

The Impact of Judicial Selection Methods on Opinion Writing Behavior

Alexandria Nagy

Abstract

Insert abstract here

Introduction

Hall examines the tension between electoral accountability and judicial independence in partisan versus nonpartisan elections. Judicial reform advocates argue that partisan judicial elections fail to promote accountability because they lack substantive issue content and meaningful evaluation of judicial performance, and they favor nonpartisan elections and retention systems under the Missouri Plan as better preserving judicial independence by encouraging voters to assess candidates based on professional qualifications. Hall's findings complicate these initiatives, indicating that nonpartisan and retention races are less insulated from partisan politics and contextual forces than assumed. Despite the absence of party labels on the ballot, statewide partisan competition and election-specific challenges significantly affect incumbent vote share. Additionally, Hall finds that partisan elections are far more substantive than assumed, generating electoral accountability by providing incentives for judges to remain responsive to voters. In these systems, greater ideological distance between voters and justices systematically reduces incumbent support, and increases in the murder rate significantly depress electoral performance. These findings suggest that the strongest

form of electoral accountability exists in partisan judicial elections.

Literature Review

Scholarly research on judicial dissent has identified several categories of factors that influence opinion-writing behavior on state supreme courts. These include institutional structures that govern court operations, electoral systems that determine how judges are selected and retained, internal court dynamics that shape collegial relations, and environmental conditions that reflect broader political and social contexts. Understanding these mechanisms is essential for evaluating how reforms to judicial selection systems affect judicial independence and accountability.

Institutional Structures and Dissent

Institutional factors that dictate court structures and internal rules exert substantial influence on dissent rates. Jaros and Canon find that states with intermediate appellate courts exhibit much higher dissent rates in their supreme courts, the probability of dissent increases with court size, and shorter judicial terms are associated with higher dissent, as judges may be more responsive to their constituencies in anticipation of reelection. These structural features operate independently of individual judicial preferences. Brace and Hall find that four institutional features shape dissent rates: the method of judicial selection, the presence of intermediate appellate courts, random opinion assignment, and sanctioning rules. The presence of a lower-level appellate court increases dissent by filtering routine cases, leaving the high court with more complex or controversial disputes. Randomly assigning cases for opinion writing promotes dissent by limiting the Chief Justice's ability to reward or

punish colleagues through discretionary assignments. Procedures such as opinion assignment, conference discussion by seniority, and voting in reverse seniority promote consensus by creating incentives that discourage open disagreement. Brace and Hall conclude that while both environmental and institutional factors matter, institutional factors are far stronger predictors of dissent. Variation in dissent rates explained by institutional factors is over six times greater than that explained by environmental variables.

Internal Court Dynamics and Consensus Formation

Beyond formal structures, internal dynamics shape the extent to which courts produce consensus or permit dissent. Jaros and Canon identify consensus-producing variables, including strong task leadership by the Chief Justice, high social integration among members, and a “norm of unanimity” that encourages judges to reach a single collective decision. Epstein, Landes, and Posner find that judges may suppress disagreement to avoid imposing costs on colleagues or themselves, suggesting that dissent aversion reflects both institutional pressures and interpersonal considerations. Hall and Windett further examine how the formal powers and institutional resources of chief justices impact dissenting behavior in state supreme courts. Chief justices often value consensus because it fosters collegiality, reinforces precedent, and enhances public confidence. They discourage dissent because it imposes costs on colleagues and has the potential to weaken precedent, confuse the law, encourage future appeals, and generate dissatisfaction on the court. By contrast, other justices resist consensus, viewing dissents as a means to preserve individual expression, influence future courts, or accommodate bargaining that produces more ambiguous opinions. These competing incentives limit the chief’s ability to enforce agreement. Hall and Windett find that dissent rates are lower

on courts where the chief possesses greater formal powers and where members have fewer institutional resources such as staff, pay, docket control, time, or insulation from electoral pressures. Member resources primarily encourage dissent by undermining the chief's authority, making formal powers less effective on resource-rich courts.

Environmental Factors and Political Context

Environmental factors reflecting broader political and social contexts also influence dissent rates. Brace and Hall find that higher urbanization, greater political competition, and higher state spending are all associated with increased dissent rates. These variables capture the complexity and diversity of state political environments, suggesting that courts in more heterogeneous and politically competitive contexts face greater pressure to accommodate divergent legal interpretations. Epstein, Landes, and Posner examine how ideology and panel composition interact with dissent incentives in federal courts. Judges appointed by presidents of different parties have systematically different ideological preferences, increasing the likelihood of disagreement on mixed panels. Their model predicts that dissent becomes more likely as ideological divisions grow, panel sizes increase, and aversion to dissent declines. These findings indicate that environmental conditions, particularly ideological heterogeneity, create structural pressures that facilitate dissent even when institutional rules encourage consensus.

Judicial Selection Systems and Electoral Accountability

The method of judicial selection represents one of the most consequential institutional factors shaping dissent rates. Brace and Hall find that judges appointed by governors

or legislatures promote consensus, while elected judges are associated with higher dissent, potentially due to the recruitment of more risk-tolerant judges. Jaros and Canon similarly find that popularly elected courts show higher dissent rates than those where judges are appointed. Hall and Windett provide a systematic comparison of dissent rates across four judicial selection systems: appointment by the governor or state legislature, the Missouri Plan in which judges are appointed and later face retention elections, nonpartisan elections in which candidates appear on the ballot without party affiliations, and partisan elections in which candidates appear with party affiliations. Courts with appointed judges exhibit highly stable dissent rates averaging roughly 10 to 12 percent between 1995 and 2009. The authors attribute this stability to longer judicial tenures, higher professionalization, and insulation from electoral pressure. Missouri Plan courts show similarly stable average dissent rates of approximately 12 to 18 percent, although there is substantial cross-state variation. States that employ nonpartisan judicial elections show significantly higher dissent rates, approximately four percentage points above those in Missouri Plan states, indicating that even when party labels are absent, electoral incentives may shape judicial behavior. The most pronounced and unstable patterns appear in states that use partisan judicial elections. In these states, dissent rates rise sharply from under 20 percent in the mid-1990s to over 40 percent by the early 2000s, accompanied by a dramatic increase in cross-state variation. Hall and Windett conclude that this volatility may indicate that something beyond judicial interpretation is driving these differences, and that judicial decision-making in these states may be influenced by reelection concerns.

