Measuring Bias and Uncertainty in Ideal Point Estimates via the Parametric Bootstrap

Jeffrey B. Lewis
Department of Political Science
University of California, Los Angeles
Los Angeles, CA 90095

Keith T. Poole
Department of Political Science
University of Houston
Houston, TX 77204-3011

20 May 2003

Abstract

Over the last 15 years a large amount of scholarship in legislative politics has used NOMINATE or other similar methods to construct measures of legislators' ideological locations. These measures are then used in subsequent analyses. Recent work in political methodology has focused on the pitfalls of using such estimates as variables in subsequent analysis without explicitly accounting for their uncertainty and possible bias (Herron and Shotts, 2003). This presents a problem for those employing NOMINATE scores because estimates of their unconditional sampling uncertainty or bias have until now been unavailable. In this paper, we present a method of forming unconditional standard error estimates and bias estimates for NOMINATE scores using the parametric bootstrap. Standard errors are estimated for the 90th Senate and the 93rd House in two dimensions and the 105th Senate in one dimension. Standard errors of first dimension placements are in the 0.03 to 0.08 range. The results are compared to those obtained using the MCMC estimator of Clinton, Jackman, and Rivers (2002) for the 105th Senate in one dimension. We also show how the bootstrap can be used to construct standard errors and confidence intervals for auxiliary quantities of interest such as ranks and the location of the median Senator.

1 Introduction

The purpose of this paper is to show a general method for obtaining standard errors, confidence intervals, and other measure of uncertainty for the ideal point estimates obtained from NOMINATE and other similar scaling procedures. The number of parameters estimated by these scaling methods is so large that conventional approaches to obtaining standard errors have proven impractical. Our approach is to use the parametric bootstrap (Efron, 1979; Efron and Tibshirani, 1993) to obtain standard errors and other measures of estimation uncertainty.

Nominal Three-step Estimation (NOMINATE) was originally developed by Poole and Rosenthal (1985; 1991; 1997) to scale U.S. Congressional roll call data. The method is based upon a probabilistic spatial voting model that utilizes a random utility function (McFadden, 1976). NOMINATE produces ideal points for the legislators and two points – one corresponding to the Yea outcome and one corresponding to the Nay outcome – for every roll call along with the parameters of the utility function. If there were 100 legislators and 500 roll calls then NOMINATE estimates 1,101 parameters in one dimension and 2,202 parameters in two dimensions using the 50,000 observed choices. In a classical maximum likelihood framework the standard errors are obtained from inverting either the information matrix or inverting the analytical Hessian matrix directly. Unfortunately, this entails inverting a very large matrix and is computationally difficult even with modern computers.

Because of these difficulties, NOMINATE only computes *conditional* standard errors. For example, the standard errors for a legislator's ideal point prameters are obtained from inverting just the information matrix for those parameters – the roll call parameters are fixed and each legislator's parameters are independent of other legislator parameters. Although computationally easy to compute, the quality of these conditional standard errors is suspect and they probably underestimate the true uncertainty. Indeed, Clinton, Jackman, and Rivers (2002) show that the standard errors from W-NOMINATE are smaller than those they derive from an MCMC approach. Unfortunately, a direct comparison is not possible because CJR use a quadratic utility function while NOMINATE is based upon a normal distribution utility function. Because of the computational intensity of the MCMC approach, it has of yet not been applied to the NOMINATE model.

We bridge this gap by applying the parametric bootstrap to W-NOMINATE as well as the Quadratic-Normal (QN) scaling procedure developed by Poole (2000; 2001). The QN procedure is based upon the quadratic utility function so that we can compare the bootstrapped standard errors from both procedures with the standard errors derived from the CJR MCMC method.

In the next section we briefly describe the bootstrap. In section three, we explain how the parametric bootstrap is applied to W-NOMINATE and QN. In section four we present the results of applying the bootstrap to roll call data from the 90th Senate and 93rd House in two dimensions. In section five, we compare the bootstrap results for NOMINATE and QN to the CJR's IDEAL model using the 105th Senate in one dimension. In section six, we use the bootstrap to calculate the uncertainty in additional quantities of interest arising from the ideal point estimation. In section seven, we conclude.

2 The Parametric Bootstrap

Excellent discussions of the bootstrap are provided in Efron and Tibshirani (1993), Hall (1985), Mooney (1996), and Young (1994). We will only briefly discuss the bootstrap here and we will focus on the less common parametric form of the bootstrap that we employ. Typically, the bootstrap is used to provide non-parametric estimates of the standard errors and confidence intervals of estimators. This can be particularly useful in cases where robustness to distributional assumptions is of great concern, the estimation is itself non-parametric, or the samples are too small to rely on asymptotic approximations. In our case, the reason to apply to the bootstrap is mainly computational convenience, though, as will be shown in Section 6, the bootstrap also allows us to estimate the uncertainty of auxiliary quantities of interest such as the location of the median legislator. Recovering the variance-covariance matrix of parameter estimates by forming and inverting the full (estimated) information matrix for roll voting models such as NOMINATE or QN is sufficiently difficult that the bootstrap is an attractive and tractable alternative.

Following Efron (1979), let $\boldsymbol{\theta}$ be a vector of parameters to be estimated and $\hat{\boldsymbol{\theta}}$ be an estimator of $\boldsymbol{\theta}$. The sampling distribution of $\hat{\boldsymbol{\theta}}$ is dependent on the joint distribution of the data. Let F be the joint cumulative distribution of the data. We can then write $\hat{\boldsymbol{\theta}}(F)$. If F was known, the sampling distribution of $\hat{\boldsymbol{\theta}}$ could be ascertained directly by analytic or simulation methods. Using simulation methods, repeated samples would be drawn from F, $\hat{\boldsymbol{\theta}}$ calculated for each sample, and features of the sampling distribution of $\hat{\boldsymbol{\theta}}$ approximated with arbitrary precision (as the number of pseudo samples grows large).[1] Efron (1979) shows that $\hat{\boldsymbol{\theta}}(\hat{F})$ can provide a good approximation of $\boldsymbol{\theta}(F)$ where $\hat{F}$ is an estimate of F based on sample data. Even in small samples approximating F by some $\hat{F}$ will in many situ ations provide excellent estimates of the sampling distribution of $\hat{\boldsymbol{\theta}}(F)$. Some asymptotic properties of $\hat{\boldsymbol{\theta}}(\hat{F})$ in fairly general (usually univariate) settings are given in Hall (1994) and cites therein.

In simple settings where the data are independently and identically distributed, the non-parametric ML estimate of the marginal distribution of each observation is simply the empirical distribution of the observations. In this case, an approximate draw from the joint distribution F can be made by sampling with replacement n draws from the observed data, where n is the number of observations . Consider the case where the data are observations on a single variable Y. Letting $\vec{y} = (y_1, y_2, \ldots, y_n)$ be a vector observations on Y, drawing a sample from $\hat{F}$ is a matter of sampling with replacement n values in turn from $\vec{y}$ where each element of $\vec{y}$ is selected with probability $1/n$ at each turn.

In cases where the data are not i.i.d., simple resampling from the data does not yield draws from the joint distribution of $\hat{F}$ and, thus, does not yield approximate draws from F. If the dependence in the data is temporal or spatial, "block" resampling schemes that draw randomly groups of adjacent observations from the data have been suggested (see Hall 1985, Hall 1994, and cites therein). In the case of roll call voting data both the rows and columns of the data matrix are dependent. Indeed, it is these dependencies that are exploited in recovering the ideal point and vote parameters. However, unlike time series or spatial data, we have no *ex ante* expectations about which elements of the vote matrix are "close" to which others.[2] Given this lack of *ex ante* information about how to structure a block resampling scheme, there is no obvious way (at least to us) of implementing the

[1]Monte Carlo experiments, for example, posit a given F and then recover the sampling distribution of $\hat{\boldsymbol{\theta}}(F)$.

[2]Obviously, elements in the same column or row will be expected to be particularly dependent.

non-parametric bootstrap in this case.[3]

On the other hand, the parametric bootstrap is easy to apply to maximum likelihood estimators such as NOMINATE or QN. In the parametric bootstrap, $\hat{F}$ is estimated directly from the likelihood itself. That is, the joint distribution of the data is approximated by the likelihood evaluated at $\hat{\boldsymbol{\theta}}$. In either QN or NOMINATE individual vote choices are independent conditional on the value of the roll call and ideal point parameters. Thus, conditional on the estimated roll call parameters and ideal points, draws from the joint distribution of the data matrix can be made by drawing from each element of the data matrix independently. Because the estimated parameters are not equivalent to the true parameters, the estimated joint distribution of the data, $\hat{F}$, differs from F, as in the non-parametric case. Note that estimating F based on the parameter estimates is similar to substituting the information matrix evaluated at the estimated parameter values (as opposed to the true values) when approximating the variance covariance matrix of ML estimators in the usual way (see Efron 1982).

By the Slutsky theorem, the parametric bootstrap estimate of F will be consistent if $\hat{\boldsymbol{\theta}}$ is consistent for $\boldsymbol{\theta}$.[4] Precise conditions under which the models described above are consistent have yet to be established. The models are known not to be consistent as the number of roll calls or the number of legislators goes to infinity, though it may be that sending the number of members, the number of votes, and the ratio of votes to members to infinity is sufficient (Londregan 2000). Thus, we cannot appeal to standard asymptotic results to establish the admissibility of the parametric bootstrap estimator in this case. However, extensive Monte Carlo experiments on both these models and similar models in psychometrics suggest that accurate and reliable estimates are obtained if the data matrix has rank of 100 (Lord 1983, Poole and Rosenthal, 1997)—the rank of the Senate roll data considered below.[5]

[3]In structural equations modeling, which also involves dependent data, a non-parametric bootstrap is possible because these models operate on the variance-covariance matrix of the data which can be simulated by simple resampling techniques (Bollen and Stein 1992).

[4]To appeal to the Slutsky theorem, certain conditions must hold. In particular, F must be a continuous function of $\boldsymbol{\theta}$.

[5]Our own Monte Carlo experiments confirm the effectiveness of the parametric bootstrap technique in this setting.

3 Applications of the Parametric Bootstrap

The parametric bootstrap is very simple conceptually. In a maximum likelihood or probabilistic framework, the first step is to compute the likelihood function of the sample. The second step is to draw, for example, 1000 samples from the likelihood density and compute for each sample the maximum likelihood estimates of the parameters of interest. Finally, the sample variances computed from these 1000 values are the estimators of the variances of the parameters (Efron and Tibshirani, 1993, ch. 6).

When applied to a scaling method such as W-NOMINATE, the first step is to run the program to convergence and then calculate the probabilities for the observed choices. This produces a legislator by roll call matrix containing the estimated probabilities for the corresponding actual roll call choices of the legislators. To draw a random sample we simply treat each probability as a weighted coin and we "flip" the coin. We do this by drawing from a uniform distribution over zero to one —$U(0,1)$— and if the random draw is less than or equal to the estimated probability then our sampled value is the observed choice. If the random draw is greater than the estimated probability, then our sampled value is the opposite of the observed choice; that is, if the observed choice is Yea then our sampled value is Nay. We then apply W-NOMINATE to this sample roll call matrix. This process is repeated 1000 times and the variances of the legislator ideal points are calculated using the 1000 estimated bootstrap configurations.

