Nonparametric Ideal-Point Estimation and Inference

Alexander Tahk

Assistant Professor, Department of Political Science, University of Wisconsin–Madison, North Hall Rm 110, 1050 Bascom Mall, Madison, WI 53706, USA. Email: atahk@wisc.edu

Abstract

Existing approaches to estimating ideal points offer no method for consistent estimation or inference without relying on strong parametric assumptions. In this paper, I introduce a nonparametric approach to ideal-point estimation and inference that goes beyond these limitations. I show that some inferences about the relative positions of two pairs of legislators can be made with minimal assumptions. This information can be combined across different possible choices of the pairs to provide estimates and perform hypothesis tests for all legislators without additional assumptions. I demonstrate the usefulness of these methods in two applications to Supreme Court data, one testing for ideological movement by a single justice and the other testing for multidimensional voting behavior in different decades.

Keywords: ideal-point estimation, hypothesis testing, nonparametric analysis, nonparametric estimation

1 Introduction

Political scientists are often interested in making inferences about the ideology of legislators based on ideal-point models. For example, scholars have sought to make inferences about party influence (Snyder and Groseclose 2000; McCarty, Poole, and Rosenthal 2001), the dimensionality of voting (e.g., Rosenthal 1992; Roberts, Smith, and Haptonstahl 2016), pivot and cartel theories of lawmaking (e.g., Chiou and Rothenberg 2003; Covington and Bargen 2004), the identity of the most liberal or conservative legislators (Clinton, Jackman, and Rivers 2004a), and ideological change over time (e.g., Epstein *et al.* 2007; Poole 2007).

The bulk of these studies have used parametric ideal-point models that specify particular functional forms for the utility functions and error distributions. It is well known that the choice of parametric assumptions can affect the estimates and inferences derived from data. This has been demonstrated in the ideal-point context (e.g., Ho and Quinn 2010; Krehbiel and Peskowitz 2015) and has led studies to different conclusions on questions such as the dimensionality of congressional voting (see Poole and Rosenthal 1991; Heckman and Snyder 1997). Yet scholars rarely have theoretical or empirical grounds for choosing a particular functional form.

The only alternative approach has been to use Optimal Classification (OC), a model-free estimation method introduced by Poole (2000). Because OC "is not a statistical model" (Poole 2000, p. 211), statistical inference is not possible.[1] OC is also not a consistent estimator of the rank order of the ideal points under quadratic utility.[2] Thus, for many purposes, OC cannot be used in place of a parametric model.

Author's note: I thank R. Michael Alvarez, Bret Hanlon, Simon Jackman, Jeff Lewis, Nolan McCarty, and Keith Poole. All mistakes are my own. Open-source software for the method proposed in this article is available at github.com/atahk/npideal as a package for the statistical software R. Replication materials for all of the results in this article are provided in the online Dataverse archive associated with this article (Tahk 2017).

1 Some papers have relied upon OC with *ad hoc* methods, such as arbitrary thresholds to the marginal proportional reduction in error (e.g., Roberts, Smith, and Haptonstahl 2016), but the lack of formal hypothesis tests and arbitrary nature of the thresholds make it difficult to draw clear conclusions.

2 See Appendix B in the supplementary material available at https://doi.org/10.1017/pan.2017.38.

Political Analysis (2018)
vol. 26:131–146
DOI: 10.1017/pan.2017.38

Published
8 March 2018

Corresponding author
Alexander Tahk

Edited by
R. Michael Alvarez

© The Author(s) 2018. Published by Cambridge University Press on behalf of the Society for Political Methodology.

The main contribution of this paper is to introduce a method for ideal-point inference and estimation under much weaker assumptions than previous approaches. It requires no parametric assumptions, making it the first approach to nonparametric ideal-point inference.[3] Its estimates of the rank order of ideal points are also consistent under the same weak, nonparametric assumptions.

This method allows for formal hypothesis tests of whether legislators have the same ordering on different groups of bills without the validity of the tests depending on knowledge of the parametric form, which is almost always unavailable. Because there are no assumptions made about the distribution of bill characteristics in the two groups, these inferences remain valid even if there are also differences in agenda control or other selection effects affecting the bills being voted on in each group. For example, this method can be applied to whether a legislator's voting behavior changed after a hypothesized point in time, such as a party switch or change in public opinion. Similarly, it can be applied to testing hypotheses regarding dimensionality by testing whether a common ideological dimension can capture voting on different groups of bills.

The model consists of two key assumptions: unidimensionality and concave utility.[4] Aside from unidimensionality, the model's assumptions are far weaker than in most ideal-point models. Not only is the model not restricted to quadratic utility or errors drawn from particular parametric families, but the utility functions need not be symmetric and the error distributions can vary across votes provided the errors follow the same distribution across legislators for each vote. The utility functions can be any concave function as long as the function is shared by all legislators, aside from differences in their ideal points.[5]

The rest of the paper is organized as follows. The next section introduces the nonparametric ideal-point model used throughout this paper. Sections 3 and 4 then present methods for estimation and inference, respectively, under this model. Section 5 explores the properties of these methods through Monte Carlo simulations. Section 6 presents two applications of the model to Supreme Court data. The last section gives a concluding discussion of the methods and their possible extensions.

2 Model

Consider a legislature of K legislators voting on N bills. Assume legislators have unidimensional spatial preferences and vote sincerely according to the stochastic utility function given by

$$u_i\left(v_{ij} = 1\right) = f\left(x_i - y_{j1}\right) + \zeta_{ij1} \tag{1}$$

$$u_i\left(v_{ij} = 0\right) = -f\left(x_i - y_{j0}\right) + \zeta_{ij0} \tag{2}$$

$$\zeta_{ij1}, \zeta_{ij0} \overset{iid}{\sim} G_j \tag{3}$$

3 The term "nonparametric" was coined by Wolfowitz, who used to refer to the situation "where the functional forms of the distributions are unknown" (Wolfowitz 1942, p. 264). While some scholars have adopted a broader use of the term that includes model-free approaches (e.g., Randles, Hettmansperger, and Casella 2004), *nonparametric inference* cannot be quite so broadly defined as statistical inference inherently requires a probabilistic model. Most definitions of nonparametric inference are similar to that given by Gibbons and Chakraborti (2011): "Nonparametric statistical inference is a collective term given to inferences that are valid under less restrictive assumptions than with classical (parametric) statistical inference."

4 The other necessary assumptions are that the stochastic components of this utility function must be independent across legislators and bills and that the error distributions and utility functions do not vary between legislators, aside from differences in their ideal points, on a single bill. These assumptions are also made by all common ideal-point models.

5 This does not mean that all other ideal-point models are a special case of this model. The NOMINATE model (e.g., Poole and Rosenthal 1991, 1997) is not a special case because it assumes a Gaussian utility, which is not globally concave. Gaussian utility poses a problem for nonparametric approaches because it implies that two legislators can have different ideal points but the same probability of voting "yea" on a bill. Most other models assume quadratic utility, (e.g., Heckman and Snyder 1997; Poole 2001; Clinton, Jackman, and Rivers 2004b), which is globally concave, and are special cases of this model.

where v_{ij} is the vote of legislator i on bill j, $v_{ij} = 1$ indicates a vote of "yea," $v_{ij} = 0$ indicates a vote of "nay," y_{j1} is the policy outcome of the bill, y_{j0} is the status quo, x_i is the ideal point of legislator, ζ_{ij1} and ζ_{ij0} are stochastic errors, and G_j is the joint distribution of ζ_{ij1} and ζ_{ij0} for bill j. Assume that f is a strictly concave function with a maximum at zero.[6] Further, assume that the errors, ζ_{ij1} and ζ_{ij0}, are independent across bills and legislators and identically distributed for each bill. Finally, assume legislators vote "yea" when indifferent.[7] Note that they only differ in their ideal points, x_i, and not in f, which is assumed to be the same for all legislators. Note also that the distribution G_j of the stochastic utility components, ζ_{ij1} and ζ_{ij0}, is assumed to be the same for all legislators for any given bill but can vary across bills.

