An Optimally Classified Supreme Court:
Teaching Optimal Classification as an Introduction to Models of Spatial Voting

Jake S. Truscott
University of Georgia
Contact: jake.truscott@uga.edu

Abstract

I illustrate existing methodological approaches for non-parametric unfolding procedures to optimally classify justices serving on the United States Supreme Court. The purpose of which is to introduce a method of spatial voting that is both intuitive and less demanding than advanced statistical approaches. I begin by providing a descriptive analysis of alternative scaling methods before introducing the computational optimal classification procedure. Next, I employ 1,197 non-unanimous cases decided by the Rehnquist Court to illustrate the optimal classification procedure in a single dimension, which I subsequently compared to a W-NOMINATE statistical approach. I ultimately conclude by noting how optimal classification provides a useful introduction to models of spatial voting.

I. Introduction

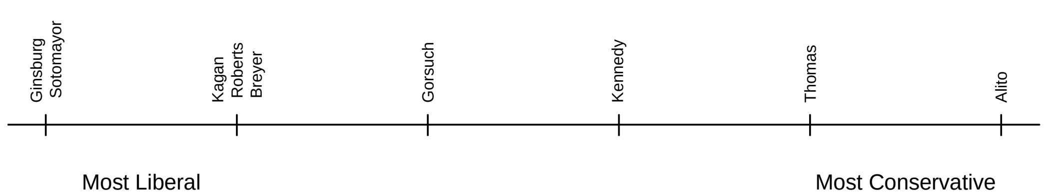
Measures of judicial ideology are routinely a subject of interest for scholars (e.g., Epstein, et al. 2007; Martin and Quinn 2002; Segal and Cover 1989). These measures are largely built on the premise that jurists can be organized into a policy space of ideological liberalism and conservatism to the same degree that we associate with legislators. With this, early attempts to frame judicial ideology – most prominently the justices of the Supreme Court – were often built as a reflection of uncomplicated measures of attitudinal voting behavior (e.g., Segal and Spaeth 1993, 2002). This generally began with a contextual reading of the Court's decisions to assess competing liberal and conservative-leaning positions. Scholars have routinely taken advantage of tools like the Spaeth, et al. Supreme Court database, which provides dispositions and a broad collection of other important information for every case ever decided by the Court. With this, a simple tally of which justices formed the majority and minority coalitions reflective of the liberal and conservative positions in the case produces individual-level voting behaviors. Expanding this to include a large sample of cases across a single (or successive set of) term(s) produces an ordinal ranking of the justices. For example, consider the following cases from the Court's 2017-2018 term (Table 1), where (1) corresponds with a justice voting for a *liberal* position in the case and (0) corresponds with the *conservative* position:

Table 1. Supreme Court Justices' Votes on 5 Non-Unanimous Cases (2017-2018)

	Collins v. Virginia	*Marinello v. US*	*Murphy v. NCAA*	*Artis v. D.C.*	*McCoy v. Louisiana*	(%) Liberalism
Roberts	1	1	0	1	1	0.800
Kennedy	1	1	0	0	1	0.600
Thomas	1	0	0	0	0	0.200
Ginsburg	1	1	1	1	1	1.000
Breyer	1	1	0	1	1	0.800
Alito	0	0	0	0	0	0.000
Sotomayor	1	1	1	1	1	1.000
Kagan	1	1	0	1	1	0.800
Gorsuch	1	1	0	0	0	0.400

From these votes, we can calculate the proportion of times that an individual justice voted for the liberal position. The results produce a rank-ordering of the justices from *Most Liberal* to *Most Conservative* (Figure 1). In essence, they replicate the same procedure that produces the multitude of interest group scores by constructing a measure of ideological preferences rooted in the proportion of times that a voter indicates support for a preferred position. A similar rank-ordering procedure is routinely reproduced by Segal and Cover using popular news editorials to discern classifications of ideology for nominees to the Supreme Court (*see* Segal and Cover 1989).

Figure 1. Supreme Court October Term 2017-2018 Liberalism



Yet, while these scores are surely intuitive, they face several empirical shortcomings. For one, there are several situations where a perfect voting record on non-unanimous decisions can produce rank-ordered ties between justices when the measure is based on the proportion of policy position support – like which we can see from Figure 1. This wouldn't be as much of a concern if not for the reality that perfect spatial voting among voters – and especially so for legislators – is notably uncommon, and even so among alike partisans (*see* Poole 2005). A second concern is that sometimes the policy positions are not neatly discernable. Perhaps the greatest misconception about the Supreme Court is the belief that every case they review concerns some hyper-salient social or political issue where neatly defined policy positions can be easily identified. In reality, a substantial portion of the Court's docket are mundane cases and controversies where the justices serve largely as mediators. For example, consider the Court's 2019 decision in *BNSF Railway Company v. Loos,* which concerned a former railroad employee's liability to pay their share of taxes owed under the Railroad Retirement Tax Act (RRTA). The case specifically questioned whether damages for lost wages that were awarded by a jury as a result of a negligence claim were considered compensation and thus subject to employment taxes. Using a legal doctrine similar to previous holdings related to the Social Security Act (e.g., *Board v. Nierotko* 1946; *United States v. Quality Stores* 2014), the Court ruled 7-2 that lost wages are indeed considered compensation and taxable under the RRTA. The overarching question that emerges in a case like this is how to neatly disseminate between the liberal and conservative positions? As cases move more towards clarifying ambiguity in the law rather than considering substantive constitutional questions, discerning policy positions surely becomes increasingly difficult. The same distinctions can be made about legislative voting. This limitation often pushes observers to measure voting behavior on policies with easily identifiable positions, which will almost surely reduce the sample size and not be reflective of aggregate voting behaviors. A final concern is that the sample set of votes in Table 1 is noticeably small (n=5). Surely, we could expand the sample set to include more votes, but a measure rooted in proportionality still runs a greater risk of ties. Nonetheless, even interest group scores like those from Americans for Democratic Action (ADA), which represents one of the most frequently used proportion-based scores for American legislators, often rely on small samples of votes. ADA only considers the 20 pieces of legislation that they deem *most important* for their organization, which of course is a remarkably small sample of the total pieces of legislation that might be voted on in a given Congress. There are other concerns, but these constitute some of the most frequently cited and outstanding.