Technically, let c_{ij} be the observed choice for the ith legislator $(i = 1, \ldots, p)$ on the jth roll call $(j = 1, \ldots, q)$ where the possible choices are Yea or Nay. In the U.S. Congress there is very little policy related abstention (Poole and Rosenthal, 1997) so we treat non-voting as missing data. Let $\hat{P}_{ijc}$ be the estimated probability for the observed choice and let ϕ be a random draw from $U(0,1)$. Let $\hat{c}_{ij}$ be the sampled choice. The sample rule is:

$$\hat{c}_{ij} = \begin{cases} c_{ij}, & \text{if } \phi \leq \hat{P}_{ijc} \\ \sim c_{ij}, & \text{if } \phi > \hat{P}_{ijc} \end{cases} \tag{1}$$

where $\sim c_{ij}$ represents the opposite choice to c_{ij}. This technique allows the underlying uncertainty to propagate through to all the estimated parameters. To see this, note that as the $\hat{P}_{ijc} \rightarrow 1$, the $\hat{c}_{ij} \rightarrow c_{ij}$, that is, sample choices become the observed choices so that the bootstrapped variances for the parameters of the model go to zero. If the fit of the model is poor, for example, if the $\hat{P}_{ijc}$ are between 0.5 and 0.7, then the bootstrapped variances for the parameters will be large.

3.1 The W-NOMINATE Model

The $\hat{P}_{ijc}$'s estimated by W-NOMINATE are based upon a standard random utility model (McFadden, 1976); namely, the utility for the observed choice is:

$$U_{ijc} = u_{ijc} + \varepsilon_{ijc} = \beta e^{-\left[\frac{\sum_{k=1}^{s} {w_k}^2 (x_{ik} - z_{jck})^2}{2}\right]} + \varepsilon_{ijc} \tag{2}$$

where u_{ijc} is the deterministic portion of the utility function, ε_{ijc} is the stochastic portion, and s is the number of dimensions. The ith legislator's ideal point on the kth dimension (k=1,. . . ,s) is x_{ik} and the policy outcome of the observed choice of the jth roll call on the kth dimension is z_{jck}. β is an overall signal-to-noise ratio common to all legislators and the first dimension weight w_1 is set to one and the weights on the second and higher dimensions are estimated.

The errors are distributed as the logarithm of the inverse of an exponential variable (Dhrymes, 1978, p. 342); namely:

$$f(\varepsilon) = e^{-\varepsilon} e^{-e^{-\varepsilon}} \text{ for } -\infty < \varepsilon < \infty \tag{3}$$

and the distribution of the difference between two errors is:

$$f(\varepsilon_{ijn} - \varepsilon_{ijy}) = \frac{e^{-(\varepsilon_{ijn} - \varepsilon_{ijy})}}{[1 + e^{-(\varepsilon_{ijn} - \varepsilon_{ijy})}]^2}.$$

Let $\varphi = \varepsilon_{ijn} - \varepsilon_{ijy}$. The W-NOMINATE choice probabilities are:

$$\begin{aligned} \hat{P}_{ijy} &= P(U_{ijy} > U_{ijn}) \\ &= P(\varepsilon_{ijn} - \varepsilon_{ijy} < u_{ijy} - u_{ijn}) \\ &= \int\limits_{-\infty}^{u_{ijn}-u_{ijy}} \frac{e^{-\varphi}}{(1+e^{-\varphi})^2} d\varphi \\ &= \frac{e^{u}_{ijy}}{e^{u_{ijy}} + e^{u_{ijn}}} \end{aligned} \tag{4}$$

and

$$\hat{P}_{ijn} = 1 - \hat{P}_{ijy}.$$

The likelihood function is:

$$L = \prod_{i=1}^{p} \prod_{j=1}^{q} \hat{P}_{ijc} \tag{5}$$

W-NOMINATE maximizes (5) subject to a set of constraints on the legislator ideal points and the roll call outcome points. The legislator ideal points and the midpoint of the two policy outcome points for each roll call are constrained to be in the interval $[-1, 1]$ in one dimension and in the unit hypersphere in two or more dimensions. In addition, W-NOMINATE estimates one dimension at a time similar to the classic eigenvector extraction algorithm. In one dimension, the most extreme legislators at opposite ends of the first dimension are always set to -1 and +1 so that there will *always* be at least one legislator at -1 and at least one legislator at +1. When additional dimensions are estimated and a legislator's second or higher dimension estimated coordinate moves her outside the unit hypersphere then a grid search is conducted on the surface of the hypersphere to find the best point. Thus, in general, no estimated coordinates are equal to 1 or -1 when more than one dimension is estimated. A similar process is performed for the roll call midpoints.

Because of this constraint structure, W-NOMINATE is technically a constrained maximum-likelihood method—it maximizes a likelihood function subject to several constraints. Some these constraints identify the scale, location, and rotation of the underlying space and relaxing those constraints would not increases the likelihood. However, constraints beyond those required for identification are imposed. These constraints have the effect of slightly lowering the $\hat{P}_{ijc}$'s relative to what they would be without the constraints and the bootstrapped standard errors for the legislators are slightly inflated as a result.

3.2 The Quadratic-Normal Model

In QN, the utility for the observed choice is:

$$U_{ijc} = u_{ijc} + \varepsilon_{ijc} = -\sum_{k=1}^{s} (x_{ik} - z_{jkc})^2 + \varepsilon_{ijc}. \tag{6}$$

The errors are normally distributed so that the distribution of the difference between two errors for legislator i on roll call j is:

$$f(\varepsilon_{ijy} - \varepsilon_{ijn}) \sim N(0, \sigma_i^2).$$

The error variance in QN is allowed to vary across legislators. However, to ensure comparability with W-NOMINATE, we constrain all the σ_i to be equal to each other. Hence our assumption about the error is:

$$f(\varepsilon_{ijy} - \varepsilon_{ijn}) \;\sim\; N(0, \sigma^2).$$

Using vector algebra, the difference between the deterministic utilities for the Yea and Nay alternatives simplifies to:

$$u_{ijy} - u_{ijn} = 2\vec{x}_i'(\vec{z}_{jy} - \vec{z}_{jn}) - (\vec{z}_{jy} + \vec{z}_{jn})'(\vec{z}_{jy} - \vec{z}_{jn}) \tag{7}$$

where $\vec{x}_i$ is legislator i's $s \times 1$ vector of coordinates, and $\vec{z}_{jy}$ and $\vec{z}_{jn}$ are the $s \times 1$ vectors of coordinates for the Yea and Nay alternatives, respectively. Let the $s \times 1$ vector $\vec{z}_{mj}$ be the midpoint of the Yea and Nay alternatives; that is:

$$2\vec{z}_{mj} = \vec{z}_{jy} + \vec{z}_{jn}.$$

If there were no voting error, a plane could 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 $\vec{n}_j$ be the $s \times 1$ normal vector for the jth roll call where $\vec{n}_j'\vec{n}_j = 1$. Because $\vec{n}_j$ and its reflection $-\vec{n}_j$ are both normal vectors, without loss of generality, let the first coordinate in the normal vector be greater than zero – $n_{j1} > 0$. The vector $\vec{z}_{jy} - \vec{z}_{jn}$, by definition, is perpendicular to the plane so that:

$$\vec{z}_{jy} - \vec{z}_{jn} = \gamma_j \vec{n}_j \tag{8}$$

where

$$\gamma_j = \left[\sum_{k=1}^{s} (z_{jky} - z_{jkn})^2\right]^{\frac{1}{2}} \text{ if } \vec{z}_{jy}'\vec{n}_j > \vec{z}_{jn}'\vec{n}_j$$

and

$$\gamma_j = -\left[\sum_{k=1}^{s} (z_{jky} - z_{jkn})^2\right]^{\frac{1}{2}} \text{ if } \vec{z}_{jy}'\vec{n}_j < \vec{z}_{jn}'\vec{n}_j.$$

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

Substituting (8) into (7) we get:

$$u_{ijy} - u_{ijn} = 2\gamma_j \vec{x}_i'\vec{n}_j - \gamma_j(\vec{z}_{jy} + \vec{z}_{jn})'\vec{n}_j = 2\gamma_j \vec{x}_i'\vec{n}_j - \vec{z}_{mj}'\vec{n}_j) = 2\gamma_j(w_i - m_j) \tag{9}$$

where w_i is the projection of the legislator's ideal point onto the line defined by $\vec{n}_j$ and its reflection $-\vec{n}_j$ and m_j is the projection of the midpoint of the Yea and Nay alternatives. Equation (9) shows that:

if $\gamma_j > 0$ and $w_i > m_j$, or

if $\gamma_j < 0$ and $w_i < m_j$, then $u_{ijy} > u_{ijn}$.

The QN choice probabilities are:

$$\begin{gathered}\hat{P}_{ijy} = P(U_{ijy} > U_{ijn}) = P\left(\varepsilon_{ijn} - \varepsilon_{ijy} < u_{ijy} - u_{ijn}\right) = \\ \Phi\left[\frac{2\gamma_j}{\sigma}(w_i - m_j)\right]\end{gathered} \tag{10}$$

and

$$\hat{P}_{ijn} = 1 - \hat{P}_{ijy}.$$

QN constrains the legislator ideal points and the roll call midpoints to lie within a unit hypersphere. The only other constraint that is imposed occurs if a perfectly classified roll call is encountered—that is, all legislators voting Yea are on one side of the cutting line and all legislators voting Nay are on the opposite side of the cutting line. In this instance, the directional distance term will want to go to infinity so that a simple constraint is imposed to keep this from happening; namely, $|\gamma_j| < 10$. As we noted above, to ensure comparability between QN and W-NOMINATE we have set all the σ_i to the same value, σ. However, since only the ratio γ_j/σ is identified, we adopt the standard convention of setting $\sigma = 1$ and focus on the γ_j in our analyses.

Before we turn to our bootstrap results note that the effect of the distance between the Yea and Nay alternatives is different in the two models – equations (4) and (10). QN and W-NOMINATE treat roll calls quite differently. In W-NOMINATE the roll call outcome points are parameters (technically, they are parameterized by half the distance between the roll call outcome coordinates and the midpoint of the two outcome coordinates). So in two dimensions, four parameters are estimated for each roll call. In contrast, in QN a roll call is parameterized by its normal vector, $\vec{n}_j$, γ_j, and m_j. In two dimensions this is also four parameters. The crucial geometric difference between W-NOMINATE and QN is that in QN the position of the cutting plane is determined by $\vec{n}_j$ and m_j. In W-NOMINATE the cutting plane is inherently linked to the two outcome coordinates. In both models the distance between the outcome coordinates cannot be disentangled from roll call specific noise (Poole and Rosenthal, 1997). If there is a very high level of error on a roll call then the distance between the two outcome coordinates will tend to decrease. In QN this effect is captured in γ_j while in W-NOMINATE it tends to bring the two outcome points close together near the edge of the unit hypersphere. Hence, with very high error roll call cutting planes in W-NOMINATE tend to be tangent to the unit hypersphere while in QN the cutting planes can be in the interior because of the γ_j parameter.

The situation is a bit different with very low error. To see this, assume a one-dimensional space and place a legislator at 1.0, the midpoint at 0.0, and the Yea and Nay alternatives at 1.0 and −1.0, respectively. With $\beta = 15$ and $w_1 = 0.5$ the utility of Yea is 15 and the utility of Nay is $15e^{-1/2}$. This produces a probability of about 0.997. The corresponding QN probability is $\Phi(4) = 0.99997$ because $\gamma_j = 2$. On a perfectly classified roll call the QN choice probabilities will all be very close to 1.0 – as the γ_j increases the choice probabilities increase. In W-NOMINATE, as the distance between the Yea and Nay alternatives increases, the choice probabilities will start to decline and eventually go to 0.5 when the distance becomes very large.

As a practical matter this analytical difference between equations (4) and (10) does not have a very large effect. As we show below, the constraints used by W-NOMINATE — in part inspired by the functional form of equation (4) (Poole and Rosenthal, 1997, Appendix A) – have more impact than the analytical difference.