The probability of legislator i voting "yea" on bill j is given by

$$\begin{aligned}\Pr\left(v_{ij} = 1\right) &= \Pr\left(u_i\left(v_{ij} = 1\right) \geq u_i\left(v_{ij} = 0\right)\right) \\ &= \Pr\left(f\left(x_i - y_{j1}\right) - f\left(x_i - y_{j0}\right) \geq \zeta_{ij0} - \zeta_{ij1}\right). \end{aligned} \tag{4}$$

Let H_j be the cumulative distribution function of $\zeta_{ij0} - \zeta_{ij1}$ $\left(H_j(z) = \Pr\left(\zeta_{ij0} - \zeta_{ij1} \leq z\right)\right)$. Thus,

$$\Pr\left(v_{ij} = 1\right) = H_j\left(f\left(x_i - y_{j1}\right) - f\left(x_i - y_{j0}\right)\right). \tag{5}$$

The results in this paper rest on the implication of this model that $\Pr\left(v_{ij} = 1\right)$ is monotonic in x_i. It is also worth noting that, if $y_{j0} \neq y_{j1}$ and H is strictly increasing, then $\Pr\left(v_{ij} = 1\right)$ is strictly monotonic in x_i. These two conditions require only that the policy outcome of the bill differs from the status quo and that the error distribution has support over the entire real line.

The following theorem establishes these results:

Theorem 1. *If f is strictly concave, $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is monotonic in z. Further, if $y_{j0} \neq y_{j1}$ and H is strictly increasing, then $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is strictly monotonic in z.*

Proof. Consider the three cases $y_{j0} = y_{j1}$, $y_{j0} < y_{j1}$, and $y_{j0} > y_{j1}$.

Case 1: $y_{j0} = y_{j1}$. In this case,

$$H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right) = H_j\left(f\left(z - y_{j0}\right) - f\left(z - y_{j0}\right)\right) = 0. \tag{6}$$

Since this constant in z, it is also monotonic in z.

Case 2: $y_{j0} < y_{j1}$. Take any $x_i < x_k$. Since f is strictly concave, $tf\left(z_1\right) + (1-t)f\left(z_2\right) < f\left(tz_1 + (1-t)z_2\right)$ for all $z_1 \neq z_2$ and $t \in (0,1)$. Let $t' = \frac{y_{j1} - y_{j0}}{x_k - x_i + y_{j1} - y_{j0}}$. Since $y_{j1} - y_{j0} > 0$ and $x_k - x_i > 0$, $t' \in (0,1)$. Thus,

$$\begin{aligned}&t'f\left(x_i - y_{j1}\right) + (1-t')f\left(x_k - y_{j0}\right) \\ &\quad < f\left(t'\left(x_i - y_{j1}\right) + (1-t')\left(x_k - y_{j0}\right)\right) \\ &\quad = f\left(\left(\frac{y_{j1} - y_{j0}}{x_k - x_i + y_{j1} - y_{j0}}\right)\left(x_i - y_{j1}\right) + \left(1 - \frac{y_{j1} - y_{j0}}{x_k - x_i + y_{j1} - y_{j0}}\right)\left(x_k - y_{j0}\right)\right) \\ &\quad = f\left(\left(\frac{y_{j1} - y_{j0}}{x_k - x_i + y_{j1} - y_{j0}}\right)\left(x_i - y_{j1}\right) + \left(\frac{x_k - x_i}{x_k - x_i + y_{j1} - y_{j0}}\right)\left(x_k - y_{j0}\right)\right) \\ &\quad = f\left(\frac{\left(y_{j1} - y_{j0}\right)\left(x_i - y_{j1}\right) + \left(x_k - x_i\right)\left(x_k - y_{j0}\right)}{x_k - x_i + y_{j1} - y_{j0}}\right)\end{aligned}$$

6 The assumption that f has a maximum at zero is not necessary for the results in this paper but is included to give the ideal points their usual meaning as a legislator's most preferred policy.

7 Whether legislators vote "yea," vote "nay," or vote randomly (with identical and probability for all legislators and independent of all other components of the model) makes no meaningful difference for the results.

$$
\begin{aligned}
&= f\left(\frac{\left(y_{j1}x_i - y_{j0}x_i - y_{j1}y_{j1} + y_{j0}y_{j1}\right) + \left(x_k x_k - x_i x_k - x_k y_{j0} + x_i y_{j0}\right)}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(\frac{y_{j1}x_i - y_{j1}y_{j1} + y_{j0}y_{j1} + x_k x_k - x_i x_k - x_k y_{j0}}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(\frac{y_{j1}x_i - y_{j1}y_{j1} + y_{j0}y_{j1} + x_k x_k - x_i x_k - x_k y_{j0} + x_k y_{j1} - x_k y_{j1}}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(\frac{y_{j1}\left(-x_k + x_i - y_{j1} + y_{j0}\right) + x_k\left(x_k - x_i + y_{j1} - y_{j0}\right)}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(\frac{x_k\left(x_k - x_i + y_{j1} - y_{j0}\right) - y_{j1}\left(x_k - x_i + y_{j1} - y_{j0}\right)}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(\frac{\left(x_k - y_{j1}\right)\left(x_k - x_i + y_{j1} - y_{j0}\right)}{x_k - x_i + y_{j1} - y_{j0}}\right) \\
&= f\left(x_k - y_{j1}\right).
\end{aligned} \tag{7}
$$

Similarly,

$$
\begin{aligned}
&(1-t')\, f\left(x_i - y_{j1}\right) + t' f\left(x_k - y_{j0}\right) \\
&\quad < f\left((1-t')\left(x_i - y_{j1}\right) + t'\left(x_k - y_{j0}\right)\right) \\
&\quad = f\left(\left(1 - \frac{y_{j0} - y_{j1}}{x_i - x_k + y_{j0} - y_{j1}}\right)\left(x_i - y_{j1}\right) + \left(\frac{y_{j0} - y_{j1}}{x_i - x_k + y_{j0} - y_{j1}}\right)\left(x_k - y_{j0}\right)\right) \\
&\quad = f\left(\left(\frac{x_i - x_k}{x_i - x_k + y_{j0} - y_{j1}}\right)\left(x_i - y_{j1}\right) + \left(\frac{y_{j0} - y_{j1}}{x_i - x_k + y_{j0} - y_{j1}}\right)\left(x_k - y_{j0}\right)\right) \\
&\quad = f\left(\left(x_i - x_k\right)\left(x_i - y_{j1}\right) + \frac{\left(y_{j0} - y_{j1}\right)\left(x_k - y_{j0}\right)}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\left(x_i x_i - x_k x_i - x_i y_{j1} + x_k y_{j1}\right) + \frac{\left(y_{j0}x_k - y_{j1}x_k - y_{j0}y_{j0} + y_{j1}y_{j0}\right)}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\frac{y_{j0}x_k - y_{j0}y_{j0} + y_{j1}y_{j0} + x_i x_i - x_k x_i - x_i y_{j1}}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\frac{y_{j0}x_k - y_{j0}y_{j0} + y_{j1}y_{j0} + x_i x_i - x_k x_i - x_i y_{j1} + x_i y_{j0} - x_i y_{j0}}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\frac{y_{j0}\left(-x_i + x_k - y_{j0} + y_{j1}\right) + x_i\left(x_i - x_k + y_{j0} - y_{j1}\right)}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\frac{x_i\left(x_i - x_k + y_{j0} - y_{j1}\right) - y_{j0}\left(x_i - x_k + y_{j0} - y_{j1}\right)}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(\frac{\left(x_i - y_{j0}\right)\left(x_i - x_k + y_{j0} - y_{j1}\right)}{x_i - x_k + y_{j0} - y_{j1}}\right) \\
&\quad = f\left(x_i - y_{j0}\right).
\end{aligned} \tag{8}
$$

