The dynamics of ideology drift among U.S. Supreme Court Justices: A functional data analysis

Xiner Zhou[1] *, Hans-Georg Müller [1]

1 Department of Statistics, University of California, Davis, Davis, CA, USA

* Corresponding author
E-mail: xezhou@ucdavis.edu (XZ)

Abstract

We study the U.S. Supreme Court dynamics by analyzing the temporal evolution of the underlying policy positions of the Supreme Court Justices as reflected by their actual voting data, using functional data analysis methods. The proposed fully flexible nonparametric method makes it possible to dissect the time-dynamics of policy positions at the level of individual Justices, as well as providing a comprehensive view of the ideology evolution over the history of Supreme Court since its establishment. In addition to quantifying individual Justice's policy positions, we uncover average changes over time and also the major patterns of change over time. Additionally, our approach allows for representing highly complex dynamic trajectories by a few principal components which complements other models of analyzing and predicting court behavior.

Introduction

In view of the dysfunction of the U.S. Congress in recent years, the importance of the Supreme Court may be more critical than ever. With the death of Supreme Court Justice Ruth Bader Ginsburg just weeks before the 2020 presidential election and Justice Stephen Breyer's recent decision to retire, the ideological direction of the Supreme Court is likely to change drastically.

The study of the ideologies of Supreme Court justices has attracted many scholars [4, 5, 8]. Due to its inherently latent nature, many approaches have been taken to quantify judicial ideologies [6, 7, 9, 11–15]. One of the most popular quantifications is the Martin-Quinn score [15], which introduced a Bayesian item response model to estimate so-called "ideal points". The literature on the voting behavior of Supreme Court justices [1–3] generally postulates that preferred policy positions of justices are the key explanatory variables of realized voting behavior. These policy positions reflect the underlying latent preferences of justices, with their voting behaviors resulting from these latent factors along with other factors pertaining to the specific cases to be decided and additional extraneous factors. For instance, the "attitudinal model" theorizes that justices vote according to their true attitudes [10]. Based on this postulated relationship between voting behavior and latent ideology, we take here an approach that utilizes the direction (conservative versus liberal) of the observed votes to infer the latent ideologies. Our approach is related to that of [7] using the percentages of liberal votes in a single policy area and that of [15] using a Bayesian item response model.

The proposed approach is based on Functional Data Analysis (FDA) [18–20], a powerful nonparametric statistical methodology that to date has not been much used in the social sciences at large but is becoming increasingly popular for the analysis of longitudinal studies or panel data (time-series cross-sectional (TSCS) data) [21–24]. In a nutshell, FDA is a methodology to model the dynamic behavior of an underlying latent stochastic process over some continuum such as time, while the available observations are noisy data, collected on a discrete and possibly incomplete grid over the domain. Specifically, it is assumed that each justice has a latent policy position that changes continuously over time. We regard this as an instantiation of a latent policy position process, and that the justice's voting behaviors are direct manifestations of this underlying latent process, possibly disturbed by some temporaneous influencing factors such as specific features of a case. A similar assumption is also utilized by [15]. The data consist of repeatedly observed votes for each justice and are binary (yes -no) on a dense temporal grid; as we demonstrate, FDA is uniquely suited for the analysis of such data.

One of the benefits of our approach is its ability to compress high-dimensional complex trends into a few variables, the functional principal components, which can then facilitate further modeling. For instance, judicial behavior might be an important explanatory factor for other political phenomena; for the purpose of predicting Supreme Court votes [16, 17], past judicial behavior can be summarized by functional principal components which can serve as features for machine learning models. Thus, the proposed approach complements existing work related to judicial ideology and voting behavior. Another benefit is that it opens the door to a suite of statistical models analyzing "trajectory" data in the social sciences.

Using the proposed methodology, we attempt to shed some light on a few questions of interest: Do the policy preferences of U.S. Supreme Court justices change over time? Can patterns be discerned in the policy preference trajectories for justices throughout their tenure and in their individual relative position in the Court? What is the overall dynamics of the Court both in the past and currently? And do justices tend to express ideology differently through their voting behavior for different issues, such as civil rights or economic activity?

The paper is organized as follows. In the next section, we briefly describe the Supreme Court Database that forms the starting point for our analysis. This is followed by a very brief introduction to functional principal component analysis and a section on results, where the proposed estimation methods and main results are presented. The paper concludes with a discussion.

Data

The Supreme Court Database (SCDB) was accessed on January 8th, 2022 [25]. The Supreme Court Database (SCDB) includes two releases: SCDB Modern and SCDB Legacy. The most recent SCDB Legacy release dated October 1st 2021, contains terms from 1791 to 1945 and contains 172,213 justice-votes records. The most recent SCDB Modern release dated September 30th 2021, contains terms from 1946 to 2020 and contains 122,754 justice-votes records. In total, 294,967 records at the justice-vote level are available between 1791 and 2020.

Since the Supreme Court was established in 1789, 115 justices have served on the Court. As of January 8, 2022, the SCDB (Modern and Legacy combined) contains votes from all justices except the most recently tenured justice Amy Coney Barrett. Since Thomas Johnson only served for 163 days, he is also excluded. Consequently, all 113 justices with voting records are included in the analysis, which thus covers essentially all votes of the justices since the establishment of the Supreme Court.

The Justice Centered data include an indicator of whether a participating justice cast a liberal or conservative vote along with the date when the vote was cast. These data form the observations on which our analysis is based.

Cases are further labeled according to larger issue areas, which include criminal procedure, civil rights, First Amendment, due process, privacy, attorneys' or governmental officials' fees or compensation, unions, economic activity, judicial power, federalism, interstate relations, federal taxation, miscellaneous, and private law. These labels make it possible to conduct a more detailed analysis, as the preferred policy positions of justices may manifest themselves differently in different case categories.

