Article

An Evolutionary View of the U.S. Supreme Court

Noah Giansiracusa (ID)

Department of Mathematical Sciences, Bentley University, 175 Forest St, Waltham, MA 02452, USA;
ngiansiracusa@bentley.edu

Abstract: The voting patterns of the nine justices on the United States Supreme Court continue to fascinate and perplex observers of the Court. While it is commonly understood that the division of the justices into a liberal branch and a conservative branch inevitably drives many case outcomes, there are finer, less transparent divisions within these two main branches that have proven difficult to extract empirically. This study imports methods from evolutionary biology to help illuminate the intricate and often overlooked branching structure of the justices' voting behavior. Specifically, phylogenetic tree estimation based on voting disagreement rates is used to extend ideal point estimation to the non-Euclidean setting of hyperbolic metrics. After introducing this framework, comparing it to one- and two-dimensional multidimensional scaling, and arguing that it flexibly captures important higher-dimensional voting behavior, a handful of potential ways to apply this tool are presented. The emphasis throughout is on interpreting these judicial trees and extracting qualitative insights from them.

Keywords: legal studies; law; judicial; supreme court; voting; branching; tree; phylogenetic; genomic; evolution

1. Introduction

When you picture the interrelationships among the nine justices sitting on the U.S. Supreme Court, what do you see? For concreteness, consider the Court in the 2018–2019 term, after President Trump's appointees Gorsuch and Kavanaugh had been seated but before Ginsburg passed away. There is a liberal wing of Sotomayor, Ginsburg, Breyer, and Kagan, and a conservative wing of Roberts, Kavanaugh, Gorsuch, Alito, and Thomas. Within the conservative wing, Roberts and Kavanaugh lean toward the center-right and Roberts in particular sometimes plays the pivotal role of the swing justice, while Gorsuch, Alito, and Thomas are more deeply entrenched in conservative territory. This expresses our intuitive political placements of these justices, and one of the staples of empirical legal studies over the years has been the use of ideal point estimation to quantify this intuition by placing the judges at precise locations along a one-dimensional political axis:

There are a variety of ways to perform ideal point estimation, as I will briefly discuss in Section 2. The layout in Figure 1 was produced using multidimensional scaling (MDS), which was pioneered by Schubert in the 1960s and 70s [1,2] then further refined by, among others, Brazill and Gronfman in 2002 [3,4]; it is based on the voting agreement rates of the justices. In this figure, we see that Sotomayor and Ginsburg are closely aligned on the far left; a bit to their right, but still deeply in liberal territory, Breyer and Kagan are paired up; Roberts and Kavanaugh find themselves closely aligned as center-right justices; on the far right we have the trio of Gorsuch, Alito, and Thomas, with just a sliver of more daylight between Alito and Thomas than between Alito and Gorsuch.

However, are there more refined interrelationships that can be deduced about the justices directly from their voting agreement rates? This is the topic I will focus on in this paper, and to do so I will rely on a tool from evolutionary biology called phylogenetic tree estimation which, despite its name, is not inherently biological—it is a general statistical method for creating a certain kind of network structure from data. The main insight of

this paper is that phylogenetic tree estimation based on voting agreement rates provides an intriguing non-Euclidean analogue of ideal point estimation that, as I will argue, is more flexible and sensitive to deviations from unidimensional political voting than the prevailing higher-dimensional alternative: planar ideal point estimation based on MDS.

Figure 1. One-dimensional ideal point estimation for the 2018–2019 Court provided by multidimensional scaling (MDS).

As a preview of what is to come, here in Figure 2 is a phylogenetic tree constructed from the same voting agreement rate data as was used for MDS in Figure 1.
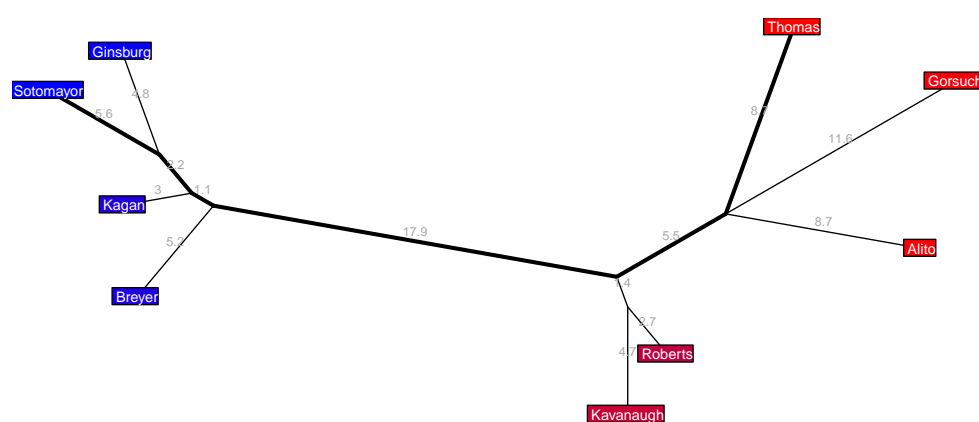
Figure 2. Phylogenetic tree estimation for the 2018–2019 Court, illustrating the main object that will be introduced and studied in this paper.

A detailed description of how this is made and what it means is given in Section 3; for now the main thing to know is that the distance between each pair of justices (approximating the percentage of cases in which they voted against each other) is the sum of edge lengths (indicated in light gray) along the unique tree-path between them, and the blue-to-red coloring reflects the one-dimensional MDS coordinates to help elucidate the relation between these two constructions.

Several simple but meaningful observations jump out here. On the left, while MDS sandwiches Breyer between Ginsburg and Kagan, in the tree we see that Breyer has split off from the liberal branch earlier than Kagan has. Consequently, there is a Kagan–Sotomayor–Ginsburg branch of the Court—a branch comprising all the female justices—that is hidden from the MDS view but immediately revealed by the tree. This is quite intriguing, as no demographic data was involved in the construction of this tree: the gender division seen here is evidently embedded in judicial voting patterns. On the right, we see that not only do Roberts and Kavanaugh lie between the four liberal justices and the three more staunchly conservative justices, but they actually have branched off together from the main horizontal path across the tree. This refines the view of them as both being center-right justices who sometimes play the role of the pivotal swing justice: it suggests that there is some ideological common ground uniting Roberts with Kavanaugh while distancing them from both political extremes of the Court. On the other hand, Kagan and Breyer also appear extremely close in the MDS picture and then some distance between them emerges in the tree—but here it happens in a different way: Kagan and Breyer each branch off separately from the main horizontal path, rather than doing so in tandem as Roberts and Kavanaugh do. Finally, on the far right, note that MDS places Gorsuch, Alito, and Thomas in very close proximity to each other, suggesting they are closely aligned and united by their conservative viewpoints—whereas the phylogenetic tree reveals that there are rather

substantial disagreement rates among them; this means they are conservative, but in three different ways, so to speak.

This captures an important higher-dimensionality to the justices' judicial reasoning. Unsurprisingly, there is much more to Supreme Court voting behavior than simply one-dimensional political alignment. Exploring the second dimension of ideal point estimation produced from MDS has been a recent topic of interest [5–8]. Two-dimensional MDS creates an axis orthogonal to the main political axis to accommodate voting behavior that deviates from the expected liberal-conservative spectrum, but it forces the same axis on all the justices. Trees, in contrast, provide an additional flexibility by allowing divisions into an unlimited number of additional locally determined directions. Trees reveal and quantify divisions (such as between Gorsuch, Alito, and Thomas on the right, and between the Sotomayor and Ginsburg on the left) without forcing these divisions to all take place in the same fixed, globally construed axis, and they reveal and quantify alliances (such as with Roberts and Kavanaugh) without forcing the common ground involved to be measured by the same global second axis.

In short, phylogenetic trees provide snapshots of the Court that refine the political pictures provided by one-dimensional ideal point estimation—and, as I hope to illustrate throughout this paper, the higher-dimensionality that trees reveal is quite distinct from that captured by two-dimensional ideal point estimation. In a sequel paper [9] I use these phylogenetic judicial trees, including their precise edge lengths, quantitatively to measure the divisions within the main two branches of the Court; in this first paper I stick to a more qualitative approach along the lines of the preceding discussion.

All the source code for this paper, written in the computer language R, is available at https://github.com/noahgian/Evolutionary-View-of-the-U.S.-Supreme-Court (accessed on 27 April 2021), and I will gladly assist anyone interested in using it and/or developing these methods further.

Outline

In Section 2, I provide a quick tour of the literature on ideal point estimation and mention some other papers that have used sophisticated mathematical/statistical tools to analyze vote patterns at the Supreme Court—in order to set the stage for, and draw some contrasts with, the approach taken in this paper.

In Section 3, I first provide some background on phylogenetic trees and then discuss what it means to estimate a tree from data and how this is done computationally. Next, I introduce the main judicial tree construction of this paper and explain how a one-dimensional political axis could be embedded in these trees. Then, I compare tree estimation to MDS ideal point estimation—both conceptually, by discussing political alignment and multidimensionality ng, and quantitatively, by comparing the residuals for these different methods of approximating voting disagreement rates with metric geometry.

In Section 4, I walk the reader through a handful of potential applications/illustrations of judicial trees. Several judicial trees, constructed from different sets of justices, are depicted and interesting features of these trees are highlighted and discussed. A tree-based alternative to the usual notion of a median justice is provided—one that I argue is more faithful to the idea of political centrality than MDS-based approaches. The tree for a group of justices is split into separate trees according issue area (civil rights, due process, etc.) to try to disentangle some of the factors that go into the tree's overall structure. An evolution of the 2010–2016 Roberts Court is presented by using judicial trees computed one case at a time to detect when important cases changed the "shape" of the Court. Finally, matrix completion is used to put all Supreme Court justices who have served since 1946 onto a single tree.

In Section 5, I conclude the paper by providing a brief discussion of how judicial tree estimation fits into the literature and a short list of possible topics for future investigations.

2. Related Work

The field of empirical legal studies aims to use data to draw insight into legal analysis, procedure, decision-making, etc. It is a large and growing field, with entire research journals devoted to it; the reader curious to get a sense of the field as a whole might consult some of the following sources: [10–14]. An important topic arising in empirical legal studies is ideal point estimation for the Supreme Court justices. Roughly speaking, this refers to any method of placing the justices in a Euclidean space $\mathbb{R}^d$ of some specified dimension $d \geq 1$ in a way that reflects the justices' attitudes and judicial behavior. Ideal point estimation also shows up in other areas of political science and social science, for instance it is commonly applied to legislative voting bodies, but I will focus on the Supreme Court justices throughout this paper.

Ideal point estimation yields visualizations of judicial ideology that help provide qualitative insights. For instance, one of my favorite data charts in any setting, which I find endlessly fascinating to inspect, is the following one tracing ideal points across the years for all justices since 1935: https://en.wikipedia.org/wiki/Martin-Quinn_score (accessed on 27 April 2021). However, ideal point estimation is also used quantitatively in a variety of empirical investigations—for instance, it has been used to identify the median justice [4,15,16] and to identify cases that yielded the most surprising voting patterns among the justices [17].

In many ways ideal point estimation, at least in the judicial setting, was launched with Schubert's 1965 book [1], but it mostly lay dormant for several decades until large legal datasets became available and computational methods for processing them were developed. There are now a handful of popular approaches to calculating ideal points, and dozens of papers on the topic. Schubert used multidimensional scaling (MDS), which I will return to in Section 3, and this has remained a popular choice when one only wishes to consider the votes of the justices. Segal and Cover developed a method that uses data external to the Court, such as newspaper editorials [18]. One of the most widely used methods (and the one that produced the captivating chart mentioned above) is the Bayesian approach of Martin and Quinn; this starts with prior assessments of each justice, based on data such as political party of the nominating president, votes in lower courts if the justice previously served as a judge, etc., then it uses Markov chain Monte Carlo methods to dynamically update the ideal points as the justices cast their votes to affirm or reverse each Supreme Court case that they hear [19]. A refinement of the Martin-Quinn Markov chain Monte Carlo Bayesian model was later introduced by Bailey; in addition to considering votes of affirm or reverse, it considers a range of data concerning concurring and dissenting opinions, references by the justices to earlier cases, and presidential and congressional positions on the Court cases [20].

Most of the work on creating and using ideal point estimates has, at least until recently, only been one-dimensional. This means each justice is placed along a single axis that is unequivocally interpreted as a liberal-conservative political spectrum. In 2007 the two-dimensional MDS coordinates for nine justices were depicted by Hook but there was no attempt to interpret the mysterious second dimension ([21], p. 254). Another ideal point estimation method was developed by Peress in 2009, and he said that interpreting the second dimension of their scores is a "difficult task", but that there is "some evidence to suggest that the second dimension captures 'judicial activism' ([22], p. 286). The first in-depth analysis of a second dimension for Supreme Court ideal point estimation came in a 2016 paper by Fischman and Jacobi and a sequel paper in 2019 by Fischman [5,6]. These two papers collected and analyzed cases for which the justices voted in an unusual pattern from a one-dimensional political perspective to see if the second dimension, as generated by MDS, could help make sense of the cases—and if, in turn, these cases could help make sense of the second MDS dimension. They discovered that the meaning of the second dimension seems to vary and depend on factors such as the issue area of the cases; a clear ideological interpretation was not always found, but in certain situations it seems to convey what they referred to as a "legalism versus pragmatism" axis. A pair of my own

papers then explored some ways of applying computational geometry to probe and extract insight from two-dimensional MDS ideal point estimates [7,8].

In a recent paper, I apply the phylogenetic tree perspective developed in the present paper to measure the distance between the two main branches of the Court, the cohesion within each of these two branches, and the overall multidimensionality of the Court; the evolution of these quantities over the past several decades is then studied [9]. As you will soon see, one can think of the judicial tree estimation put forth here as a variant of ideal point estimation: the justices are still placed in a geometric space in which the distances between them reflect voting disagreement rates, and hence some form of ideological opposition, but here the geometric space is no longer Euclidean—it is a tree, which is an example of a hyperbolic space.