Adding these two results together yields

$$
f\left(x_i - y_{j1}\right) + f\left(x_k - y_{j0}\right) < f\left(x_k - y_{j1}\right) + f\left(x_i - y_{j0}\right). \tag{9}
$$

So,

$$
f\left(x_i - y_{j1}\right) - f\left(x_i - y_{j0}\right) < f\left(x_k - y_{j1}\right) - f\left(x_k - y_{j0}\right). \tag{10}
$$

Since H is nondecreasing,

$$H_j\left(f\left(x_i - y_{j1}\right) - f\left(x_i - y_{j0}\right)\right) \leq H_j\left(f\left(x_k - y_{j1}\right) - f\left(x_k - y_{j0}\right)\right) \quad (11)$$

with strict inequality holding if H is strictly increasing. So, by definition of monotonicity, $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is monotonic in z with strict monotonicity if H is strictly increasing.

Case 3: $y_{j0} < y_{j1}$. Take any $x_i < x_k$. The same argument as in Case 2 with y_{j1} and y_{j0} swapped establishes that

$$H_j\left(f\left(x_i - y_{j0}\right) - f\left(x_i - y_{j1}\right)\right) \leq H_j\left(f\left(x_k - y_{j0}\right) - f\left(x_k - y_{j1}\right)\right). \quad (12)$$

Thus, negating both sides of the inequality yields

$$H_j\left(f\left(x_i - y_{j1}\right) - f\left(x_i - y_{j0}\right)\right) \geq H_j\left(f\left(x_k - y_{j1}\right) - f\left(x_k - y_{j0}\right)\right) \quad (13)$$

with strict inequality holding if H is strictly increasing. So, by definition of monotonicity, $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is monotonic in z with strict monotonicity if H is strictly increasing.

Therefore, in all cases, $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is monotonic in z. If $y_{j0} \neq y_{j1}$ and H is strictly increasing, then $H_j\left(f\left(z - y_{j1}\right) - f\left(z - y_{j0}\right)\right)$ is strictly monotonic in z. □

This model bares some similarity to a nonparametric item-response theory model developed by Mokken (1971): the monotone homogeneity model. The crucial difference is that the above model assumes that the probability of a certain vote choice is merely a monotonic function of a legislator's ideal point and can, thus, be either nondecreasing or nonincreasing. In contrast, the monotone homogeneity model assumes *a priori* knowledge of which answer is correct and, therefore, nondecreasing in a student's ability. Unfortunately, such a simple approach is not applicable to the more general problem of nonparametric ideal-point estimation, where researchers cannot be presumed to know whether a "yea" vote is the more "conservative" or "liberal" vote choice.[8] Mokken's (1971) approach also requires an absence of missing values, further limiting its relevance to analyzing roll-call data.

3 Estimation

Under the model described in the previous section, knowing which of two legislators is to the right of the other does not tell us which is more likely to vote "yea" on a given bill. This is in sharp contrast to the monotone heterogeneity model, where knowing which student has higher ability directly implies which is more likely to answer a question correctly.

Most approaches to ideal-point estimation have addressed the need to know the polarity of each bill and other bill-specific parameters by simultaneously estimating these parameters.[9] Introducing these parameters can complicate both estimation and inference. For example,

8 The monotone homogeneity model leads to essentially the same method used by Americans for Democratic Action to compute Liberal Quotients, commonly called ADA scores. ADA scores give the percentage of 20 key votes on which a legislator voted with the ADA's position. This depends upon the ADA to identify the "liberal" position on each of the twenty votes and the decision to treat missed votes as votes against the ADA's position.

9 Heckman and Snyder (1997) is the most notable exception.

maximum likelihood estimators may be inconsistent when such parameters are included (Neyman and Scott 1948). Further, even aside from these complications, estimating the relationship between ideal points and vote probabilities is impractical here because the parameter space is infinite.

In this paper, I adopt a different approach that avoids estimating bill-specific parameters. The key insight for both estimation and inference is that, while we cannot learn much from the votes of a pair of legislators alone, the votes of two pairs can be informative. For any two pairs of legislators, if we condition on each pair containing one "yea" vote and one "nay" vote, then the conditional probability that the more conservative of the legislators in each pair voted together and the more liberal legislators in each pair voted together is greater than the probability that the more liberal of the legislators of each pair voted with the more conservative of the other.

More formally, if $x_i < x_k$ and $x_\ell < x_m$, then either $\Pr(v_{kj} = 1) - \Pr(v_{ij} = 1) \geq 0$ and $\Pr(v_{mj} = 1) - \Pr(v_{\ell j} = 1) \geq 0$ or $\Pr(v_{kj} = 1) - \Pr(v_{ij} = 1) \leq 0$ and $\Pr(v_{mj} = 1) - \Pr(v_{\ell j} = 1) \leq 0$. For notational convenience, let $p_i = \Pr(v_{ij} = 1)$ and $q_i = \Pr(v_{ij} = 0) = 1 - \Pr(v_{ij} = 1)$. In either case,

$$
\begin{aligned}
&0 \leq (p_k - p_i)(p_m - p_\ell) && (14)\\
&0 \leq (p_k - p_i p_k - p_i + p_i p_k)(p_m - p_\ell p_m - p_\ell + p_\ell p_m) && (15)\\
&0 \leq ((1 - p_i) p_k - p_i (1 - p_k))((1 - p_\ell) p_m - p_\ell (1 - p_m)) && (16)\\
&0 \leq (1 - p_i) p_k ((1 - p_\ell) p_m - p_\ell (1 - p_m)) - p_i (1 - p_k)((1 - p_\ell) p_m - p_\ell (1 - p_m)) && (17)\\
&0 \leq q_i p_k (q_\ell p_m - p_\ell q_m) - p_i q_k (q_\ell p_m - p_\ell q_m). && (18)
\end{aligned}
$$

So,

$$
\begin{aligned}
p_i q_k (q_\ell p_m - p_\ell q_m) &\leq q_i p_k (q_\ell p_m - p_\ell q_m) && (19)\\
p_i q_k q_\ell p_m - p_i q_k p_\ell q_m &\leq q_i p_k q_\ell p_m - q_i p_k p_\ell q_m && (20)\\
p_i q_k q_\ell p_m + q_i p_k p_\ell q_m &\leq p_i q_k p_\ell q_\ell + q_i p_k q_\ell p_m. && (21)
\end{aligned}
$$

By independence of v_{ij}, v_{kj}, $v_{\ell j}$, and v_{mj}, this implies

$$
\begin{aligned}
\Pr(v_{ij} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}) &\leq \Pr(v_{kj} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}) && (22)\\
\frac{\Pr(v_{ij} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj})}{\Pr(v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj})} &\leq \frac{\Pr(v_{kj} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj})}{\Pr(v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj})} && (23)\\
\Pr(v_{ij} = v_{mj} \mid v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}) &\leq \Pr(v_{kj} = v_{mj} \mid v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}). && (24)
\end{aligned}
$$