Methods

Basic approach

The main methodology for our analysis is based on Functional Data Analysis (FDA) [19, 26]. In FDA, curve data are viewed as sample paths of a continuous time stochastic process over some continuum, usually over a common time interval. In the present context, a sample path $X(t)$ corresponds to a Justice's latent policy preference process, or ideology process, as it evolves over time within a fixed common time interval $\mathcal{T}$.

Since we are interested in the evolution of Justices' ideology, in a preprocessing step we take the origin of time $t = 0$ for each Justice as the time at which a Justice was appointed. Accordingly, the trajectories for all Justices are defined on the same time interval $\mathcal{T}$ and the ideology process is viewed as evolving over "time since tenure". The observed end point of this continuum, however, varies from Justice to Justice, because Justices serve on the Court for different lengths of time; the end of their tenure is random and may be due to resignation or death. The length of service on the Court for the 106 non-incumbent Justices ranged from William O. Douglas's 36 years and 211 days to the 163-day tenure of Thomas Johnson. As of January 9, 2022, the length of service for the nine incumbent Justices ranges from Clarence Thomas' 30 years and 78 days to Amy Coney Barrett's 1 year and 74 days. The median, first quantile and third quantile of length of service are 24, 16, and 30 years, respectively. Due to the rich information available through SCDB, we were able to select the relatively long time interval $\mathcal{T} = [0, 35]$ years since appointment as the common time interval for the analysis.

It is then of interest to study how the Justice's policy preferences $X(t)$ change throughout the tenure of the Justice on the Court on the time interval $\mathcal{T} = [0, 35]$ years since appointment. One may always visualize the results in calendar time if desired.

A challenge is that only William Douglas had a tenure period of more than 35 years, and 8 other Justices had tenures that are just shy of 35 years and for these Justices their voting behavior is observable over the entire time interval $\mathcal{T} = [0, 35]$ and thus they have completely observed functional data. For all the other Justices with less than 35 years of tenure, the functional data are partially observed as data towards the right end of the time interval are not available for these Justices, and the shorter their tenure period is the more data are missing.

The Supreme Court Database features data that consist of $\{(t_{ij}, Y_{ij}) : i = 1 \ldots 113,\ j = 1 \ldots, m_i\}$, where there are 113 Justices and the i-th Justice has m_i recorded votes. Here t_{ij} refers to the time when the j-th decision is recorded for the i-th Justice, measured in terms of days since appointment of the i-th Justice. The Y_{ij} denote the i-th Justice's votes at decision time t_{ij}, with $Y_{ij} = 1$ if the Justice casts a conservative vote and $Y_{ij} = 0$ if the Justice casts a liberal vote. To address the issue of unequal tenure periods and thus partially observed functional data, we use a statistical model that connects the latent policy position process to the actually observed data for each Justice.

The observed votes at day t_{ij} for the i−th Justice are assumed to follow a Bernoulli distribution with probability

$$p_i(t_{ij}) = P(Y_{ij} = 1|T = t_{ij}), \tag{1}$$

where the binary observed response at t_{ij} is the result of the Justice's preferred policy position and additional noise stems from the nature of particular cases or other exogenous or subjective variables that are unknown. The preferred policy position can be considered a latent trait that is expected to be stable in the short-term but may change over time smoothly. This is reflected by our assumption that the functions $p_i(t)$ are continuously differentiable.

The link between $p_i(t)$ and the desired ideology process $X_i(t)$ can be modeled by hypothesizing that ideology of Justices is the key explanatory variable for realized voting behavior while other factors pertaining to the specific cases to be decided and additional extraneous factors may play an additional role. This hypothesis is adopted by many scholars [1–3, 7, 10, 15]. Based on this relationship between voting behavior and latent ideology, we build the following model to infer the latent ideology process $X_i(t)$ from the observable decisions $p_i(t_{ij})$.

Then under the latent variable model framework, we link the probability of an observation of a "1" outcome and the latent (Gaussian) process via a logit transformation as follows,

$$\text{logit}(p_i(t_{ij})) = \log \frac{p_i(t_{ij})}{1 - p_i(t_{ij})} = X_i(t_{ij}) + e_{ij}, \quad t_{ij} \in \mathcal{T}, \tag{2}$$

where errors e_{ij} denote local aberrations from the smooth underlying processes $X_i(t)$. The logistic transform has the effect to transform the functions p_i which are restricted by $0 < p_i(t) < 1$ to $X_i(t)$ =logit$(p_i(t))$, which are unrestricted real-valued functions as required by FDA methodology. This approach does not rely on any assumption about the time-varying or constant nature of the ideology of Supreme Court Justices over the time domain This methodology embodies the principle of "letting the data speak for themselves" by imposing only minimal assumptions. Thus it is ideally suited to provide empirical evidence for the debate whether judicial preferences are constant or changing over time [27–32].

To summarize, the estimation of the latent policy position processes $\{X_i(t) : i = 1 \ldots n, t \in \mathcal{T}\}$ follows two steps. The first step is to transform the binary observations $\{(t_{ij}, Y_i(t)) : i = 1, \ldots n,\ j = 1, \ldots, m_i\}$ into functional data $X_i(t) = \text{logit}(p_i(t)) : i = 1 \ldots n\}$. In a second step, we apply Functional Principal Component Analysis (FPCA) to the functional data logit$(p_i(t))$, aiming to estimate the underlying latent policy processes $X_i(t)$ for each Justice.