Electoral Reform and Strategic Behavior

Renberg employs synthetic control analysis to examine the effects of removing partisan identification from voter guides and reelection ballots on opinion-writing behavior in state supreme courts, measured by the frequency of dissenting opinions.¹ This shift reflects a strategic response to electoral reform, as justices seek to enhance their reelection prospects. Courts of last resort often hear high-profile cases, and dissenting opinions can signal ideology to political elites, interest groups, the media, and the public. Unlike majority opinions, which are typically products of compromise and deliberation, dissents allow justices to communicate their views clearly and cultivate a reputation for principled independence among court observers. When partisan labels are present, justices may moderate their behavior to align with party ideals. Removing these labels allows them to author dissents in areas where their legal ideology diverges from partisan expectations.

Methods

This study adopts a mixed-methods research design that integrates quantitative statistical modeling with qualitative textual analysis to assess the effects of institutional reform on judicial opinion writing. The quantitative component employs the synthetic control method, originally developed by Alberto Abadie and Javier Gardeazabal and later extended by Abadie, Alexis Diamond, and Jens Hainmueller, to estimate the causal effects of aggregate institutional interventions.² The method constructs a counterfactual by forming a weighted

¹Kristen M. Renberg, “The Impact of Retention Systems on Judicial Behavior: A Synthetic Controls Analysis of State Supreme Courts,” *The Justice System Journal* 41, no. 4 (2020): 292–312, <https://www.jstor.org/stable/27224753>.

²Alberto Abadie and Javier Gardeazabal, “The Economic Costs of Conflict: A Case Study of the Basque Country,” *American Economic Review* 93, no. 1 (2003): 113–32, <https://doi.org/10.1257/000282803321455188>;

composite of control units, or donor pool, selected to closely replicate the treated unit’s pretreatment characteristics and project its trajectory had the intervention not occurred.

In the present study, the aggregate intervention of interest is the method of judicial selection. The treated unit is the North Carolina Supreme Court, which shifted from nonpartisan to partisan judicial elections following a statutory reform in 2017, with the first partisan contest held on November 6, 2018. A synthetic North Carolina Supreme Court is constructed using a donor pool of pretreatment covariates collected from state supreme courts that have historically employed nonpartisan judicial elections, including Arkansas, Georgia, Kentucky, Minnesota, Montana, Oregon, and Wisconsin. The synthetic controls approach is designed to approximate how judicial dissent would have evolved on the North Carolina Supreme Court in the absence of this institutional reform. Renberg finds that synthetic control methods provide a credible estimate of the causal effect of institutional changes on opinion-writing behavior.

The statistical aspect of the study employs an original dataset of majority and dissenting opinions authored by justices of the North Carolina Supreme Court. The dependent variable, dissent rate, is defined as the annual proportion of published majority opinions that are accompanied by at least one dissenting opinion. Reflecting the institutional scope of the reform and its uniform application to all justices serving on the bench, the unit of analysis is at the court level. Dissent rates are calculated by dividing the number of majority opinions that include a dissent by the total number of published majority opinions issued by the court in a given year. Data related to the were collected from LexisNexis for the period

Alberto Abadie et al., “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” *Journal of the American Statistical Association* 105, no. 490 (2010): 493–505, <https://doi.org/10.1198/jasa.2009.ap08746>.

spanning November 6, 2012, to November 6, 2024. This twelve-year window captures six years before and six years after North Carolina’s transition from nonpartisan to partisan judicial elections. Dissenting opinions were identified using LexisNexis’s DissentBy search function, incorporating the surnames of all justices who served on the court during the study period. The total number of published majority opinions was obtained using LexisNexis’s Find by Source function.

Pretreatment Characteristics

Table 1: Pretreatment Characteristics

Variable	Operationalization	Source Initiative
Term Length	Count of years between judicial elections.	Ballotpedia
Number of Justices	Count of justices on the bench.	Ballotpedia
Professionalization Score	Professionalism index based on the Court Statistics Project.	Squire (2021)
Campaign Fundraising	Amount, in dollars, of reported campaign funds raised.	Open Secrets
Number of Capital Punishment Cases	Count of capital punishment cases reviewed.	Lexis Nexis
Number of Lower Court Capital Punishment Cases	Count of capital punishment cases reviewed in the previous year by lower courts.	Lexis Nexis
Number of Published Opinions	Count of reported opinions.	Lexis Nexis
Single-Member Election District	Valued at one if the state uses single-member districts and valued at zero otherwise.	Ballotpedia
Percent of Docket Concerning Criminal Procedure	Number of reported criminal procedure cases divided by the number of reported cases.	Lexis Nexis
Competitive Election	Valued at one if the average victory margin in the most recent election was less than 10% and zero otherwise.	Ballotpedia
Ideological Spread	Absolute difference in cfscores between the most and least liberal judge in the court.	Bonica (2024)
State Citizen Ideology	Mean ideological self-placement of all respondents in a state, aggregated from individual ANES survey responses.	ANES Time Series Study (2024)
State Government Ideology	Mean of first-dimension NOMINATE scores for congressional members (House and Senate)	Voteview (2026)

Each donor court is characterized by a vector of pretreatment covariates that capture

political, legal, and institutional context. All pretreatment covariates are measured annually for each donor court across the study period from 2012 to 2024. Structural features of judicial selection systems, including term length, number of justices on the bench, single-member versus multimember election districts, and electoral competitiveness, were obtained from Ballotpedia.³ Campaign finance data, operationalized as the total amount of funds raised by candidates in the most recent election for each year in the treatment window, were collected from OpenSecrets.⁴ These data incorporate information originally compiled by the Follow the Money Institute, which merged with OpenSecrets in 2019.⁵

Court professionalization scores are drawn from Squire and Butcher’s 2019 index, an updated version of the 2008 professionalism measure that incorporates staff size, judicial pay, and docket control based on the Court Statistics Project’s classification of mandatory and discretionary jurisdiction.⁶ Given the relative stability of these institutional characteristics, the 2019 values are assumed to remain substantively unchanged through 2024. Caseload characteristics, including the number of published opinions, capital punishment appeals reviewed by the court, capital punishment cases resolved by lower courts in the previous year, and the percentage of the docket devoted to criminal procedure, were collected from LexisNexis.⁷ The ideological spread of each court is measured using Bonica’s common-space campaign finance scores (CFscores). These scores are drawn from the 2024 Database on

³“State Supreme Courts,” in *Ballotpedia*, https://ballotpedia.org/State_supreme_courts, n.d.