Thus, if we consider only the votes where $v_{ij} \neq v_{kj}$ and $v_{\ell j} \neq v_{mj}$, the expected count of votes on which $v_{kj} = v_{mj}$ is greater than or equal to the expected count of votes on which $v_{kj} = v_{mj}$. Let

$$
v^j_{im|k\ell} = \begin{cases} 1 & \text{if } v_{kj} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj} \\ 0 & \text{otherwise} \end{cases} \tag{25}
$$

so that $v^j_{im|k\ell}$ indicates that legislators i and m voted together and legislators k and ℓ voted together but each pair voted differently from the other on bill j. Under weak regularity conditions,[10] the law of large numbers implies that

$$\lim_{N\to\infty} \Pr\left(\sum_{j=1}^{N} v^j_{km|i\ell} > \sum_{j=1}^{N} v^j_{im|k\ell}\right) = 1. \tag{26}$$

Thus, if we consider only bills for which legislators i and k voted differently and legislators ℓ and m voted differently, the fraction of these in which legislators k and m voted together should be greater than $\frac{1}{2}$.

Let $\prec$ be the binary relation giving the true ordering of the ideal points where $i \prec k$ indicates that legislator i is to the left of legislator k (i.e., $x_i < x_k$). Let $\hat{\prec}$ be the binary relation giving the estimated ordering relation where $i \hat{\prec} k$ indicates that legislator i is estimated to be to the left of legislator k. If $\hat{\prec}$ is a consistent estimator of $\prec$, then

$$\sum_{j=1}^{N} v^j_{km|i\ell} > \sum_{j=1}^{N} v^j_{im|k\ell} \tag{27}$$

must imply $i \hat{\prec} k \iff \ell \hat{\prec} m$ for sufficiently large N and all $i \neq k \neq \ell \neq m$. Thus, a consistent estimator will minimize $\sum_{j=1}^{N} v^j_{km|i\ell} - v^j_{im|k\ell}$ for sufficiently large N and all $i \neq k \neq \ell \neq m$ such that $i \hat{\prec} k \wedge \ell \hat{\prec} m$. Therefore, it also minimizes the sum of this quantity over all possible groups of legislators satisfying $i \hat{\prec} k \wedge \ell \hat{\prec} m$,

$$\sum_{i\neq k\neq \ell\neq m} \sum_{j=1}^{N} \left(v^j_{km|i\ell} - v^j_{im|k\ell}\right) \mathbb{1}\left(i \hat{\prec} k \wedge \ell \hat{\prec} m\right) \tag{28}$$

where $\mathbb{1}(x)$ is the indicator function ($\mathbb{1}(x) = 1$ if x is true and 0 otherwise) for sufficiently large N.

This suggests choosing as an estimator the relation $\hat{\prec}$ that minimizes Formula (28). Identifiability also requires constraining the order of two legislators. Given this constraint and the necessary regularity condition (see footnote 25), this estimator is unique, at least for sufficiently large N, and therefore consistent.

Computing this estimator is difficult. There is no apparent alternative to an exhaustive search, but this requires considering $\frac{1}{2}K!$ possible orderings for K legislators and summing over $2\binom{K}{4}$ choices of legislators satisfying $i \hat{\prec} k \wedge \ell \hat{\prec} m$. Thus, computing Formula (28) for $\frac{1}{2}K!$ possible orderings requires computing the inner sum of Formula (28) $K!\binom{K}{4}$ times.[11] This is feasible for a sufficiently small number of legislators, but cases where $n > 10$ seem impractical.[12]

Thus, for larger numbers of legislators, a different approach is necessary. However, if we choose two legislators, ℓ and m, and consider only the votes in which they disagree, then the remaining $K - 2$ can be ordered based on the fraction of the votes in which each voted with m instead of ℓ. As was the case when we considered two pairs of legislators, the expectation of the fraction of

10 It is sufficient here to assume that

$$0 < \liminf_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \left(\Pr\left(v_{kj} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}\right) - \Pr\left(v_{ij} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}\right)\right).$$

11 Computing the inner sum of Formula (28) for a given set of four legislators requires N computations, but can be computed and stored for each of the $\frac{K!}{(K-4)!}$ possible choices of four legislators rather than recomputed for each possible orderings of the legislators. Thus, since $\frac{K!}{(K-4)!} \sim K^4$, it requires $O\left(NK^4\right)$ computations. Considering $\frac{1}{2}K!$ possible orderings and summing $2\binom{K}{4}$ quantities for each ordering is $O\left(K!\binom{K}{4}\right)$ because $n! \sim \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}$ by Sterling's formula and $\binom{n}{4} \sim n^4$. The overall number of computations is the sum of each of these, and is, thus, $O(\max(K^{K+\frac{9}{2}}e^{-K}, NK^4))$.

12 Assuming each of these can be computed in 10^{-5} seconds, this would require roughly one and a half days with eleven legislators and one and a half years with thirteen legislators.

the votes in which each voted with m is increasing in their ideal points if $\ell \prec m$ and decreasing if $m \prec \ell$.

If we repeat this process for each possible pair of legislators, we then only need to combine these orderings of each of the $\binom{K}{2}$ sets of legislators into a single ordering for all legislators. Doing so can be done by creating a $\binom{K}{2} \times K$ matrix with each row giving a numeric ranking of a different set of $K - 2$ with the two held-out legislators given the mean rank, then mean-centering each row, and finally performing a singular value decomposition (SVD). If all rankings on each subset of $\binom{K}{2}$ legislators are correct, as will be true asymptotically, the order of the values in the first right-singular vector (i.e., the one corresponding to the largest singular value) should agree with the true ordering of the ideal points up to a reversal. Further, if it is too burdensome to compute rankings for all $\binom{K}{2}$ groups of legislators, this can be replaced by a smaller number of groups of $K - 2$ chosen at random.

Although this estimator is also consistent under the same regularity conditions, it can only be based on roll-call items with no missing votes. In contrast, the alignment of two pairs of legislators can be computed using only the votes of those four legislators. Thus, the original estimator proposed in this paper is able to use any roll-call items in which at least four legislators voted. Since roll-call data often contain few or no items without any missing votes, the SVD-based approach as outlined above is not feasible but can be modified by computing the fraction of times one legislator agrees with another with missing values imputed by the mean agreement on nonmissing votes. This estimator need not be consistent but is likely to return similar results to the other, consistent estimators provided the fraction of missing votes is not too large.

4 Inference

Suppose we have two groups of roll-call items and wish to test whether a single ideological dimension could explain voting on both groups—or, equivalently, we wish to test whether the legislators' ideal points maintained the same ordering on both items. The same approach used for estimation of choosing two pairs of legislators and focusing on those votes where each pair disagrees also provides a natural approach for performing inference. Indeed, the pairs-of-pairs approach is, in some ways, a more natural fit to inference than estimation and will provide a means to test the equal-ordering hypothesis without needing to estimate an ordering or to impose identifiability constraints.

Consider a two-sample extension of the model in Section 2 in which K legislators vote on two separate groups of bills, the first consisting of N_A items and the second of N_B. Let A be the set of indices of the first group, such that bill j is in group A if and only if $j \in A$ and let B be the set of indices for group B, where A and B are disjoint. Assume the order of the legislators' ideal points is given by the binary relation $\prec_A$ for the first group of N_A bills and $\prec_B$ for the second.[13]

Finally, although most of the results in this section can be derived without this assumption, it will simplify matters to assume that the bill-specific parameters (y_{j1}, y_{j0}, and G_j) are jointly independently and identically distributed within each group, with distinct distributions for the bill parameters for each group and with no restrictions placed on the forms of these distributions. Thus, the bill parameters for one group may be drawn from a quite different distribution than the bill parameters in the other.