Converting binary observations to functional data

To convert the binary observations $\{(t_{ij}, Y_i(t_{ij}) : i = 1, \ldots, n,\ j = 1, \ldots, m_i\}$ into functional data, the starting point is to obtain smooth probability functions in time t, over the time period for which one has data for an individual Justice. To this end, we observe that $p_i(t) = P(Y_i(t) = 1) = E[I\{Y_i(T) = 1\}|T = t]$ and that this conditional expectation can be viewed as a regression function over the time domain, an estimate of which can then be obtained by scatterplot smoothing. For this smoothing step, we adopt Local Linear Smoothing (LLS) [33, 34] to obtain a continuous estimated probability function $\hat{p}_i(t)$ for each Justice, where a smoothing bandwidth of $h = 365$ days was used to borrow information from neighboring cases within one year. The choice avoids situations where there are too few cases at the Court during a shorter period of time, which would lead to highly variable estimates and yields practically interpretable results. Choosing different smoothing bandwidths led to similar results.

To facilitate the subsequent application of the logit transform, which requires $\hat{p}_i(t)$ to be strictly larger than 0 and smaller than 1, we introduce a small threshold $\rho > 0$ such that the function estimates $\hat{p}_i(t)$ are always in the interval $[\rho, 1-\rho]$. This can be achieved by setting values that fall outside this interval to equal the closer one of these boundary points, where $\rho = 0.001$ was chosen as this value was found to be adequate to shrink the values $\hat{p}_i(t)$ away from 0 or 1 but only by a negligible amount. We then obtained the set of measurements of the underlying latent trajectories as

$$Z_{ij} = \text{logit}(\hat{p}_i(t_{ij})) : i = 1, \dots, n, \ j = 1, \dots, n_i. \tag{3}$$

These are considered measurements of the underlying unknown trajectories $X_i(t)$ at time points t_{ij} that may carry noise due to aberrations from the smooth underlying trajectories when the vote was taken. While we assume here that the combination of a smoothing step followed by a logit transformation leads to potentially still noisy measurements of an underlying smooth process, which is vindicated by the practical success of this approach, nonparametric alternatives where the link function is not specified could also be considered [35].

After these pre-processing steps, the resulting data are $\{(t_{ij}, Z_{ij}) : i = 1, \dots, n, \ j = 1, \dots, m_i\}$. These are assumed to be related to the underlying latent policy preference processes through

$$Z_{ij} = X_i(t_{ij}) + \epsilon_{ij}, \quad i = 1, \dots, n, \ j = 1, \dots, n_i. \tag{4}$$

Here the errors ϵ_{ij} are assumed to be mean zero finite variance (Gaussian) random variables with $EZ_{ij}^2 = \sigma_\epsilon^2$ that reflect estimation errors and noisy oscillations that are not part of the smooth latent trajectories X_i.

Functional principal component analysis (FPCA)

Our goal in this step is to estimate the latent policy position processes $\{X_i(t) : i = 1 \dots 113\}$ on a domain $\mathcal{T} = [0, 35]$ years since tenure by applying Functional Principal Component Analysis (FPCA) to the observed data $\{(t_{ij}, Z_{ij}) : i = 1, \dots, n, \ j = 1, \dots, m_i\}$ which can be considered as noisy realization of the underlying ideology processes.

FPCA is based on the eigendecomposition of the Hilbert-Schmidt linear operator with kernel $C(s,t) = \text{cov}(X(s), X(t))$, leading to the decomposition $C(s,t) = \sum_{j \geq 1} \lambda_j \phi_j(s) \phi_j(t)$ with eigenvalues $\lambda_j > 0$, $\lambda_1 > \lambda_2 > \dots$ and a sequence of orthonormal eigenfunctions ϕ_j, $j \geq 1$. Under mild assumptions, this entails the Karhunen-Loève representation of trajectories X_i [26] given by

$$X_i(t) = \mu(t) + \sum_{j=1}^{\infty} \xi_{ij} \phi_j(t), \tag{5}$$

where $\xi_{i1}, \xi_{i2}, \dots$ are mean zero uncorrelated functional principal components (FPCs) with explicit representation as integrals (inner products) $\xi_{ij} = \int (X_(t) - \mu(t)) \phi_j(t) dt$, and with variances $\text{var}(\xi_{ij}) = \lambda_j$, $j = 1, 2, \dots$.

In practice, the expansion is approximated by only including the first K components in the sum on the right hand side, where K is typically chosen to achieve a large fraction of the variation explained, most commonly FVE= .95 or FVE= .99, where here we choose the latter. The mean function $\mu(t)$ and the eigenfunctions, eigenvalues and principal components can be estimated with the PACE algorithm [36–38] which is available in the R package fdapace [39].

It is important to note that for the PACE implementation of FPCA one does not require data for individual justices to be available on the entire domain $[0, T]$, rather it is required that the pairs of times where repeated measurements are obtained for the same justice when plotted against each other, for all justices combined, will densely fill the square $[0,T]^2$, which can be ascertained by a domain plot [36]. If this is satisfied, then assembling the voting data for justices with both longer and shorter tenures will still lead to consistent estimates of the eigencomponents of the FPCA over the entire domain $[0,T]$, as long as one has data from a sufficient number of justices whose tenure exceeds $[0,T]$.

Once the eigencomponents on the entire interval $[0,T]$ have been obtained, for K selected components, the estimated FPCs can be obtained as follows.

$$\hat{\xi}_{ik} = \widehat{E}\left[\xi_{ik} \mid \tilde{\boldsymbol{Y}}_i\right] = \hat{\lambda}_k \hat{\boldsymbol{\phi}}_{ik}^T \widehat{\boldsymbol{\Sigma}}_{Y_i}^{-1}\left(\tilde{\boldsymbol{Y}}_i - \hat{\boldsymbol{\mu}}_i\right), \quad k = 1, \ldots, K, \tag{6}$$

where $\widetilde{\boldsymbol{Y}}_i$ is a vector containing the voting data Z_{ij} Eq (4) for the i-th justice evaluated at the times t_{ij} when this justice voted on a case, with $\widetilde{\boldsymbol{\mu}}_i$ denoting the overall means, also evaluated at the times t_{ij}. Furthermore, $\boldsymbol{\Sigma}_{Z_i} = \mathrm{cov}(\tilde{\boldsymbol{Z}}_i) = \mathrm{cov}(\widetilde{\boldsymbol{X}}_i) + \sigma^2 \boldsymbol{I}_{N_i}$ is the covariance matrix of the observed data Z_{ij} at time points t_{ij}. To implement the PACE algorithm we used the R package fdapace [39].