4 Parametric Bootstrap Results for the 90th Senate and 93rd House

Although we obtain bootstrap estimates of the means and standard deviations of all the parameters, we focus our analysis on the legislator ideal points because they are used in a wide variety of secondary analyses by many researchers. Let $\hat{\mathbf{X}}$ be the $p \times s$ matrix of legislator coordinates estimated by either W-NOMINATE or QN. Let $h = 1, \ldots, m$ be the number of bootstrap trials and let $\mathbf{X}_h$ be the $p \times s$ matrix of legislator coordinates estimated on the hth bootstrap trial. The legislator and roll call coordinates are only identified up to an arbitrary rotation in the s-dimensional space. This arbitrary rotation must be removed to ensure that the bootstrapping process produces accurate estimates of the standard deviations of the parameters. In particular, we assume that:

$$\hat{\mathbf{X}} = \mathbf{X}_h \mathbf{V} + \mathbf{E}. \tag{11}$$

$\mathbf{V}$ is an $s \times s$ matrix such that $\mathbf{V}'\mathbf{V} = \mathbf{V}\mathbf{V}' = \mathbf{I}_s$ where $\mathbf{I}_s$ is an $s \times s$ identity matrix and $\mathbf{E}$ is a $p \times s$ matrix of errors. In psychometrics, equation (11) is known as the *orthogonal procrustes problem*. We use Schonemann's (1966) solution to remove the arbitrary rotation, $\mathbf{V}$. Note that we are *rigidly rotating* $\mathbf{X}_h$, we are not altering the estimated points vis a vis one another in any way. Consequently, in our discussion below we will simply denote the hth bootstrap trial matrix as $\mathbf{X}_h$ to avoid notational clutter.

For our first example we apply the parametric bootstrap to W-NOMINATE and QN for the 90th Senate (1967-68). We performed 1000 bootstrap trials as described above and computed the means and standard deviations of all the estimated parameters. For example, for the ith legislator on the kth dimension the mean of the bootstrap trials is:

$$\bar{x}_{ik} = \frac{\sum_{h=1}^{m} x_{hik}}{m} \tag{12}$$

where $m = 1000$ is the number of trials and x_{hik} is the estimated coordinate on the hth trial. The corresponding standard deviation is:

$$\mathrm{SE}\left(x_{ik}\right) = \sqrt{\frac{\sum_{h=1}^{m} \left(x_{hik} - \hat{x}_{ik}\right)^2}{m-1}} \tag{13}$$

where $\hat{x}_{ik}$ is the coordinate estimated by W-NOMINATE or QN.

We take a conservative approach and use the estimated coordinate, $\hat{x}_{ik}$, rather than the mean of the bootstrap trials, $\bar{x}_{ik}$, as our "sample mean" in our calculation of the standard deviation. This inflates the standard deviations somewhat but we feel it is better to err on the safe side and not *underreport* the standard deviations.

4.1 The 90th Senate in Two Dimensions

Figure 1 shows the estimated ideal points for the 90th Senate from W-NOMINATE along with the bootstrapped standard errors. The cross hairs through the ideal points show the 95 percent confidence intervals. The standard errors are small even for the second dimension. On the first dimension, 97 of 101 standard errors were 0.09 or less and on the second dimension, 90 of 101 standard errors were 0.10 or less. Given that W-NOMINATE constrains the legislator ideal points to lie within the unit circle, these standard errors are small relative to the estimated ideal points. For most Senators the correlation between the estimated first and second dimension coordinates is very low. Normal theory confidence ellipses are shown for Senators whose first dimension coordinate is correlated with their second dimension coordinate at $|\rho| > 0.15$. This correlation is a consequence of the constraint that legislator ideal points lie within the unit circle.

There is very little evidence of bias. Regressing the bootstrapped mean ideal points on the estimated ideal points yields an r-square of 0.999 for the first dimension (bootstrap_1st = -.017 + 1.015*W-Nom_1st) and an r-square of .996 on the second dimension (bootstrap_2nd = .005 + 1.038*W-Nom_2nd).

The upper left-hand panel of Figure 2 shows the bootstrapped standard errors for the two dimensions graphed against one another by Senator. As expected, the standard errors on the first dimension are smaller than the second. Only seven Senators had larger standard errors on the first dimension than they had on the second. Charles Goodell (R-NY) is a notable outlier on both dimensions. He was appointed to the Senate on 10 September 1968 to replace Robert F. Kennedy. Goodell only voted on 39 roll calls of which only 31 were scalable (2.5 percent in minority or better).

The remaining two panels in the first row of Figure 2 show the conditional standard errors from W-NOMINATE versus the bootstrapped standard errors by dimension. As expected, the conditional standard errors are smaller than the bootstrapped standard errors especially on the first dimension. On the first dimension the conditional standard errors are about half the magnitude of the bootstrapped standard errors while on the second dimension the magnitude difference is not as large.

The second row of panels in Figure 2 display the Senator coordinates on the two dimensions versus the respective bootstrapped standard errors. The unit circle constraint shows up clearly in the plots. Senators near the edges of the space have smaller standard errors.

Figures 3 and 4 show the bootstrapping results for QN applied to the 90th Senate. Figure 3 shows the estimated ideal points for the 90th Senate from QN along with the bootstrapped standard errors. The cross hairs through the ideal points show the 95 percent confidence intervals. The standard errors for QN are smaller than those for W-NOMINATE. On the first dimension, 100 of 101 standard errors were 0.09 or less and on the second dimension, 92 of 101 standard errors were 0.10 or less. As was the case for W-NOMINATE, for most Senators the correlation between the estimated first and second dimension coordinates is

Estimated Legislator Locations from the W-NOMINATE Model, 90th Senate

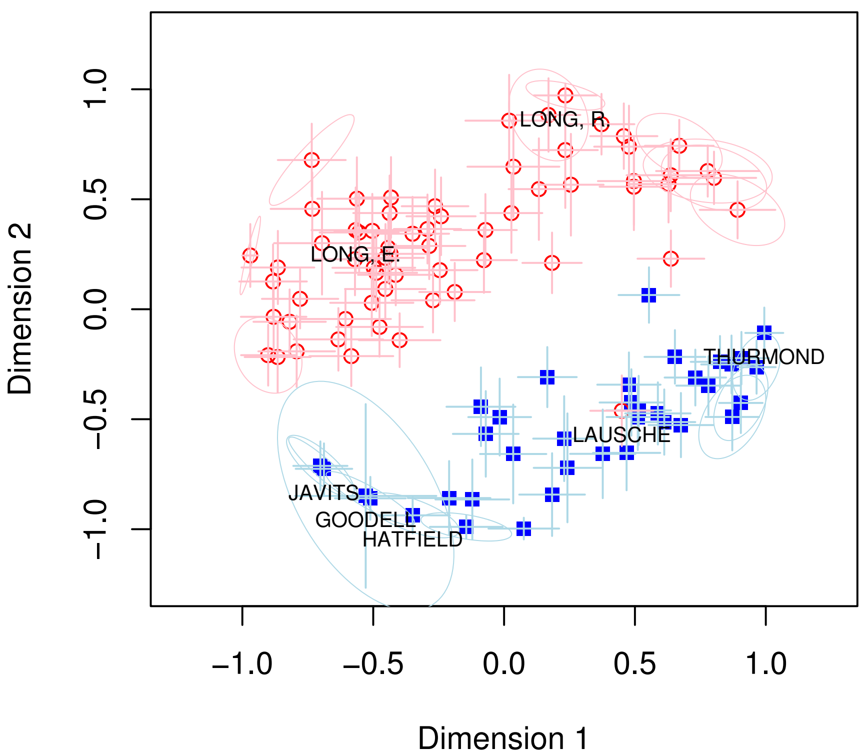


Figure 1: Shows the estimated legislator locations in two dimension based on roll calls taken the 90th Senate. Squares represent Democrats. Circles represent Republicans. The vertical and horizontal lines through each estimate show the 95 percent confidence intervals for each coordinate of a given legislator's position. For most Senators the correlation between the estimated first and second dimension estimates is very low. Normal theory confidence ellipses are shown of Senators whose first dimension coordinate is correlated with their second dimension coordinates at $|\rho| > 0.15$. This correlation results from the identifying constraint that Senators be located on the unit circle.

Bootstrapped and Conditional Standard Error Estimates from the W-NOMINATE Model, 90th Senate

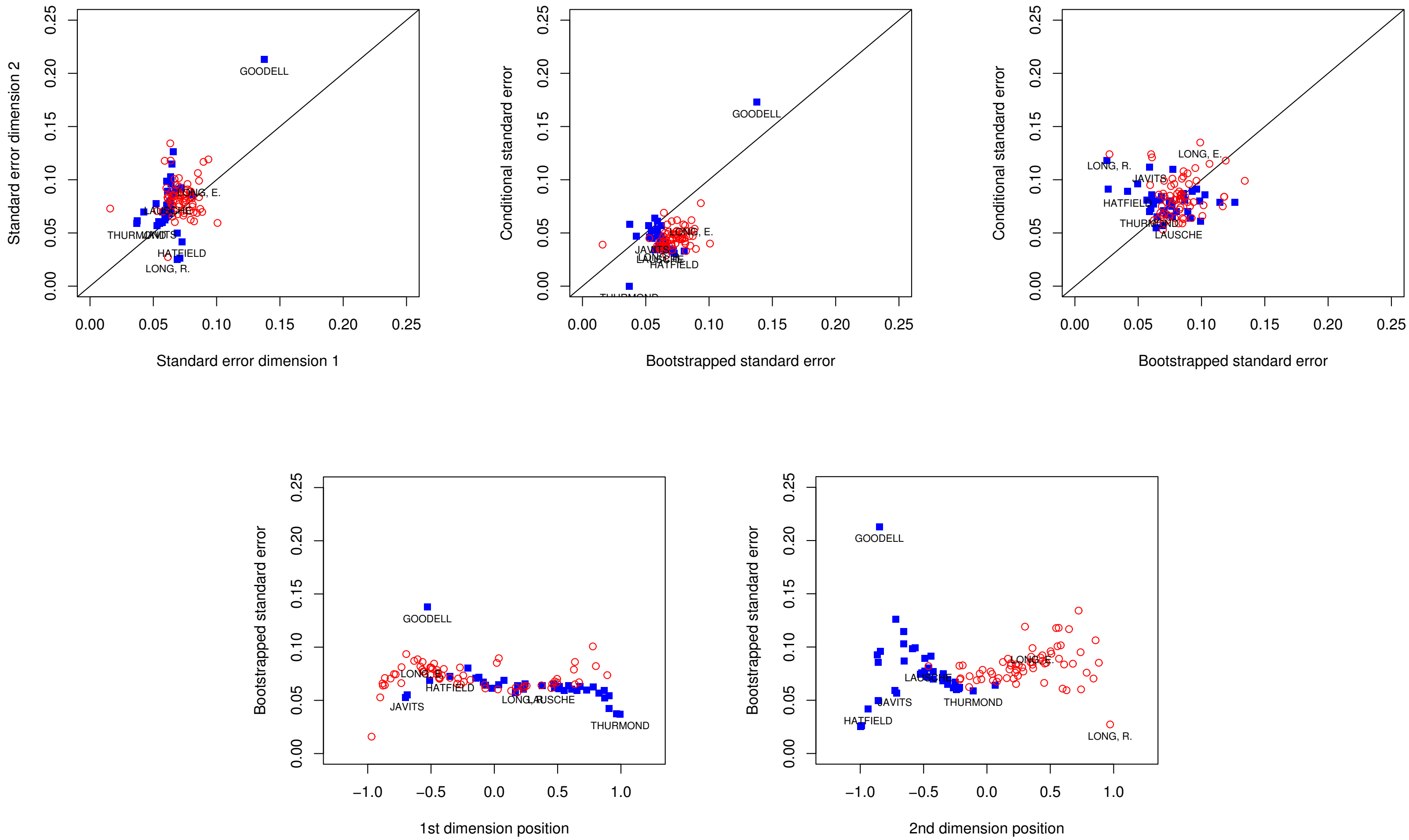


Figure 2: The first of panels shows bootstrapped standard error estimates plotted against each other and against the conditional standard error estimates that take the estimated roll call parameters as known. As expected, the conditional standard errors understate the degree of uncertainty in most cases, particularly for legislators with extreme positions. The second row of panels plots the bootstrap estimates of the standard errors of the scores aga inst the standard error estimates that condition on the estimated roll-call parameters.

very low. Normal theory confidence ellipses are shown for Senators whose first dimension coordinate is correlated with their second dimension coordinate at $|\rho| > .30$. This correlation is a consequence of the constraint that legislator ideal points lie within the unit circle.