With this, scholars have moved towards alternative measures derived from scaling procedures. The emergence and intrigue associated with Poole and Rosenthal's seminal works on NOMINATE (e.g., Poole and Rosenthal 1985, 2000, 2001) motivated a surge in the social sciences to develop scaled representations of ideology. Among these, a handful of dominant measures of Supreme Court ideology began to emerge. These include, but are not entirely limited to, Adam Bonica's DIME common space (2019), Giles, Hettinger, and Pepper (GHP) scores (2001), Michael Bailey's Bridged Common Space Ideal Points (2007), and Martin-Quinn scores (2002). While Bonica's and Bailey's scores are rooted in constructing a common space for cross-institutional actors, Martin-Quinn and GHP remain solely in the domain of judicial actors.[1] Though each approach their estimation marginally different, the core concepts tend to remain consistent: employing a Bayesian (or an alike statistical) approach to estimate ideal points rooted in observable voting behaviors that are representative of an underlying dimension of voter preferences.

However, estimating ideal points using a Bayesian approach – for example, an Item Response Theory (IRT) model – can be a daunting experience for those first being introduced to

[1] *Note:* Though they can be, and have been, translated into a common space reflective of Poole-Rosenthal's NOMINATE scale through Epstein, et al.'s (2007) work on the judicial common space.

the theories of spatial voting. Yet, an introduction to this subset of political science does not actually require advanced statistical skills. While the next section might employ some technical language, the procedure for employing optimal classification – especially in a single dimension – only requires a basic understanding of addition and subtraction.

II. Manual Optimal Classification

Optimal classification (OC) in the context of parliamentary voting considers a simple question: When comparing a group of legislators' roll call voting behaviors, what is the rank-ordering of the members that optimally reduces the number of classification errors? In essence, if we consider the voting behavior of legislators across a term of voting, how can we classify their rank ordering in such a way that we cannot improve the maximum classification of the scale? This work aims to reserve these core considerations of voting but transfer the behavior of focus (i.e., legislative roll call voting) to justices of the Supreme Court.

Before outlining the optimal classification procedure, it is necessary to consider the underlying parameters and assumptions, which are noticeably flexible and undemanding. The OC scaling method represents a class of nonmetric unfolding procedures. As Poole (2005) notes, "*It is 'unfolding' in that the roll calls are treated as preferential choice data and parameters for individuals (legislators) and stimuli (roll calls) are being estimated. It is 'nonmetric' in that no assumptions are made about the parametric form of the legislators' true preference functions other than that they are symmetric and single-peaked"* (p. 46-47). This work will also assume that the justices vote deterministically, as opposed to probabilistically.[2] This means that the

---

[2] *Note:* Though Poole (2005) does illustrate that the procedure can be classified probabilistically with the incorporation of random error such that a voter's overall utility for voting *yea* is the sum of a deterministic utility and random error. Formally, a justice *i* would vote *yea*$_{(y)}$ on case *j* if U_{ijy} > U_{ijn} where $U_{ijy} = u_{ijy} + \varepsilon_{ijy}$

justices will vote for the alternative closest to them in a policy space as utility maximizers. Formally, if a justice$_{(i)}$ is deciding between voting liberally$_{(l)}$ or conservatively$_{(c)}$ on a case$_{(j)}$,[3] justice *i* will vote liberally if $U_{ijl} > U_{ijc}$.

Compared to the techniques employed for interest group scores described previously, optimal classification corrects for several shortcomings. First, deriving policy positions on votes are no longer required. Rather than asking whether a justice supported a liberal or conservative position, we simply need to know how they voted in comparison to the other justices who voted on the same cases. In essence, the process is rooted in comparative voting on alike cases rather than simply discerning their proportional support of a position. In a similar vein, it is now possible to observe the aggregate voting record rather than a small sample that is rooted either in an interest group's perception of *important* cases or those where it is easy to derive policy positions. Finally, without needing to know the policy position and having access to a much larger sample size, this almost surely reduces the propensity of rank-ordered ties. That's not to say that it entirely eliminates the possibility of perfect voting – though it is very unlikely with enough votes, but an optimal rank ordering assures that no two justices will occupy the same position.