We then insert the estimated FPCs obtained from Eq (6) into the representation Eq (7), which yields the estimated trajectory $\widehat{X}_i(t)$ on $[0,T]$ for the i-th justice.

The guiding principle is to gain strength by pooling the data from all justices when representing the trajectory of individual justices. When choosing K components, substituting these estimates leads to the representation

$$\widehat{X}_i(t) := \hat{\mu}(t) + \sum_{j=1}^{K} \hat{\xi}_{ij} \hat{\phi}_j(t), \tag{7}$$

where $\widehat{X}_i(t)$ is the estimated ideology process for ith justice from time of appointment to up to 35 years of tenure.

If the tenure period of a justice is $[0,S]$, $S < 35$, then the estimator $\widehat{X}_i(t)$ presented in Eq (7) on $[S,35]$ is the predicted ideology trajectory for this justice, where the end point S corresponds to current time, and all times $t > S$ are in the future relative to the tenure period of the justice for whom the future ideology trajectory is to be predicted. The prediction relies on pooling information from that specific justice as well as others. The estimator given in Eq (7) is based on the best linear unbiased prediction principle and there is additional justification for these predictors if the (transformed) trajectories X_i are Gaussian processes. The predicted trajectories are unbiased if K is sufficiently large [40]. The principle of the prediction is to take the data from the justices that are observed on the entire time domain $[0,35]$ or at least a longer domain and to infer from those data the future voting behavior of a justice for whom data are only observed on a subset $[0,S]$ of the total domain $[0,35]$. This device has been recently also used for the prediction of COVID-19 case trajectories [41].

Interpretation of the mean and eigencomponents

The estimated mean function $\hat{\mu}(t)$ represents the average ideology process on the domain [0,35] years for the 113 Supreme Court justices. The estimated eigenfunctions $\hat{\phi}_j(t), j = 1 \ldots K$ represent the leading patterns of variation in the ideology processes extracted from the observed voting behavior. They reflect the modes of variation, i.e. the major ways if ideology change over time for the justices. A flat eigenfunction indicates no change over time in the direction of this eigenfunction. The estimated FPCs $\hat{\xi}_{ij}$ are justice-specific in contrast to mean and eigenfunctions, which reflect the entire population of justices. They indicate in which way the ideology trajectory of the i-th justice moves along the variation over time in the direction of the jth estimated eigenfunction $\hat{\phi}_j(t)$, with large positive value reflecting a stronger alignment in the corresponding eigendirection and a negative value reflecting the opposite pattern of variation.

Dimension reduction

The representation in Eq (7) establishes a one-to-one correspondence between the latent processes $X_i(t)$ and K-dimensionalc random vectors that consist of the FPCs $\xi_{i1} \dots \xi_{iK}$. It is through this correspondence we achieve dimension reduction for the original highly complex trajectory data to a random vector often of 2 to 4 dimensions. The resulting K-vector $(\xi_{i1}, \dots, \xi_{iK})$ thus represents the trajectory for the i-th justice and can then be used for other statistical analysis or machine learning models.

Prediction of future ideology process

The PACE principle is to pool the data to gain insights into the general modes of variation (which are determined by the eigenfunctions and show the main directions of variation). Eq (7) gives the inferred ideology process over the complete 35 years after tenure. For those justices whose voting data are only available on $[0, S], S < 35$, it can also be used to infer their future likely ideology processes, which is our prediction of potential (and never observable for those with short tenure) trajectories. It can be easily transformed into probability scale (between 0 and 1) as the likelihood of conservative votes, for more transparent interpretation, by applying the expit transformation which is the inverse of logit, i.e.,

$$\hat{p}_i(t) = \text{expit}(\widehat{X}_i(t)) = \frac{\exp(\widehat{X}_i(t))}{1 + \exp(\widehat{X}_i(t))}, \quad t \in [0, T], \tag{8}$$

Results

Time evolution of ideology of Justices

After the pre-processing smoothing step, we have a sample of 113 curves $\{\hat{p}_i(t) : i = 1 \dots 113\}$, each representing the observed proportion of conservative votes as a function of time for one Justice. These curves are displayed in Fig 1 according to calendar year and in Fig 2 according to year since appointment. The overall picture suggests that, prior to the 1940s, Justices tended to vote similarly as their observed voting behaviors were closely aligned; after the 1940s, the discrepancies in terms of voting behavior between Justices increased substantially and currently a clear divide between Democratic and Republican appointed Justices began to emerge.

Fig 1. Pre-smoothed curves $\hat{p}_i(t)$, $i = 1, \dots, 113$ against calendar year. Here $\hat{p}_i(t)$ represent the observed proportion of conservative votes as a function of calendar time for the i-th Justice. The curves are color coded by the nominating president's party affiliation: red stands for a Justice nominated by a Republican president, blue stands for a Justice nominated by a Democratic president, and gray stands for a Justice nominated by a third party president. The currently active Justices are highlighted.