⁴“OpenSecrets,” in *OpenSecrets*, <https://www.opensecrets.org/>, n.d.

⁵*Home - FollowTheMoney.org*, <https://www.followthemoney.org/>, n.d.

⁶Peverill Squire and Jordan Butcher, “An Update to the Squire State Court of Last Resort Professionalization Index,” *State Politics & Policy Quarterly* 21, no. 3 (2021): 326–33, <https://doi.org/10.1017/spq.2020.7>; Peverill Squire, “Measuring the Professionalization of U. S. State Courts of Last Resort,” *State Politics & Policy Quarterly* 8, no. 3 (2008): 223–38, <https://www.jstor.org/stable/40421611>.

⁷*NexisUni*, <https://advance-lexis-com.turing.library.northwestern.edu/firsttime?crd=151f04a4-41f3-4caa-9fd6-1ad2972659d7>; LexisNexis, n.d.

Ideology, Money in Politics, and Elections (DIME).⁸

The original measures used in Renberg’s analysis for state-level citizen and government ideology, developed by Berry et al., are unavailable for 2012–2024.⁹ State citizen ideology is therefore approximated using aggregated responses from the 2024 American National Election Studies (ANES) Time Series Study, which reports state-level mean ideological self-placement every four years, aligned with election cycles. For state government ideology, NOMINATE scores provide a consistent, spatially comparable measure of elite ideology that strongly correlates with Berry et al.’s original measure.¹⁰ Accordingly, state government ideology is operationalized as the average first-dimension NOMINATE score of all congressional members (House and Senate) in each state.¹¹

Word Embedding Model for Evidence Minus Intuition (EMI) Analysis

To construct a dictionary for Evidence Minus Intuition (EMI) literary analysis of judicial opinions, I trained a Word2Vec model on the Collaborative Open Legal Data (COLD) Cases dataset. COLD Cases is a comprehensive corpus of 8.3 million United States legal decisions maintained by the Harvard Library Innovation Lab in collaboration with the Free Law Project. This dataset consolidates semantic information from CourtListener’s bulk data, including the text of majority and dissenting opinions, head matter, and substantive

⁸*Database on Ideology, Money in Politics, and Elections (DIME): Public Version 4.0 / Stanford Libraries Social Science Data Collection*, <https://data.stanford.edu/dime>, n.d.

⁹William D. Berry et al., “Measuring Citizen and Government Ideology in the American States, 1960-93,” *American Journal of Political Science* 42, no. 1 (1998): 327–48, <https://doi.org/10.2307/2991759>.

¹⁰*Measuring Citizen and Government Ideology in the U.S. States: A Re-appraisal - Northwestern University*, <https://search.library.northwestern.edu>, n.d.

¹¹Jeffrey B. Lewis et al., “Voteview: Congressional Roll-Call Votes Database,” in *Voteview: Congressional Roll-Call Votes Database*, <https://voteview.com/>, 2026.

metadata, into standardized records with extraneous data removed.

The Word2Vec model was trained using the Gensim implementation, provided by the Gensim Python library and widely utilized for natural language processing (NLP), to learn vector representations of words based on their contextual usage within supreme court opinions. Word2Vec employs a shallow neural network architecture, a relatively simple machine learning model with a single hidden layer, to embed words in a lower-dimensional vector space where semantically similar words occupy proximate positions. The embedding process converts words into numerical vectors, or lists of numbers. For example, the word “justice” might be represented as $[0.23, -0.45, 0.78, \dots]$. Each vector has a set number of components, or “dimensions” (typically 100–300), which together capture patterns of how words appear in relation to one another. These dimensions do not correspond to individual words; rather, each dimension encodes an abstract feature learned from word co-occurrence patterns. In this vector space, words that appear in similar contexts—such as “justice” and “judge”—are located near each other. In this vector space, semantically similar words occupy proximate positions, enabling the model to capture meaningful relationships between vocabulary terms present in supreme court opinions. This approach addresses limitations of traditional bag-of-words models, which lose information about word order and fail to capture semantic relationships. The resulting word embeddings serve as the foundation for the EMI dictionary, enabling systematic identification of language patterns in judicial opinions based on data-driven representations of legal terminology.