Before turning to judicial tree estimation, there are some more papers worth mentioning that bear some relevance to the present paper and help flesh out the growing landscape of interesting mathematically-driven investigations into the voting behavior of the Supreme Court justices. In 2003, Sirovich employed singular value decomposition (a cousin of MDS) and information-theoretic entropy to deduce that the ostensibly nine-dimensional voting of the Supreme Court justices behaves more like it is four- or five-dimensional [23], but this interpretation was soon called into question by Edelman [24]. Edelman, in separate paper written collaboratively with Chen, used game theory to estimate the voting power of each justice [25].

Guimerà and Sales-Pardo used methods from network science to study the relationships between the justices by measuring how accurately each justice's votes can be predicted from the votes of the remaining justices in each case [26]. Sullivan-Georgakopoulos-Georgakopoulos introduced a measure of the "fluidity of judicial coalitions" for each natural court by looking at the diversity of justices forming the majority in 5-to-4 cases [27]; while not stated as such, I believe their notion of fluidity also quantifies the multidimensionality of each natural court in some sense, since purely one-dimensional political voting would yield only two possible 5-to-4 majorities. In fact, Brams-Camilo-Franz used voting patterns to estimate the likelihood that each collection of justices will form a coalition (the majority or minority in a case), and they noted that their estimates were much more compatible with the real data than the coalitions predicted by one-dimensional voting models [28]; their methods also find "smaller subcoalitions that form on their way to becoming" a coalition, which might be related to the sub-branches within the main liberal/conservative branches that are revealed in this paper. Lauderdale and Clark considered the multidimensionality of the Court, particularly in the context of median justices, by splitting cases along issue areas [29]; in essence, this creates a higher-dimensional Euclidean space in which to do ideal point estimation.

While the preceding papers draw primarily from mathematics and statistics, and the present paper inherits tools from computational biology, a fascinating trilogy of papers by Lee and collaborators takes inspiration from another source: physics. In the first paper [30], an important shift in perspective that in some sense extends from Sirovich's 2003 paper is taken: rather than focusing on the matrix of pairwise voting disagreement rates between the justices (as is done here and in the MDS-based papers, among others) one should use the matrix of pairwise correlations between the vote vectors—meaning the vector of 1 s and −1 s representing each justice's affirm or reject vote in each case (suitably symmetrized). Of course, this is only the beginning: the authors show that the simplest possible (in the sense of maximum entropy) model compatible with these pairwise correlations can be constructed from the Ising model of a magnet, and that this correlation-based model does an impressively good job of generating vote patterns that are similar to the real ones from the Court. Moreover, by studying this model as a proxy for the nine justices and drawing further inspiration from statistical physics and magnets in particular, the authors provide some remarkable new insight into certain phenomena, both theoretical and observed, pertaining to judicial voting behavior. One should keep in mind, however, that this is a model in the sense of physics—a simplification of a real-world phenomenon (the vote-

generating nine-justice Supreme Court, in this case) down to a mathematical object defined by functions and equations—rather than a model in the more human-centric intuitive sense, such as a picture of the justices and their relations that we can easily hold in our mind. Ideal point estimation, and the judicial tree estimation introduced in this paper, are (in my opinion) not so much about making predictions as they are about conceptualizing behavior in as simple a way as possible that still captures some of the intrinsic complexity inherent in the matter.

The next paper in the statistical physics Supreme Court trilogy by Lee et al. (this one solo-authored) [31] applies the same method of creating a maximum entropy model compatible with pairwise correlations among the vote vectors of the justices—but this the time the full set of 36 justices from 1946 to 2016 are used so many pairs of correlations are missing and the model fills them in. I will return to this paper in Section 4.5 when I use a different form of matrix completion to put all justices since 1946 together on a single judicial tree. The third paper in the trilogy [32] uses the maximum entropy model to investigate properties of a statistical replacement for the notion of the median voter.

3. Tree Estimation

In this section, I explain what phylogenetic trees are and how they are constructed from Supreme Court data.

3.1. Phylogenetic Trees

While the idea of a "tree of life" has ancient origins, Charles Darwin sketched one of the first evolutionary trees in 1837 and then in their seminal 1859 treatise The Origin of Species, he included a more detailed illustration; see Figure 3.
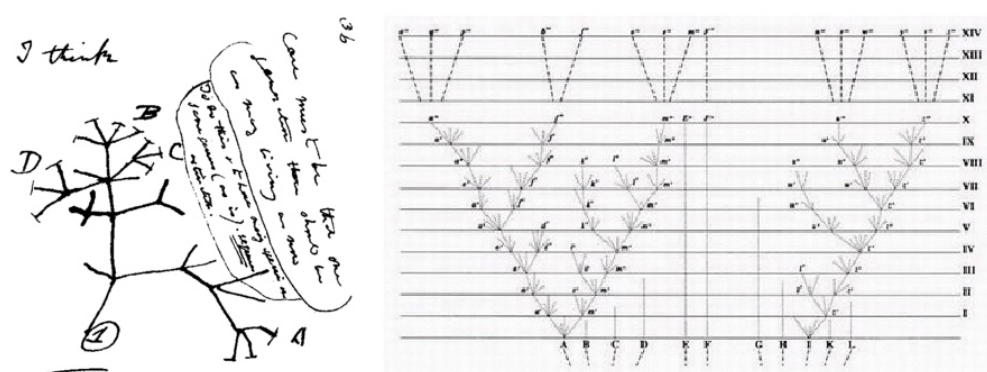
Figure 3. Evolutionary trees (precursors to the modern phylogenetic tree) drawn by Charles Darwin in 1837 (**left**) and 1859 (**right**).

This laid the foundations for the notion of a phylogenetic tree that rose to prominence in evolutionary biology throughout the 20th century. (Here is one indication of this prominence: in 2014, *Nature* published a list of the top 100 most-cited scientific papers of all time, and at #20 was Saitou and Nei's 1987 paper on the Neighbor-Joining method [33] for phylogenetic tree estimation that I will briefly mention in Section 3.2, and at #41 was Felsenstein's 1985 paper [34] showing how to use bootstrapping to place confidence intervals on phylogenetic tree estimates.) Despite this biological origin and nomenclature, a phylogenetic tree really is just a mathematical object: it is an edge-weighted tree with labeled leaves. More precisely:

Definition 1. *An unrooted phylogenetic tree on m taxa is a graph (i.e., collection of edges and nodes) that is a tree (i.e., has no loops) with m labeled leaves (called the "taxa") and a non-negative length assigned to each edge.*

(For us, the taxa will be the Supreme Court justices.) This data equips the taxa with the structure of a finite metric space; simply put, it yields a notion of distance between any pair of taxa. Indeed, in a tree there is a unique path between any pair of leaves, and the length of this path (defined as the sum of edge lengths in the path) is the distance between the two leaves. This metric is hyperbolic.

3.2. Distance-Based Tree Estimation

In the biological setting, the taxa are often different species and phylogenetic tree estimation is the task of creating a tree on the taxa such that the distances (as measured in the tree) between the pairs of taxa are as close as possible to the genomic distances between the species. Remarkably, this process makes sense for any objects, not just species, and one need not make any assumptions on the original notion of distance between the objects that the tree aims to approximate. Let me state this more precisely in mathematical terms.

Definition 2. *Given an integer $m \geq 2$ and real numbers $d_{ij} \geq 0$ for $1 \leq i, j \leq m$, phylogenetic tree estimation is the optimization problem of finding an unrooted phylogenetic tree on m taxa such that the leaf-to-leaf path lengths t_{ij} minimize the sum of squares $\sum(d_{ij} - t_{ij})^2$.*

We usually assume that the original distances form a premetric: $d_{ij} = d_{ji}$ and $d_{ii} = 0$. However, crucially, we do not assume that the triangle inequality holds—and in practice it will not. Phylogenetic tree estimation can be thought of as a particular way of approximating a finite premetric with a (hyperbolic) metric.

Technically, what I have described here is known as distance-based tree estimation. There are also maximum parsimony and maximum likelihood approaches, but these, I believe, make more sense in the original biological setting than they do in general. Even for distance-based tree estimation there are a variety of algorithms (among the most popular are neighbor joining [33] and minimum evolution [35]), and there is a considerable literature investigating them. However, this is because in biological settings phylogenetic trees tend to be quite large so a direct ordinary least square minimization approach is computationally infeasible. Indeed, this is a linear least squares problem within each tree topology (that is, if we fix the graph but search through the different possible edge lengths)—but the challenge is that the number of tree topologies is exponential in the number of leaves. Thus, the popular algorithms all take a heuristic approach by replacing the least squares optimization objective with something more computationally attainable. In the judicial setting of this paper, however, I will mostly be dealing with only nine taxa at a time (since there are nine justices on the bench of the Supreme Court) so these heuristic algorithms are not needed—it is straightforward to directly compute ordinary least squares minimizing trees.

3.3. Judicial Trees

In order to estimate a phylogenetic tree for a collection of justices, we need a premetric on the justices conveying some notion of ideological distance between them. This is also what is needed to do MDS ideal point estimation, so I follow the same overall approach as the scholars cited in Section 2 who have used MDS to study the Supreme Court. The basic idea is to use disagreement rates among the justices based on their affirm/reverse votes in all the cases they hear. Let me now make this more precise.

The first step is collecting data, and fortunately this has already been done for us: the Supreme Court Database (SCDB) is a widely used dataset in empirical legal studies that contains over two hundred pieces of information about each case decided by the Court [36].

Definition 3. *For any two justices, call them i and j, let their agreement rate a_{ij} be the number of cases in which the justices either both voted in the majority or both voted in the minority (as indicated by the binary* `majority` *variable in the SCDB) divided by the total number of cases in which both justices cast votes. Let $d_{ij} := 1 - s_{ij}$ denote the disagreement rate between the two justices.*

It is important to note that the `majority` variable only considers the disposition of each justice's vote (affirm or reverse); it does not consider whether they wrote a separate concurring or dissenting opinion, nor does it attempt the more subjective task of classifying the political orientation of the decision.

In legal parlance, a natural court is any maximal span of time in which no justices either joined or departed the Court. A fully staffed natural court has nine justices, and the disagreement rates d_{ij} among them clearly form a premetric. The judicial tree I construct for this collection of justices is the ordinary least squares minimizer for this premetric. (In an earlier version of this paper, to compute the (dis)agreement rates I used only the cases heard during the natural court in question. One of the anonymous referees suggested instead using all cases voted on by each pair of justices, as this provides a richer proxy for their judicial compatibility. I now fully agree: to measure how frequently two justices disagreed with each other, we should consider all cases in which they either agreed or disagreed with each other—regardless of who else sat on the bench at the time).

Now a few words about visualizing these judicial trees. As with the plot of any graph, the precise locations of the nodes are arbitrary and immaterial—all that matters when drawing a phylogenetic tree is that the physical lengths of the edges are proportional to the mathematical edge lengths. In other words, you should not read anything into the angles you see in the pictures. I will stick to a convention of printing edge lengths in light gray and as percentage points (rounded to the nearest tenth). For instance, in Figure 2 we see that the path-distance from Ginsburg to Sotomayor is $4.8 + 5.6 = 10.4$, which means they vote against each other in roughly 10.4% of cases (I say "roughly" here because the leaf-to-leaf tree path lengths are only approximations to the actual disagreement rates), while the path from Ginsburg to Thomas has length $4.8 + 2.2 + 1.1 + 17.9 + 5.5 + 8.7 = 40.2$. Another convention I use throughout the paper is to color the leaf labels according to the one-dimensional MDS coordinates—meaning that in addition to using the disagreement rates d_{ij} for tree estimation, I also use them for one-dimensional MDS ideal point estimation and then map this axis onto a blue-red color palette. For the psychological convenience of the reader, I always use the blue end of the palette for what is obviously the liberal direction in MDS (that this MDS axis has a political interpretation is discussed shortly) and I rotate the tree so that the liberal justices are on the left and the conservative justices are on the right.

There is one further visual decoration I add to the tree plots in this paper, which is a bolding of a certain special path in the tree. This involves a concept that I have not seen in the literature, and which I do not think is particularly relevant for phylogenetic trees in general, but for judicial trees I believe it is relevant and meaningful. Let me formally introduce this now.

Definition 4. *The major axis in a phylogenetic tree estimated from Supreme Court voting disagreement rates is the path between the two justices whose voting disagreement rate is maximal.*

In other words, this is the sequence of edges connecting the two justices who voted against each other most often (and who are, in turn, typically at the extreme ends of the one-dimensional MDS configuration). The major axis is frequently, though not always, equal to the longest path in the tree; the only reason these two paths sometimes differ is because leaf-to-leaf tree path lengths are only approximations to the original voting disagreement rates. Throughout this paper I plot the major axis in bold.

You can think of the major axis as something like a one-dimensional political axis that runs through the tree: the closer justices are to the major axis, the more their voting behavior is driven by one-dimensional politics, and the location where each justice's edge or branch joins the major axis gives a rough sense of their position on the liberal-conservative spectrum. Here is why. If it were possible to place the justices along a line in such a way that their pairwise distances perfectly match their voting disagreement rates, then this linear arrangement would clearly be the one-dimensional MDS ideal point estimate. Simultaneously, the judicial tree constructed in this situation would be a tree equal to its

own major axis that exactly matches the one-dimensional configuration, since this tree would be the unique one whose residuals $d_{ij} - t_{ij}$ are all zero. However, it has already been established empirically in the above-cited works on one-dimensional MDS ideal point estimation that the single line which most closely captures voting disagreement rates is the liberal-conservative political axis. In other words, if judicial voting rates were perfectly matched by one-dimensional politics then the tree would simply be the major axis (representing the liberal-conservative axis), so branches extending away from the major axis should generally be interpreted as voting behavior that deviates from purely one-dimensional politics. With this in mind, we can even coordinatize the major axis and the justices' locations along it.