Fig 2. Pre-smoothed curves $\hat{p}_i(t)$, $i = 1, \ldots, 113$ against year since appointment of the Justices. Here the $\hat{p}_i(t)$ represent the observed proportion of conservative votes as a function of time since appointment for the i-th Justice. The curves are color coded by the nominating president's party affiliation: red stands for a Justice nominated by a Republican president, blue stands for a Justice nominated by a Democratic president, and gray stands for a Justice nominated by a third party president. The currently active Justices are highlighted.

Fig 3. Average ideology trajectory and first two eigenfunctions A: the estimated average ideology trajectory, with shaded region corresponding to 95% pointwise bootstrap confidence intervals for the actual mean trajectory; B: the first and second eigenfunctions $\phi_1(t)$ and $\phi_2(t)$; C: the first mode of variation $\mu(t) \pm k\sqrt{\lambda_1}\phi_1(t)$ for $k = 0, 1, 2$ where $\sqrt{\lambda_1}$ is the standard deviation of the FPC scores corresponding to the first eigenfunction ; D: the second mode of variation $\mu(t) \pm k\sqrt{\lambda_2}\phi_2(t)$ for $k = 0, 1, 2$ where $\sqrt{\lambda_2}$ is the standard deviation of the FPC scores corresponding to the second eigenfunction.

Ideology dynamics through FPCA

The average ideology process and main patterns of change

Fig 3 A shows the mean ideology trajectory across all Justices, in the logit scale, so that a probability of 0.5 for a conservative vote corresponds to 0. This suggests that, on average, Justices start as centrists for the first 15 years, but then are subject to a slight tendency to a more liberal ideology. This supports the argument that the Supreme Court Justices' ideology does change over time.

The major archetypes of dynamics are extracted by FPCA and presented in the first two eigenfunctions in Fig 3 B. The first two eigenfunctions ϕ_1, ϕ_2 explain 97.3% of the total variation, with ϕ_1 accounting for 91% and ϕ_2 accounting for 6.3%. Like multivaraite principal component analysis, FPCA projects high-dimensional curve data into a low dimensional space, with each dimension representing a major pattern of change that explains the complex dynamics in the data.

The first eigenfunction ϕ_1, in Fig 3 B represents the first such dimension, and when multiplied with positive factors indicates consistently more liberal positions than the average. The large fraction of variance explained by this first eigencomponent shows that by far the largest source of variation is indeed a basic conservative or liberal policy preference, along with the tendency that this preference becomes more expressed as the tenure of a justice wears on.

The second eigenfunction ϕ_2 represents the second dimension, that is, a time-dynamic trend toward the opposite side of policy preference from the starting point after 10 years followed by a moderate swing back toward the original position after 20 years. This component thus reflects a moderate swing dynamic pattern. The second eigenfunction mainly captures how policy preferences change, in contrast to the the first eigenfunction, which captures the basic levels of liberalism or conservatism with a deepening trend.

To see how the average trajectory, the first and second eigenfunctions interact to lead to the manifest different dynamics, it is helpful to visualize the modes of variation in Fig 3 C and D. Fig 3 C shows the first mode of variation, $\mu(t) \pm \sqrt{\lambda_1}\phi_1(t)$, which is comprised of hypothetical trajectories that lie one standard deviation away from the average trajectory in the dimension represented by the first eigenfunction; and $\mu(t) \pm 2\sqrt{\lambda_1}\phi_1(t)$, lying two standard deviations away from the average trajectory. Clearly, without any additional dynamic provided by the second dimension, trajectories would mostly differ in the overall levels of liberalism or conservatism throughout, where these levels are measured in terms of differences from the average level and intensify slightly over time as tenure progresses. Fig 2 D shows the second mode of variation: $\mu(t) \pm \sqrt{\lambda_2}\phi_2(t)$ shows hypothetical trajectories one standard deviation away from the average trajectory in the dimension represented by the second eigenfunction; and $\mu(t) \pm 2\sqrt{\lambda_2}\phi_2(t)$ such trajectories situated two standard deviations away. The effect of this second dimension is trending toward the opposite side of policy preference from the starting point after 10 years followed by a moderate swing back toward the original position after 20 years.

A unique feature of the FDA approach and especially FPCA is to decompose the observed trajectories into the distinct modes of variation and thus dissect the main drivers of the observed trends. The orthogonal eigenfunctions thus serve to elucidate different aspects of the observed ideology dynamics. The principle of FPCA is to decompose complex pattern of changes into a few mutually orthogonal patterns. This approach is completely data-driven and one is not limited to any pre-conceived pattern. While the average trajectory reflects the average pattern of change that applies to all Justices equally, the first two FPCs encompass the two major patterns of individual variation across Justices, where the trends encapsulated in the eigencomponents can be understood as random effects in addition to the fixed effect represented by the average trajectory.

Justice-specific functional principal components

For most Justices the two patterns of variation reflected by the modes of variation are both present, where the first component explains by far most of the variation and thus dominates, while the dynamic pattern reflectedby the second component is less noticeable. The scatterplot of the FPC scores in Fig 4 reveals which pattern a Justice largely follows.

Fig 4. First and second Functional Principal Component (FPC) scores. Estimated FPC scores were obtained as per Eq (6) and visualize the FPC-space representation of ideology trajectories, where coordinates represent the amount of deviation from the average trajectory in the direction of the first and second eigenfunctions. Color coded by the nominating president's party (blue for Democratic, red for Republican). The Justices that are mentioned specifically are highlighted in yellow. (A) exhibits all 113 Justices, with 3 obvious outliers: Thomas Todd, James Wilson, and John Jay. (B) exhibits all non-outlier Justices before 1900 with annotated names. (C) exhibits all non-outlier Justices after 1900 with annotated names.

Different regions in the FPC space represent different ideology dynamics. FPC scores of Justices with consistently more liberal disposition are located in the right half-plane, and those with consistently more conservative disposition are located in the left half-plane. Justices with a shift toward a more conservative ideology over the course of their tenure in addition to the average trend are located on the upper half-plane, and those with a shift toward a more liberal ideology are located on the lower half-plane.