At its core, the OC process in a single dimension consists of two sequentially repeated algorithms: a *cutting point* procedure and a *legislative* procedure.[4] Poole (2000, 2005) illustrates this procedure of one-dimensional maximum classification scaling (the *Janice* algorithm) with exceptional precision, and I will review these procedural steps. However, I will begin to

[3] *Note:* I classify a *liberal* vote as the equivalent of voting *yea* (or 1) in a parliamentary voting structure. However, the identification choice is reflexive – i.e., identifying a *conservative* vote as *yea* does not alter the classification.
[4] *Note:* Or a *cutting plane* procedure in two-dimensions.

supplement the key terminology from Poole's application of the procedure to legislators with its counterparts for classifying justices of the Supreme Court. Terms in italics and parentheses will denote the terminology in the context of parliamentary voting.

1. Let *p* (1…,p) denote the number of justices (*legislators*), *q* (1…,q) denotes the number of non-unanimous cases (*roll calls*) decided by the Court, and *s* (1…,s) denotes the number of voting dimensions. In this context, the number of dimensions will be fixed at one. The resulting *p x s* matrix (*p x 1,* where *s* = 1*)* contains the ideal points of each justice denoted x_i**.** The first step requires the estimation of starting values for x_i, which will be iteratively improved upon throughout the classification procedure. To do so, calculate an agreement score matrix whereby the *agreement score* between two justices is the proportion of times that they jointly vote the same way across cases. The agreement matrix is subsequently squared and double-centered. The resulting eigenvectors will produce the starting values for each justice (x_i) and classify them as a rank order. Following this procedure is useful because it provides an objective set of starting values reflective of the population's voting behavior. However, this technical approach is not an absolute requirement. Considering that the underlying premise of OC is to optimize classification from an initial rank-ordering, the initial ordering does not need to be perfect, but it should nonetheless be rooted in expectations better than simply guessing. For example, we could surely presume that Justice Ruth Bader Ginsburg would have a liberal-leaning voting record while Justice Samuel Alito would be conservative-leaning. Yet, filling the remainder of the space, at least as a starting point for classification, should ideally be rooted in an analysis of their proportional agreement and the distance between justices on alike cases rather than simple inference.

2. Fixing initial ideal points for x_i, the next step is to employ Poole's *Janice* algorithm by locating the optimal cutting point ordering. For every case *j*, there will be one cut point *n* that optimizes the classification procedure. This means that for every case *j*, there exists a cutting line that separates justices who would vote *liberally* from those who vote *conservatively* at a point that maximizes the number of correct classifications. This point would also reflect where the number of incorrect classifications – i.e., instances where a justice would predictively vote *liberally* but instead votes *conservatively* (or visa-versa) – are minimized. Notice too that the classification method does not deter the possibility of two (or more) optimized cut points. In this case, the algorithm will select the cut point closest to the center of the rank order. To illustrate, Table 2 offers the initial cut points for each of the original (5) cases. The justices' initial starting points x_i were strategically selected in such a way as to increase the number of initial classification errors.

Table 2. Initial Cut Points of Sample Supreme Court Cases (2017-2018 Term)

Case	Gi	Ka	S	Ke	B	T	R	Go	A	Errors
Collins	1	1	1	1	1	1	1	1	0	0
Marinello	1	1	1	1	1	0	1	1	0	1
Murphy	1	0	1	0	0	0	0	0	0	1
Artis	1	1	1	0	1	0	1	0	0	2
McCoy	1	1	1	1	1	0	1	0	0	1

Murphy *Artis* *McCoy* *Collins*
Marinello

Note: Gi= Ginsburg, Ka= Kagan, S= Sotomayor, B= Breyer, R= Roberts, Go=Gorsuch, A= Alito, T= Thomas, and Ke=Kennedy.

3. With such an abundance of initial classification errors, the final original step would be to construct cut points that minimize the number of classification errors using the *Janice* algorithm. This process should be repeated until the number of classification errors is

minimized. The resulting rank ordering with optimal cut-points is illustrated in Table 3. To reiterate, there is no assurance that this procedure will produce a rank order where the row classifications produce zero errors. Instead, the objective is to optimize the classification in such a way that the errors are minimized.

Table 3. Optimal Cut Points of Sample Supreme Court Cases (2017 Term)

Case	Gi	S	Ka	B	R	Ke	Go	T	A	Errors
Collins	1	1	1	1	1	1	1	1	0	0
Marinello	1	1	1	1	1	1	1	0	0	0
Murphy	1	1	0	0	0	0	0	0	0	0
Artis	1	1	1	1	1	0	0	0	0	0
McCoy	1	1	1	1	1	1	0	0	0	0

Murphy *Artis* *McCoy* *Marinello* *Collins*

Note: Gi= Ginsburg, Ka= Kagan, S= Sotomayor, B= Breyer, R= Roberts, Go=Gorsuch, A= Alito, T= Thomas, and Ke=Kennedy.