loss function

Analysis

Dissent Rates

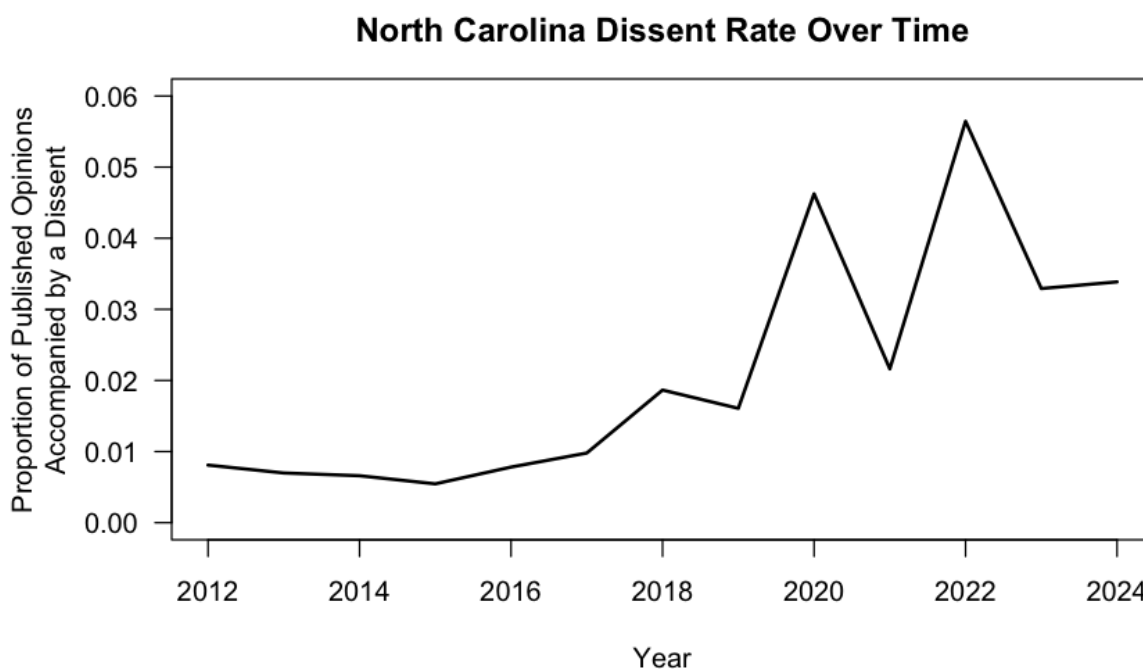


Figure 1: Dissent Rates in North Carolina Supreme Court

This synthetic control analysis follows Renberg’s methodological approach to examine the impact of judicial electoral reform on dissenting opinion rates in North Carolina’s Supreme Court.¹² Complete data for all variables and courts analyzed in this study are available in the replication code. North Carolina’s dissent rate remained consistently low from 2012 to 2017, never exceeding 1%. The rate increased notably in 2018 to 1.9% and reached 4.5% in 2020. Coinciding with periodic court closures across the United States related to the onset of the COVID-19 pandemic, dissent rates declined to 2.2% in 2021. Subsequently, the rate rose to a recorded maximum of 5.6% in 2022 before declining to approximately 3.1% by 2023 and 2024. In contrast, other courts in the sample exhibited more variable patterns. Wisconsin

¹²Renberg, “The Impact of Retention Systems on Judicial Behavior.”

consistently maintained the highest dissent rates throughout the period, with rates frequently exceeding 50% and reaching 88% in 2016. This could be due to a high volume of cases with unpublished opinions not captured in the data collection process, resulting in an artificially inflated dissent rate when calculated against only published opinions. Arkansas also showed considerable volatility, with dissent rates ranging from 7.7% in 2012 to peaks above 50% in 2018-2020. States such as Georgia and Minnesota maintained relatively stable and low dissent rates throughout the observation period, typically remaining below 10%.

In order to evaluate if these dissent rate values are consistent with established expectations, it is necessary to address reproducibility concerns. Renberg does not explicitly report baseline dissent rates for any of the courts analyzed. Although she presents treatment effects numerically (a 7.5% increase in Arkansas, 5.1% in Tennessee, and 20.8% in Mississippi), these figures are not anchored to specific baseline values. Consequently, readers must rely on graphical representations to infer initial dissent rates, which introduces substantial ambiguity. Renberg's figures label the y-axis as "Percent of Opinions" while displaying decimal values ranging from 0.05 to 0.25, creating confusion about whether the values represent proportions (e.g., $0.15 = 15\%$) or percentages (e.g., $0.15 = 0.15\%$). Standard practice would interpret these as proportions, suggesting dissent rates ranged from 14 to 19% on the Arkansas Supreme Court between 1997 and 2012, from 5 to 22% on the Tennessee Supreme Court between 1985 and 2002, and 8 to 46% on the Mississippi Supreme Court between 1986 and 2000. These reported dissent rates are notably high relative to existing judicial opinion-writing behavior literature. By contrast, the literal interpretation would indicate rates of 0.14 to 0.19 percent, which would be notably low. The former interpretation appears correct given Renberg's description of Mississippi experiencing "an increase of over 30% for some years following

electoral reform,” and her reported average treatment effects.

Regardless of which interpretation was intended, these dissent rates could not be reproduced using publicly available court data. Multiple attempts were made to replicate these figures using data from CourtListener and LexisNexis. In CourtListener, replication efforts included filtering by court and employing advanced query techniques and operators, including combined-opinion, unanimous-opinion, lead-opinion, plurality-opinion, concurrence-opinion, in-part-opinion, dissent, addendum, remittitur, rehearing, on-the-merits, and on-motion-to-strike. Replication attempts using LexisNexis utilized queries with the OpinionBy and DissentBy fields, as well as the “find by source” option.

Replication efforts also followed Hall and Windett’s 2013 study.¹³ Table 2 presents three sets of values generated using different LexisNexis query strategies. The first set, labeled “Hall & Windett,” reproduces the values from the original study’s published table. For the second set, labeled “OpinionBy & DissentBy,” I constructed LexisNexis queries using the Opinionby and Dissentby functions with the names of justices serving on each court between 1995 and 2010 in an attempt to replicate Hall and Windett’s original values. The third set, labeled “AND NOT,” reflects a more restrictive query approach. Because Hall and Windett claim to focus their data on “the subset of data most likely to interest judicial politics scholars: those cases in which a state supreme court rendered a final judgment with a full written opinion,” this final query used AND NOT operators to exclude entries containing “dismiss!”, “petition for review”, or “motion for.” As Table 3 demonstrates, despite these systematic efforts to replicate the original methodology using various LexisNexis search

¹³Matthew E. K. Hall and Jason Harold Windett, “New Data on State Supreme Court Cases,” *State Politics & Policy Quarterly* 13, no. 4 (2013): 427–45, <https://www.jstor.org/stable/24710959>.

strategies, the resulting dissent rates could not be reproduced at levels comparable to those reported by Hall and Windett or Renberg. The variation across query methods is particularly pronounced in certain jurisdictions, such as Oregon and West Virginia, where dissent rates differ substantially depending on the search approach employed.