Clearly interpretable patterns emerge for Justices whose FPCs are situated near the main axes. Trajectories of Justices whose FPCs are situated close to the x-axis generally follow the pattern depicted in the first mode of variation as shown in Fig 4 C. Trajectories of Justices whose FPC scores are situated close to the y-axis generally follow the pattern depicted for the second mode of variation in Fig 4 D. For example, Felix Frankfurter started as more liberal but had a drastic shift toward a conservative ideology after 10 years. This pattern is confirmed in Fig 8

The Justices with scores near the diagonal or off-diagonal regions exhibit a mix of dynamics that adds to the average trajectory, and show dynamic patterns related to both first and second eigenfunctions. For example, William Rehnquist has very negative FPC1 and FPC2, which means his trajectory is a mix of "consistently more conservative than the average" and "a shift toward liberal direction after 10 years since appointment". This pattern can be confirmed in Fig 8. An interesting outlier is Thomas Todd who has a very negative FPC1 but a very positive FPC2, which means his trajectory would be a mix of "consistently more conservative than the average" and "a shift toward conservative direction after 10 years since tenure". This pattern is confirmed in Fig 5.

The separation between Republicans and Democrats is clearly visible, where Republicans are clustered on the left half-plane, and Democrats on the right half-plane. This is as expected. We ascertained these effects by simple linear regression models, by regressing ξ_1 and ξ_2 on party affiliation of the appointing president (Republican or Democrat), and calendar year at tenure. The coefficients and associated significance from the two models are reported in Table 1. Only party affiliation (Republican or Democrat) is significantly associated with ξ_1, with Republicans more likely to associated with the pattern "consistently more conservative position than the average", again as expected.

Table 1. Results from Linear regression models

	FPC1 ξ_1	FPC2 ξ_2
Republican	-35.597***	-2.726
Year afer appointment	0.080	0.007
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2. Linear regression models are used to assess which factors are associated with the observed segregation, by regressing FPC1 ξ_1 and FPC2 ξ_2 on party affiliation (Republican or Democrat), Justice's age at appointment, and calendar year at appointment. Coefficients and associated 95% confidence intervals are reported.

Reconstruction of latent policy position processes

The inferred latent ideology processes $\{\hat{X}_i(t) : i = 1 \dots 113\}$ for the 113 Justices, as per Eq (7), transformed into probability scale, as per Eq (8), are shown in Figs 5 - 9 chronically by the time of tenure.

Fig 5. The inferred latent ideology processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8), from James Wilson to John Catron. The estimated latent ideology processes $\{\text{expit}\left(\widehat{X}_i(t)\right) : i = 1 \dots 113\}$ is shown in black and the observed $\hat{p}_i(t)$ is shown in green.

Fig 6. The inferred latent ideology processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8), from John McKinley to David Brewer. The estimated latent ideology processes $\{\text{expit}\left(\widehat{X}_i(t)\right) : i = 1 \dots 113\}$ is shown in black and the observed $\hat{p}_i(t)$ is shown in green.

The current Roberts court

It is of great interest to discuss the ideology makeup of the Roberts Court. Before Justice Ginsburg passed away, the ranking of the policy preferences of the Justices on

Fig 7. The inferred latent ideology processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8), from Henry Brown to Hugo Black. The estimated latent ideology processes $\{\text{expit}\left(\widehat{X}_i(t)\right) : i = 1 \ldots 113\}$ is shown in black and the observed $\hat{p}_i(t)$ is shown in green. Justice Charles Hughes stepped down as the Associate Justice at 1916 and appointed as the Chief Justice at 1930, as a result, there were no observed votes from him between 1916 and 1930, which explains the gap in the observed $\hat{p}_i(t)$.

Fig 8. The inferred latent ideology processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8), from Stanley Reed to John Stevens. The estimated latent ideology processes $\{\text{expit}\left(\widehat{X}_i(t)\right) : i = 1 \ldots 113\}$ is shown in black and the observed $\hat{p}_i(t)$ is shown in green.

Fig 9. The inferred latent ideology processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8), from Sandra Conner to Brett Kavanaugh. The estimated latent ideology processes $\{\text{expit}\left(\widehat{X}_i(t)\right) : i = 1 \ldots 113\}$ is shown in black and the observed $\hat{p}_i(t)$ is shown in green.

the current and recent court from most conservative to most liberal was: Thomas, Alito, Gorsuch, Kavanaugh, Roberts, Breyer, Kagan, Ginsburg and Sotomayor, as indicated by the FPC1 values in Fig 4. Though their policy position trajectories have been relatively stable, except for Chief Justice Roberts who exhibits a tendency to shift toward moderate, their views may still vary both over time and over different issue areas.

Specifically, Cases before the court were labeled according to the following case categories: Criminal procedure, civil rights, First Amendment, due process, privacy, attorneys' or governmental officials' fees or compensation, unions, economic activity, judicial power, federalism, interstate relation, federal taxation, miscellaneous, and private law. The numbers of cases and relative frequencies for each issue area in the aggregate are as follows: Criminal procedure (27,781, 11%), civil rights (23,834, 9.42%), First Amendment (7,175, 2.84%), due process (9,842, 3.89%), privacy (1,220, 0.482%), attorneys' or governmental officials' fees or compensation (2,967, 1.17%), unions (4,705, 1.86%), economic activity (73,684, 29.1%), judicial power (48,239, 19.1%), federalism (7,899, 3.12%), interstate relations (2,383, 0.942%), federal taxation (13,036, 5.15%), miscellaneous (994, 0.393%), and private law (28,248, 11.2%). These labels make it possible to conduct a more detailed analysis. The reconstructed latent policy position processes for selected issue areas are shown in Fig 10.