I aim to construct a test of the null hypothesis that the legislators have the same ordering on both groups—that is, that $i \prec_A k \Leftrightarrow i \prec_B k$ for all i and k.

Let

$$\pi^{(A)}_{km|i\ell} = E_A \left[\Pr \left(v_{ij} = v_{mj} \mid v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj} \right) \right] \tag{29}$$

13 As in the previous section, $\prec_A$ and $\prec_B$ can be defined in terms of the legislators' ideal points $x_i^{(A)}$ and $x_i^{(B)}$ on the respective sets by $i \prec_A k \Leftrightarrow x_i^{(A)} < x_k^{(A)}$ and $i \prec_B k \Leftrightarrow x_i^{(B)} < x_k^{(B)}$.

indicate the marginal probability that $\Pr\left(v_{kj} = v_{mj} \mid v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}\right)$ marginalized over the distribution of the bill parameters for group A, with $\pi^{(B)}_{km|i\ell}$ defined analogously for group B. Thus, given the number of votes in which $v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}$, the conditional distribution for the number of votes in which v_{kj} also equals v_{mj} (i.e., the sum of $v^j_{km|i\ell}$ over A) is the sum of $\sum_{j\in A} v^j_{km|i\ell} + v^j_{im|k\ell}$ trials each with success probability $\pi^{(A)}_{km|i\ell}$, where, as in the previous section, $v^j_{km|i\ell} = \mathbb{1}\left(v_{kj} = v_{mj} \wedge v_{ij} \neq v_{kj} \wedge v_{\ell j} \neq v_{mj}\right)$. Thus, the distribution is given by

$$\sum_{j\in A} v^j_{km|i\ell} \sim \text{Binomial}\left(\sum_{j\in A} v^j_{km|i\ell} + v^j_{im|k\ell}, \pi^{(A)}_{km|i\ell}\right). \tag{30}$$

Similarly,

$$\sum_{j\in B} v^j_{km|i\ell} \sim \text{Binomial}\left(\sum_{j\in B} v^j_{km|i\ell} + v^j_{im|k\ell}, \pi^{(B)}_{km|i\ell}\right). \tag{31}$$

If $i \prec_A k$ and $\ell \prec_A m$. Under the null hypothesis, $i \prec_A k \Leftrightarrow i \prec_B k$, so $i \prec_B k$ and $\ell \prec_B m$. From the previous section, this implies $\pi^{(A)}_{km|i\ell} \geq \frac{1}{2}$ and $\pi^{(B)}_{km|i\ell} \geq \frac{1}{2}$. The same holds if $k \prec_A i$ and $m \prec_A \ell$. If, instead, $k \prec_A i$ and $\ell \prec_A m$ or if $i \prec_A k$ and $m \prec_A \ell$, then $\pi^{(A)}_{km|i\ell} \leq \frac{1}{2}$ and $\pi^{(B)}_{km|i\ell} \leq \frac{1}{2}$. Thus, either $\pi^{(A)}_{km|i\ell} \geq \frac{1}{2}$ and $\pi^{(B)}_{km|i\ell} \geq \frac{1}{2}$ or $\pi^{(A)}_{km|i\ell} \leq \frac{1}{2}$ and $\pi^{(B)}_{km|i\ell} \leq \frac{1}{2}$. We, thus, form a p-value for a test of the hypothesis that $\left(\pi^{(A)}_{km|i\ell} \geq \frac{1}{2} \wedge \pi^{(B)}_{km|i\ell} \geq \frac{1}{2}\right) \vee \left(\pi^{(A)}_{km|i\ell} \leq \frac{1}{2} \wedge \pi^{(B)}_{km|i\ell} \leq \frac{1}{2}\right)$. Call this value $\rho_{ik\ell m}$.

We want to combine the values of $\rho_{ik\ell m}$ across all possible choices of the four indices $ik\ell m$. When combining p-values for disjoint groups of four legislators, the p-values are independent. Thus, the p-values can be aggregated via Fisher's method. This gives $-2\sum\log\left(\rho_{ik\ell m}\right)$, which, under the null hypothesis, will be stochastically dominated by a chi-squared distribution with degrees of freedom equal to twice the number of p-values being aggregated.

If p-values are computed for two groups of legislators that overlap, the p-values are not independent. However, the expectation of the sum of these p-values is unaffected by the lack of independence. Thus, $E\left[\sum\left(-2\log\left(\rho_{ik\ell m}\right) - 1\right)\right] \leq 0$ under the null hypothesis. We can then compute the sample distribution of $\sum\left(-2\log\left(\rho_{ik\ell m}\right) - 1\right)$ using a stratified nonparametric bootstrap with stratification of the two groups, A and B.[14] From this, we can test the hypothesis that $E\left[\sum\left(-2\log\left(\rho_{ik\ell m}\right) - 1\right)\right] \leq 0$, which then also serves as a test of the hypothesis that the orderings of the two groups are the same.

This test as constructed has relatively low power in cases where $\pi^{(A)}_{km|i\ell}, \pi^{(B)}_{km|i\ell} \ll \frac{1}{2}$ or $\pi^{(A)}_{km|i\ell}, \pi^{(B)}_{km|i\ell} \gg \frac{1}{2}$. Under the null hypotheses, all p-values have at least 50% probability of being greater than $\frac{1}{2}$. However, in cases where $\pi^{(A)}_{km|i\ell}, \pi^{(B)}_{km|i\ell} \ll \frac{1}{2}$ or $\pi^{(A)}_{km|i\ell}, \pi^{(B)}_{km|i\ell} \gg \frac{1}{2}$, p-values may be particularly conservative and have at least 50% probability of being greater than $\frac{1}{2}$. To improve its power, we can only include p-values in the sum for groups of legislators where $\sum_{j\in A} v^j_{km|i\ell} > \sum_{j\in A} v^j_{im|k\ell}$ and $\sum_{j\in B} v^j_{km|i\ell} < \sum_{j\in B} v^j_{im|k\ell}$ or where $\sum_{j\in A} v^j_{km|i\ell} < \sum_{j\in A} v^j_{im|k\ell}$ and $\sum_{j\in B} v^j_{km|i\ell} > \sum_{j\in B} v^j_{im|k\ell}$. These are the groups of legislators where the null hypothesis seems most suspect.

Of course, since we are selecting only a subset of p-values, their distribution will no longer be uniform under the null hypothesis as the p-value is not independent of its inclusion. Thus, without adjustment, we can no longer expect these p-values to be uniformly distributed under the null hypothesis. However, if we replace these p-values with p-values computed conditional on their inclusion—that is, based on the conditional distribution of a test statistic conditional on $\sum_{j\in A} v^j_{km|i\ell} > \sum_{j\in A} v^j_{im|k\ell}$ and $\sum_{j\in B} v^j_{km|i\ell} < \sum_{j\in B} v^j_{im|k\ell}$ or $\sum_{j\in A} v^j_{km|i\ell} < \sum_{j\in A} v^j_{im|k\ell}$ and $\sum_{j\in B} v^j_{km|i\ell} > \sum_{j\in B} v^j_{im|k\ell}$—then the included p-values will again be uniformly distributed (or stochastically dominated by a uniform distribution) under the null hypothesis.