Court	Hall & Windett Opinions	Hall & Windett Dissents	OpinionBy Opinions	DissentBy Dissents	AND NOT Opinions	AND NOT Dissents
NC	13241	154	15660	159	10687	156
OR	1155	143	1074	325	331	91
WV	2218	765	1253	678	351	163
MO	1874	267	3271	533	484	75
MA	2808	172	2731	534	780	152
ME	3233	227	2464	408	1002	136
IL	2302	616	1685	630	399	141
ID	2119	244	2141	439	551	110
HI	2086	282	1768	389	661	86
AZ	1546	167	2802	241	336	39
AK	2506	272	1923	269	543	66
AL	5185	1423	4474	1418	1626	563

Table 2: Opinions and dissents by court and method

Court	Hall & Windett	OpinionBy & DissentBy	AND NOT
NC	0.0116	0.0102	0.0146
OR	0.1238	0.3026	0.2749
WV	0.3449	0.5411	0.4644
MO	0.1425	0.1629	0.1550
MA	0.0613	0.1955	0.1949
ME	0.0702	0.1656	0.1357
IL	0.2676	0.3739	0.3534
ID	0.1151	0.2050	0.1996
HI	0.1352	0.2200	0.1301
AZ	0.1080	0.0860	0.1161
AK	0.1085	0.1399	0.1215
AL	0.2744	0.3169	0.3462

Table 3: Dissent rates by court and method

Data availability further constrained efforts to replicate Renberg’s findings. The State

Supreme Court Data Project remains the premier database for state supreme court research, but it covers only four years between 1995 and 1998. In 2013, Hall and Windett sought to address this deficiency by presenting a new dataset of state supreme courts ostensibly made publicly available to the scholarly community.¹⁴ Unfortunately, their data are not accessible online due to a broken link on their website, and their Python code involves web scraping expressly prohibited by LexisNexis’s terms of service. Although LexisNexis provides broader access to published opinions, its proprietary nature limits transparency and reproducibility for researchers without expensive institutional-level subscriptions. CourtListener, one of the few open-source repositories available for state supreme court research, contains substantially less comprehensive coverage of published opinions than LexisNexis.

Consequently, even when replication attempts rely on the most comprehensive publicly accessible sources, data limitations introduce uncertainty into dissent rate estimates and complicate direct comparison with studies based on non-public or undisclosed datasets. The FAIR principles aim to make scientific data Findable, Accessible, Interoperable, and Reusable.¹⁵ Data are findable when they have persistent identifiers and rich metadata. They are accessible when they can be retrieved through standardized protocols. They are interoperable when they use standardized formats and vocabularies that allow integration with other datasets. They are reusable when they are well-described, with clear provenance, context, and licensing information. Together, these principles promote transparency, reproducibility, and efficient reuse of scientific data. Neither Hall and Windett’s nor Renberg’s studies meet these standards, as their underlying data and replication codes are not publicly available. This

¹⁴Hall and Windett, “New Data on State Supreme Court Cases.”

¹⁵Mark D. Wilkinson et al., “The FAIR Guiding Principles for Scientific Data Management and Stewardship,” *Scientific Data* 3, no. 1 (2016): 160018, <https://doi.org/10.1038/sdata.2016.18>.

prevents independent verification of their results and limits cumulative knowledge building in the study of judicial opinion-writing behavior.

Synthetic Control

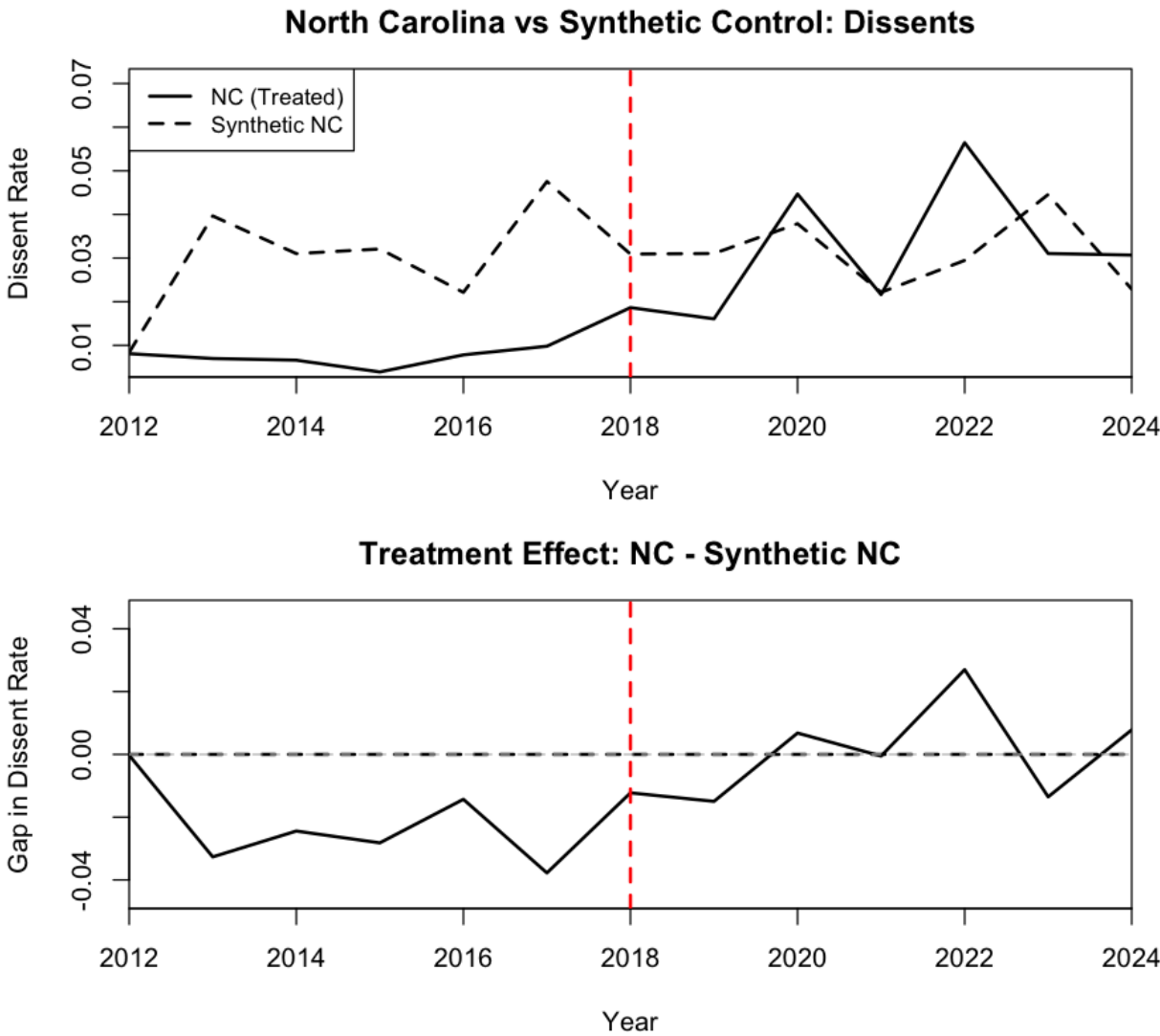


Figure 2: Estimated impact of judicial reform on dissenting behavior

The synthetic control method constructs a counterfactual North Carolina Supreme Court using Minnesota (100% weight) to match North Carolina’s pretreatment characteristics. The algorithm’s selection of Minnesota alone, rather than a weighted composite of all control

courts, suggests that Minnesota provides the best single-court match to North Carolina’s pretreatment trajectory among available donor courts. The algorithm assigned zero weight to all other potential donor courts, including Arkansas, Georgia, Kentucky, Minnesota, Montana, Oregon, and Wisconsin. The analysis yields a null result: the estimated average treatment effect on dissents is 0.000%, indicating no detectable change in dissenting behavior following North Carolina’s transition to partisan elections in 2018.