Rather than focusing on the between-Justice difference in terms of overall policy position trajectories, the Justice-area-specific ideology trajectories, estimated by applying the FPCA method discussed in the Method Section to issue-specific votes, enable the study of within-Justice differences, which is clearly present. The data suggests that there exists substantial variability even for the same Justice across different issues. For each Justice, his or her disposition revealed by votes for civil right, criminal procedure and judicial power cases are more closely aligned with their overall disposition. The Court is relatively more homogeneous regarding economic activity cases, where all nine Justices are closer to the moderate center. The cases with most divergent views seem to be those on federalism and first amendment. Justices Gorsuch and Kavanaugh have extremely conservative voting patterns on federalism cases, while the rest of the Court is mostly liberal, including other Republican Justices. There are more unexpected turns for first amendment cases, for instance, Justices Roberts and Alito have drifted towards a higher percentage of liberal decisions in recent years, and also Justice Gorsuch has been surprisingly liberal in his votes. Such area-specific analysis can assist to identify the drivers of the potential ideology shifts of the Justices.

Fig 10. The inferred latent policy position processes in probability scale $\{\mathbf{expit}\left(\widehat{X}_i(t)\right)\}$. The inferred latent policy position processes in probability scale $\{\text{expit}\left(\widehat{X}_i(t)\right)\}$, as per Eq (7) and Eq (8) of current Justices, estimated using votes stratified by selected issue areas: criminal procedure, civil rights, First Amendment, economic activity and judicial power, federalism.

With the confirmation of Justice Amy Coney Barrett, there is a possible prospect of a shift given the new conservative 6-3 majority. After examining the past, can one utilize the historical data to predict what the ideological makeup of the Court would be like in the next 5 years? Our method has the ability to predict future ideology process for those Justices that have served less than 35 years and are considered as partially observed functional data, as detailed in the Method Section.

According to our model's prediction, as shown in Fig 11, it suggests the following possibilities. The liberal camp – Justices Sotomayor, Breyer and Kagan – will likely to stay where they are, with around 60%-70% of liberal decisions. The solid conservative camp – Justices Thomas and Alito – will also likely to stay where they are, with around 60%-70% of conservative decisions. Justices Gorsuch and Kavanaugh are projected to be moderately conservative. Chief Justice Roberts is projected to position almost at the center of the ideology space, with just a slight conservative inclination. The data-driven prediction should be viewed as suggestion and with caution as the models may not capture case contents and legal changes. Another wild card is the voting behavior of Justice Barrett, although early indicators put her into the same cluster as Justices Alito and Thomas. It also suggests the following ideology position to hold for the near future, from the most conservative to the most liberal (except Justice Barrett): Clarence Thomas, Samuel Alito, Neil Gorush, Brett Kavanaugh, John Roberts, Elena Kagan, Stephen Breyer, and Sonia Sotomayor. With Justice Stephen Breyer's decision to retire at the end of the current term (October term 2021), it gives Democrats opportunity to replace the liberal justice and maintain the current 6-3 conservative majority in the court and also opens up possibilities of ideological shift of the court.

Fig 11. Predicted ideology trajectories in the next 5 years (2021-2015) for the current Court (except Justice Barrett). Color coded by the nominating president's party (blue for Democratic, red for Republican). The solid lines represents estimated ideology trajectories corresponding to time periods during which there are observed voting data, while the dashed lines represents prediction of future ideology trajectories that have yet not realized and based on past voting data.

Discussion

The ideology dynamics of the Supreme Court Justices is of great interest especially under current circumstances when inequality and injustice in race, health, education and in other social aspects are at the frontline of societal discourse.

We have developed a functional data approach to analyze the latent policy position process using observable voting behaviors and show that this yields substantial insights in the time dynamics of policy preferences of the Justices. Our approach is based on Functional Data Analysis (FDA), which is increasingly popular for the analysis of longitudinal studies or panel data (time- series cross-sectional (TSCS) data) and provides a suite of statistical models for analyzing dynamic behavior of stochastic processes, such as the ideology process. It also makes it possible to compress high-dimensional trajectory data into a few concise variables (FPC scores) which are directly interpretable in conjunction with the eigenfunctions, but are also useful for further statistical analysis. For instance, judicial behavior might be an important explanatory factor for other political phenomena, or for the purpose of predicting Supreme Court votes. Our approach complements many works related to judicial ideology and voting behavior.

The proposed approach has the advantages that it is truly dynamic in the sense that it can recover the whole trajectory of the latent policy position process for each Justice, with weak requirements on the available data which allows for discretely observed, noise contaminated data; it provides predictions for individual Justice's ideological trajectories. Fast and timely updates of trajectories are available whenever new votes become available and the approach facilitates the discovery of patterns of change over time. It does not require any a priori restrictive model assumptions or specification of presumed patterns. It is also possible to quantify an individual Justice's trajectory by just specifying two functional principal component scores, where the first score corresponds to a mostly static and the second to a dynamic change component. This also provides a substantial dimension reduction of the otherwise unwieldy longitudinal patterns.

Implementation is straightforward through the established R package fdapace [39]. The methodology more generally provides useful tools to the political science community for the study of time-dynamic phenomena.

Acknowledgments

We would like to thank the reviewers for their helpful comments. This research has been supported in part by NSF grant DMS-2014626.

References

1. Segal JA, Spaeth HJ. The Supreme Court and the attitudinal model revisited. Cambridge University Press; 2002.

2. Eskridge Jr WN. Overriding Supreme Court statutory interpretation decisions. Yale Law Journal. 1991; p. 331–455.

3. Epstein L, Knight J. The choices Justices make. Sage; 1997.

4. Schubert GA. The judicial mind: The attitudes and ideologies of Supreme Court Justices, 1946-1963. Northwestern University Press; 1965.