III. Optimal Classification in R

The steps listed above from Poole's *Spatial Models of Parliamentary Voting* (2005) provide a step-by-step illustration of optimal classification. Using a sample of binary choice votes and a basic understanding of addition and subtraction, the application is noticeably simple. However, as one might expect, optimally classifying voters – even in a single dimension – becomes increasingly difficult as the number of voters (*i*) and cases (*j*) increase. Optimally classifying nine justices of the Supreme Court across five cases in a single term might only require a few minutes of effort, but hundreds of cases across multiple successive terms is a challenging task. Luckily, Poole and others in the field have translated the intuition behind the optimal classification procedure to convenient and adaptable CRAN packages in R. I will illustrate these easy-to-use packages to optimally classify the Supreme Court by focusing on every non-unanimous case decided during the Rehnquist Court between 1986 and 2005.[5]

III. A. Upload Packages

A collection of CRAN packages will be used to run optimal classification in R – the primary of which will be the *oc* package from Poole, Lewis, Lo, and Carroll.[6] Before moving any further, be sure to install each and upload their respective CRAN libraries.

```
#install.packages(“pscl”)
#install.packages(“gdata”)
#install.packages(“oc”)
#install.packages(“wnominate”)
#library(pscl)
#library(gdata)
#library(oc)
#library(wnominate)
```

III. B. Preparation

Begin by constructing an empty data frame in R where the number of columns represents the number of cases, and the rows represent the number of Supreme Court justices. I was able to collect data from the Spaeth, et al. Supreme Court database for every non-unanimous case decided by the Court between 1790 and 2020 – though we will disaggregate to only include the Rehnquist-era cases in subsequent steps. Across the history of the court, this amounted to 8,629 non-unanimous cases (column value) and 113 justices (row value). The next step is to fill the frame with the vote each justice offered across each case. Subsequently construct a vector of justice names (*justice_name)* that will be used to assign row names. Specific column names (e.g.,

[5] *Note*: The full R-script will be provided in Appendix A.
[6] *Note:* The primary package was removed from the official CRAN repository, but archived versions and guides can be accessed at: https://cran.r-project.org/src/contrib/Archive/oc/

official case names or titles) can also be applied for each case, though I opted to simply assign a numerical value to each column.

```
##Create Data Frame ##
#oc <- matrix(justice_name, nrow = 113, ncol=8629)
#oc[1,1] #<- looks at column 1 row 1 (JJAY Case 1)
#rownames(oc) <-justice_name
#colnames(oc) <- c(1:8629)
#justice_name = (#Insert justice names separated by quotes and
 commas#)
#rownames(oc) <-justice_name
#colnames(oc) <- c(1:8629)
#oc[oc == justice_name] <- 9 #The original macro code left the
replacement cell value as the justice's name ID if they did not vote
on the case. This replaces each situation where the cell value is the
justice's name with the number 9 (Did Not Vote)
```

III. C. Measurement in One-Dimension

With the primary data frame constructed and the packages uploaded, we can now optimally classify the Rehnquist Court (1986-2005). The first step is to disaggregate the mass of cases from across the Court's history by selecting only cases from the Rehnquist Court. To do this, create an object in R that you define as the columns corresponding with the cases from that period. Doing so produces a new object where each column vector represents a case along with each row corresponding with a justice's vote.

```
#rehnquist <- oc[,c(6803:8000)]
```

The next step is to define polarity. As noted previously, optimal classification can iteratively optimize the classification from an initial rank ordering. However, just as NOMINATE requires starting positions (priors) to define the space, optimal classification requires a basic assumption of polarity as a means to define the directionality of the scale. This can be provided by denoting which justice could theoretically represent the most liberal or conservative member of the Court. Assuming you know each case's disposition and policy positions, a simple way to derive this might be to just measure the proportion of times a justice voted liberally or conservatively across a large sample of cases. Another might simply be selecting a member based on the inference that history defines them as the most partisan justice on the bench during that given era. Regardless, selecting a justice whose preferences align (very) near or at the end of the spectrum is important to derive the correct inferences from the resulting rank ordering. For the Rehnquist Court, I defined Justice Clarence Thomas as the most conservative member. When the rank order is produced, I will be able to reference the location of Justice Thomas to infer the directionality of the scale – i.e., dependent upon where Thomas is ordered, I can infer that to be the conservative wing of the scale.

```
#REST1_rehnquist <- 105 #Note: Thomas is the 105th row of the
justice_name list
```

Once polarity has been defined, the next step is to define a *roll call* object for running optimal classification. This requires the coordination of two commands: the *roll call* command in Simon Jackson's *pscl* CRAN-Package[7] and the *oc* command from Poole, et al.'s *oc* CRAN-Package. *Roll call* creates an object class in R for binary choice data to analyze legislative voting, though we can define the voting data from Supreme Court decisions for an easy translation.

```
#rcd_rehnquist <- rollcall(data=rehnquist, yea=1, nay=0, missing=,
notInLegis=9, desc="Rehnquist Court", legis.names=justice_name,
vote.names=colnames(oc))
```

[7] For help and reference, please visit: https://cran.r-project.org/web/packages/pscl/pscl.pdf

Data represents the object of disaggregated cases from the Rehnquist Court, *yea* represents the liberal voting position (1),[8] *nay* represents the conservative voting position (0), and *notInLegis* (9) is used to code justices who did not vote on the case because they were not on the bench at that time. Especially since Chief Justices will experience turnover of associate justices during their tenures, it is important to rank order based solely on how the justices voted on cases where they actually participated. *Legis.name* represents the vector list names of the justices (*justice_name*) and *vote.names* represent the column names coinciding with each case from the period. The final step is to optimally classify the justices using the parameters defined in the *roll call* object.[9]

```
#oc_result_rehnquist <- oc(rcd_rehnquist, dims=1,
 polarity=REST1_rehnquist, minvotes=20, lop=0.005)
```