Model Fit and Parallel Trends Assumption

Mean Squared Prediction Error (MSPE) measures the average squared difference between the treated unit’s actual outcome and the synthetic control’s predicted outcome. The synthetic control method rests on the parallel trends assumption that the treated and synthetic units would have continued on the same trajectory absent treatment. The pretreatment MSPE from 2012-2017 is 0.0007, indicating that Minnesota’s dissent trajectory closely tracked North Carolina’s before treatment. The post-treatment MSPE from 2018-2024 is 0.0002, representing a substantial decrease. The post-to-pre ratio of 0.29 indicates that the synthetic control’s predictive ability actually improved after 2018. This improvement in post-treatment fit suggests that changes occurring in North Carolina following electoral reform also occurred in Minnesota. Thus, the synthetic control yielded a near-zero treatment effect estimate.

Pretreatment Covariate Balance

Pretreatment covariate balance is essential for establishing the validity and robustness of the synthetic control method because it determines whether the synthetic control unit credibly represents the counterfactual outcome for the treated unit. The fundamental logic of

SCM rests on the assumption that, if a weighted combination of control units closely matches the treated unit on observable characteristics prior to treatment, it will also approximate how the treated unit would have behaved in the absence of treatment. Poor pretreatment balance undermines this core assumption, indicating that the synthetic control is constructed from fundamentally dissimilar units and therefore cannot serve as a valid counterfactual.

Table 4: Covariate Balance Between Treated and Synthetic North Carolina

Covariate	Treated (NC)	Synthetic NC	Sample Mean
Campaign Finance	1,348,421	177,336	381,739
Court Professionalization (2019)	0.609	0.610	0.612
Capital Appeals	0.167	0.333	2.548
Lower Court Capital Appeals (Lag 1)	1.000	0.500	1.548
Criminal Procedure Docket	0.458	0.129	0.519
Ideological Spread	2.283	1.964	1.291
Citizen Ideology	3.903	4.161	3.933
Government Ideology	0.240	-0.093	0.147
Term Length	8.000	6.000	7.429
Number of Justices	7.000	7.000	7.286
Election Structure	0.000	0.000	0.143
Electoral Competition	0.667	0.000	0.167
Published Opinions	1,057	657	224

The importance of covariate balance becomes clear when comparing the imbalances in Renberg’s analysis with those in the present study. In Renberg’s analysis, structural and institutional covariates are generally well-balanced, with minimal differences in term length, closely matched single-member election districts, near-identical professionalization scores, and reasonably aligned ideological measures. However, substantial imbalances emerge in operational and caseload variables. Synthetic controls overestimate published opinions by 92-494%, producing two to six times more opinions than the treated courts. Criminal procedure dockets are underestimated by approximately 50% in Arkansas and Mississippi, and capital punishment caseloads diverge sharply, with treated courts handling 161% more

cases in Arkansas and 252% more in Mississippi. Because these variables directly reflect judicial workload and case composition, such imbalances undermine the validity of Renberg’s synthetic controls. This point that is not addressed at length in her paper.

In the present study, the synthetic control achieves near-perfect balance on court professionalization (0.610 vs. 0.609) and exactly matches both the number of justices and the election structure. Term length shows reasonable balance, differing by only two years. However, substantial imbalances persist on key covariates. Campaign finance exhibits a 660% difference (1.35 million vs. 177,336), criminal procedure dockets diverge markedly (0.458 vs. 0.129), and capital appeals also differ notably (0.167 vs. 0.333). Most critically, electoral competition demonstrates complete imbalance: North Carolina’s value of 0.667 reflects highly competitive elections, whereas Minnesota’s 0.000 indicates no competitive elections during the study period. This discrepancy directly undermines the theoretical mechanism under investigation, as competitive elections are hypothesized to drive dissenting behavior.

Standardized difference calculations further highlight these issues: 11 of 14 covariates exhibit imbalances exceeding the 0.25 threshold. The most severe imbalances occur in published opinions (3.59 SD), campaign finance (2.12 SD), ideological spread (1.59 SD), electoral competitiveness (1.33 SD), dissent rate (-0.78 SD), and capital appeals (-0.67 SD). Only three covariates can be considered balanced: lagged lower court capital appeals (-0.18 SD), citizen ideology (-0.09 SD), and court professionalization (-0.03 SD). These general imbalances raise fundamental questions about whether the synthetic control adequately captures the political and institutional pressures shaping judicial opinion-writing in North Carolina. The analysis cannot overcome the absence of a suitable counterfactual court that simultaneously reflects North Carolina’s competitive elections, campaign finance environment,

ideological composition, and caseload characteristics.

Additionally, the algorithm relies entirely on Minnesota as the sole donor court, assigning it 100% weight while excluding all other potential donors. This raises a fundamental concern: Minnesota is merely the least unsuitable option among courts that are otherwise incomparable. While the pretreatment MSPE of 0.0007 may suggest a close fit, it obscures the reality that no available donor court adequately reflects North Carolina’s institutional characteristics. By assigning full weight to a single court, the algorithm abandons the method’s core advantage of combining multiple control units to construct a robust counterfactual. This comparability problem likely stems from structural features that distinguish North Carolina’s judicial system. Unlike most states, North Carolina’s Supreme Court exercises both mandatory and discretionary jurisdiction over certain cases that bypass the Court of Appeals entirely. The state constitution grants direct appellate jurisdiction over constitutional questions, decisions striking down state statutes, and other matters of substantial public importance. This bifurcated structure may inflate North Carolina’s published opinion volume. As shown in Table 2, North Carolina produced 1,057 opinions annually during the pretreatment period, compared to Minnesota’s 657 and a sample mean of 224. The resulting standardized difference of 3.59 standard deviations, the largest imbalance in the dataset, reflects fundamental structural differences in how cases reach the North Carolina supreme court. These findings raise concerns about the use of SCM to study judicial opinion-writing behavior in state supreme courts. Institutional characteristics vary so widely across states that they can violate SCM assumptions.