Definition 5. *Identify the major axis of a judicial tree with the line segment $[0, a] \subset \mathbb{R}$, where a is the total length of the major axis, by placing the liberal end of the axis at 0. The political position of each justice is the point on this line segment corresponding to the node where the path from the justice first meets the major axis.*

For example, in Figure 2 Sotomayor has political position 0, then Ginsburg is at 5.6, Kagan is at 7.8, Breyer is at 8.9, Kavanaugh and Roberts are both at 26.8, Alito and Gorsuch are both at 32.3, and Thomas is at 41. Now, not too much import should be placed on these values at either end of the interval, because a small change in disagreement rates could, for instance, switch Gorsuch and Thomas as the conservative end of the major axis, but aside from the justices at the ends of the Court I believe these values are meaningful in terms of providing an order and relative distribution of the justices on the liberal-conservative spectrum.

3.4. Comparing to MDS

Multidimensional scaling [37,38] takes a finite premetric $d_{ij}, 1 \leq i, j \leq m$, and produces a configuration of m points in $\mathbb{R}^d$, where the dimension $d \geq 1$ is specified by the user, such that the Euclidean distances between the pairs of points approximate the premetric values as closely as possible. There are different flavors of MDS that come from optimizing different objective functions (that is, different ways of measuring the "as closely as possible" part of the process). A particularly popular form is called classical MDS, and this appears to be what is typically used in legal and other political science settings—so in what follows I will refer to it simply as MDS without further stipulation. One important property of this MDS is that it builds up the dimensions one at a time: the first coordinates of the m points are found first, then they are held constant while the second coordinates are found, etc. Put another way: MDS is compatible with the coordinate linear projection $\mathbb{R}^d \to \mathbb{R}^{d'}$ for any $d' < d$. In particular, the x-axis in the recent investigations of two-dimensional MDS ideal point estimation for the Supreme Court has the same interpretation (and even the same values) as the single axis in the more traditional one-dimensional MDS ideal point estimation—and, with remarkable consistency, this axis seems to capture politics in the sense of a liberal-to-conservative spectrum [5], pp. 1693–1694.

Since both MDS and phylogenetic tree estimation provide a metric approximation to the premetric of judicial disagreement rates, it is helpful to compare them. Let $d_{ij}^{(1)}, d_{ij}^{(2)}$, and d_{ij}^{tree} denote the approximations to the disagreement rates d_{ij} provided by one-dimensional MDS, two-dimensional MDS, and tree estimation, respectively. The residuals for one-dimensional MDS are the differences $d_{ij} - d_{ij}^{(1)}$, and similarly for the other methods. In Figure 4, we see the histograms of the absolute values of the residuals for one-dimensional MDS (green), two-dimensional MDS (red), and tree estimation (blue) computed for the same set of justices used in Figure 2 (which is the most recent fully staffed natural court in the latest release of the SCDB).
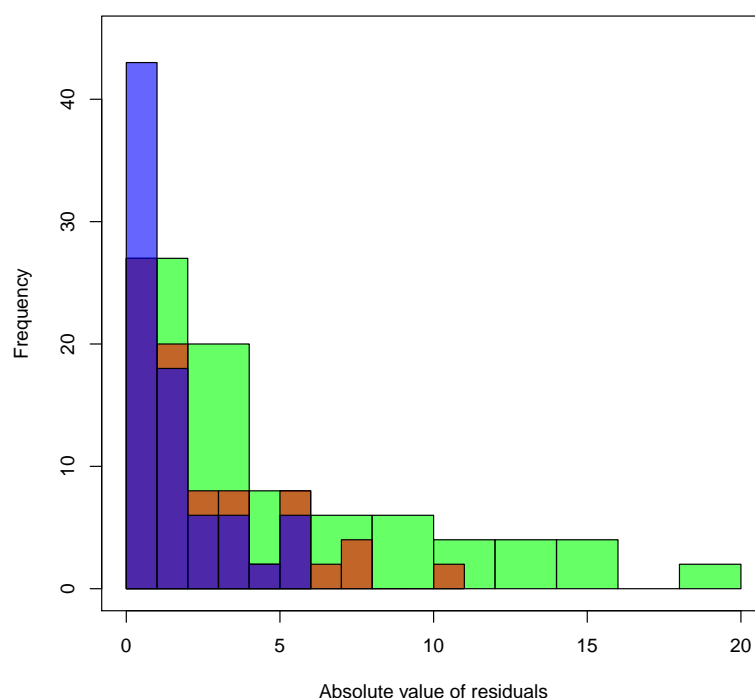
Figure 4. Histograms of the absolute values of the residuals for one-dimensional MDS (green), two-dimensional MDS (red), and tree estimation (blue) for the justices on the 2018–2019 Court.

Unsurprisingly two-dimensional MDS has smaller residuals than one-dimensional MDS, but intriguingly we see that tree estimation has even smaller residuals here. The maximum absolute residual for one-dimensional MDS is 19.3%, for two-dimensional MDS it is 10.6%, and for tree estimation it is only 5.7%. Rounding to the nearest integer, the sum of squares of the residuals for one-dimensional MDS is 4038, for two-dimensional MDS it is 997, and for tree estimation it is 399. Thus, in terms of residuals, we see that there is a sharp improvement from one-dimensional MDS to two-dimensional MDS, then a further (more modest but still substantial) improvement from two-dimensional MDS to trees. However, is this behavior specific to this particular group of justices? To answer this, let us turn to Figure 5 which shows the sum of the squares of the residuals and the maximum absolute residual for all three methods for all the natural courts included in the modern release of the SCDB.

That tree estimation produces more accurate approximations than one-dimensional MDS is automatic and obvious: every one-dimensional configuration can be viewed as a tree in which all the non-leaf nodes have degree two (meaning the tree is just a sequence of edges connected end-to-end with the justice labels placed along it). What is interesting is that trees have smaller sum of residual squares than two-dimensional MDS for 77% of the natural courts considered here and smaller maximum residuals in 92% of the natural courts. At least intuitively, I believe this is because trees allow for branching into an unlimited number of unspecified directions (such is the nature of hyperbolic space), whereas two-dimensional MDS constrains all movement in non-political directions to take place in a single additional dimension. (By "non-political" directions I really mean directions other than the primary liberal-to-conservative axis. For simplicity I often call that axis the political axis, but nobody doubts that political ideology is more than one-dimensional, so it is a slight abuse of speech—that I will nonetheless make throughout this paper—to refer to the liberal-to-conservative axis as the political axis.) However, accuracy is not the only goal here—interpretability is of utmost importance, and the rest of this paper is largely devoted to exploring how to interpret judicial trees and some of the insights that can be extracted from them.
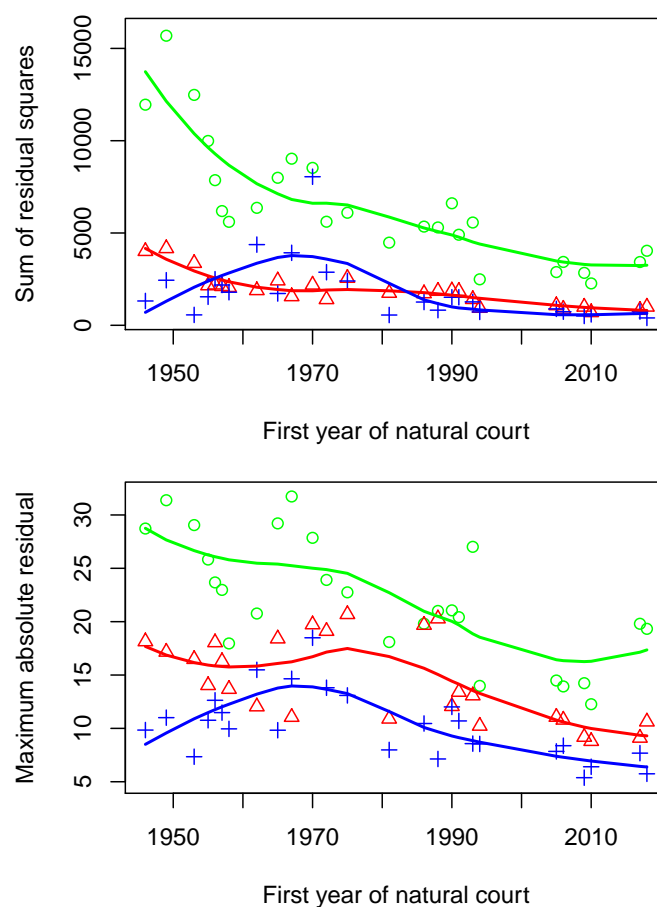
Figure 5. Sum of squares of residuals (**top**) and maximum absolute value of residuals (**bottom**) across all natural courts in the modern SCDB (plotted at the first year of each natural court), together with a smooth trend line provided by local polynomial regression, for one-dimensional MDS (green), two-dimensional MDS (red), and tree estimation (blue).

With that in mind, here is another perspective on the matter. While justices often vote along political lines, we know that there are many cases where they vote in ways that are incompatible with a purely one-dimensional political model [5,17]. This causes one-dimensional ideal point estimates that are based on voting disagreement rates to pick up some of this higher-dimensional behavior. Consequently, such ideal point estimates do not faithfully reflect the liberal-conservative positions of the justices. Two-dimensional MDS is able reduce the residuals compared to one-dimensional MDS by adjusting the vertical placement of the justices, but it does not alter the horizontal placement of the justices—so the x-axis is still not purely political in the liberal-conservative sense, it carries precisely the same distortions from higher-dimensional voting as one-dimensional MDS. In sum, the x-axis is primarily political as scholars have observed, but not entirely so—it is distorted by the lack of unidimensionality ng disagreement rates. I propose the tree-based political position (in the technical sense defined above) as an alternative way to quantify the liberal-conservative spectrum. I cannot prove that it is more accurate than the x-axis in MDS because there is no underlying truth to compare to, but philosophically speaking I believe it is reasonable because trees allow the major axis to capture one dimension while all the branches extending off the major axis independently capture all the higher-dimensional behavior—as opposed to MDS where the dimensions are more entangled due to the compatibility with coordinate linear projection mentioned earlier.

4. Applications

The following subsections explore and apply this judicial tree construction in several different ways. First, the tree associated to a handful of natural courts in the modern SCDB is depicted and discussed; I try to highlight the interesting features of these trees and use this as an opportunity to illustrate how I believe judicial trees can be interpreted (Section 4.1). Next, I provide a tree-based alternative to the usual notion of a median justice based on MDS coordinates—one that I argue is more faithful to the idea of political centrality (Section 4.2). I then split the judicial tree for a natural court into a forest, one tree for each issue area (civil rights, due process, etc.), to try to disentangle some of the factors that go into the tree's overall shape (Section 4.3). By computing trees for a natural court one case at a time, a dynamic evolution of the Court can be provided; this idea is illustrated by walking the reader through some of the key cases that shaped the 2010-2016 Roberts 4 Court (Section 4.4). Finally, matrix completion is used to construct a massive judicial tree with every Supreme Court justice who has served since 1946 (Section 4.5).

4.1. Snapshots of the Court

Here I look at, and comment on, a few particularly intriguing judicial trees to show the diversity of configurations that have occurred over the years and to practice interpreting them. In each case the method described in Section 3 is used: the `majority` variable in the SCDB is used to calculate voting disagreement rates for all pairs of justices in a given natural court (based on all cases heard jointly by each pair of justices, not just the cases during the given natural court) then the ordinary least squares minimizing phylogenetic tree is computed and plotted with the justices colored by the one-dimensional MDS coordinates coming from the same disagreement rates, and with the major axis (the path between the most frequently disagreeing justices) in bold. The most recent natural court in the SCDB (know as the Roberts 7 Court) is depicted in Figure 2, so we'll proceed reverse chronologically from there.

The long 2010–2016 Roberts 4 Court in Figure 6 has a clear liberal wing with Sotomayor paired up with Ginsburg and Kagan paired up with Breyer (each pair only voting against each other about 10% of the time); a conservative wing with Thomas, Alito, and Scalia (with approximate disagreement rates among them ranging from 13.6% to 16.4%); Kennedy, a notorious swing vote during this era, sits near the center, veering ever so slightly to the right, with only a small deviation from the major axis, and a little further along to their right sits Roberts, who also only extends from the major axis by a small amount and who also is known for some important swing votes during these years—most notably, perhaps, when he surprised many pundits by siding with the liberals in the 5-to-4 case *National Federation of Independent Businesses v. Sebelius* (2012) upholding the individual mandate of Obama's Affordable Care Act on the grounds of the Constitution's Taxing and Spending Clause.
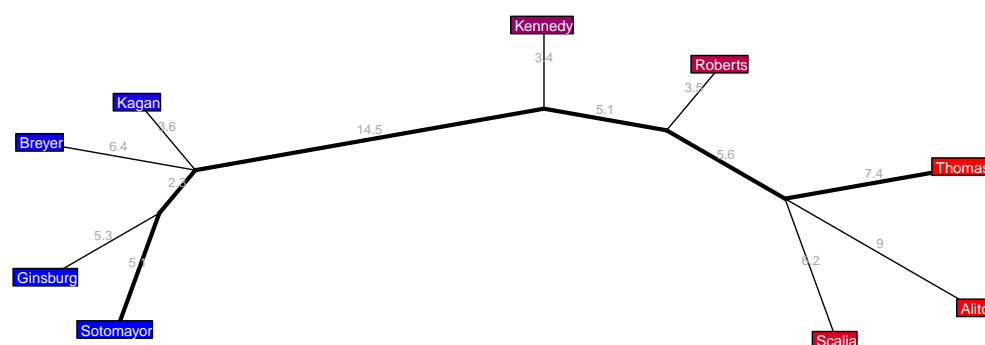

Figure 6. Judicial tree for the 2010–2016 Roberts 4 Court.