The *oc* command produces several noteworthy results beyond just the rank ordering of the justices themselves, such as the optimal classification of the roll calls (cases), the identification of the number of dimensions classified, and the eigenvalues, fits, and identification of the object class. Further, the object also lists classification performance metrics within the legislators (justices) and roll calls (cases) subsets. This includes the number of correctly and incorrectly classified justices and cases using the *Janice* (and *Edith*) algorithm, which is surely useful for gauging the accuracy of the rank order.[10] It is always important to reference these measures before continuing because regardless of the number of cases or justices included in the

[8] *Note:* Defining *yea* as liberal and *nay* as conservative was an arbitrary decision. Using the opposite inference (*yea = conservative* and *nay* = liberal) will not alter the results.
[9] Note: *minvotes* represent the minimum number of votes that a justice needs to have offered to be included in the dataset. If that threshold is not met, they will not be included in the optimal ordering.
[10] *Note:* Performance metrics can be accessed for the justices via oc_result_rehnquist$legislators, and for the cases via oc_result_rehnquist$rollcalls

data, the *oc* package will always produce a rank order, assuming that the command is constructed correctly. However, if there is a substantial degree of incorrectly classified voters and votes, then the accuracy of the ordering is surely questionable. There is no pre-defined test or threshold for determining what degree of inaccuracy means the ordering is questionable, but it is important to ensure that the rank ordering isn't overwhelmingly plagued by inaccurate classifications. Table 4 reproduces the classification performance metrics from the optimally classified Rehnquist Court, which illustrates a substantially greater degree of correct classifications than incorrect.

Table 4. Performance Classification Metrics of Optimally Classified Rehnquist Court

Justice	Rank	Correct Yea	Wrong Yea	Wrong Nay	Correct Nay
WJBrennan	2	178	4	7	189
BRWhite	8	462	35	6	84
TMarshall	1	197	2	3	250
HABlackmun	4	358	47	8	229
LFPowell	7	93	9	2	9
WHRehnquist	14	738	3	142	299
JPStevens	3	534	61	75	512
SDOConnor	10	893	94	25	170
AScalia	12	777	44	61	300
AMKennedy	11	795	59	68	139
DHSouter	9	574	62	3	163
CThomas	13	452	17	28	232
RBGinsburg	6	388	23	7	176
SGBreyer	5	314	29	48	148
%		93.24	6.76	14.27	85.73

The rank values in Table 4 represent the optimally classified ordering of the Rehnquist Court. However, a few additional steps are required before plotting. Most importantly, the non-necessary justices need to be defined and removed – i.e., justices who were not on the Court during the analyzed period. This is not always necessary if the original dataset only included specific justices that coinhabited the same period, but R will not automatically drop justices from the *justice_name* object that we used to define the *roll call* command. Since every justice across the Court's history is included in the initial calculation, begin by creating an object that represents the rank order. Second, because non-necessary justices will be ranked as *NA* in the *oc* object (*oc_result_rehnquist)*, run a replace command that substitutes the value with a large numerical value that we will omit from the scale in the succeeding plot.

```
#x_rehnquist <- (oc_result_rehnquist$legislators[,7])
#x_rehnquist <- ifelse(is.na(x_rehnquist),999,x_rehnquist)
```

Next, we can begin to build the parameters for the plot. Many of the elements included here are not required and can be adjusted based on the desire and needs of the author, but I've come to believe that this produces the cleanest illustrations.

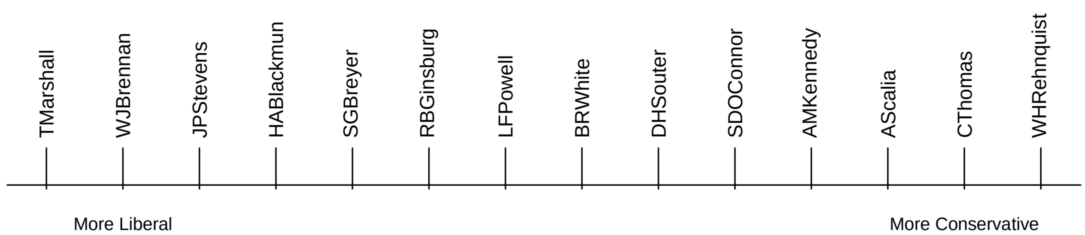
```
#Marker labels for the plot
#marker_rehnquist <- justice_name
#marker_missing_rehnquist <-ifelse(is.na(marker_rehnquist),1,0)
#Height of the marker labels
#height_rehnquist <-rep(1, 113)
#height_rehnquist[] <-1.10
#X-axis label
lablist <- as.vector(c("More Liberal", "More Conservative"))
#Plot construction
#plot(x_rehnquist,height_rehnquist,type="n",
       main="Polarity of United States Supreme Court \n (Rehnquist
 Court, 1986-2005)",
       xlab="",
       ylab="",
       xlim=c(1,14),
       ylim=c(1,1.5),cex=1.5,font=2, axes=FALSE)
text(seq(2, 13, by=11), par("usr")[3] - 0.05, labels = lablist, srt =
 0, pos = 1, xpd = TRUE, cex=0.75)
abline(h=0.99)
# Plot Markers and Add Justice Names
# i <- 1
  while (i <= length(x_rehnquist)){
    if(x_rehnquist[i] < 99){
```

```
      segments(x_rehnquist[i], 0, x_rehnquist[i],
height_rehnquist[i])
            if(marker_missing_rehnquist[i] != 1){

text(x_rehnquist[i],height_rehnquist[i]+.03,labels=justice_name[i],col
="black", cex= 0.85, srt=90, adj=0)
      }
    }
  i<-i+1
  }
```

This will produce Figure 2, which illustrates the optimal classification of the Rehnquist Court.

