") in contrast to the distribution function of the normal which must be calculated through a series expansion, made the logit model especially attractive. However, there is no longer any reason to use the logit distribution in most simple problems because computer memory is now so plentiful that the distribution function of the normal can be calculated to any practical level of precision required and simply stored in memory (the corresponding logs and portions of the partial derivatives can also be stored in memory). Performing a table look-up is fast and simple to accomplish.

[10] Reinforcing this conclusion is the pattern of eigenvalues from a Heckman-Snyder decomposition. The first 10 eigenvalues are: 25.22, 4.85, 1.16, 0.95, 0.83, 0.68, 0.65, 0.61, 0.52, 0.50.

Another measure is the pattern of eigenvalues from the double-centered legislator-by-legislator agreement score matrix. Technically, given a matrix of squared distances, double-centering is subtracting from each entry in the matrix the mean of the row, the mean of the column, and adding the mean of the matrix. This has the effect of removing the squared terms from the matrix leaving just the cross-product matrix. It also reduces the rank of the matrix by one (see Young and Householder, 1938; Ross and Cliff, 1964). For example, let A be an n by s matrix of coordinates. Let $\mathbf{diag(AA')}$ be the n length vector of diagonal terms of $\mathbf{AA'}$ and let $\mathbf{J_n}$ be an n length vector of ones. The matrix of squared distances can be written as: $\mathbf{diag(AA')J_n' - 2AA' + J_n diag(AA')'}$. Double-centering eliminates the squared terms leaving only the cross product term. The first ten eigenvalues are: 32.56, 7.37, 2.39, 1.54, 1.33, 1.13, 1.03, .96, .82, .78.

[11] The Southern states are the 11 states of the Confederacy plus Kentucky and Oklahoma. This is the definition used by *Congressional Quarterly* and the definition used throughout Poole and Rosenthal (1997).

[12] To calculate the fit between two legislator configurations, the legislator coordinates in Figure 3 were rotated to best match the NOMINATE, Heckman-Snyder, and optimal classification configurations using Schonemann's (1966) technique. The r-squares between the corresponding legislator coordinates on the first and second dimensions were .986 and .977 respectively for NOMINATE, .959 and .941 respectively for Heckman-Snyder, and .980 and .919 respectively for optimal classification.