The 1993–1994 Rehnquist 6 Court in Figure 7 has a clear conservative coalition of five justices and a liberal coalition of four justices. There are two things that I find striking here. First, O'Connor, an influential swing voter [39], extends from the major axis at the

exact same center-right location as the swing voting Kennedy—yet there is a substantial distance between them (14.9%) and no common edge connecting them to the major axis as there was with Roberts and Kavanaugh in Figure 2. This suggests that these two swing-voting justices are united in their political orientation but that their frequent disagreements with the rest of the conservative coalition tend to occur in distinct directions from each other, meaning they do not have much ideological common ground aside from their center-right political position. Second, the four-justice liberal coalition is remarkably spread out: these justices are all distant from the conservative justices, but additionally there are considerable distances between each pair of liberal justices—much more so than between the conservatives here and much more so than between the liberals in the two previous trees we have looked (Figures 2 and 6). The liberal justices here seem to be frequently disagreeing with each other in ways more complicated than just degree of political extremity—suggested in the tree by the relatively large distances between them in contrast to the more modest separation between their political positions along the major axis.
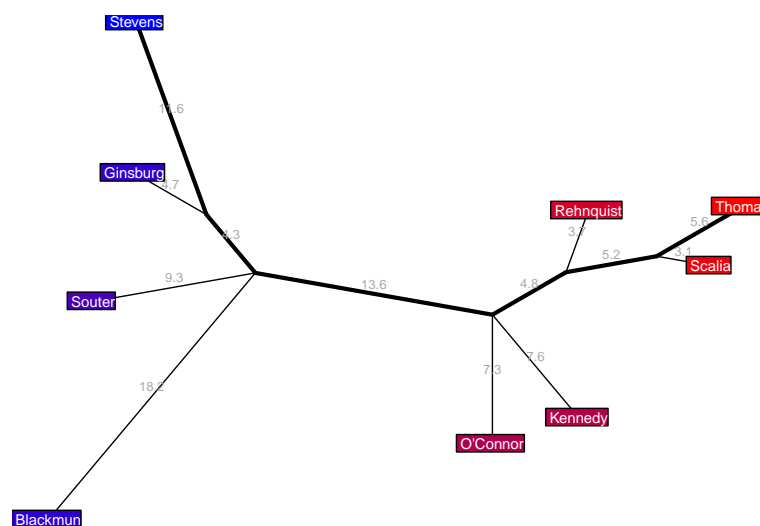
**Figure 7.** Judicial tree for the 1993–1994 Rehnquist 6 Court.

The 1975–1981 Burger 6 Court in Figure 8 is quite remarkable: the Court has essentially fractured into a tripolar configuration. Marshall and Brennan are closely united on the liberal side, and joining them from a considerable distance at a center-left political position is Stevens—who not only is a sizable distance of 11.9% to the right of them, but he also extends a sizable 10.5% from the major axis, indicating disagreements with the liberals that occur in both the standard political direction and an additional unspecified direction. On the right we have a fairly standard configuration of three conservative justices, but quite unusually in the center of the Court we have a trio of justices—White, Blackmun, and Stewart—who all join the major axis at the same exact location yet the approximate disagreement rates between them are rather large, ranging from 24.7% to 25.6%. I think it would be fascinating to look further into the ideology and judicial behavior of these three centrist justices to see what sets them apart from the rest of the Court and from each other, but a deep dive into the cases they have voted on is beyond my expertise.
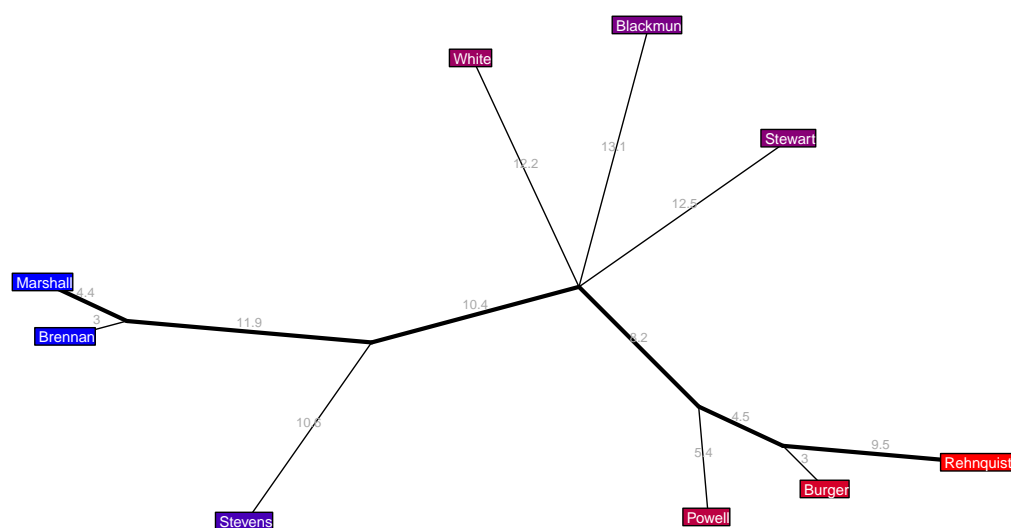
Figure 8. Judicial tree for the 1975–1981 Burger 6 Court.

I am including the 1967–1969 Warren 10 Court in Figure 9 as one of our snapshots mostly because it is so bizarre. The three conservative justices are easy to recognize, but beyond this things are quite scrambled. Douglas is the most liberal justice here (and indeed he was once called "the most doctrinaire and committed civil libertarian ever to sit on the court" [40]), but one step to their right along the major axis is a node sprouting out five center-left justices. Two of these, Warren and Black, are each on their own edge leading directly to the major axis (and Black's length 16.4% edge is sizable!), while the other three justices of this center-left quintet—Marshall, Brennan, and Fortas—form a curious trio who are all very close to each and separated from the major axis by a sizable length 8.8% edge. The interpretation here is that these five center-left justices all have the same political position and their frequent disagreements occur in three different directions orthogonal to this liberal-conservative axis: one direction for Warren, a separate direction for Black, and a further separate direction that is common to all of Marshall, Brennan, and Fortas. It is difficult to figure out what these directions could be, but if nothing else this is an example where there is clearly much more to the story than one-dimensional ideal point estimation can convey.
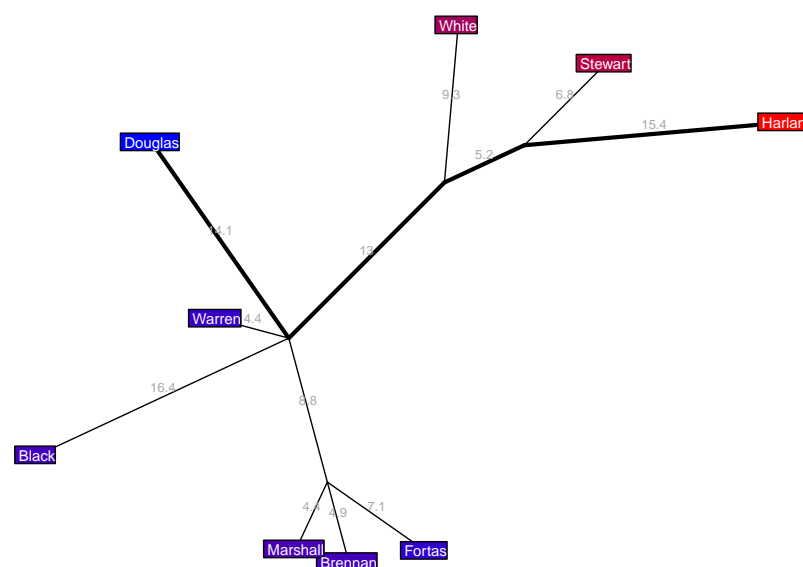
Figure 9. Judicial tree for the 1967–1969 Warren 10 Court.

In the 1967–1969 Warren 10 Court in Figure 10 we have a clear liberal coalition of three justices, a clear far-right pair of Frankfurter and Harlan, a centrist Clark, but then the center-right trio of Reed, Minton, and Burton is intriguing: all three of these justices have the same political position, but Reed and Minton are further united by a length 3.6% edge sprouting from the major axis. This suggests that these three justices all have similar center-right political orientations but that there is some ideological common ground uniting Reed and Minton that is not shared by Burton.
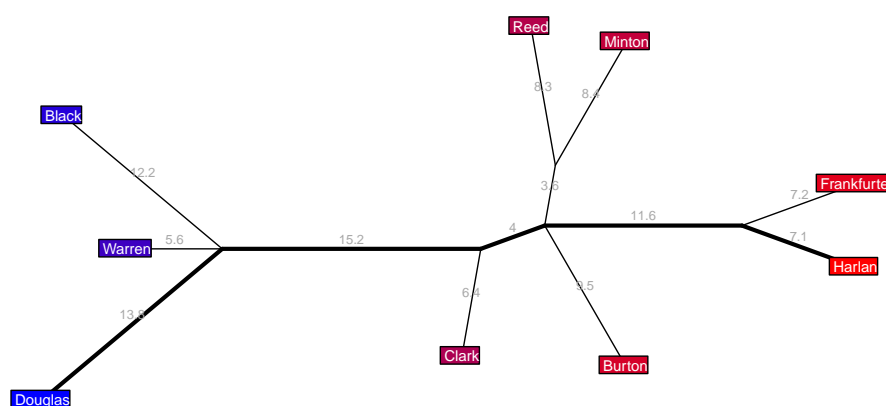

Figure 10. Judicial tree for the 1955–1956 Warren 3 Court.

Finally, going all the way back to the 1946–1949 Vinson 1 Court, we see in Figure 11 a very strongly bipolar layout more typical of recent times. There is a fairly cohesive four-justice liberal coalition and a less cohesive five-justice conservative coalition and a sizable gap between the two coalitions.
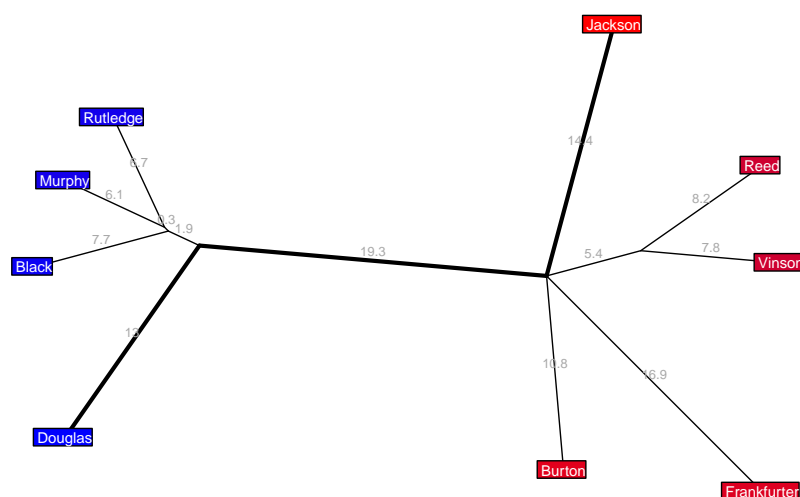

Figure 11. Judicial tree for the 1946–1949 Vinson 1 Court.

4.2. Central Justices

The related though meaningfully distinct notions of the "median justice," the "power" of a justice, and a "swing vote" have interested scholars for many years. The field of empirical legal studies has produced a range of data-driven methods for quantifying and studying these concepts [4,5,15,16,25,32,39,41,42]. The phylogenetic tree perspective of the Court developed in this paper yields a new network-theoretic way of describing which justices are most central to the Court. One particularly novel aspect of this approach is that it looks not just at whether there is an individual justice sitting at an influential central position but also whether there is a pair or even a trio of justices who collectively play the central role. It also attempts to extract political centrality from voting agreement

rates, which as we have discussed earlier is not quite the same as measuring centrality directly from the one-dimensional MDS ideal point estimates. Let me start with the formal definition.

Definition 6. *In the tree estimated for a given nine-justice natural court, if there is a justice j such that removing the node connecting j to the major axis results in two connected components not containing j, each with four justices, then j is central in this natural court. If there is a pair of justices, j_1, j_2, that are connected to the major axis at the same node and removing this node results in two connected components not containing them, one with four justices and one with three justices, then the pair j_1, j_2 is central. If there is a triple of justices j_1, j_2, j_3 connected to the major axis all at the same node and removing this node results in two connected components not containing these justices, each with three justices, then the triple j_1, j_2, j_3 is central.*

This is a lot to process, but it's really just capturing whether there are justices who sit as the median of the Court as measured by political position along the major axis. Indeed, the definition can be rephrased as follows: j is central if four justices have political positions to the left and four justices have political positions to the right; j_1, j_2 are central if their political positions coincide and there are three justices with political positions on one side and four on the other; j_1, j_2, j_3 are central if their political positions coincide and there are three justices on either side of them in terms of political position. You might recall that I cautioned the reader from putting too much import on political position near the ends of the spectrum; fortunately this does not impact this notion of centrality which really only depends on the overall balance of justices on either side rather than the detailed order and position of the justices at the ends.

The usual definition of a median justice is that their one-dimensional ideal point estimate has four justices on either side, so this tree-based definition is replacing one-dimensional ideal point estimates with political position along the major axis (fitting with the theme of this paper that the major axis might better capture the liberal-conservative spectrum than the *x*-coordinate in MDS) while simultaneously allowing for "ties" because the nature of trees means that justices often share political positions with each other and express their disagreements in other directions captured by branches extending from the major axis. Scholars have defined a "super median" justice by measuring the gap in one-dimensional ideal point estimates between the median justice and the nearest neighbors [16]—a justice who sits in the middle of the Court with a large gap from them to the liberal and conservative coalitions tends to play an outsize role as the pivotal swing voter—and if desired one could immediately adapt this notion to tree-based centrality as well. Let us take a look at some examples by inspecting the judicial trees pictured so far.

- In Figure 2, Roberts and Kavanaugh form a central pair. I think of this as follows. If the court behaved in a purely one-dimensional political way, and if the major axis does reflect this political axis, then depending on the political focal point of the case we'd expect Roberts and Kavanaugh to either both side with the liberals or both side with the conservatives, and either way they are guaranteed to supply the majority as they do so. In other words, politically they should swing together. Of course, the court does not behave in a purely one-dimensional way, but in more than one dimension there is not really a notion of median justice (for instance, in [5] medians are computed separately for the two axes), so my tree-based definition here is attempting to capture and convey how swing voters would behave if everything were driven by one-dimensional politics.
- Figure 6 is simpler: Kennedy is the lone central justice. The strength of their position here, in the sense of super medians, is determined by their nearest neighbor, namely Roberts. This makes sense: if Roberts were more firmly allied with the other three conservative justices then Kennedy could freely swing the majority in political cases as he pleases, but Robert's proximity to Kennedy means that there will still be cases where Roberts swings over to the liberal side (as we saw with the aforementioned case on the

Affordable Care Act) and when this happens the liberals have a majority of five votes already so Kennedy is unable to swing the majority back over to the conservatives.