Figure 1 and Figure 3 are very similar. QN and W-NOMINATE recover essentially the same configuration of ideal points. Regressing the first dimension W-NOMINATE coordinates on the first dimension QN coordinates produces an r-square of .982 (W-NOM_1st = -.025 + 1.316*QN_1st) and an r-square of .958 for the corresponding second dimension coordinates (W-NOM_2nd = .010 + 1.324*QN_2nd). The fact that the W-NOMINATE configuration is slightly "inflated" vis a vis the QN configuration is due to W-NOMINATE setting the most extreme legislators at opposite ends of the first dimension to –1 and +1. In a one-dimensional scaling there will *always* be at least one legislator at –1 and at least one legislator at +1. When a second dimension is estimated, some of these legislators at or near –1 or +1 may end up on the rim of the circle. QN estimates both dimensions simultaneously and only constrains legislators to lie within the unit hypersphere. Hence, a typical QN configuration will not be as "inflated" as the corresponding configuration from W-NOMINATE. To make the results more comparable across methods, the QN results are re-scaled such that the most extreme members on each end of each dimension are placed at -1 or 1.[6]

There is very little evidence of bias. Regressing the bootstrapped mean ideal points on the estimated ideal points yields an r-square of .998 for the first dimension (bootstrap_1st = -.002 + 1.039*QN_1st) and an r-square of .996 on the second dimension (bootstrap_2nd = .004 + 1.078*QN_2nd).

The upper left-hand panel of Figure 4 shows the bootstrapped standard errors graphed against one another by Senator. Just as with W-NOMINATE, the QN standard errors on the first dimension are smaller than the second. Once again, Charles Goodell (R-NY) is a notable outlier on both dimensions.

The next two panels of Figure 4 show the conditional standard errors from QN versus the bootstrapped standard errors by dimension. As expected, the conditional standard errors are smaller than the bootstrapped standard errors. On both dimensions the conditional

[6]This is accomplished by simple linear transformations which are applied to the ML estimates and to the bootstrapped standard errors.

Estimated Legislator Locations from the Quadratic Normal Model, 90th Senate

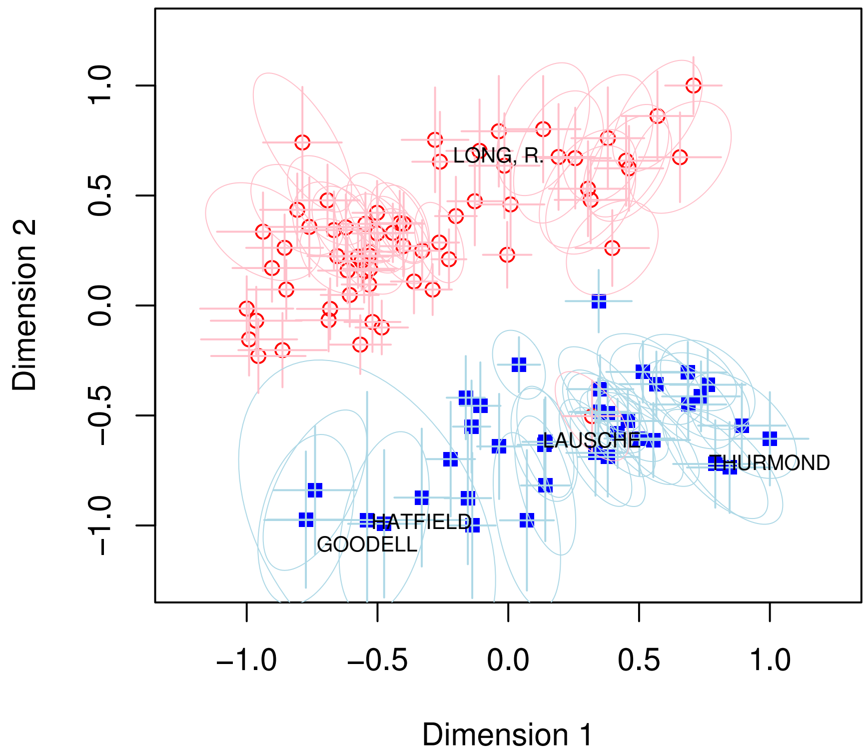


Figure 3: Shows the estimated legislator locations in two dimension based on roll calls taken in the 90th Senate. Squares represent Democrats. Circles represent Republicans. The vertical and horizontal lines through each estimate show the 95 percent confidence intervals for each coordinate of a given legislator's position. For most Senators, the correlation between the estimated first and second dimension estimates is very low. Normal theory 95 percent confidence ellipses are shown of Senators whose first dimension coordinate is correlated with their second dimension coordinates at $|\rho| > 0.3$.

standard errors are about half the magnitude of the bootstrapped standard errors.

The bottom row of panels in Figure 4 display the Senator coordinates on the two dimensions versus the respective bootstrapped standard errors. The patterns in Figure 4 for QN are just the opposite of those for W-NOMINATE in Figure 2. In W-NOMINATE, legislators are constrained to lie on the –1 to +1 interval when the first dimension is estimated. Because W-NOMINATE estimates one dimension at a time this has the effect that legislators who are extreme on the first dimension tend to lie near the unit circle when the second dimension is estimated. Since they cannot wander out of the unit circle when the second dimension is estimated extremists have little "wiggle room" in the W-NOMINATE framework. In contrast, in QN the dimensions are estimated simultaneously. The different constraint structure in QN means that extremists have more "wiggle room". Hence, legislators furthest from the center (recall that we have scaled the QN coordinates to −1 to +1 for graphical purposes only!) tend to have slightly larger standard errors. However, this difference between the two procedures is not really that great. Note that the standard errors are small for both procedures especially for the bulk of legislators who are not extremists.

4.2 The 93rd House In Two Dimensions

Figure 5 shows the estimated ideal points for the 93rd House (1973-74) from W-NOMINATE along with the bootstrapped standard errors in the same format as Figures 1 and 3. The pattern of ideal points is very similar to the 90th Senate. Both Congresses occurred during the three political party period that lasted roughly from 1937 into the 1980s. Voting in Congress was strongly two-dimensional during this period. The second dimension was produced by the split in the Democratic party between Northerners and Southerners over Civil Rights. With the passage of the Civil Rights laws in 1964, 1965, and 1967, this split gradually disappeared by the 1980s (Poole and Rosenthal, 1997; 2001).

The upper left-hand panel of Figure 6 shows the bootstrapped standard errors for the two dimensions graphed against one another by Represenative. The bootstrapped standard errors on the second dimension are about 1.5 times larger than those for the first dimension and the pattern is very similar to that shown in the corresponding panel of Figure 2 for the 90th Senate.

Bootstrapped and Conditional Standard Error Estimates from the Quadratic Normal Model, 90th Senate

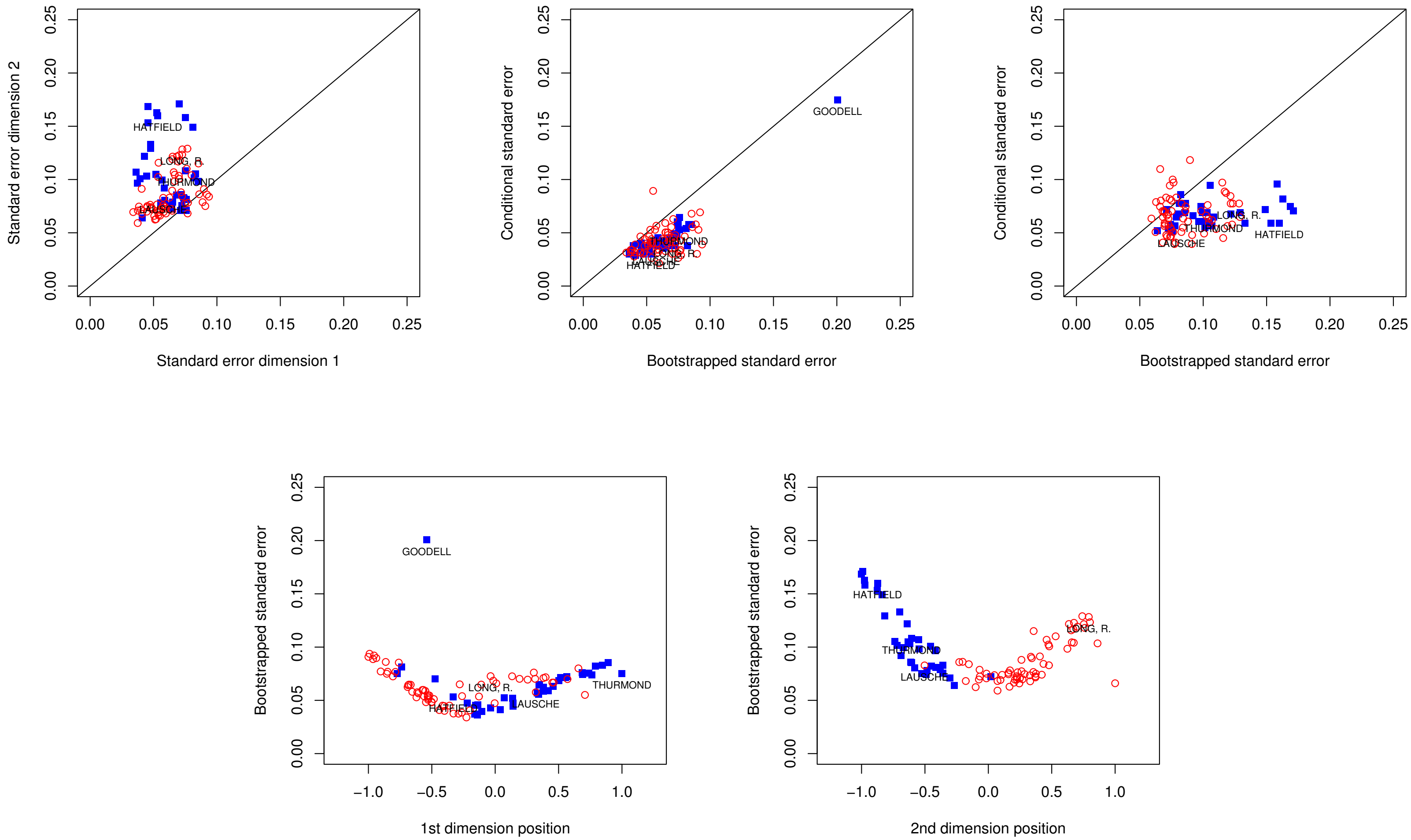


Figure 4: The first rows of panels shows bootstrapped standard error estimates plotted against each other and against the conditional standard error estimates that take the estimated roll call parameters as known. As expected, the conditional standard errors understate the degree of uncertainty in most cases, particularly for legislators with extreme positions. Not shown on right panel is Senator Goodell whose coordinates would be $(0.30, 0.71)$. The second row of panels plots the bootstrap estimates of the standard errors of the QN scores against the estimated Senator locations. Not shown on the right panel is Senator Goodell whose coordinates in that plot would be $(-0.97, 0.30)$.