14 An outline of the algorithm for this procedure is provided in Appendix A in the supplementary material.

In some cases, we may wish to test whether a particular legislator changed location relative to any other legislator regardless of whether other legislators changed order. Doing so requires identifying two other legislators who can safely be assumed to have the same ordering in both groups. With this, we can compute the test statistic as above, but only varying k with i equal to the legislator whose movement is to be tested and ℓ and m equal to the legislators identified as nonmoving. The test statistic is, thus, $\sum_k \left(-2\log\left(\rho_{ik\ell m}\right) - 1\right)$ with nonpositive expectation under the null hypothesis. Its distribution can again be estimated via the stratified bootstrap.

5 Monte Carlo Studies

I conducted a series of Monte Carlo simulations to test the properties of the estimator and hypothesis test developed in the previous sections. For all simulations, I use a legislature of size $K = 9$ with legislator i having ideal point $x_i = i$. In each simulation, I generate probabilities of voting "yea" according to one of two possible distributions.

The first possible distribution for vote probabilities, which I will term the "uniform bill distribution," the probability of voting in the conservative direction is $\Phi\left(2\left(x_i - \alpha_j\right)\right)$ where Φ is the cumulative distribution function of a standard normal distribution, each α_j is drawn independently from a uniform distribution over the ideal points (x_i), the conservative direction is a "yea" vote with probability $\frac{1}{2}$ and is independent of any other random variables in the model. This distribution is uniform in the sense that the point at which the probability of voting "yea" is exactly $\frac{1}{2}$ occurs at the ideal point of each legislator with equal probability.

The second possible distribution, which I will term the "lopsided bill distribution," the probability of legislator i voting in the conservative direction is $1-\left(\frac{1}{2}\right)^i$. Since $x_i = i$, this probability is strictly increasing in the ideal point of a legislator. As in the previous case, the probability that the conservative direction is a "yea" vote is also randomly determined with $\frac{1}{2}$ probability of each. In this scenario, the bills being considered are both located toward one end of the spectrum, making the votes of the legislators at the other end more predictable. OC may be less likely to perform well in this scenario because the cutting points—that is, the points separating voters more likely to vote "yea" from those more likely to vote "nay"—occur only at one end of the spectrum.

5.1 Estimation

Table 1 shows the mean estimates and mean squared error based on the uniform bill distribution with 100 votes. Both the new estimator and the OC estimator recover the true rank of the ideal points quite well, although the new estimator has slightly lower mean squared error.

Table 2 shows the analogous results based on the lopsided bill distribution. Here, the legislators further to the right are all likely to vote together. Thus, it is difficult to estimate their relative positions. Both estimators do show higher mean squared error. However, while the new estimator and OC estimators performed relatively similarly with the uniform bill distribution, the new estimator performs significantly better here, with a mean squared error over 80% smaller.

Finally, Table 3 shows mean squared error based on data with different numbers of votes. Again, using the uniform bill distribution, both models perform roughly similarly. However, for data generated using the lopsided bill distribution, the new estimator performs noticeably better. Further, whereas the mean squared error for the new estimator decreases as more votes are added, the mean squared error of OC appears to *increase* slightly as the number of votes increases.

5.2 Inference

To test the inference, I simulate data for two groups of bills, with vote probabilities for each generated from the uniform bill distribution. For the first simulation, the legislators maintain the same ordering on both sets of bills. Thus, the null hypothesis is true, so the test should not reject at the 5% significance level in more than 5% of simulations. Note that, given the weak

Table 1. Estimate means for the uniform bill distribution.

	Mean prob. of cons. vote	True rank	Mean of estimated rank	
			New	OC
Legislator 1	0.12	1	1.11	1.07
Legislator 2	0.20	2	1.95	2.00
Legislator 3	0.29	3	2.97	2.99
Legislator 4	0.39	4	3.99	4.01
Legislator 5	0.50	5	5.00	5.00
Legislator 6	0.61	6	6.01	5.99
Legislator 7	0.71	7	7.02	7.02
Legislator 8	0.81	8	8.05	8.01
Legislator 9	0.88	9	8.89	8.91
Mean Sq. Error			0.88	1.14

Note: $N = 100$ votes. Based on 10,000 Monte Carlo simulations. “New” indicates the estimator discussed in Section 3.

Table 2. Estimate means for lopsided bill distributions.

	Mean prob. of cons. vote	True rank	Mean of estimated rank	
			New	OC
Legislator 1	0.500	1	1.15	1.01
Legislator 2	0.250	2	2.14	5.40
Legislator 3	0.125	3	3.14	5.54
Legislator 4	0.063	4	4.27	5.35
Legislator 5	0.031	5	5.40	5.47
Legislator 6	0.016	6	6.42	5.59
Legislator 7	0.008	7	7.11	5.58
Legislator 8	0.004	8	7.58	5.53
Legislator 9	0.002	9	7.79	5.52
Mean Sq. Error			13.76	78.07

Note: $N = 100$ votes. Based on 10,000 Monte Carlo simulations. “New” indicates the estimator discussed in Section 3.

Table 3. Mean squared error by number of votes

	Uniform		Lopsided	
Num. votes	New	OC	New	OC
50	2.18	2.57	25.30	75.90
100	0.87	1.17	13.79	78.32
500	0.01	0.01	3.65	82.42
1000	0.00	0.00	1.89	83.41