- In Figure 7, we have Kennedy and O'Connor as a central pair, which fits a common perception of them both as influential swing-voting justices [39].
- The curious tripolar judicial tree in Figure 8 has White, Blackmun, and Stewart as a central trio—and a quite strong one due to the large gaps between their political position and that of the other justices on either side of them.
- The scrambled tree in Figure 9 has no central justice(s)—and indeed even by inspecting this tree manually it is difficult to tell where the balance of power lies in it. Note that by contrast one-dimensional ideal point estimation forces a median justice to always exist, which in my opinion does not make much sense here given how multidimensional and bizarrely balanced this Court is.
- Figure 10 is far less scrambled and more straightforward than the previous one, yet it too has no central justice(s). Clark is not central because siding with the liberals is not enough to give them the majority, and the Reed-Minton-Burton trio is not central because they have two justices to their right and four to their left rather than three and three. I must admit that my definition of central here seems a bit questionable because this trio could grant the majority to either side by swinging all together (since two plus three reaches the crucial threshold of five) so perhaps one might want to enlarge the definition of a central trio to include these 4-3-2 configurations, but I will resist doing so because that just does not seem central enough to merit the name.
- Figure 11 is actually the same tree topology as the previous figure (the only difference is the edge lengths), so here too there is no central justice(s).

The central justices for all the nine-justice natural courts in the modern SCDB are depicted below in Table 1.

It is curious that all the natural courts after 1972 have at least one central justice, whereas only four of the prior eleven do. In general, the first half of these natural courts seem to exhibit more unusual balances of power and chaotic looking configurations whereas throughout the second half of these natural courts the structure seems to settle down more and more. This was seen in some of the snapshots from Section 4.1, and it's also seen in Figure 5 where for instance all three methods (one-dimensional MDS, two-dimensional MDS, and tree estimation) have maximal absolute residues that steadily trend downward from the 1970s onward. This settling of the Court over the past fifty years is also one of the main topics I look into quantifying in the sequel paper [9].

The names in this table of central justices have a lot of overlap with the other empirical methods suggested in the literature for measuring median and swing justices (it is interesting to compare in particular with (Table 4 in [15]) where there is considerable though not complete agreement with the year-by-year analysis provided by Martin-Quinn scores). There is no way to say which method is "best" because there is no underlying truth to compare to; instead, each method gives a slightly different perspective on the issue. That said, it is reassuring that the different methods do not give wildly varying results.

4.3. Issue Areas

The SCDB includes a variable indicating the "issue area" of each case. The values of this variable are somewhat subjective and far from unequivocal—often there are multiple issues involved in a case and it is not clear how to choose just one of them, and sometimes the main issue in a case does not neatly align with one of the 14 possible values of this variable. Nonetheless, in the aggregate this provides a useful way to divide the cases. For a given natural court, one can produce a separate judicial tree for each different issue area, although for some issue areas there will not be enough cases to have rich disagreement rates among the justices, especially between the justices who did not overlap on the Court for very long. Let us take a look (see Figures 12 and 13) at some of the interesting trees in the issue area forest for the most recent natural court in the SCDB, the 2018–2020 Roberts 7 Court pictured in Figure 2.

Table 1. The tree-based central justice(s) for each fully staffed natural court in the modern SCDB.

Natural Court	Central Justice(s)
1946–1949	
1949–1953	
1953–1954	
1955–1956	
1956–1957	Clark
1957–1958	Clark
1958–1962	Clark
1962–1965	
1965–1967	Black
1967–1969	
1970–1971	
1972–1975	Blackmun–Stewart–White
1975–1981	Blackmun–Stewart–White
1981–1986	White
1986–1987	Powell–White
1988–1990	White
1990–1991	White
1991–1993	Kennedy–O'Connor
1993–1994	Kennedy–O'Connor
1994–2005	Kennedy–O'Connor
2005–2006	Kennedy–O'Connor
2006–2009	Kennedy
2009–2010	Kennedy
2010–2016	Kennedy
2017–2018	Kennedy
2018–2020	Kavanugh–Roberts

Criminal procedure cases look fairly similar to the full judicial tree in Figure 2 except that one prominent difference is Gorsuch has moved from the Court's far right to its center, where he extends a considerable distance from the major axis. Furthermore, Kavanaugh has moved from their center-right position with Roberts and now sits at the center with Gorsuch in terms of political position, though he shows no common ground with Gorsuch beyond this. Turning to Civil Rights cases, the liberals have not changed much but the conservatives have some intriguing positions: here Kavanaugh is a far-right justice whereas Gorsuch is centrist, and curiously Roberts is pretty far to the right in terms of political position but as their purple shading indicates he is still center-right in the one-dimensional MDS estimate.

What I find most striking about the First Amendment cases is that there is very little extension off the major axis anywhere in the tree. Based on the earlier discussions about what the major axis means, I interpret this as suggesting that these justices' votes in First Amendment cases are very much captured by one-dimensional politics. This is mostly true of Due Process cases as well, though to a lesser extent. It is also interesting to note that in Due Process cases we have Gorsuch at the Court's center, leaning just slightly to the right, whereas Kavanaugh is again the far right justice in these cases (as he was with Civil Rights and First Amendment cases), which is quite distinct from their overall center-right position in this natural court. Privacy cases are also mostly one-dimensional, the main small exception being some branching off the major axis by Thomas and Alito, suggesting that the small amount of higher-dimensional disagreement that occurs in these cases is among the different conservative viewpoints.
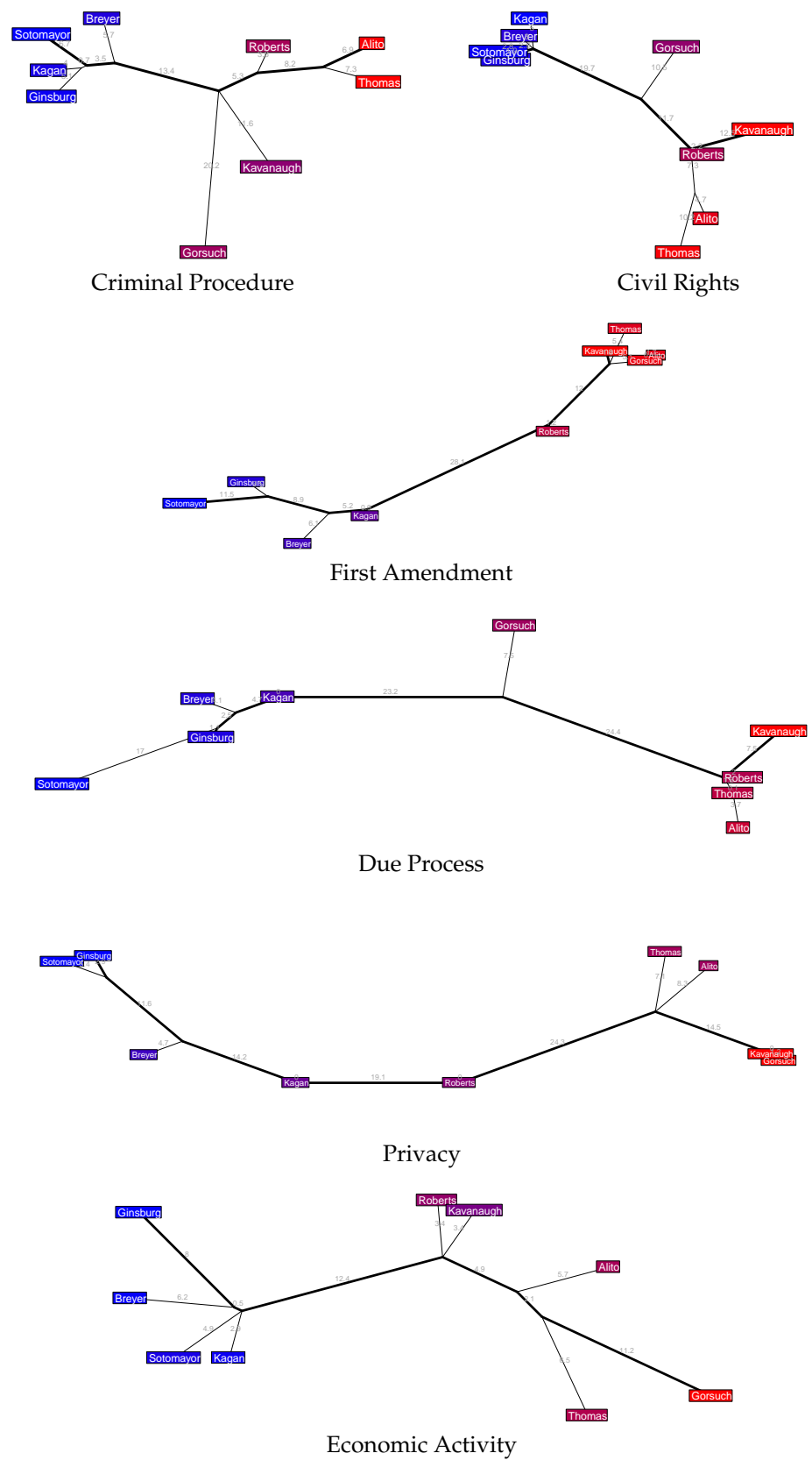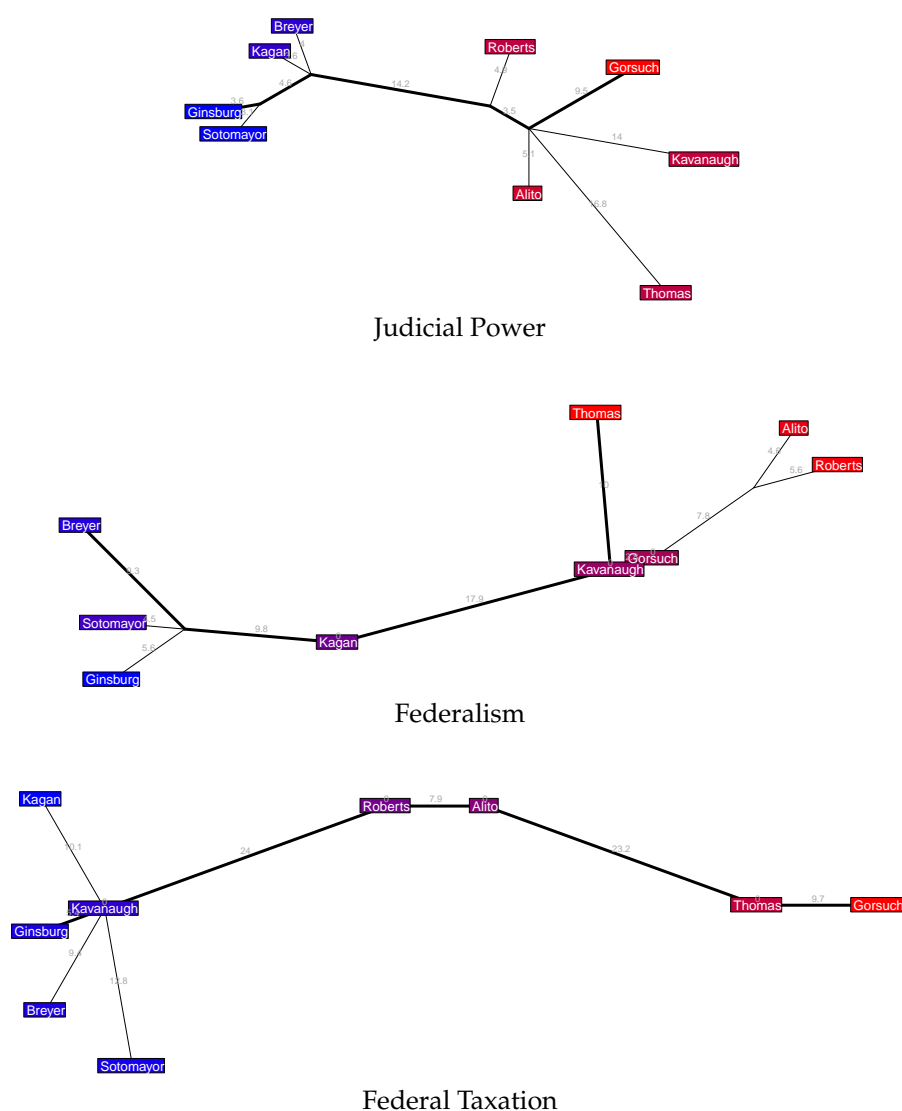
Figure 12. Judicial trees for the 2018–2020 Roberts 7 Court, separated by issue area.

Figure 13. More judicial trees for the 2018–2020 Roberts 7 Court, separated by issue area.

The tree for Economic Activity cases is very similar to the overall tree for this natural court, though curiously MDS seems to consider Alito and Thomas closer to the center of the Court here than the tree does (and also closer to the center than MDS does when combining all case issue areas). The tree for Judicial Power cases shows a fairly tight cohesion and little branching from the major axis among the liberal justices—but the conservative justices are fairly spread out, and mostly in non-political dimensions (by which, as usual, I mean dimensions beyond the primary liberal-conservative spectrum measured by the major axis). The Federalism tree also shows an interesting schism on the Court's right side: Thomas, Alito, and Roberts are all colored staunchly conservative by the MDS estimate, but there is a sizable disagreement rate between Thomas and the pair of Alito-Roberts, who vote together much more often. The tree for Federal Taxation cases is quite striking: there is almost complete unidimensionality from the Court's center justices (Roberts and Alito) to its right-most justices (Thomas and Gorsuch), whereas on the left we have a fractured liberal coalition that branches into many directions and which, perhaps most surprisingly, includes Kavanaugh at its center.