Sensitivity Analysis

A robust causal estimate should remain stable across reasonable variations in model specification. Conducted by systematically excluding each donor court one at a time and recalculating the treatment effect, the leave-one-out analysis reveals severe instability that undermines confidence in the estimate. Excluding each potential donor produces treatment effects ranging from -0.1887 (Minnesota excluded) to -0.0875 (Wisconsin excluded), with a mean of -0.1617 and standard deviation of 0.0382. Every leave-one-out specification produces a negative treatment effect, yet the optimal weighted combination yields exactly zero. This pattern indicates that the zero treatment effect is an artifact of the specific weighting scheme rather than a robust finding.

Several additional robustness checks illuminate the fragility of causal inference in this setting. First, comparing optimal synthetic control weights to equal weights reveals no difference. Both produce an estimated average treatment effect (ATE) of 0.00, indicating no mean difference in dissent rates between North Carolina and the synthetic control across all post-treatment years. This equivalence suggests that the sophisticated weighting algorithm provides no advantage over a simple average of control courts, indicating that no weighting scheme can adequately balance the severe covariate imbalances.

Placebo tests that apply the synthetic control method to pre-treatment years (e.g., 2014, 2015, 2016, 2017) as if each were the “treatment” year reveal systematic differences between North Carolina and its synthetic version. These estimated placebo effects on dissent rates range from -2.62 to -3.78%, with an average of -2.92%. The actual 2018 treatment effect of 0.00% differs significantly from these pre-treatment placebo effects ($p = 0.042$),

suggesting that the treatment may have had a significant effect in 2018. However, this should be interpreted with caution, as the discontinuity may reflect measurement error, changes in case composition, or other factors unrelated to electoral reform, particularly given the low baseline dissent rates and small absolute magnitudes involved. Relatedly, examining year-by-year treatment effects reveals substantial temporal instability. The estimated effect of partisan elections on dissent rates varies dramatically across post-treatment years: 2018 (-0.0123), 2019 (-0.0150), 2020 (+0.0068), 2021 (-0.0005), 2022 (+0.0270), 2023 (-0.0135), 2024 (+0.0078). These estimates range from -0.015 to +0.027 with a standard deviation of 0.0152—nearly as large as the effects themselves. If partisan elections truly increased dissenting behavior through electoral accountability mechanisms, we would expect a sustained, consistent effect on dissent rates following the 2018 reform. Instead, the effects fluctuate in both direction and magnitude across years, suggesting that any apparent treatment effects likely reflect random variation, case-specific factors, or temporary fluctuations. Additionally, COVID-19 disruptions beginning in 2020—including court closures, delayed proceedings, and shifts in case composition—may have introduced temporal shocks that affect the analysis independently of electoral reform.

The instability documented through sensitivity analysis and robustness checks, combined with persistent covariate imbalances on theoretically central variables, collectively indicate that the synthetic control method fails to produce credible counterfactuals for North Carolina. Rather than providing evidence for or against a causal relationship between electoral reform and dissenting behavior, the findings highlight fundamental limitations in applying synthetic control methods to judicial contexts where comparable control units are scarce, outcome variables exhibit low baseline prevalence and high volatility, theoretically central

covariates vary dramatically across potential donors, and treatment effects are hypothesized to be small relative to measurement error and secular trends.

Textual Analysis

Note: This analysis remains in its preliminary stages. My qualitative findings suggest that dissent rates do not constitute an adequate proxy for either judicial independence or the behavioral effects of electoral institutions, and that the synthetic control method is ill-suited for modeling Supreme Court opinion-writing given the substantial institutional variation across courts. Moreover, dissent rates alone offer limited insight into how electoral reform shapes judicial behavior, as the prevalence of dissent merely captures whether justices disagree with the majority, rather than the depth, substance, or reasoning underlying those disagreements. A more promising approach lies in examining the full text of judicial opinions to assess how electoral reform influences the political tenor and substantive content of judicial reasoning. This textual analysis can reveal whether justices elected in partisan systems are more likely to invoke partisan considerations in their written justifications.

The Evidence Minus Intuition framework distinguishes between two modes of judicial reasoning: Evidence-based language emphasizes empirical facts, data, precedent, and concrete legal analysis, and intuition-based language relies on normative appeals, abstract principles, moral reasoning, and political considerations. By identifying clusters of words associated with each mode, the model enables the quantification of the extent to which judicial opinion-writing reflects more intuitive versus more evidentiary reasoning across different judicial selection systems.

The figure above shows the Word2Vec model’s identification of words semantically

Semantic Corner with Global Inset Map (PCA)