Estimated Legislator Locations from the W-NOMINATE Model, 93rd House

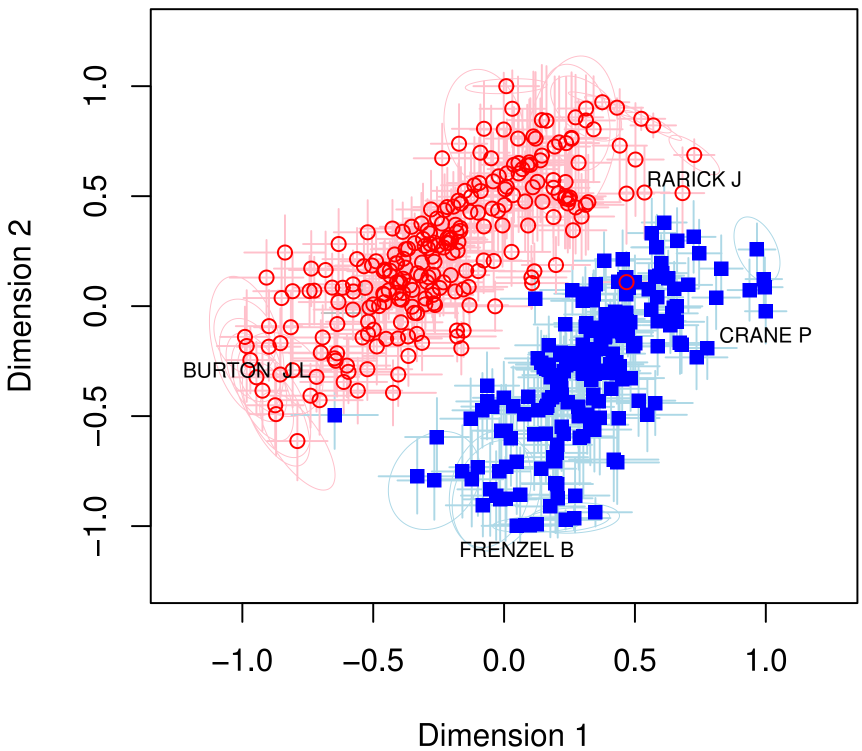


Figure 5: Shows the estimated legislator locations in two dimension based on roll calls taken the 93rd House. Squares represent Democrats. Circles represent Republicans. The vertical and horizontal lines through each estimate show the 95 percent confidence intervals for each coordinate of a given legislator's position. For most Representatives, the correlation between the estimated first and second dimension estimates is very low. Normal theory 95 percent confidence ellipses are shown of Representatives whose first dimension coordinate is correlated with their second dimension coordinates at $|\rho| > 0.15$. This correlation results from the identifying constraint that Representatives be located on the unit circle.

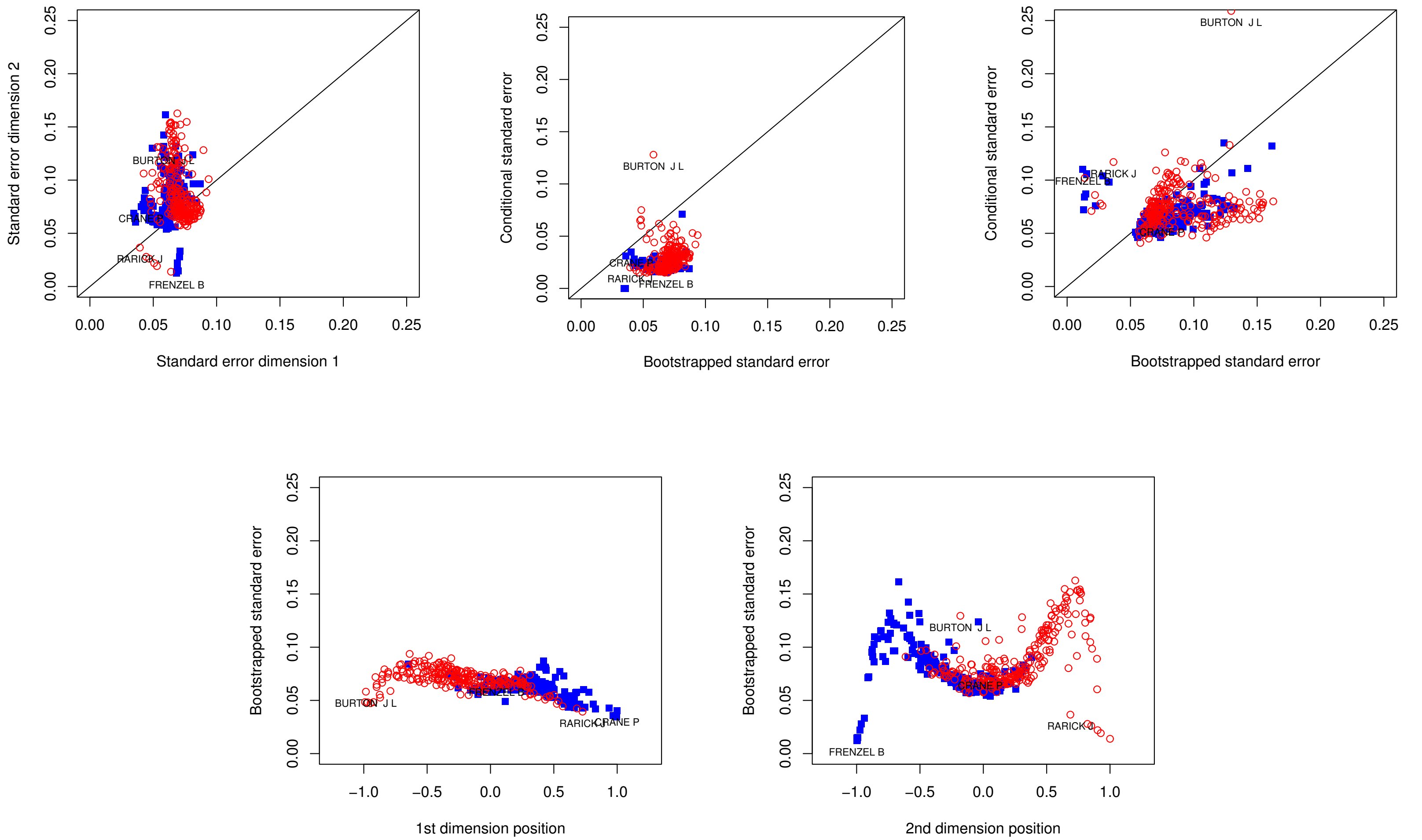
The next two panels of Figure 6 show the conditional standard errors from W-NOMINATE versus the bootstrapped standard errors by dimension. The patterns here are very similar to those in Figure 4 for the 90th Senate. The conditional standard errors are smaller than the bootstrapped standard errors especially on the first dimension.

The bottom row of panels in Figure 6 display the coordinates of the Representatives on the two dimensions versus the respective bootstrapped standard errors. The unit circle constraint shows up clearly in the plots. Representatives near the edges of the space have smaller standard errors.

Figures 7 and 8 show the bootstrapping results for QN applied to the 93rd House. Figure 7 shows the ideal points along with the bootstrapped standard errors. Consistent with our comments above, note that the second dimension standard errors are larger than those for the first dimension.

Figure 7 shows plots of the bootstrapped standard errors against each other by dimension and plots them by dimension against the corresponding conditional standard errors. Once again the conditional standard errors are smaller than the bootstrapped standard errors.

Bootstrapped and Conditional Standard Error Estimates from the W-NOMINATE Model, 93rd House



First Dimension Second Dimension

Figure 6: The first row of panels shows bootstrapped standard error estimates plotted against each other and against the conditional standard error estimates that take the estimated roll call parameters as known. As expected, the conditional standard errors understate the degree of uncertainty in most cases, particularly for legislators with extreme positions. The second row of panels plots the bootstrap estimates of the standard errors of the W-NOMINATE scores against the estimated Representative locations on each dimension.

Estimated legislator locations from the Quadratic Normal Model, 93rd House

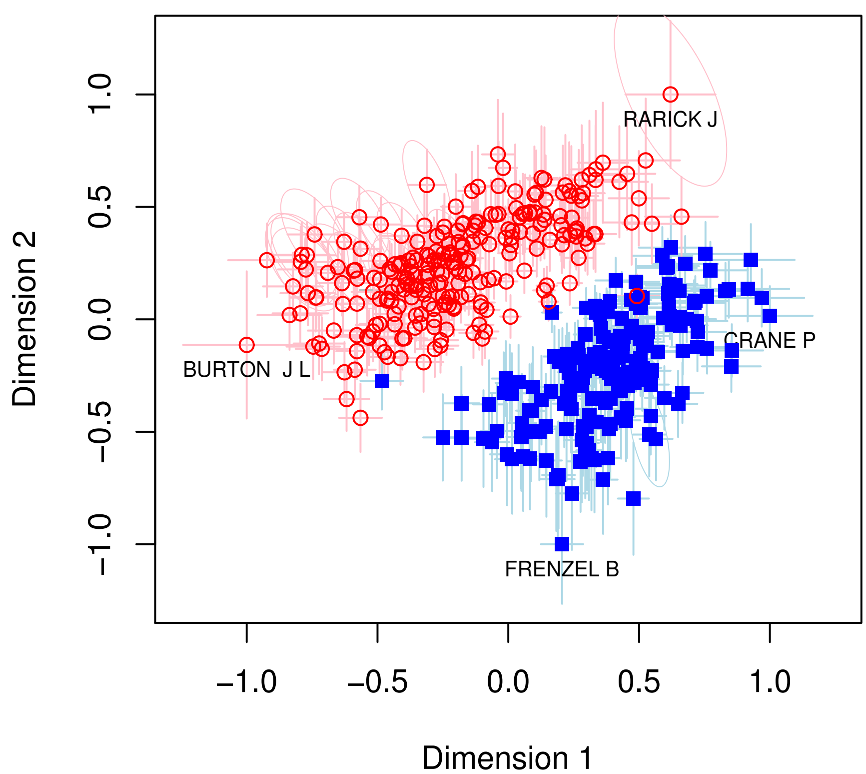


Figure 7: Shows the estimated legislator locations in two dimension based on roll calls taken the 93rd House. Squares represent Democrats. Circles represent Republicans.