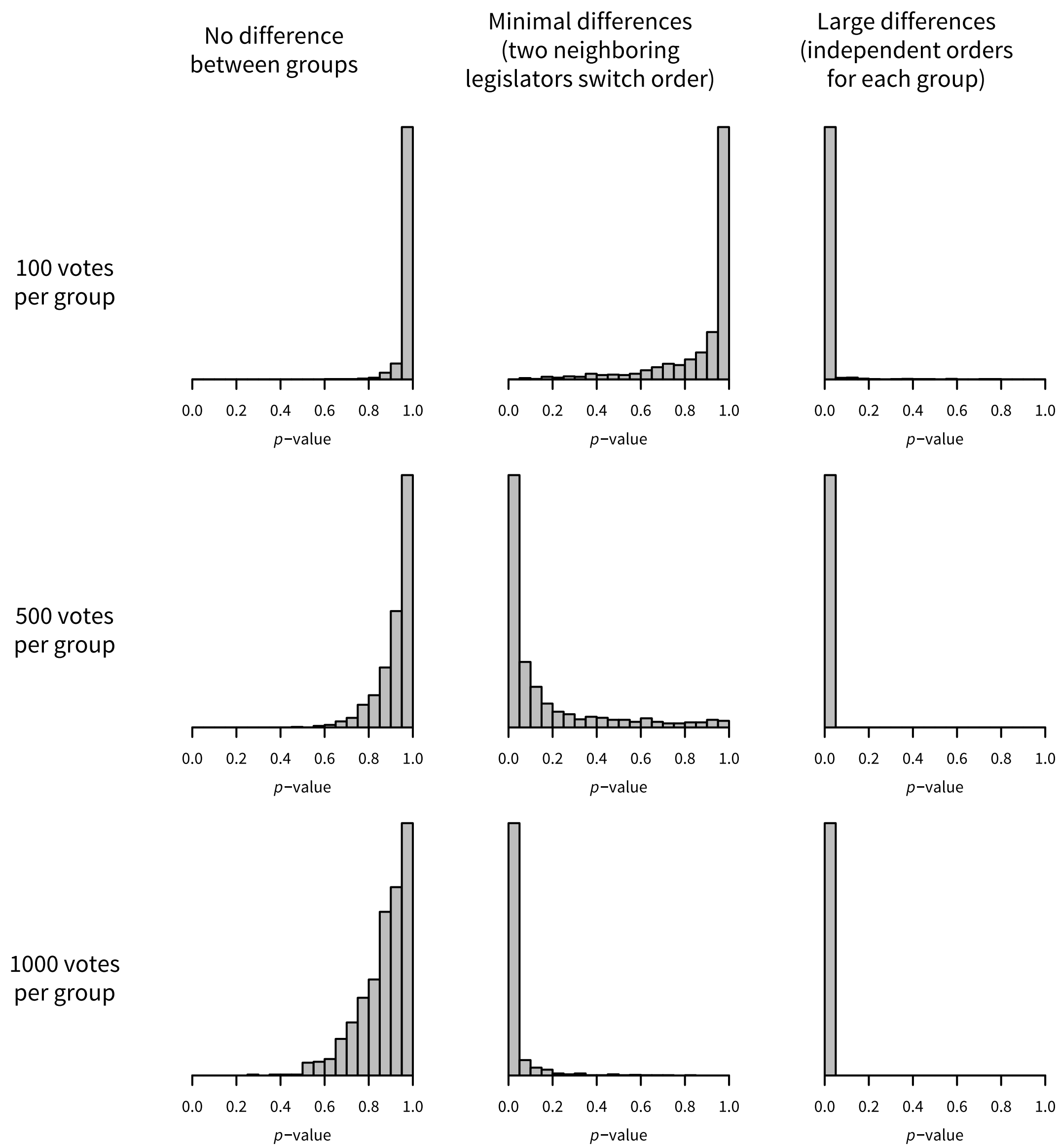

Figure 1. Histograms of p-values for test of same ordering from Monte Carlo simulations.

assumptions underlying the test, it is likely to be conservative and reject in significantly less than 5% of simulations.

In the next simulation, I choose two neighboring legislators at random and swap their order, which provides as minimal a violation of the null hypothesis as possible. Only one pair of legislators is in a different order in one group of votes as in the other. All other pairs maintain the same ordering. Thus, the null hypothesis is false and can be correctly rejected, but, given that the change is relatively small, then it may be difficult to detect.

Finally, I simulate data in which the ordering of the legislators for the first group of bills is unrelated to the ordering for the second. Here, each possible ordering of the legislators is equally likely, with the ordering chosen for the first group of bills independent of the ordering chosen for the second.

The results mirror expectations. Figure 1 shows histograms of p-values for each of the three scenarios discussed above and for each of 100, 500, and 1000 votes per group. The p-values in the first column, where the null hypothesis is true, tend to be greater than would be expected under a uniform distribution and the probability of rejection under the null hypothesis is below the nominal level of the test. This indicates the test is more conservative than possible, which is unsurprising.

In the second column, the null hypothesis is false but the orderings differ as little as possible. Here, the power of the test is quite low with 100 votes. However, as the number over voter increases, so does the power of the test. By 1000 votes, the probability of rejecting the null at the 5% significance level is 86.9%.

In the third column, the null hypothesis is false and a legislator's ideology on one dimension is independent of her ideology on the other. Here, even with only 100 votes per group, the probability of rejecting the null is estimated to be 97.4%. With at least 500 votes per group, the power is over 99.9%.

6 Supreme Court

6.1 Did Justice Blackmun move to the left?

As a second example, consider the question of whether Justice Harry Blackmun's ideology changed relative to those of the other justices during the 1970s. Blackmun is often hypothesized to have became more liberal during this period. While the tests developed in this paper cannot test this hypothesis in an absolute test, we can test whether Blackmun moved relative to the other justices. For the purposes of this example, I test specifically whether Blackmun's position relative to the other justices during the Nixon and Ford administrations was the same as his position during the Carter and Reagan administrations.

First, we can test whether a single ordering can explain voting both before 1977 and after. Using the test developed in Section 4, we can strongly reject this hypothesis ($p = 0.0005$). Thus, the data are incompatible with the justices having the same ordering in both time periods.

But this does not reveal whether Blackmun in particular changed. Choosing two justices to anchor the space permits a test of whether Blackmun in particular changed relative to any of the remaining justices. I use Justice Marshall and Chief Justice Burger. Note that this does not require assuming that Marshall and Burger do not move relative to the other justices. It only requires assuming that Burger is more conservative than Marshall in both time periods. Again, this test strongly rejects a null hypothesis that Blackmun remained on the same side of each of the remaining justices ($p < 0.0001$).

Of course, this does not establish whether it was Blackmun's ideology that changed or whether one or more of the other justices moved relative to Blackmun. However, we can test the hypothesis that the remaining justices maintained the same ordering across both time periods. With Blackmun excluded, the p-value becomes $p = 0.71$. Thus, we see no evidence that the ordering of the remaining justices changed during this time period.

6.2 Is the Supreme Court multidimensional?

First, consider the question of whether the Supreme Court is multidimensional. Using data from the Supreme Court Database (Spaeth *et al.* 2015), I will use the hypothesis test developed in Section 4 to test whether a common dimension can explain voting on pairs of issue areas. I aggregate cases into four groups based on the Supreme Court Database's coding of issue area: (1) *criminal procedure*, which includes only cases coded as "criminal procedure," (2) *civil liberties*, which includes cases coded as "civil rights," "First Amendment," "due process," and "privacy," (3) *economics*, which includes cases coded as "economic activity," "unions," and "federal taxation," and (4) *other*, which includes all other cases. Because the criminal procedure issue area is sometimes grouped together with the civil liberties issue areas and these issue areas are the ones "conventionally associated with the standard left–right splits on the Court" (Bailey 2013), I test whether a single ordering could explain voting on criminal procedure and civil liberties cases,

Table 4. Tests of the hypothesis that the ideal points of Supreme Court justices are in the same order on cases on two different issue areas.

Comparison	1950s	1960s	1970s	1980s	1990s	2000s
Crim. proc. & civ. lib. v. economics	0.547	0.001	0.140	0.900	0.399	0.908
Criminal procedure v. economics	0.789	0.040	0.066	0.894	0.675	1.000
Civil liberties v. economics	0.746	0.007	0.655	0.968	0.739	0.907
Criminal procedure v. civil liberties	1.000	0.720	0.993	0.395	0.985	0.991

Note: Data from the Supreme Court Database (Spaeth *et al.* 2015). *Civil liberties* combines "civil rights," "First Amendment," "due process," and "privacy" issue areas. *Economics* combines "economic activity," "unions," and "federal taxation" issue areas.

treated as a single group, and on economics for each decade beginning with the 1950s.[15] I then similarly test whether a single ordering could explain voting on criminal procedure and economics, on civil liberties and economics, and on criminal procedure and civil liberties.

Table 4 shows the *p*-value computed for each comparison in each decade. Several comparisons are significant during the 1960s. Tests for the comparison between economics and criminal procedure, between economics and civil liberties, and between economics and both criminal procedure and civil liberties reject the null hypothesis that the ideal points have a common ordering during the 1960s at the 5% significance level. The test comparing criminal procedure and economics also nearly rejects for the 1970s ($p = 0.066$).

The *p*-values reported in the table make no adjustment for the fact that tests are performed for six different decades. However, even when the *p*-values for each comparison are adjusted using the Holm procedure so that the family wise error rate—that is, the collective probability of falsely rejecting the null across all six decades—for each comparison across all six decades is less than 5%, the null hypothesis can be rejected for the comparisons between economics and civil liberties (adjusted *p*-value of 0.039) and between economics and both criminal procedure and civil liberties (adjusted *p*-value of 0.009).[16]

Thus, the extent to which the Court is multidimensional—that is, the extent to which different orderings are needed to explain Supreme Court voting on different issue areas—may depend on the time period in question. For the 1960s, we can reject the hypothesis that the justices had the same ordering on cases involving economics as on criminal procedure and civil liberties. On the other hand, there is no evidence that multiple dimensions are needed for other decades. Note that, while the number of cases the Court considered was slightly lower during the 1990s and 2000s, this alone does not appear to explain the difference, both because there were roughly equal numbers of cases considered for each issue area during the 1960s and 1980s and because the significance of the results persist for the 1960s even when using shorter, five-year periods of time.