Can we use these issue area trees to help understand why Roberts and Kavanaugh have some "common ground" in the original Figure 2 in the sense of the common edge that connects them to the major axis? I am not sure, to be honest, but this seems like a question worth pursuing in the future.
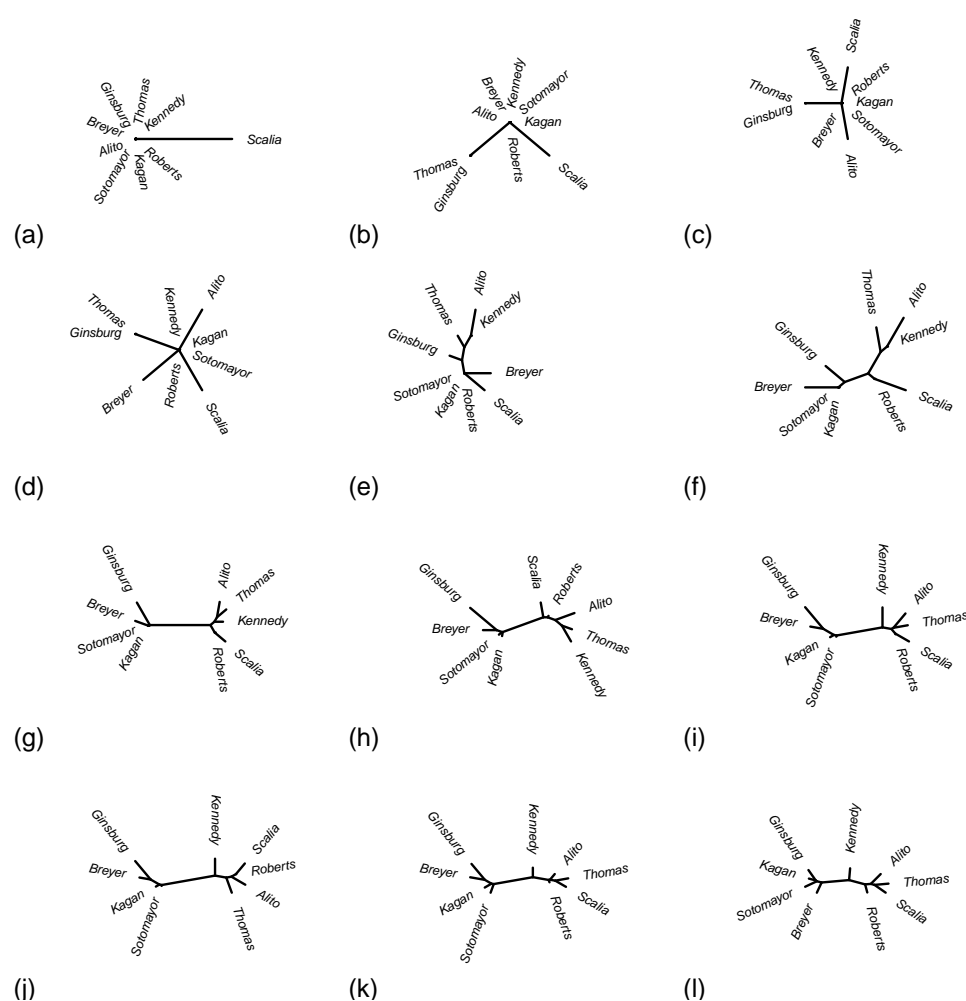
4.4. Evolution of the Court

Having seen the phylogenetic snapshots of various natural courts, a question that inevitably comes to mind is how this structure emerges—or, put another way, how natural courts evolve over time. One way to approach this is as follows. We can compute disagreement rates for the justices in a natural court one case at a time, using only cases heard during that natural court, and each time we update the disagreement rates we can compute a new judicial tree. At first the tree will vary wildly because each new case has a sizable impact on the disagreement rates, but over time the rates settle down and the tree converges to its mature state. This final state of the tree is not quite the same as the snapshots pictured earlier because those used disagreement rates based on all the cases heard by each pair of justices, not just the cases during the given natural court, but nonetheless this iterative approach gives a reasonable sense of how judicial trees take shape. This process is more interesting for the longer natural courts, so to illustrate it I will apply it here to the 2010–2016 Roberts 4 Court from Figure 6.

Often the judicial tree does not change substantively as additional cases are included one at a time, or justices may swap positions back and forth multiple times before settling down, so I will just highlight the cases where interesting changes occur to the tree that seem more than just temporary vacillations. Admittedly this is a subjective matter, but that's ok; I view this process mostly as a tool for uncovering interesting cases that help tell the story of each natural court's development. The highlighted judicial trees are depicted below in Figure 14; they do not have all the additional annotations that I normally include because they would not be very readable on such small illustrations of these rather compact trees (and some of them, such as the MDS-coloring of the justice labels, do not really make sense early on). In discussing how the tree changes at each step of this process, I will draw from the legal summaries in Oyez [43] to help describe the cases that led to these junctures in the tree's evolution.

The case *Ransom v. FIA Card Services, N.A.* (2011) resulted in an 8-to-1 decision, with Kagan writing her first opinion for the Court and Scalia providing the lone dissent. At issue here was whether the projected disposable income for a Chapter 13 bankruptcy plan could allow a vehicle ownership deduction only if the debtor was making payments on the vehicle. The majority opinion ruled that this was indeed so. Scalia's dissent stemmed from a disagreement over the interpretation of a particular word in the Bankruptcy Abuse Prevention and Consumer Protection Act of 2005. He questioned the Court's view that a word in a statutory text that does not seem to add any meaning should be construed in a manner so that it does, quipping: "The canon against superfluity is not a canon against verbosity." This 8-to-1 vote, since it is the first non-unanimous case in the natural court under consideration (according to the order of the cases in the SCDB) results in the phylogenetic tree in Figure 14a where a single edge connects the Scalia leaf to a leaf labelled by the other eight justices.

The next change, seen in Figure 14b, is that Ginsburg and Thomas—an unlikely duo—branch off together from the node that was previously labelled by eight justices. This happened because of the 7-to-2 case *CSX Transportation, Inc. v. Alabama Department of Revenue* (2011), which had a majority opinion by Kagan and a dissent by Thomas that was joined by Ginsburg. The railroad transportation company CSX had challenged the state of Alabama's sales and use tax on diesel fuel, as a violation of the Railroad Revitalization and Regulatory Reform Act of 1976, but the suit was dismissed by a lower court. The Supreme Court ruled that a railroad company could indeed challenge such a tax that applies to rail carriers while exempting competitors in the transportation industry. The dissent agreed that Alabama's tax was potentially discriminatory and that if it did single out railroads among general commercial and industrial taxpayers then the CSX suit would be valid, but in the view of Thomas and Ginsburg this particular tax did not satisfy this criterion.

Figure 14. Evolution of the 2010–2016 Roberts 4 Court as seen by tracking the judicial trees as cases are included one at a time. The trees pictured here show the junctures where the tree undergoes a substantive change.

Alito was the next to split off from the central node, as Figure 14c illustrates, and he did so because of the 8-to-1 case *Snyder v. Phelps* (2011). The family of a deceased soldier, Snyder, had been awarded USD 5 million in damages after members of the cult-like Westboro Baptist Church picketed Snyder's funeral with extremely vulgar and offensive signs. The circuit courts held that this judgement violated the First Amendment's protection of religious expression, and the Supreme Court's majority opinion penned by Roberts agreed. Alito's lone dissent asserted poignantly: "Our profound national commitment to free and open debate is not a license for the vicious verbal assault that occurred in this case." Breyer subsequently split off from the central node, as Figure 14d shows, due to the 8-to-1 vote in *Milner v. Department of the Navy* (2011) where the Court supported a Freedom of Information Act request for internal Navy documents discussing the effects of explosions at several locations caused by the Navy. Breyer's lone dissent ominously asserted: "I would let sleeping legal dogs lie."

The next phylogenetic development that occurred, depicted in Figure 14e, is more interesting. The 6-to-3 vote in *Skinner v. Switzer* (2011) had a rather unusual majority of Breyer, Ginsburg, Kagan, Roberts, Scalia, and Sotomayor. This majority ruled that a convicted prisoner seeking access to biological evidence for DNA testing can assert that claim in a civil rights action under a federal statute on civil action for deprivation of rights. Thomas' dissent, joined by Alito and Kennedy, stretched the phylogenetic tree out, so to speak, by pulling Kennedy and Thomas out along Alito's branch and separating Thomas from Ginsburg. While in some sense the tree is really starting to take shape

here—for instance, it now has four interior nodes instead of just one—the justices are quite scrambled as the non-unanimous cases leading to this phylogenetic tree have been, for whatever reason, heavily skewed towards oddball (in the sense of [17]) majority and minority coalitions. In other words, this particular natural court evidently started rather chaotically before eventually settling down into voting patterns that align more closely with political ideology.

Turning now to Figure 14f, we see a much more familiar state of this Court: a clear liberal branch with Breyer, Ginsburg, Kagan, and Sotomayor and conservative branch with Alito, Kennedy, Roberts, Scalia, and Thomas. The biggest difference between this and the final phylogenetic structure of the Court is that here Kennedy is deeply embedded in the conservative branch. Nevertheless, how did the Court suddenly reach such a mature configuration? The 5-to-4 vote in *Connick v. Thompson* (2011) split right down the liberal-conservative divide, with Kennedy joining the conservative majority. It is somewhat striking that this was the 14th case decided by this natural court (again, according to the order in the SCDB) yet the first 5-to-4 decision—although the very next two cases, *Arizona Christian School Tuition Organization v. Winn* (2011) and *Cullen v. Pinholster* (2011), were both 5-to-4 with the exact same majority as *Connick*. The politically divided vote in *Connick* was enough to pull the phylogenetic tree apart into two very distinct and ideologically consistent branches, then the immediately following cases *Arizona* and *Cullen* both extended and solidified this transformation.

In Figure 14g we have very little movement in the liberal branch, but in the conservative branch Kennedy has effectively swapped positions with Thomas, placing himself closer to the root of the conservative branch. This occurred because of the case *Brown v. Plata* (2011), which had a majority consisting of Breyer, Ginsburg, Kagan, Kennedy, and Sotomayor. This was thus also a 5-to-4 case splitting along political lines, but this time Kennedy swung over to the liberals, and he did so while penning the majority opinion. The case concerned a class-action lawsuit alleging that California prisons violated the Eighth Amendment's ban on cruel and unusual punishment. A special panel of three federal judges determined that serious overcrowding was the primary cause for such a violation and ordered the release of roughly forty-thousand inmates to decrease the inmate population to a number closer to the design capacity of these California prisons. The question before the Supreme Court was whether such a court order was in violation of the Prison Litigation Reform Act, and Kennedy's majority opinion said no, thereby affirming the special panel in the face of Scalia's strongly worded objection declaring its order as "perhaps the most radical injunction issued by a court in our Nation's history." Kennedy's swing to the liberals here is reflected phylogenetically by their placement slightly closer to the center of the tree. Since at this point in the Court's development he has sided with the conservatives in several 5-to-4 cases and only with the liberals in one 5-to-4 case, it makes sense that he is still placed squarely on the conservative branch.

The tree in Figure 14h is in some sense a step backwards: there is again little movement in the liberal branch, but here Scalia has moved to the root of the conservative branch and Kennedy has moved out onto a more extreme branch with Thomas and Alito. There were several non-unanimous cases since the previously discussed one—namely, an Alito-Ginsburg minority in *Fowler v. United States* (2011), a Kennedy-Thomas minority in *Camreta v. Greene* (2011), a solitary Kennedy minority in *Global-Tech Appliances, Inc. v. SEB S.A.* (2011), and a Breyer-Ginsburg minority in *Board of Trustees of the Leland Stanford Junior University v. Roche Molecular Systems, Inc.* (2011)—but the phylogenetic movement here was driven primarily by *Sykes v. United States* (2011), which had an oddball minority of Scalia, Ginsburg, and Kagan. The question underlying this case was whether resisting arrest counts as a violent felony under the Armed Career Criminal Act (ACCA). Kennedy wrote the majority opinion affirming this, and Thomas wrote a concurring opinion in which he mostly agreed with the Court but took issue with its interpretation of a particular legal test that does not appear in the ACCA. The oddball nature of the minority is largely explained by the fact that it took the form of one dissent by Scalia and a separate, strongly contrasting

dissent by Kagan that was joined by Ginsburg. Scalia inveighed against the Court for repeatedly attempting to distinguish violent felonies from non-violent ones and succeeding only in issuing an "ad hoc judgment that will sow further confusion," whereas Kagan and Ginsburg disagreed with the majority opinion because the particular petitioner in the case being heard was convicted only of simple vehicular flight rather than any flight offense involving aggressive or dangerous activity.

Next, coming to Figure 14i, we have Kennedy moving strongly to their rightful place as the gatekeeper to the conservative branch of the Court. This was the result of *Janus Capital Group, Inc. v. First Derivative Traders* (2011), which was 5-to-4 splitting down political lines with Kennedy siding with the conservatives, *Davis v. United States* (2011), which had a Breyer-Ginsburg minority, and most importantly *J.D.B. v. North Carolina* (2011), which was another 5-to-4 along political lines but with Kennedy siding with the liberals. This latter case, where the majority decided that courts should consider the age of a juvenile suspect in determining what "custody" means in the context of Miranda rights, helped solidify Kennedy's role as a swing vote at the center-right of the Court—and he would remain there for the rest of this natural court's development.

The trees in Figure 14j,k,l show only minor jostling among the justices within each of the two main political branches of the Court. These finer movements are the result of a large number of cases and accumulate over the course of years. During this period Breyer establishes himself as the gatekeeper to the liberal branch, thus creating a smaller more strongly liberal branch of female justices, while Kennedy moves closer to the center of the Court and Roberts establishes himself as the gatekeeper to the conservative branch. One case that contributed to Ginsburg, Kagan, and Sotomayor branching off together was the 6-to-3 decision in *J. McIntyre Machinery, Ltd. v. Nicastro* (2011) where these three justices formed the minority. The question before the Court was whether a consumer could sue a foreign manufacturer in state court over a product that the manufacturer marketed and sold through an exclusive distributor in the United States. The majority said no, on the grounds of due process rights of the defendant. The dissenting trio found the majority's view too antiquated and believed it would make it too easy for manufacturers to unfairly protect themselves by strategically utilizing independent distributors.