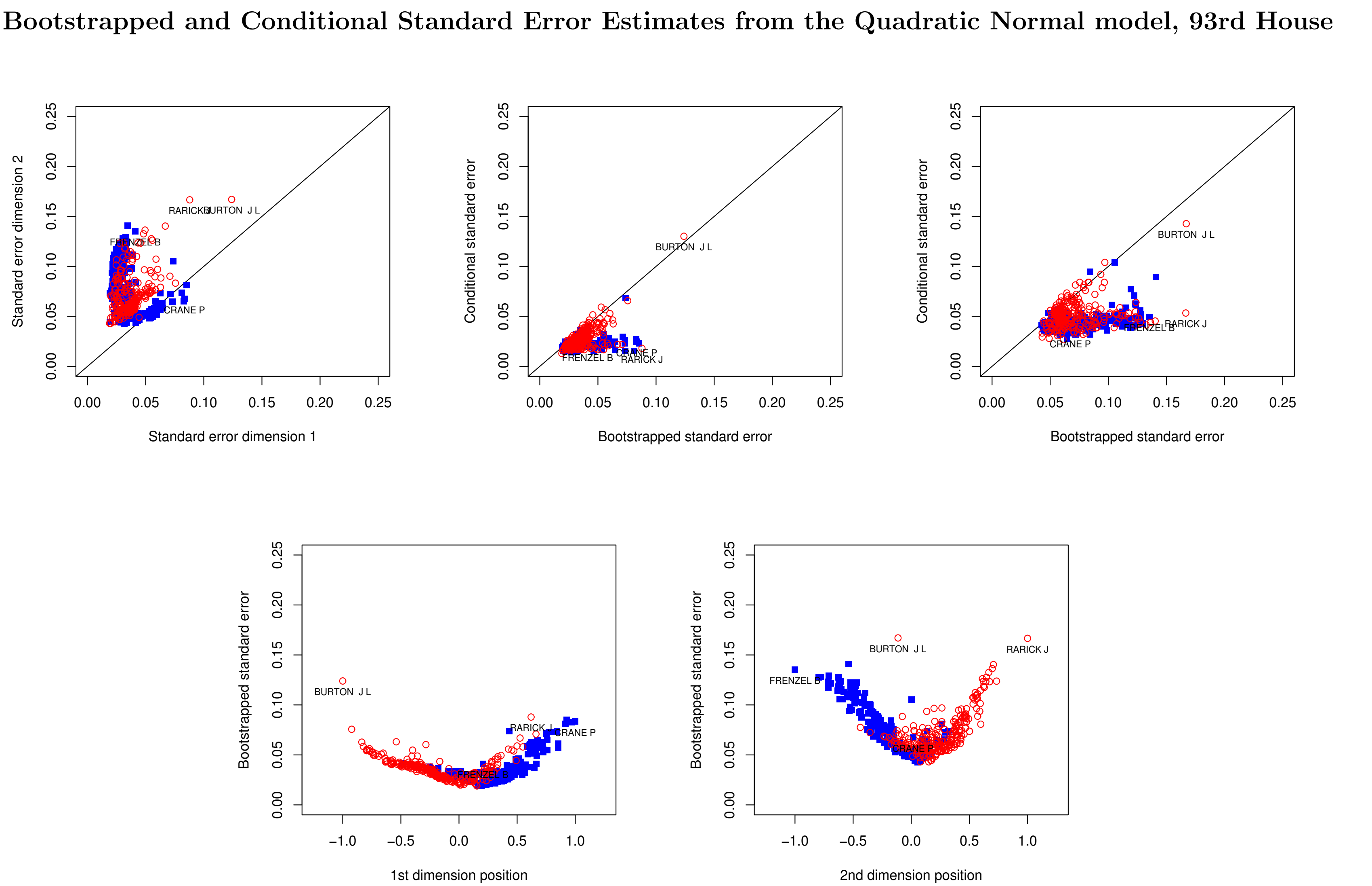


Figure 8: The first row of panels plots the bootstrap estimates of the standard errors of the QN scores for each dimension against each other and against the conditional standard error estimates. The second row of panels plots the bootstrap estimates of the standard errors of the QN scores against the estimated Senator locations.

Finally, Figure 8 shows the Representative coordinates on the two dimensions versus the respective bootstrapped standard errors. The pattern here is similar to that for the 90th Senate shown in Figure 4. Representatives furthest from the center (again recall that we have scaled the QN coordinates to lie between -1 and +1 for graphical purposes only) tend to have larger standard errors.

5 Comparing W-NOMINATE, QN and IDEAL

In this section we apply the parametric bootstrap to the 105th Senate roll calls in one dimension. In this setting, the bootstrapped NOMINATE and QN estimator can be compared to the IDEAL model of Clinton, Jackman, and Rivers (2002). The IDEAL model begins with the same random utility model as QN. However, IDEAL is a Bayesian estimator that is estimated using Markov Chain Monte Carlo (MCMC).[7] The standard implementation of IDEAL as previously described in Jackman (2000a, 2000b) identifies the ideal point distribution by imposing a $N(0,1)$ prior over the distribution of ideal points. This prior establishes the scale and location of the issue dimension. However, as we will see, using the prior to identify the model in this way leads the uncertainty of the estimates to be somewhat overstated.

In order to make the recovered issue space more comparable across models, we rescale the QN and IDEAL results such that the most extreme members on each side of the space are located at -1 and 1.[8] In this way, the intuitive underlying metric of the scale is the same, scores from -1 to 1 and the standard errors are then roughly comparable across methods.

Figure 9 plots the results of the three methods against one another. Looking first at the estimates themselves, we find a very high level of agreement among all three methods. All the points in the plots fall very near to the 45 degree line and there are few differences in the rank order recovered by each method. This result is consistent with previous comparisons of roll scaling techniques (see Heckman and Snyder 1997; Poole and Rosenthal, 1997; Poole, 2000; 2001). Two things are striking about the standard errors estimates. First, all three methods

[7]IDEAL can be implemented in settings with more than a single dimension. However, because of very different identifying restrictions, comparisons between NOMINATE or QN and IDEAL in more than one dimension are difficult.

[8]In QN this involves rescaling the ML estimates and then applying that same rescaling to each bootstrap sample. In IDEAL it involves rescaling (by the same linear transformation) all of the posterior draws such that the mean posterior position of the two most extreme members are -1 and 1.

Comparisons of One-dimensional Ideal Point Estimates and their Standard Errors, 105th Senate

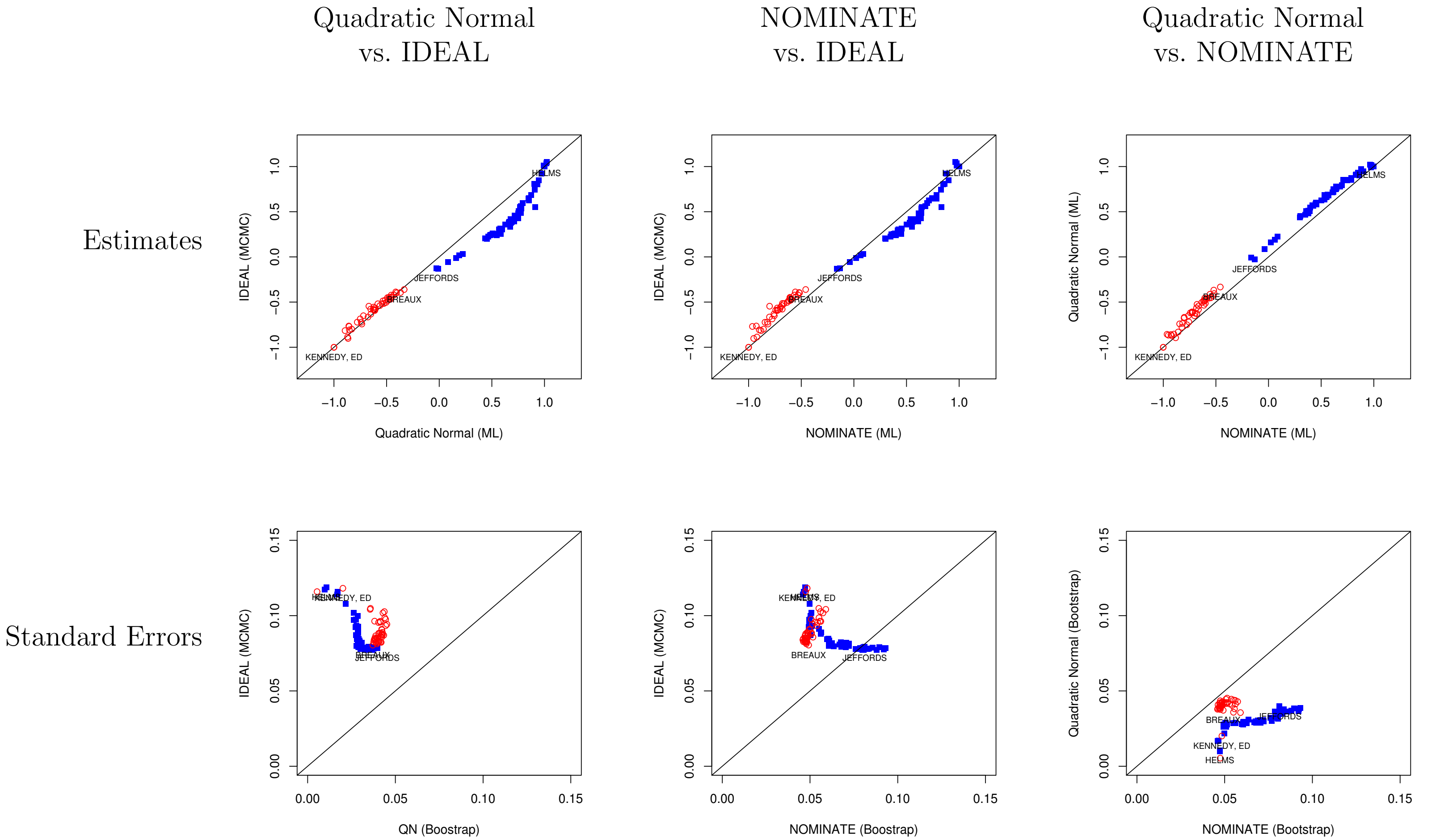


Figure 9: Each panels shows the estimated location or standard error of that location by two of three estimation techniques.

Correlation Among Ideal-Point Estimates and Among Their Standard Errors from Three Different Estimators, 105th Senate

	Quadratic Normal and IDEAL	NOMINATE and IDEAL	NOMINATE and Quadratic Normal
Estimate	0.988	0.991	0.998
Standard error	-0.443	-0.554	-0.010

Table 1: Shows the correlations among the ideal point estimates and standard errors of Senators' ideal points as estimated by W-NOMINATE, QN, and IDEAL. For IDEAL, the "standard errors" are posterior standard deviations.

yield very similar overall estimates of uncertainty. The posterior standard deviations of IDEAL and the bootstrapped standard errors generally fall between 0.03 and 0.12. However, there is less agreement as to which Senators are more or less reliably measured. Table 1 reports and the correlations among the point estimates and standard errors. The point estimates all correlate at over 0.99. Consistent with what is seen in the graph, the correlation among the standard error estimates is zero in case of NOMINATE versus QN and about -0.5 when IDEAL standard errors are compared to either NOMINATE or QN. There is also a tendency for the IDEAL estimate to exceed the QN and the NOMINATE standard errors; most of the points on the left two panels of the bottom row of Figure 9 fall above the 45 degree line.