5. Schubert G. The judicial mind revisited: Psychometric analysis of Supreme Court ideology. 1974;.

6. Segal JA, Cover AD. Ideological values and the votes of US Supreme Court Justices. The American Political Science Review. 1989; p. 557–565.

7. Epstein L, Walker TG, Dixon WJ. The Supreme Court and criminal Justice disputes: A neo-institutional perspective. American Journal of Political Science. 1989; p. 825–841.

8. Rohde DW, Spaeth HJ. Supreme Court Decision Making. American Political Science Review. 1978; p. 283–284.

9. Tate CN, Handberg R. Time binding and theory building in personal attribute models of Supreme Court voting behavior, 1916-88. American Journal of Political Science. 1991; p. 460–480.

10. Segal JA, Spaeth HJ The Supreme Court and the attitudinal model. Cambridge University Press New York. 1993.

11. Poole KT, Rosenthal H. Congress: A political-economic history of roll call voting. Oxford University Press on Demand; 2000.

12. Londregan J. Estimating legislators' preferred points. Political Analysis. 1999;8(1):35–56.

13. Jackman S. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. Political Analysis. 2001;9(3):227–241.

14. Bailey M, Chang KH. Comparing presidents, senators, and Justices: interinstitutional preference estimation. Journal of Law, Economics, and Organization. 2001;17(2):477–506.

15. Martin AD, Quinn KM. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999. Political Analysis. 2002;10(2):134–153. doi:10.1093/pan/10.2.134.

16. Katz DM, Bommarito MJ, Blackman J. A general approach for predicting the behavior of the Supreme Court of the United States. PloS one. 2017;12(4):e0174698.

17. Katz DM, Bommarito MJ, Blackman J. Justice blocks and predictability of us supreme court votes. PloS one. 2011;6(11):e27188.

18. Ramsay JO. Functional data analysis. Encyclopedia of Statistical Sciences. 2004;4.

19. Wang JL, Chiou JM, Müller HG. Functional Data Analysis. Annual Review of Statistics and Its Application. 2016;3(1):257–295.

20. Ullah S, Finch CF. Applications of functional data analysis: A systematic review. BMC medical research methodology. 2013;13(1):43.

21. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. 2005;100(470):577–590.

22. Sentürk D, Müller HG. Functional varying coefficient models for longitudinal data. 2010;105:1256–1264.

23. Liu S, Zhou Y, Palumbo R, Wang JL. Dynamical correlation: A new method for quantifying synchrony with multivariate intensive longitudinal data. Psychological Methods. 2016;21(3):291–308.

24. Zhang X, Wang JL. Optimal weighting schemes for longitudinal and functional data. Statistics & Probability Letters. 2018;138:165–170.

25. Harold Spaeth LE. 2020 Supreme Court Database, Version 2020 Release 1. URL: `http://Supremecourtdatabase.org`; 2020.

26. Hsing T, Eubank R. Theoretical foundations of functional data analysis, with an introduction to linear operators. vol. 997. John Wiley & Sons; 2015.

27. Segal JA, Epstein L, Cameron CM, Spaeth HJ. Ideological Values and the Votes of U.S. Supreme Court Justices Revisited. The Journal of Politics. 1995;57(3):812–823.

28. Epstein L, Martin AD, Quinn KM, Segal JA. Ideological drift among Supreme Court Justices: Who, when, and how important. Northwestern University Law Review. 2007;101(4):1483–1542.

29. Martin AD, Quinn KM. Assessing preference change on the U.S. Supreme Court. Journal of Law, Economics, and Organization. 2007;23(2):365–385.

30. Bailey MA, Maltzman F. Does Legal Doctrine Matter? Unpacking Law and Policy Preferences on the U.S. Supreme Court. American Political Science Review. 2008;102(3):369–384.

31. Bartels BL, O'Geen AJ. The Nature of Legal Change on the U.S. Supreme Court: Jurisprudential Regimes Theory and Its Alternatives. American Journal of Political Science. 2014;59:880–895.

32. Klemonski D, Natarajan O, Studnitzer S, Sommers P. Ideological Reversal among Supreme Court Justices. Open Journal of Social Sciences. 2017;5(7):290–299.

33. Müller HG. Weighted local regression and kernel methods for nonparametric curve fitting. 1987;82(397):231–238.

34. Fan J, Gijbels I. Local polynomial modelling and its applications: monographs on statistics and applied probability 66. vol. 66. CRC Press; 1996.

35. Hall P, Müller HG, Yao F. Modeling sparse generalized longitudinal observations with latent Gaussian processes. Journal of the Royal Statistical Society B. 2008;70:730–723.

36. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. Journal of the American Statistical Association. 2005;100(470):577–590.

37. Chen K, Zhang X, Petersen A, Müller HG. Quantifying Infinite-Dimensional Data: Functional Data Analysis in Action. Statistics in Biosciences. 2015; p. 1–23.

38. Wang JL, Chiou JM, Müller HG. Functional Data Analysis. Annual Review of Statistics and its Application. 2016;3:257–295.

39. Chen Y, Carroll C, Dai X, Fan J, Hadjipantelis PZ, Han K, et al.. fdapace: Functional Data Analysis and Empirical Dynamics; 2020.

40. Yao F, Müller HG, Wang JL. Functional linear regression analysis for longitudinal data. The Annals of Statistics. 2005; p. 2873–2903.

41. Carroll C, Bhattacharjee S, Chen Y, Dubey P, Fan J, Gajardo A, Zhou, X, Müller HG, Wang JL. Time Dynamics of COVID-19. Scientific reports. 2020;.