An important case that pulled Roberts out of the deeply conservative branch populated by Alito, Scalia, and Thomas was the much-discussed *National Federation of Independent Businesses v. Sebelius* (2012), mentioned briefly before in this paper, where Roberts joined the liberals—Breyer, Kagan, Ginsburg, and Sotomayor—to form the majority in this salient 5-to-4 case that upheld the individual mandate in the Affordable Care Act. In fact, it was at precisely this point in this natural court's chronology that the phylogenetic tree reached its final shape shown in Figure 14l. Indeed, from here onward the edge lengths in the tree adjusted slightly but no substantive changes to the tree occurred. To get a sense of how long exactly it took the phylogenetic tree to settle into its final shape, note that the first non-unanimous case in this natural court, *Ransom*, was argued on 4 October 2010, and *National Federation* was argued in 26–28 March 2012, whereas the last non-unanimous case argued before this Court, *Montanile v. Board of Trustees of the National Elevator Industrial Health Benefit Plan* (2016), was argued on 9 November 2015.

While this phylogenetic view of the evolution of the Court likely does not provide insight into any individual cases, hopefully the reader sees that it can at least point to interesting individual cases where important shifts in the Court occurred—as well as providing an intriguing overview of the general movement of the justices relative to each other throughout the duration of the natural court being analyzed.

4.5. Across the Years

Having seen various uses of phylogenetic trees to study a particular natural court, a question that might come to mind is whether the interrelationships among justices across different natural courts can also be studied. Intuitively, if Justice A decided many cases with Justice B, and Justice B decided many cases with Justice C, then even if Justice A and

Justice C never sat on the bench at the same time, it seems reasonable that an estimate of how frequently they would have voted together should be possible. Furthermore, indeed mathematicians have tools for doing exactly this. In this subsection I explore this task by seeking to create a phylogenetic tree with one leaf for each justice who has served since 1946, the first year of the modern SCDB. It should be noted, however, that the methods described below apply to any collection of justices, so scholars interested in seeing how smaller groups of justices are interrelated can certainly do so using the same techniques.

The disagreement rates d_{ij} used to estimate judicial trees can be viewed as entries of a symmetric $n \times n$ matrix, where here $n = 38$ is the number of justices since 1946. This matrix, shown in Figure 15a, has rows and columns indexed by the justices, and the *ij*-entry is missing whenever justice i and justice j did not overlap for any cases. Matrix completion and statistical imputation refer to a variety of algorithms for filling in the missing entries of a matrix. Obviously there is no "correct" way to fill in missing entries—guessing how frequently non-overlapping justices would have voted against each other is a hypothetical gambit!—but each algorithm proceeds to fill in the missing entries based on some assumption about how the data is structured or some property of the incomplete matrix that should be preserved or some minimality property that distinguishes the completed matrix among all possible completions, etc. I will just pick a popular general purpose method that uses nuclear-norm regularization to find a low-rank matrix completion (as implemented in the R package `softImpute` with the function `softImpute` using the SVD algorithm type [44]): see Figure 15b. I make no claims that this is the right approach to matrix completion in the present setting, I simply want to use a popular method to see what it leads to.

Now that we have a completed matrix, we can view all the entries as voting disagreement rates (the original ones real, the filled in ones hypothetical) and apply tree estimation to produce a judicial tree as before, though now with 38 justices. Here the least squares tree estimate is computationally much harder to produce than for all the nine-justice trees in this paper, so instead of directly solving the optimization problem I first use the popular Neighbor-Joining algorithm to produce an approximate tree and then apply the heuristic algorithm `optim.phylo.ls` in the R package `phytools` [45] to iteratively adjust this tree closer to a least squares minimum until it converges. This process is not guaranteed to result in the global least squares minimum, but for a tree of this modest (by biological standards) size it usually does so, and indeed I always got the same final result when perturbing the initial conditions here, so I suspect it is indeed the global minimum. The resulting judicial tree is pictured in Figure 16.

The (hypothetical) most frequently disagreeing justices, as determined by the matrix completion, who therefore span the major axis in this tree are Douglas on the left and Thomas on the right. Recall that Douglas was earlier described as "the most doctrinaire and committed civil libertarian ever to sit on the court" [40]. One striking feature of this tree is that the major axis appears quite prominent. There is a lot of extension off it by individual justices, but not many sizable branches carrying multiple justices. I am apprehensive to say that this means the Court's aggregate behavior over the years has been driven largely by politics, because there is just too many tenuous components to that assertion, but if nothing else this suggests that the few branches we do see in this picture might be interesting and worth investigating. Let us walk through them one at at time. In what follows, I put the years each justice served on the Supreme Court in parentheses.

Starting from the left, we have a length 1.1% edge common to Murphy (1940–1949), Brennan (1956–1990), and Rutledge (1943–1949), and at the same political position along the major axis we have a length 1.6% edge common to Fortas (1965–1969) and Marshall (1967–1991). The two edges are so short that I do not read much into them. However, a bit further along the major axis there is another quintet of justices who share a political position but with a more interesting branching structure. Breyer (1994–) is attached directly at the major axis, then a length 0.7% edge leads to Kagan (2010–) splitting off, then after a length 1.2% edge Ginsburg (1993–2020) splits off, and finally a length 1.1% edge leads to a

split between Sotomayor (2009–) and Stevens (1975–2010). If we view the branch point at the major axis as one end of a new axis (a "semimajor" axis, if you will) then Stevens is the opposite end of this axis, a full $10.3 + 1.1 + 1.2 + 0.7 = 13.3\%$ away, and the other four justices have the following coordinates along this axis: Breyer is 0, Kagan is 0.7, Ginsburg is 1.9, Sotomayor is 3. Is there an ideological interpretation to this axis—meaning something that unites these justices and is roughly quantified by their branch points along it? I do not know. Perhaps it would be an interesting question for a legal scholar to look into. That said, it is curious to note that these five justices almost all served on the Court simultaneously: the earliest departure was Stevens on 29 June 2010, and the latest arrival was Kagan on 7 August 2010. For ten years after that summer of 2010, and for one year prior to it, four of these five justices served concurrently.
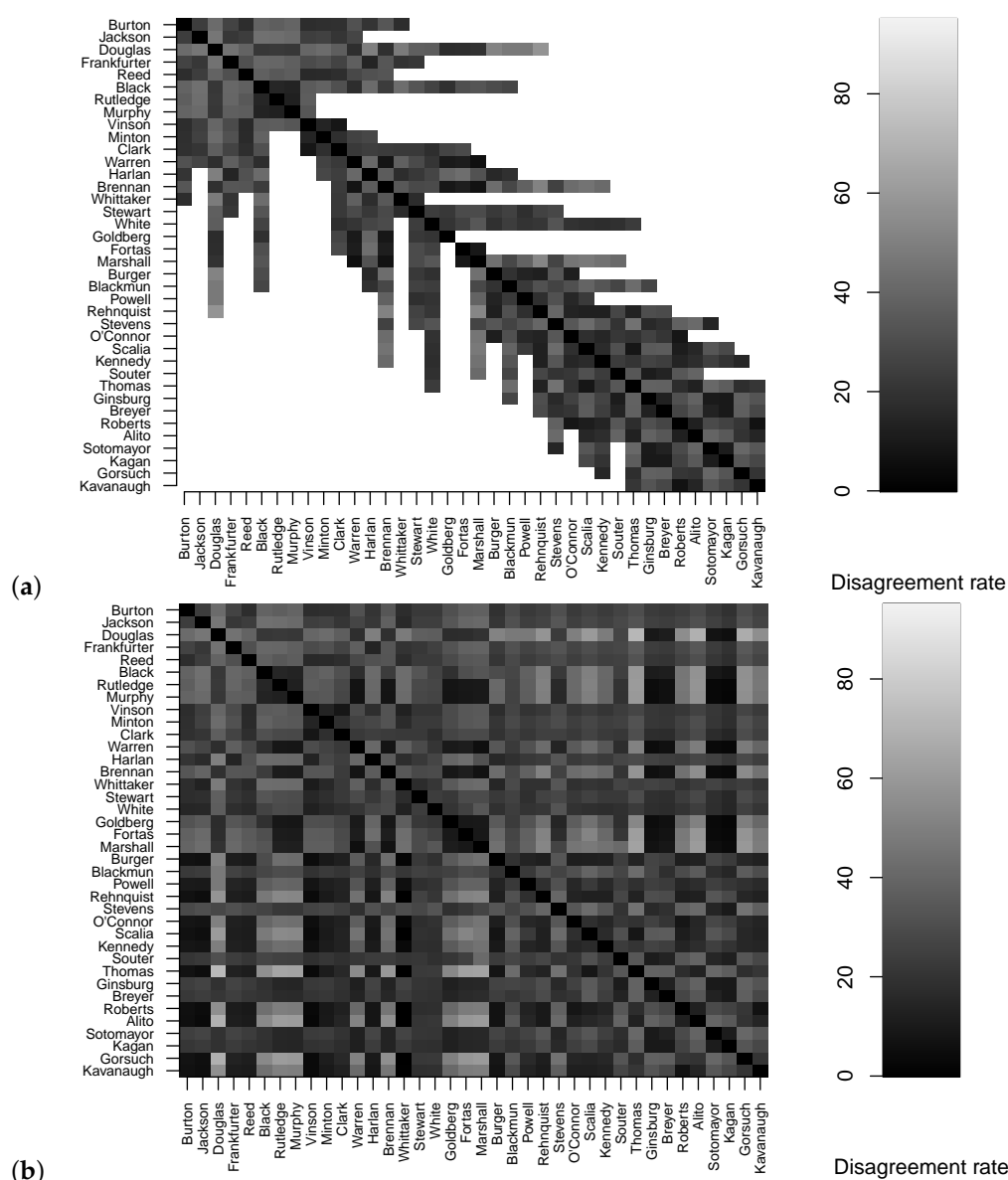
Figure 15. The disagreement matrix for the full 1946–2020 SCDB (**a**) before completion and (**b**) after softImpute matrix completion. Missing entries are plotted white, otherwise lighter shades indicate higher disagreement percentages among the justices (so black means the two justices always agree and light gray that they usually disagree).
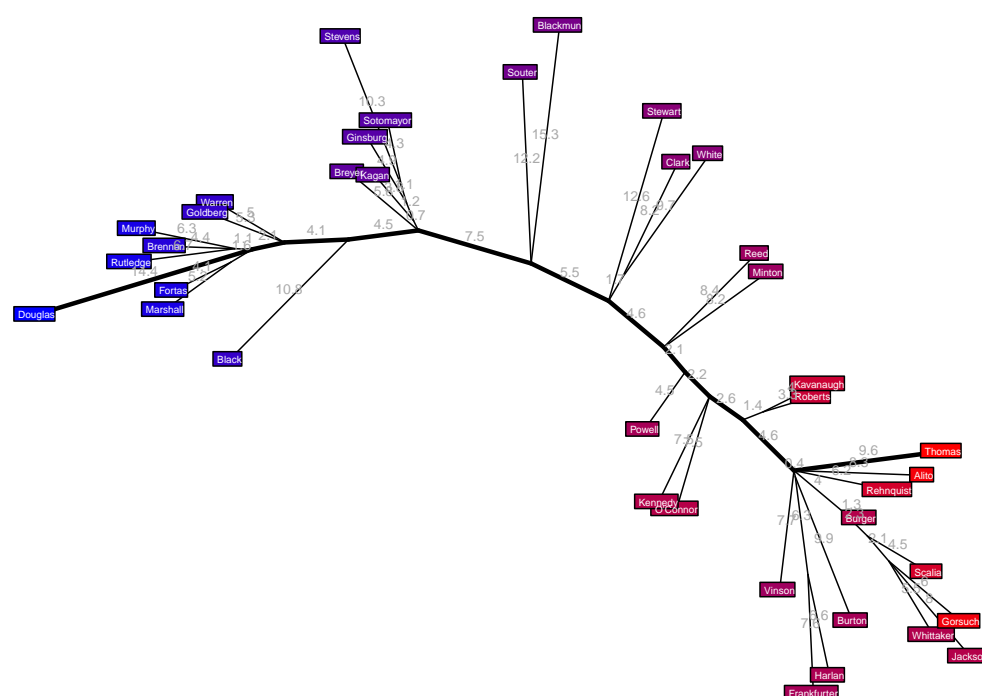
Figure 16. Judicial tree with all 38 justices who served on the Court between 1946 and 2020 computed from the low-rank matrix completion in Figure 15b.

The next branch from the major axis occurs where Stewart (1958–1981) extends directly from the major axis but at the same political position Clark (1949–1967) and White (1962–1993) split from each other after a length 1.7% edge. Further along, we have Roberts (2005–) and Kavanaugh (2018–) paired up, but then things get really interesting at the next junction. Ignoring a barely noticeable length 0.4% edge, all the remaining twelve justices split off from the same node. One of these is Thomas (1991–), whose edge is the final segment of the political axis, and four of them branch directly off from the major axis, so let me turn attention to the remaining seven justices. The seven-year overlapping Harlan (1955–1971) and Frankfurter (1939–1962) have paired off together with a considerable length 6.3% edge in common to the major axis. Lastly, there is a curious branch consisting of a length 4% edge after which Burger (1969–1986) splits off, then a length 2.3% edge after which Scalia (1986–2016) splits off, then a length 2.1% edge leading to a three-way split between the chronologically diverse trio of Gorsuch (2017–), Whittaker (1957–1962), and Jackson (1941–1954). Once again, we could think of the furthest justice among these latter five, namely Jackson, as the end of some kind of semimajor axis here in conservative territory, but what this quantifies, and what the sizable length 6.3% common edge to Harlan and Frankfurter represent, are beyond me.