A particularly difficult feature of the ideal point problem is that the information about the space that is theoretically recoverable are the relative positions of the Senators measured up to a constant of proportionality. For each of these models, however, the sampling/posterior distributions include not only variation in the relative distances, but also some difference in the choice of scale. This is easiest to see in the case of IDEAL where the average correlation among the posterior distributions of each Senator's location is 0.66. The reason for this this is that the model does not sufficiently nail down the location of the scale. Uncertainty that is associated with the estimate of each Senator's ideal point is arising not only from uncertainty in the relative distances between the members, but also from uncertainty in the location of the scale itself. However, the location of the scale is arbitrary and the uncertainty that arises

Standard Errors by Estimated Locations for Each of Three Methods

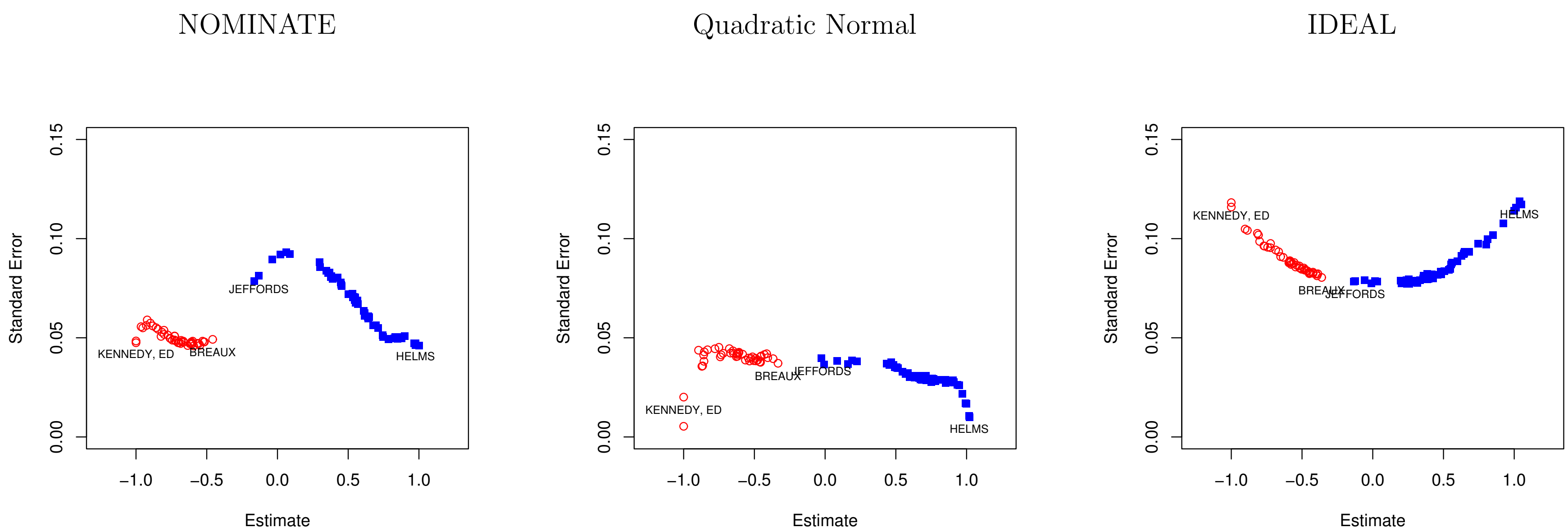


Figure 10: Plots the standard errors against the estimated legislator locations. The QN and IDEAL estimates have been rescaled to be comparable with NOMINATE as described in the text.

from it should not be included in the posterior variability of the estimates. By including this uncertainty, IDEAL overstates the degree of uncertainty in what we are trying to infer, the *relative* positions of the Senators. Note that including this scale uncertainty in the posterior would not effect confidence intervals for the relative distances between members. Indeed, one way to "purge" this scale uncertainty would be to present results in which each ideal point is measured as its distances from some reference legislator's ideal point.[9] This reference legislator would have no estimation uncertainty in her location! The fact, that we could choose any legislator to be the reference legislator makes it clear that the way that the underlying space is identified will have a dramatic effect on exactly where the uncertainty in estimates will be concentrated.

Figure 10 plots the estimated standard errors against the ideal point estimates for each of the methods. For NOMINATE and QN which bound the support the ideal point distribution and the location of the roll call cutpoints, those on the extremes have very small standard errors (and as we will see in Section 6 asymmetric sampling distributions).[10] On the other hand, IDEAL places no hard constraints on the support of the issue space. In this case, the location of the those on the ends is difficult. There are few bills whose cutpoints will be estimated to fall further from the center than their positions, and the location of those cutpoints will also be quite uncertain. For all three models those in the center of the distribution of the ideal point distribution are measured with greater precision. This is because there is a great deal of information in the data to discriminate among their positions (a large number of cutpoints that fall between them). In NOMINATE and QN, the uncertainty is greatest in the central cluster of each party's caucus because there are fewer cutpoints that fall among them and because they are not so extreme as to have their uncertainty bounded by the restriction that the ideal points and cut points must lie on the -1 to 1 interval. Also, the zero correlation between the NOMINATE and QN standard errors occurs despite the same general pattern shown in Figure 10, because Democrats have typically smaller standard

[9]Note that this discussion assumes that sufficient identifying restrictions are built into the model that only location invariance and not spread invariance is violated.

[10]These results differ from the two dimensional results because in two dimensions, the restriction that all Senators fall in the unit circle causes the less precisely measured second dimension to affect the first dimension locations of extremists. Note that it is extremists whose first dimension and second dimension positions are correlated as indicated by the confidence ellipses in figures 1, 3, 5, and 7.

Uncertainty in ideal point estimates, 105th Senate

	Average SE	Standard deviation of scores	"Signal-to-noise ratio"
NOMINATE	0.059	0.671	11.37
Quadratic-Normal	0.033	0.740	22.30
IDEAL	0.073	0.555	7.60
IDEAL (normalized)	0.043	0.555	12.90

Table 2: Average standard errors are the average of the bootstrapped standard error estimates across Senators for the NOMINATE and Quadratic Normal models and the average posterior standard deviations for the IDEAL and IDEAL (normalized) models. The second columns shows the standard deviation of the estimated scores across the 100 Senators. The "Signal-to-noise ratio" is the ratio of the first column to the second.

errors than Republicans in the NOMINATE model and typically higher standard errors than Republicans in the QN model. We conjecture that this difference results from the differing utility functions.

As shown in Table 2, the average standard errors are smallest for NOMINATE and larger for IDEAL and QN. Even though the range of the ideal point estimates have been normalize to -1, 1, there is still some variation in the standard deviations in the estimates across methods as seen in the second column of the table. The third column shows a rough "signal to noise" ratio for the estimates, by dividing the standard deviation in the point estimates (the signal) by the average standard error of the estimates (the noise). Even with these adjustments, however, the table does not allow a direct comparison of the uncertainty across methods. For example, IDEAL loads a considerable amount of uncertainty in the location of few extreme members, while QN and NOMINATE spread the uncertainty more evenly across members. However, making the direct comparison one might conclude from the fact that IDEAL's posterior standard deviations are larger than the bootstrapped standard errors of NOMINATE that perhaps the bootstrap is understating the true uncertainty in NOMINATE. However, as mentioned above, much of uncertainty revealed in IDEAL's posterior distributions results from uncertainty in the location of the scale. When this scale uncertainty is purged by normalizing each posterior draw to have mean 0, the average IDEAL standard error falls from 0.07 to 0.04 as seen in the "IDEAL (normalized)" row of table 2. In the next section we make more direct comparisons of uncertainty in the models by analyzing estimates of rank rather than interval scale position. By looking at ranks (an inherently relative measure of location), much of the apparent difference in the uncertainty associated with each member's location across estimators vanishes.

6 Beyond Standard Errors: Bootstrapping Auxiliary Quantities of Interest

One particularly useful aspect of the bootstrap is that it allows us to quickly and easily compute confidence intervals and other measures of uncertainty for many of the auxiliary quantities of interest that can be inferred from these models. As an example, we consider the location of the median voter in the chamber, the location of the filibuster pivot, and the identity of the median and filibuster pivot. We also present confidence intervals for the ordinal ranking of the members along the issue dimension.

The bootstrap provides an estimate of the complete sampling distribution of the model parameters. Figure 11 shows histograms of the (bootstrapped) sampling distribution of the ideal point estimates of five Senators. The figure shows how the constraints in NOMINATE and QN limit the variability of the estimates of extremists such as Kennedy and Ashcroft both of whom have asymmetric sampling distributions under QN and NOMINATE. On the other hand, the posterior distributions from IDEAL show the greatest uncertainty in the locations of the extremists. The comparison of the histograms for the raw and normalized IDEAL posteriors show dramatically the effect of scale uncertainty.

Figure 12 shows plots the estimated rank position of each Senator on the horizontal axis and the length of the 95 percent confidence interval for that estimated rank position on the vertical axis. The bootstrap confidence intervals for the ranks are easily computed. For each bootstrap sample, the ideal point estimates are ranked. The 0.025 and 0.975 quantiles of each Senator's rank position across the 1000 bootstrapped samples are taken as lower and upper bounds of a 95 percent confidence interval for each Senator's rank position. Because ranks are inherently scale-free and relative, we see much less variation across the three methods in Figure 12 than is shown in Figure 11. In terms of ranks, all three methods provide striking similar estimates of uncertainty. This finding demonstrates how variation in the estimation uncertainty across Senators' locations is largely a function of the constraints that must be

Sampling or Posterior Distributions for Three Ideal Point Models, 105th Senate

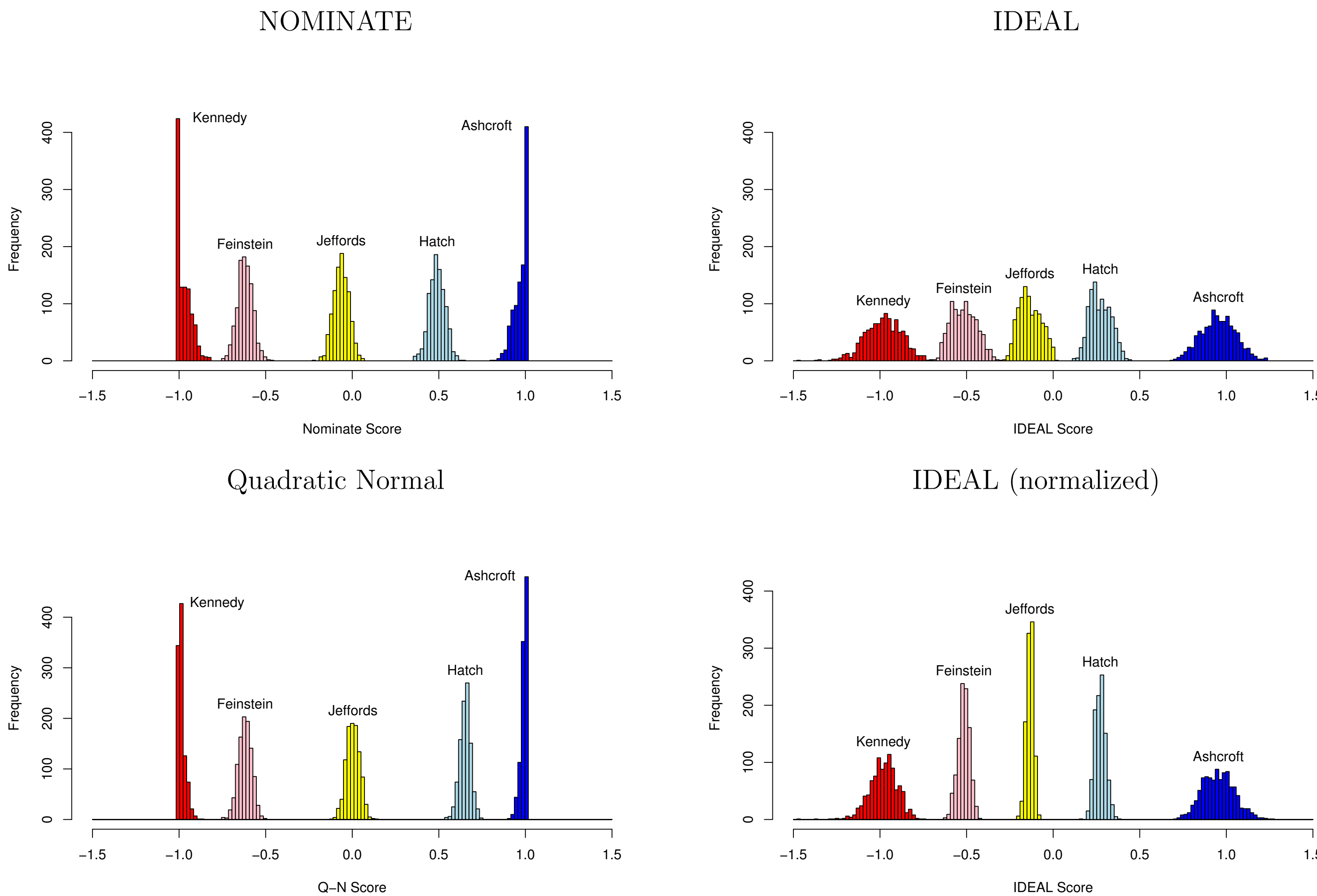


Figure 11: Shows the estimated sampling distributions of the ideal point estimates of five members of the 105th Senate based on three models/estimators. The normalized IDEAL model de-means each posterior draw across Senators as described in the text.