7 Conclusion

The nonparametric methods developed in this paper require only that voting is unidimensional, that the shape of the deterministic component of utility is concave and identical for all voters aside from differences in their ideal points, and that the errors are independent and identically distributed across legislators. The underlying model makes no parametric assumptions about the shape of the utility functions or error distribution.

Under these modeling assumptions, the estimator is consistent. In Monte Carlo simulations, it outperformed OC in terms of mean squared error by small margins in one case and larger

15 In each test, I use voting data from all justices who served on the Court for a majority of the years in question.

16 Controlling for the family wise error rate across all 24 hypothesis tests using the Holm procedure leaves the comparison between economics and both criminal procedure and civil liberties in the 1960s as the only significant relationship (adjusted *p*-value of 0.034).

margins in another. In some scenarios where OC is inconsistent, this is a significant improvement. Nonetheless, in most scenarios, both estimators are likely to return relatively similar results.

The method for nonparametric inference is more novel. Given two groups of bills, this method permits researchers to test the hypothesis that legislators' ideal points have the same ordering on one group as on the other. This is not quite the same as testing whether the legislators have the same ideal points, as the ideal points could differ while still maintaining the same order. Nonetheless, a change in the order of the ideal points does imply a change in the ideal points themselves.

There are also several reasons why testing whether the orderings are the same can be preferable. Recovering cardinal information about ideal points requires the sort of stronger assumptions provided by parametric models, but researchers are rarely confident in the accuracy of these assumptions. Previous work has suggested using only ordinal information from ideal-point models because cardinal information can be very sensitive to parametric assumptions (Ho and Quinn 2010). Further, changes in the ordering may be substantively meaningful while at least small changes in the ideal points may not be.

The primary limiting assumption of the model in this paper is of unidimensionality. The remaining assumptions are significantly weaker than those adopted in other approaches to ideal-point estimation or inference. Most existing research on ideal-point estimation, such as Poole and Rosenthal (1991), Heckman and Snyder (1997), and Clinton, Jackman, and Rivers (2004b), requires assuming a particular parametric form for the utility functions and error distributions. The only other approaches, such as Poole (2000), do not use a stochastic model, removing the possibility of statistical inference. As far as I am aware, this paper is the first to provide an approach to provably consistent estimation and valid statistical inference for the order of ideal points without requiring knowledge of the parametric form of the utility functions or error distributions.

While the model cannot easily be extended to include a multidimensional Euclidean policy space, as this would render rank order a meaningful concept, other approaches to dealing with a single set of bills with multiple dimensions are possible. I suggest following the direction of Lauderdale and Clark (2012) and Bonica (2011) as well as nonparametric item-response theory by assuming legislators have unidimensional preferences for any bill but for which their ordering may change from one bill to another. Thus, one might be able to divide a set of bills into subsets that involve different dimensions, with legislators having different dimensions on each and the determination of which bills belong to the same subset perhaps informed by the sort of hypothesis test of equal ordering discussed in this paper. I leave this extension as a subject for future research.

Supplementary material

For supplementary material accompanying this paper, please visit
https://doi.org/10.1017/pan.2017.38.

References

Bailey, M. A. 2013. Is todays court the most conservative in sixty years? Challenges and opportunities in measuring judicial preferences. *Journal of Politics* 75(3):821–834.

Bonica, Adam. 2011 *Ideology and Interests in American Politics*. Ph.D. thesis, New York University.

Chiou, F.-Y., and Lawrence S. Rothenberg. 2003. When pivotal politics meets partisan politics. *American Journal of Political Science* 47(3):503–522.

Clinton, J., Simon D. Jackman, and Douglas Rivers. 2004a. The most liberal senator? Analyzing and interpreting congressional roll calls. *Political Science and Politics* 37(4):805–811.

Clinton, J., Simon D. Jackman, and Douglas Rivers. 2004b. The statistical analysis of roll call data. *American Political Science Review* 98(2):355–370.

Covington, Cary R., and A. A. Bargen. 2004. Comparing floor-dominated and party-dominated explanations of policy change in the House of Representatives. *Journal of Politics* 66(4):1069–1088.

PA

Epstein, Lee, Andrew D. Martin, Kevin M. Quinn, and Jeffrey A. Segal. 2007. Ideological drift among supreme court justices: Who, when, and how important? *Northwestern University Law Review* 101(4):1483–1541.

Gibbons, Jean Dickinson, and Subhabrata Chakraborti. 2011. Nonparametric statistical inference. In *International encyclopedia of statistical science*, ed. Miodrag Lovric. Heidelberg: Springer, pp. 977–979.

Heckman, James J., and James M. Snyder Jr. 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *RAND Journal of Economics* 28:S142–S189.

Ho, Daniel, and Kevin M. Quinn. 2010. How not to lie with judicial votes: Misconceptions, measurement, and models. *California Law Review* 98(3):813–876.

Krehbiel, Keith, and Zachary Peskowitz. 2015. Legislative organization and ideal-point bias. *Journal of Theoretical Politics* 27(4):673–703.

Lauderdale, Benjamin E., and Tom S. Clark. 2012. The Supreme Court's many median justices. *American Political Science Review* 106(4):847–866.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2001. The hunt for party discipline in Congress. *American Political Science Review* 95(3):673–687.

Mokken, Robert J. 1971. *A theory and procedure of scale analysis: With applications in political research*. Berlin: Walter de Gruyter.

Neyman, Jerzy, and Elizabeth L. Scott. 1948. Consistent estimates based on partially consistent observations. *Econometrica* 16(1):1–32.

Poole, Keith T. 2000. Nonparametric 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. 2007. Changing minds? Not in Congress! *Public Choice* 131(3–4):435–451.

Poole, Keith T., and Howard Rosenthal. 1991. Patterns of congressional voting. *American Journal of Political Science* 35(1):228–278.

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

Randles, Ronald H., Thomas P. Hettmansperger, and George Casella. 2004. Introduction to the special issue: Nonparametric statistics. *Statistical Science* 19(4):561–561.

Roberts, Jason M., S. S. Smith, and S. R. Haptonstahl. 2016. The dimensionality of congressional voting reconsidered. *American Politics Research* 44(5):1–22.

Rosenthal, Howard. 1992. The unidimensional Congress is not the result of selective gatekeeping. *American Journal of Political Science* 36(1):31–36.

Snyder, James M. Jr., and Tim Groseclose. 2000. Estimating party influence in congressional roll-call voting. *American Journal of Political Science* 44(2):193–211.

Spaeth, Harold J., Lee Epstein, Martin Andrew D., Jeffrey A. Segal, Theodore J. Ruger, and Sara C. Benesh. 2015 Supreme Court Database. Version 2015 Release 01. August 17. http://Supremecourtdatabase.org.

Tahk, Alexander. 2017. Replication data for: Nonparametric ideal-point estimation and inference. doi:10.7910/DVN/WIRN6R, Harvard Dataverse, V1, UNF:6:v+zAWAFABTnFjZGegcS2iA==.

Wolfowitz, Jacob. 1942. Additive partition functions and a class of statistical hypotheses. *Annals of Mathematical Statistics* 13(3):247–279.