Recall from Section 2 that Lee in [31] created an incomplete matrix of the pairwise correlations among the voting vectors for all 36 justices on the Court between 1946 and 2016 and then using methods from statistical mechanics constructed the maximum-entropy model compatible with these correlations. This model produces vote patterns quite similar to what has been observed historically, and in doing so it reveals many interesting phenomena about the Court, especially regarding consensus and formation of dissenting blocs. The author shows that their model, which has around 1000 parameters, fits the data better than a standard model in political science that has over 100,000 parameters. The judicial tree in Figure 16 is a different kind of model: while it was created from the same data, it does not generate new data than can be compared with the original voting data. Instead, it is a model more in an intuitive sense of a visualization that helps the reader conceptualize the voting behavior of the court (it basically only has as many parameters as edges in the tree, which is 60, or perhaps thrice this if you separately count as parameters the length and two incident vertices for each edge). That said, it is interesting to note that Douglas is

the far-left justice in the judicial tree here defining the liberal end of its major axis, and Lee found that Douglas also plays a special role in the maximum entropy model: there are only two negative correlations in the model, and both of them involve Douglas. (A negative correlation means that one justice voting to affirm in a case actually increases the odds of the other justice voting to reverse.) In Lee's model, Douglas has a negative correlation with both Burger and Rehnquist, both of whom are deeply in conservative territory in my tree model but not especially distinguished in any way that I can see.

Before concluding this discussion of justices across the years, I would like to mention that a different network-theoretic visualization of all justices between the years 1956 and 2005, also based on voting disagreement rates but not involving trees, is shown in Chart 3 on page 253 of [21].

5. Discussion

The Supreme Court Database [36] has fueld a vast amount of empirical legal scholarship. While the coding of many substantive variables (such as issue, decision direction, legal provisions considered by the court) are somewhat subjective and have consequently drawn skepticism from some scholars, the majority variable—indicating whether each justice is in the majority or the minority based on their vote to affirm or reverse—is thankfully clear cut. For over two decades now this majority variable has driven quantitative investigations into the voting patterns and behavior of the justices. There are essentially two different ways the values in the majority variable have been summarized into a numerical distance between each pair of justices: by computing the correlation between the vectors of 1s and -1s representing the votes, and by computing the disagreement rate (the fraction of cases in which the two justices voted in opposite directions).

There are also essentially two different types of models of the justices that have been built from these notions of judicial distance. Various researchers have used probabilistic methods to create a statistical model that is capable of generating new synthetic votes that can be compared to the actual historical votes—and if the synthetic votes look sufficiently similar to the real ones (usually gauged by how closely the synthetic judicial distances match the actual ones), then one can probe the statistical model to see how it behaves and hopefully also use it to explain why certain observed phenomena on the Court have occurred. Examples of this include the network science models of Guimerà and Sales-Pardo [26], the "fallback" models of Brams–Camilo–Franz [28], and the maximum entropy models of Lee et al. [30–32]. These models map the messy real-world complexity of judicial minds onto a simpler mathematical world that can be captured in a computer, but they are still far too complex to be grasped by human cognition—not in the sense that we do not understand the math involved, but in the sense that the models do not give us a simple mental picture of the Court.

The other type of model scholars have produced from judicial distances goes the extra mile in terms of simplifying vote patterns down to something we can all visualize—at the expense, unsurprisingly, of no longer being able to produce realistic vote patterns or to make statistical predictions about judicial behavior. In short, the first type of model helps us study the Court in detail, while the second type of model helps us develop a coarse-grained intuition for judicial relationships over the years. The predominate approach to this second type of model is ideal point estimation. Without exception, as far as I can determine, this has always placed the justices in a Euclidean space, meaning a space with coordinate axes (usually one, sometimes two, occasionally more) that are independent and orthogonal and extend infinitely in both directions. However, an important recognition in recent years is that much data, especially data concerning social behavior, is better captured by non-Euclidean spaces—especially spaces that form networks (for a modern instance of this, consider the influential geometric deep learning [46] that powers many algorithms at places like Facebook and Twitter).

One particularly simple type of network is a tree, which is characterized by the property that there is a unique path between any two nodes. There are a range of computational

tools for constructing and studying trees that developed largely in evolutionary biology over the past few decades. Mathematically, trees are finite models of hyperbolic space, and in a sense they are the simplest examples of non-Euclidean space. I do not claim that they are the best way to picture the justices, but I do believe they are an appealing compromise between the simplicity of Euclidean space and the flexibility of more exotic geometries. Accordingly, I consider this paper an introduction to extending ideal point estimation to the simplest possible non-Euclidean setting. I have argued in this paper that doing so gives some additional flexibility that even multidimensional ideal point estimation does not provide, since trees can branch in as many directions as they need rather than having a predetermined number of globally defined directions. Furthermore, while interpreting trees is admittedly more subtle than interpreting points on a line (the most common setting for ideal point estimation), I hope to have conveyed in this paper that trees are still relatively simple and easy to grasp conceptually, and by walking through some examples and potential applications I hope to have helped start the reader down the path of understanding and using them in this judicial setting.

Future Work/Questions

- Instead of using disagreement rates, one could use correlations (or cosine similarity), suitably translated to be non-negative, as the premetric input for tree estimation (and for MDS as well). How does this compare?
- Recall from Section 2 that the Bayesian ideal point estimation method of Bailey [20] uses, in addition to affirm/reverse votes, a variety of data including the network of concurring/dissenting opinions and references by the justices to earlier cases. Could these same items be incorporated into the pairwise distances between the justices to enrich the disagreement rates used for tree estimation?
- Recall that in Section 4.3 we produced a separate tree for the different case issue areas. One of the anonymous referees suggested that one could first assemble the disagreement matrices for each issue area into a tensor (image stacking the disagreement rate matrices for each issue area on top of each other). Is there a richer way to extract a geometric model of the justices (Euclidean or tree-based) from this tensor than simply treating each issue area independently?
- I have argued in this paper that the major axis is essentially the one-dimensional liberal-conservative spectrum for the justices, based on the observation that when one-dimensional MDS coordinates very closely approximate the disagreement rates, the judicial tree very nearly equals its own major axis. It would be helpful to take a more empirical approach to complement this more theoretical justification.
- In a related vein, it would be interesting to try to develop an interpretation for some of the prominent branches off the major axis that we have seen in this paper, perhaps by taking a similar approach to Fischman's work on interpreting the second MDS dimension [5,6].
- An influential paper by Kemp and Tenenbaum from 2008 [47] presents a methodology for creating structural representations of data in an unsupervised manner—that is, things like trees and networks and Euclidean embeddings emerge directly from the data rather than being specified in advance as a choice. It would be interesting to apply this flexible framework to judicial votes to see if the judicial trees introduced here emerge naturally, and what other structural representations of this data arise.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Schubert, G. *The Judicial Mind*; Northwestern University Press: Evanston, IL, USA, 1965.
2. Schubert, G. *The Judicial Mind Revisited*; Oxford University Press: London, UK, 1974.
3. Brazill, T.; Grofman, B. Factor analysis versus multi-dimensional scaling: Binary choice roll-call voting and the US Supreme Court. *Soc. Netw.* **2002**, *24*, 201–229. [CrossRef]
4. Brazill, T.; Grofman, B. Identifying the median justice on the Supreme Court through multidimensional scaling: Analysis of 'natural courts' 1953–1991. *Public Choice* **2002**, *112*, 55–79.
5. Fischman, J.; Jacobi, T. The Second Dimension of the Supreme Court. *William Mary Law Rev.* **2016**, *57*, 1671–1715. [CrossRef]
6. Fischman, J. Politics and Authority in the U.S. Supreme Court. *Cornell Law Rev.* **2019**, *104*, 1513–1592.
7. Giansiracusa, N.; Ricciardi, C. Spatial analysis of U.S. Supreme Court 5-to-4 decisions. *arXiv* **2018**, arXiv:1804.08059.
8. Giansiracusa, N.; Ricciardi, C. Computational geometry and the U.S. Supreme Court. *Math. Social Sci.* **2019**, *98*, 1–9. [CrossRef]
9. Giansiracusa, N. Branching on the Bench: Quantifying Division in the Supreme Court with Trees. *SSRN* **2021**. [CrossRef]
10. Boyd, C. In Defense of Empirical Legal Studies. *Buff. Law Rev.* **2015**, *63*, 363–377.
11. Epstein, L.; Martin, A. *An Introduction to Empirical Legal Research*; Oxford University Press: Oxford, UK, 2014.
12. Katz, D. Quantitative legal prediction—or—how I learned to stop worrying and start preparing for the data driven future of the legal services industry. *Emory Law J.* **2013**, *62*, 909–966.
13. Segal, J.; Spaeth, H. *The Supreme Court and the Attitudinal Model Revisited*; Cambridge University Press: Cambridge, UK, 2002.
14. Shapiro, C. Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court. *Hastings Law J.* **2008**, *160*, 477–539.
15. Martin, A.; Quinn, K.; Epstein, L. The Median Justice on the United States Supreme Court. *N. C. Law Rev.* **2005**, *83*, 1275–1321.
16. Epstein, L.; Jacobi, T. Super Medians. *Stan. L. Rev.* **2010**, *61*, 37–99.
17. Edelman, P.; Klein, D.; Lindquist, S. Measuring Deviations from Expected Voting Patterns on Collegial Courts. *J. Empir. Leg. Stud.* **2008**, *5*, 819–852. [CrossRef]
18. Segal, J.; Cover, A. Ideological Values and the Votes of U.S. Supreme Court Justices. *Am. Political Sci. Rev.* **1989**, *83*, 557–565. [CrossRef]
19. Martin, A.; Quinn, K. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court 1953–1999. *Political Anal.* **2002**, *10*, 134–153. [CrossRef]
20. Bailey, M.; Maltzman, F. *The Constrained Court: Law, Politics, and the Decisions Justices Make*; Princeton University Press: Princeton, NJ, USA, 2011.
21. Hook, P. The Aggregate Harmony Metric and a Statistical and Visual Contextualization of the Rehnquist Court: 50 Years of Data. *Const. Comment.* **2007**, *24*, 221–264. [CrossRef]
22. Peress, M. Small Chamber Ideal Point Estimation. *Political Anal.* **2009**, *17*, 276–290. [CrossRef]
23. Sirovich, L. A pattern analysis of the second Rehnquist U.S. Supreme Court. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7432–7437. [CrossRef]
24. Edelman, P. The Dimension of the Supreme Court. *Const. Comment.* **2004**, *20*, 557–570.
25. Edelman, P.; Chen, J. The Most Dangerous Justice: The Supreme Court at the Bar of Mathematics. *South. Calif. Law Rev.* **1997**, *70*, 63–111.
26. Guimerà, R.; Sales-Pardo, M. Justice Blocks and Predictability of U.S. Supreme Court Votes. *PLoS ONE* **2011**, *6*, e27188. [CrossRef]
27. Sullivan, F., Jr.; Georgakopoulos, N.; Georgakopoulos, D. The Fluidity of Judicial Coalitions: A surprising look at coalitions within the supreme courts of the United States and Indiana. *Judicature* **2016**, *100*, 34–44.
28. Brams, S.; Camilo, G.; Franz, A. Coalition formation on the U.S. Supreme Court: 1969–2009. *Public Choice* **2014**, *158*, 525–539. [CrossRef]
29. Lauderdale, C.; Clark, T. The Supreme Court's Many Median Justices. *Am. Political Sci. Rev.* **2012**, *106*, 847–866. [CrossRef]
30. Lee, E.; Broedersz, C.; Bialek, W. Statistical Mechanics of the US Supreme Court. *J. Stat. Phys.* **2015**, *160*, 275–301. [CrossRef]
31. Lee, E. Partisan Intuition Belies Strong, Institutional Consensus and Wide Zipf's Law for Voting Blocs in US Supreme Court. *J. Stat. Phys.* **2018**, *173*, 1722–1733. [CrossRef]
32. Lee, E.; Katz, D.; Bommarito, M.; Ginsparg, P. Sensitivity of collective outcomes identifies pivotal components. *J. R. Soc. Interface* **2020**, *17*, 20190873. [CrossRef] [PubMed]
33. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [PubMed]
34. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *4*, 783–791. [CrossRef]
35. Desper, R.; Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J. Comput. Biol.* **2002**, *9*, 687–705. [CrossRef]
36. Spaeth, H.; Epstein, L.; Ruger, T.; Whittington, K.; Segal, J.; Martin, A. Supreme Court Database. Available online: http://scdb.wustl.edu/ (accessed on 27 April 2021).
37. Cox, M.; Cox, T. Multidimensional Scaling. In *Handbook of Data Visualization*; Springer: Berlin, Germany, 2008.
38. Borg, I.; Groenen, P. *Modern Multidimensional Scaling: Theory and Applications*; Springer: New York, NY, USA, 2005.
39. Lowenthal, D.; Palmer, B. Justice Sandra Day O'Connor: The World's Most Powerful Jurist. *Univ. Md. Law J. Race Relig. Gend. Cl.* **2004**, *4*, 211–239.

40. The Law: The Court's Uncompromising Libertarian. Available online: http://content.time.com/time/subscriber/article/0,33009,913732,00.html (accessed on 27 April 2021).
41. Blasecki, J. Justice Lewis F. Powell: Swing Voter or Staunch Conservative? *J. Politics* **1990**, *52*, 530–547. [CrossRef]
42. Enns, P.; Wohlfarth, P. The Swing Justice. *J. Politics* **2013**, *75*, 1089–1107. [CrossRef]
43. Oyez. Available online: www.oyez.org/cases/ (accessed on 27 April 2021).
44. Mazumder, R.; Hastie, T.; Tibshirani, R. Spectral Regularization Algorithms for Learning Large Incomplete Matrices. *J. Mach. Learn. Res.* **2010**, *11*, 2287–2322. [PubMed]
45. Revell, L. phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **2012**, *3*, 217–223. [CrossRef]
46. Bronstein, M.; Bruna, J.; LeCun, Y.; Szlam, A.; Vandergheynst, P. Geometric deep learning: Going beyond Euclidean data. *IEEE Signal Process. Mag.* **2017**, *34*, 18–42. [CrossRef]
47. Kemp, C.; Tenenbaum, J. The discovery of structural form. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10687–10692. [CrossRef]