Figure 2. Optimal Classification of the Rehnquist Court (1986-2005)



Some initial observations that emerge from Figure 2 are that the wings of the scale – both liberal and conservative – appear to coincide with historical interpretations of the justices' ideological leanings during the Rehnquist era. Justice Thurgood Marshall, who founded the NAACP Legal Defense and Education Fund, served as lead counsel in landmark civil rights cases like *Brown v. Board of Education* (1954), and was historically championed for his liberal activism on the bench is firmly classified as the Court's most liberal member, even though he would only serve until 1991. Perhaps surprisingly, Chief Justice Rehnquist was classified as the most conservative member rather than Justice Thomas, who I had originally expected to fit that distinction. However, this result highlights an interesting benefit of optimal classification that was described previously. Namely, preliminary assumptions about the voters (or even their votes) do not need to be perfectly operationalized from the onset. Instead, we can rely on optimal classification to iteratively optimize the rank order that is reflective of the justices' comparative voting behaviors.

Perhaps with the exception of Justice Ginsburg's location as a comparable moderate when historically she is seen as more liberal, the remaining locations of the justices in the rank ordering appear very neatly reflective of historical accounts. Justice Blackmun, who is historically regarded as a justice who became increasingly more likely to abandon his initial conservative preferences as his tenure continued, is indeed found to be more liberal than conservative. We can make a similar remark of Justice O'Connor who, while the first woman on the Court and responsible for landmark legislation like upholding the core protections of female reproductive rights through *Planned Parenthood v. Casey*'s (1992) "undue burden" standard, was nonetheless considered a reliable conservative for much of her tenure. It is possible to offer similar distinctions to virtually every justice on the scale, which at minimum speaks to its accuracy, at least in terms of normative historical assumptions.

IV. Comparing Optimal Classification to Statistical Techniques

As mentioned previously, there are definitive benefits for using the optimal classification approach as an introduction to spatial voting models. For one, no prior knowledge of a case's (or roll call's) disposition is entirely necessary. For the Rehnquist Court example, I was able to code yea and nay as liberal or conservative voting positions, but this knowledge is not a necessary component for optimally classifying. We could have just as easily reversed the meaning of yea and nay to conservative versus liberal, or alternatively just denoted it as being a majority-coalition vote or a minority-coalition vote. So long as we know how the justices vote reflective of each other, we can optimally classify. Further, and perhaps importantly, the computational methodology is noticeably easier to comprehend than an advanced statistical approach like those used in Bayesian models like Martin-Quinn (2002) or NOMINATE (1985, 2001, 2001). So much to a point where optimally classifying in a single dimension could theoretically be done by hand, though R provides an easy-to-use suite of packages and commands to drastically expedite the process for larger sets of voters and votes.

However, while there are benefits in terms of ease, there are likewise some potential downsides. Namely, the benefit of using a computational approach as opposed to a statistical one means that the technique produces an ordinal ranking, rather than a cardinal ranking. Further, because the procedure is non-parametric, there is no incorporation of error to derive confidence intervals.[11] As such, while we might know that Justice Thomas was more conservative than Justice Brennan based on Figure 2, the means of interpreting their ideological differences are rooted in reference rather than distance. To illustrate, we can use optimal classification to denote that Justice Thomas was more conservative than Justices Scalia and Powell, who in turn were both more conservative than Justice Brennan. We cannot say that Justice Thomas was 11 units more conservative than Justice Brennan. In essence, the scale means nothing in terms of cardinal distance – it is purely an optimal rank ordering. Yet, if the goal is an introduction to measurement or simply to derive a normative understanding of the justices' comparative ideological alignments, then optimal classification serves as a worthwhile option. Alternatively, if cardinality is a necessary concern, then a statistically driven approach is likely required.

[11] Note: There is a way to incorporate random error into optimal classification using bootstrapping, though I do not explore it here. For a more in-depth discussion, *see* Poole 2005.

Another natural question that might emerge is how optimal classification actually compares to these more advanced scaling techniques. Poole and Rosenthal (2001) offer some keen observations on this topic. Analyzing legislators who served in Congress between 1789 and 1998, they note improvements on correct classification by optimal classification compared to DW-NOMINATE (Dynamic, Weighted, Nominal Three-Step Estimation).[12] They note how the optimal classification of the data "improves correct classification over DW-NOMINATE by about 2% for the House, and about 3% for the Senate" (p. 10). However, this is not entirely surprising considering that NOMINATE aims to maximize a likelihood function rather than optimize correct classifications. Even then, they still produce similar legislator configurations that correlate above 95% in both chambers.