Size of Rank Confidence Interval versus Estimated Rank for Three Ideal Point Models, 105th Senate

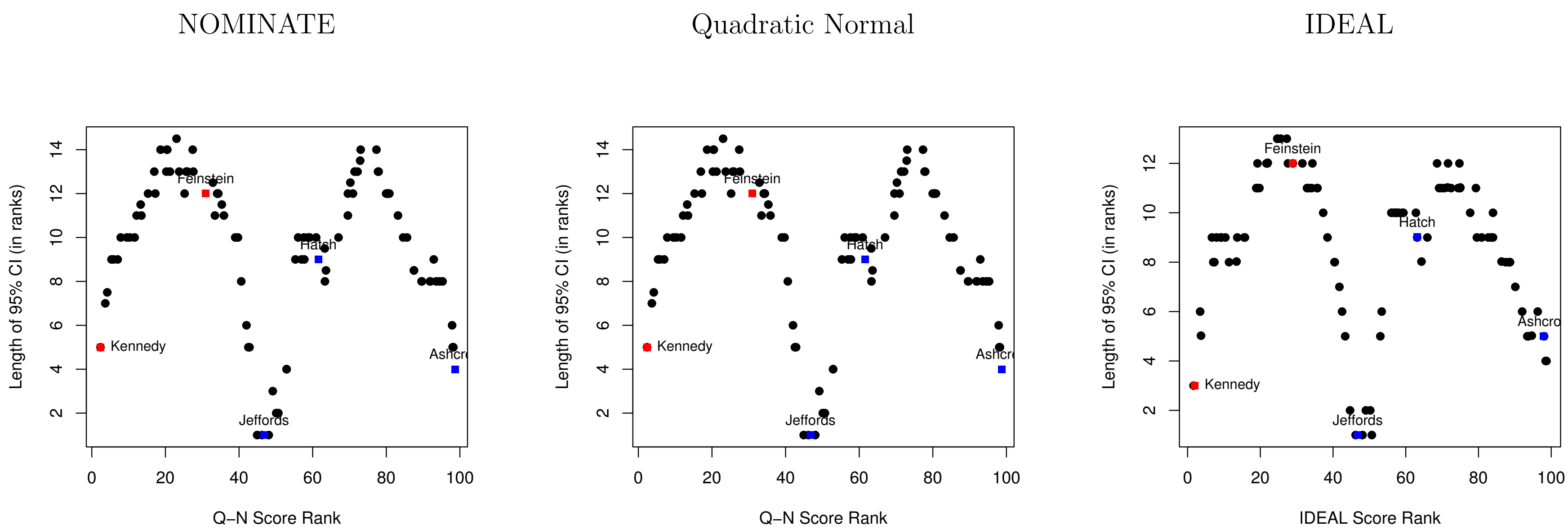


Figure 12: Plots the estimated rank position of each Senator in the 105th Senate against the length of the bootstrapped confidence interval of each Senator's rank position estimate. In all cases, the moderates and extremists have very small confidence intervals while confidence intervals are much larger for those between the center and the wings.

Location of the Median Senator, 105th Senate

Method	Estimate	S.E.	95 percent CI
NOMINATE	0.14	0.04	(0.06,0.22)
QN	0.26	0.03	(0.20,0.32)
IDEAL	-0.00	0.06	(-0.09,0.11)
IDEAL (normalized)	-0.00	0.02	(-0.03,0.03)

Table 3: Shows the estimated median Senator's location for each of three models.

imposed in order to identify the scale of the issue dimension. Once the choice of scale is removed (as when ranks are considered), the more fundamental variation is revealed. Those at the end of the scale can be said with great confidence to be at the ends, those in the middle are similarly pinned down. More difficult to disentangle are the those liberal and conservatives members located near the median of each party's caucus.

Table 3 shows the position of the estimated chamber median for the 105th Senate.[11] The bootstrapped standard errors are calculated by finding the position of the median Senator in each bootstrap estimate and then taking the standard deviation across those medians. Note that the standard error of this estimate is smaller than the average standard error for the individual members. This is true both because the median is among the members whose ideal point is estimated to be quite small, and because the identity of the median is not fixed across bootstrap estimates (or posterior draws in the case of IDEAL). Thus, in samples in which a potential median voter is estimated to have a relatively more extreme position than other nearby members, some other Senator will be estimated to be the median. The large difference between the normalized and unnormalized IDEAL estimates arises because the common scale uncertainty which we found in each member's location also resides in the estimate of the location of the median. When this scale uncertainty is purged, the uncertainty is the location of the median falls by two-thirds.

Similar results are found for the location of the filibuster pivot and ar shown in table 4. Interestingly, while there is greater uncertainty in the location of members close to the estimated filibuster pivot (the location of the 40th most liberal member), there are many

[11] Londregan and Snyder (1994) use a similar bootstrap procedure to test for committee outliers. However, their estimates of Senators locations are based on interest groups rating scores.

Location of the Filibuster Pivot, 105th Senate

Method	Estimate	S.E.	95 percent CI
NOMINATE	-0.52	0.04	(-0.59,-0.45)
QN	-0.47	0.03	(-0.54,-0.44)
IDEAL	-0.42	0.07	(-0.53,0.29)
IDEAL (normalized)	-0.41	0.02	(-0.46,-0.38)

Table 4: Shows the estimated location and standard error of the filibuster pivot in the 105th Senate.

Who was the Median Senator in the 105th Senate?

	NOMINATE	QN	IDEAL
Collins	0.595	0.728	0.609
D'Amato	0.243	0.201	0.301
Snowe	0.157	0.068	0.089
Chafee	0.005	0.003	0.001

Table 5: Table shows for NOMINATE and QN the bootstrapped sampling distribution over the identity of the median Senator in the 105th Senate. For the Bayesian IDEAL model, each value is the posterior probability that a given member is the median. $*$ = less than 0.001.

more members who are in close proximity to it. Thus, the standard error of the estimated pivot position is similar to that of the estimated median. Again we see a large reduction in the standard error of the IDEAL estimate when it is purged of scale uncertainty.

Tables 5 and 6 show the sampling distributions over the identities of the median and filibuster pivot Senator. For the Bayesian IDEAL estimator this can be directly interpreted as the posterior probability that a given Senator was in fact the median or the pivot. For the frequentist QN and NOMINATE models, we cannot correctly make the same interpretation. Nor, however, can these probabilities be construed as p-values for tests of the hypothesis that a given Senator is the median or filibuster pivot because these probabilities are conditional on the estimated model and not on the validity of the null hypothesis. However, they do give us a measure of confidence in the assertion that a particular member was indeed the filibuster pivot or median.

Who was the Filibuster Pivot in the 105th Senate?

	NOMINATE	QN	IDEAL
Bryan	*	*	0.016
Kohl	*	0.013	0.048
Bob Kerrey	0.011	0.004	0.002
Biden	0.012	0.004	0.008
Moynihan	0.022	0.027	0.031
Reid	0.027	0.014	0.015
Robb	0.031	0.064	0.079
Ford	0.033	0.061	0.092
Lieberman	0.056	0.093	0.121
Landrieu	0.068	0.018	0.005
Bob Graham	0.103	0.173	0.193
Hollings	0.161	0.198	0.226
Baucus	0.188	0.234	0.145
Byrd	0.213	0.075	0.019

Table 6: Table shows for NOMINATE and QN the bootstrapped sampling distribution over the identity of the filibuster pivot Senator in the 105th Senate. For the Bayesian IDEAL model, each value is the posterior probability that a given member is the filibuster pivot. * = less than 0.001.

7 Conclusion

Our results show that the parametric bootstrap is an effective method to obtain standard errors for the parameters of ideal point estimation methods. We show that in general the standard errors for NOMINATE and QN scores are relatively small. We also show how the bootstrap can be used to provide estimates of the many quantities of interest associated with ideal point estimates.

In future work we will describe how bootstrapped estimates can be used to correct for measurement error in regression models in which NOMINATE scores are used as dependent or independent variables.[12] We also plan to make bootstrapped NOMINATE estimates available to researchers, so that in future work scholars can incorporate the call of methodologists such as Herron and Schotts (2003) to explicitly account for the estimation uncertainty in the variables that are used in subsequent analysis.

[12]Brownstone and Valletta (1996) take as similar approach in a related area.

References

Bollen, Kenneth A. and Stein, Robert A., 1992. "Bootstrapping goodness -of-fit Measures in Structural Equation Models." *Sociological Methods and Research*, 21:205229, 1992.

Brownstone, David and Robert G. Valletta. 1996. "Modeling Earnings Measurement Error: A Multiple Imputation Approach." *Review of Economics and Statistics*, 78(4):705–717.

Clinton, Joshua, Simon D. Jackman, and Douglas Rivers. 2002. "The Statistical Analysis of Roll Call Data." Working Paper, Stanford University.

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

Efron, Bradley. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics*, 7:1-26.

Efron, Bradley, 1982. "Maximum Likelihood and Decision Theory." *Annals of Statistics*, 10:340–356.

Efron, Bradley and Robert J. Tibshirnai. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Hall, Peter. 1985. "Resampling a Coverage Pattern." *Stochastic Processes and their Applications*, 20:231-246.

Hall, Peter. 1994. "Methodology and Theory for the Bootstrap," Chapter 39 in R. F. Engle and D. L. McFadden, eds. *Handbook of Econometrics*. Vol. IV. New York: Elsevier, 2342-2379.

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.

Herron, Michael C. and Kenneth W. Shotts. 2003. "Using Ecological Inference Point Estimates as Dependent Variables in Second-Stage Linear Regressions." *Political Analysis*, 11(1):44-64.

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

Jackman, Simon D. 2000b. "Estimation and Inference are 'Missing Data' Problems: Unifying Social Science Statistics via Bayesian Simulation." *Political Analysis*, 8:307-332.

Jackman, Simon D. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference and Model Checking." *Political Analysis*, 9:227-241.

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

Londregan, John and James M. Snyder. 1994. "Comparing Committee and Floor Preferences." *Legislative Studies Quarterly*, 19(2):233–266.

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

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

McFadden, Daniel. 1976. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement*, 5:363-390.

Mooney, Chrisopher Z. 1996. "Bootstrap Statistical Inference: Examples and Evaluations of Political Science." *American Journal of Political Science*, 40(2):570–602.

Poole, Keith T. 2000. "Non-Parametric Unfolding of Binary Choice Data." *Political Analysis*, 8(3):211-237.

Poole, Keith T. 2001. "The Geometry of Multidimensional Quadratic Utility in Models of Parliamentary Roll Call Voting." *Political Analysis*, 9(3):211-226.

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. 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- 26.

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

Young, G. Alastair. 1994. "Bootstrap: More than a Stab in the Dark?" *Statistical Science*, 9(3):382–415.