To test this on Supreme Court data, I compare the optimal classification results to W-NOMINATE (Weighted, Nominal Three-Step Estimation) using the same data from the Rehnquist Court. Both approaches incorporate basic assumptions of voting. Specifically, they both assume that voters vote sincerely reflective of single-peaked and symmetric preferences. Likewise, neither require a complex understanding of the voting positions they're measuring nor will either push legislators to the extremes of the scale. However, while the underlying goal of scaling votes and voters in a space is understandably similar to optimal classification, there are a few important distinctions. Perhaps most importantly, W-NOMINATE employs a probabilistic

[12] *Note:* DW-NOMINATE is a dynamic representation of the W-NOMINATE framework, which was developed by McCarty, Poole, and Rosenthal (1997) and allows for cross-comparison of voters in a common space using a bundling of issues constraint based generally on perceptions of voter liberalism or conservatism. This is especially useful when trying to measure the dynamics of ideology among actors who never voted in the same period. As Poole and Rosenthal (2001) illustrate, "*We can thus claim that Jesse Helms is more conservative than Robert Taft, Sr. even though they never served in the Senate together*" (p. 8).

voting model using the logistic distribution function. Further, it incorporates a utility function with both a random and deterministic set of components that are distributed normally. [13]

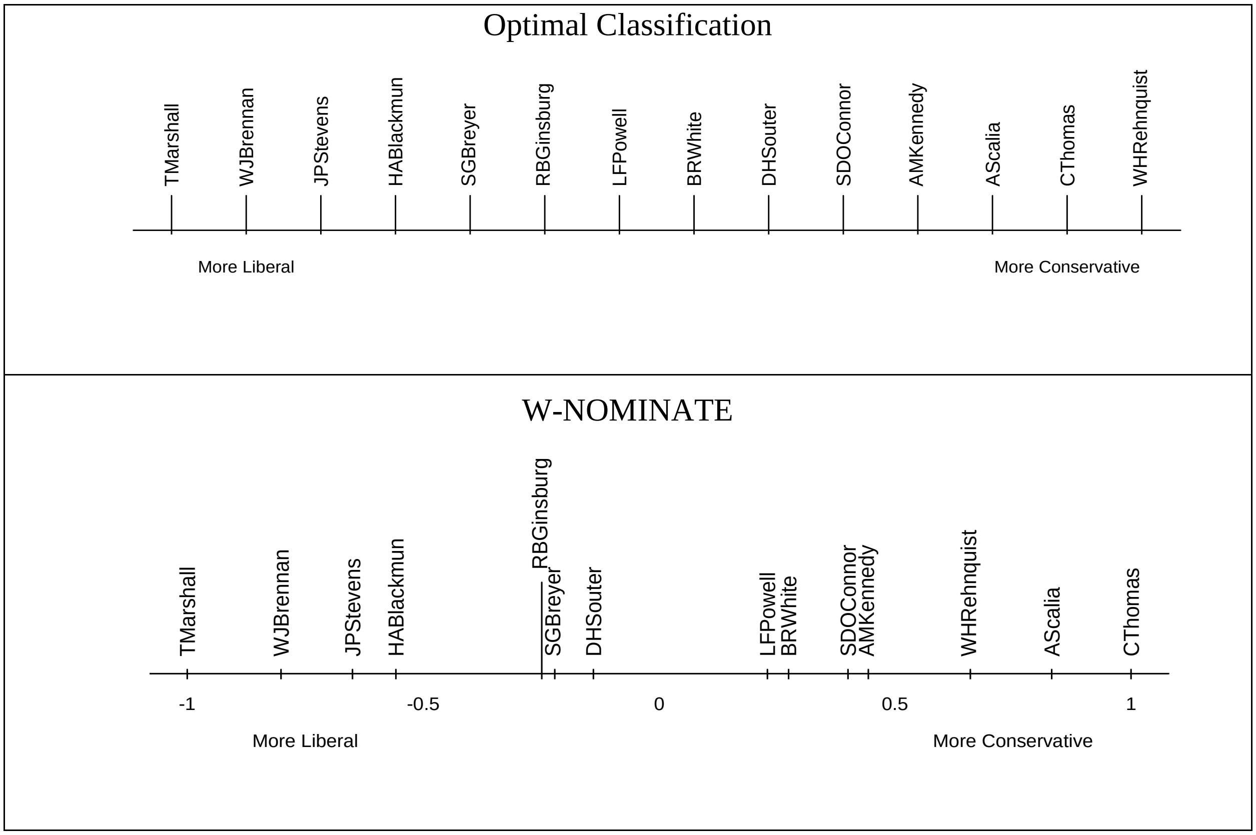
I reproduce the optimal classification rank ordering of the Rehnquist Court from Figure 2 below in Figure 3 alongside the results using W-NOMINATE.[14] There are two immediate distinctions to note between the measures. First, W-NOMINATE produces cardinality in the ordering of the justices. They are no longer spaced in an even rank ordering. Rather, they are separated by cardinal distances where we can derive the justices' ideological differences based on a numerical representation. For example, where we could originally say that Justice (Thurgood) Marshall was more liberal than Justice Stevens, we can now offer a numerical value to that distance of (0.351) units in the space.[15]

[13] For more information on the NOMINATE methodology, *see* Poole and Rosenthal (1985) and Poole (2005).

[14] *Note:* The R-Code for the W-NOMINATE estimation can be found in Appendix A.

[15] *Note:* In the example, Justice Marshall's ideal point was measured as (-1.0) while Justice Stevens's was (-0.649).

Figure 3. Optimal Classification and W-NOMINATE of the Rehnquist Court (1986-2005)



The second distinction to note is that there are some nominal differences in the location of the justices relative to each other. For example, the W-NOMINATE measure denotes Justice Thomas as the most conservative member rather than Chief Justice Rehnquist. In fact, with the exceptions of Justices Marshall, Brennan, Stevens, and Blackmun, there are some nominal shifts for all justices. This all being considered, the nominal differences between the two are not entirely surprising. As noted previously, NOMINATE focuses on maximizing the likelihood function rather than optimally classifying the rank order. More emphasis should instead be placed on the correlation between the relative rankings of the justices in both spaces, which is found here to be at approximately *p* = 74%.

Barring some nominal differences in relative positioning, the next question should be how they compare in terms of classification. Again, it should be expected that optimal classification performs better at correctly classifying legislators – and this expectation is fulfilled here with Supreme Court justices. Table 5 recreates the performance classification metrics from optimal classification from Table 4 and compares it directly to W-NOMINATE.

Table 5. Comparing the Optimal Classification to W-NOMINATE (Rehnquist Court)

Measure	Optimal Classification	W-NOMINATE	Comparison
Correct Yea	93.24	89.28	+3.96
Wrong Yea	6.76	10.72	-3.96
Correct Nay	14.27	17.04	-2.77
Wrong Nay	85.73	82.96	+2.77

V. Discussion

This research note aimed to accomplish two overarching goals. First, it aimed to offer a less-demanding introduction to spatial voting models through the optimal classification procedure and applying it to a subsection of political science beyond Poole's original intent of legislative roll call behaviors. Specifically, I offered a detailed description of the optimal classification procedure, including both its manual application and a computational demonstration of its power with larger populations of data using over 1,700 cases from the Rehnquist Court (1986-2005). The purpose of which was to illustrate to students experiencing their first introduction to spatial voting models that computational methods like optimal classification do not require advanced statistical knowledge to produce accurate rank orderings of voters. In some cases, all that is needed is a pencil and a sturdy eraser.

Second, I aimed to illustrate that the computational results obtained from optimal classification are not wholly inferior to more-advanced methods like NOMINATE. Using the Rehnquist Court, I was able to illustrate that optimal classification actually improved the classification procedure. Granted, the goal of NOMINATE is to maximize the likelihood function rather than the optimal ordering. Likewise, optimal classification does present a handful substantive downsides. For one, the rank ordering produced is ordinal rather than cardinal – meaning that we cannot derive numerical representations of the differences separating voters like we can with one of the statistical approaches. That being said, if the goal is to receive an introduction to spatial models of voting, or alternatively if the desire is simply to be able to derive which members of a legislature, court, or another group of voters compare to each other, optimal classification surely provides substantive benefits. It is my hope that this brief work removes some of the hesitance among emerging scholars to test the waters of spatial voting.

References

Bailey, M. A. (2007). Comparable preference estimates across time and institutions for the court, congress, and presidency. *American Journal of Political Science*, *51*(3), 433-448

BNSF Railway Co. v. Loos, 586 U.S. ___ (2019)

Bonica, A. (2019). Database on ideology, money in politics, and elections (DIME)

Brown v. Board of Education of Topeka, 347 U.S. 483 (1954)

Epstein, L., Martin, A. D., Segal, J. A., & Westerland, C. (2007). The judicial common space. *The Journal of Law, Economics, and Organization*, *23*(2), 303-325.

Giles, M.W., Hettinger, V.A., & Peppers, T.C. (2001). Picking Federal Judges: A Note on Policy and Partisan Selection Agendas. *Political Research Quarterly, 54*(September), 623-641.

Martin, A.D., & Quinn, K.M. (2002). Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis, 10*(2), 134-153.

Planned Parenthood of Southeastern Pennsylvania v. Casey, 505 U.S. 833 (1992)

Poole, K. T. (2005). *Spatial models of parliamentary voting*. Cambridge University Press.

Poole, K. T., & Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science*, 357-384.

—— (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press on Demand.

—— (2001). D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly*, 5-29.

Segal, J.A., & Cover, A. (1989). Ideological Values and the Votes of Supreme Court Justices. *American Political Science Review, 83*(June), 557-565.

Segal, J. A., & Spaeth, H. J. (1993). *The Supreme Court and the attitudinal model* (Vol. 65).

New York: Cambridge University Press.

—— (2002). *The Supreme Court and the Attitudinal Model Revisited*. Cambridge University Press.

Social Security Board v. Nierotko, 327 U.S. 358 (1946)

Spaeth, H.J., & Epstein, L., Martin, A.D., Segal, J.A., Ruger, T.J., & Benesh, S. (2020). 2020 Supreme Court Database, Version 2021. Release 01.
URL: http://Supremecourtdatabase.org

United States v. Quality Stores, Inc., 572 U.S. 141 (2014)