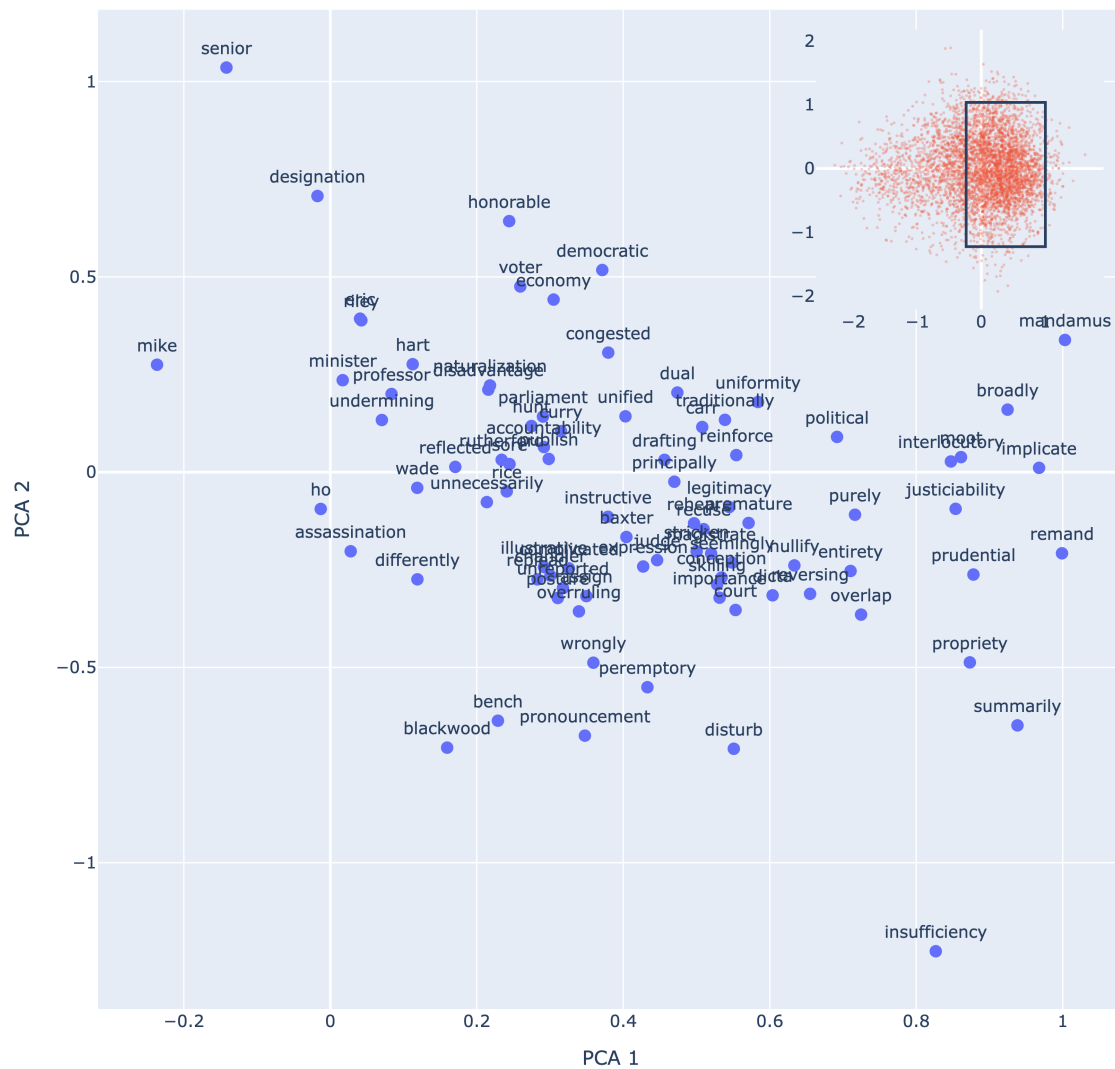


Figure 3: Semantically similar words to “partisan”

proximate to “partisan” in the ColdCases corpus. These vocabulary terms signal intuitive rather than evidence-based language, including “unqualified,” “flatly,” “untenable,” “controversial,” “expression,” “democratic,” “uniformity,” “voter,” “legitimacy,” “undermining,” “assassination,” “reinforce,” and “disadvantage.” I will use these words to inform my dictionary for intuition-based language because they signal rhetorical moves that invoke abstract principles (legitimacy, democracy), normative judgments (untenable, unqualified), and political considerations (voter, partisan) rather than procedural principles or empirical evidence. If partisan elections incentivize justices to signal their ideological positions to voters, we would expect an increase in intuition-based language in opinions authored in states with partisan selection methods or states that have experienced electoral reform in that direction. Full construction of the EMI dictionary using Word2Vec remains in progress.

Conclusion

References

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program.” *Journal of the American Statistical Association* 105, no. 490 (2010): 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>.
- Abadie, Alberto, and Javier Gardeazabal. “The Economic Costs of Conflict: A Case Study of the Basque Country.” *American Economic Review* 93, no. 1 (2003): 113–32. <https://doi.org/10.1257/00028280321455188>.
- Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. “Measuring Citizen and Government Ideology in the American States, 1960-93.” *American Journal of Political Science* 42, no. 1 (1998): 327–48. <https://doi.org/10.2307/2991759>.
- Database on Ideology, Money in Politics, and Elections (DIME): Public Version 4.0 / Stanford Libraries Social Science Data Collection*. <https://data.stanford.edu/dime>, n.d.
- Hall, Matthew E. K., and Jason Harold Windett. “New Data on State Supreme Court Cases.” *State Politics & Policy Quarterly* 13, no. 4 (2013): 427–45. <https://www.jstor.org/stable/24710959>.
- Home - FollowTheMoney.org*. <https://www.followthemoney.org/>, n.d.
- Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. “Voteview: Congressional Roll-Call Votes Database.” In *Voteview: Congressional Roll-Call Votes Database*. <https://voteview.com/>, 2026.
- Measuring Citizen and Government Ideology in the U.S. States: A Re-appraisal - Northwestern University*. <https://search.library.northwestern.edu>, n.d.
- NexisUni*. <https://advance-lexis-com.turing.library.northwestern.edu/firsttime?crid=151f04a4-41f3-4caa-9fd6-1ad2972659d7>; LexisNexis, n.d.
- “OpenSecrets.” In *OpenSecrets*. <https://www.opensecrets.org/>, n.d.
- Renberg, Kristen M. “The Impact of Retention Systems on Judicial Behavior: A Synthetic Controls Analysis of State Supreme Courts.” *The Justice System Journal* 41, no. 4 (2020): 292–312. <https://www.jstor.org/stable/27224753>.
- Squire, Peverill. “Measuring the Professionalization of U. S. State Courts of Last Resort.” *State Politics & Policy Quarterly* 8, no. 3 (2008): 223–38. <https://www.jstor.org/stable/40421611>.
- Squire, Peverill, and Jordan Butcher. “An Update to the Squire State Court of Last Resort

Professionalization Index.” *State Politics & Policy Quarterly* 21, no. 3 (2021): 326–33. <https://doi.org/10.1017/spq.2020.7>.

“State Supreme Courts.” In *Ballotpedia*. https://ballotpedia.org/State_supreme_courts, n.d.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3, no. 1 (2016): 160018. <https://doi.org/10.1038/sdata.2016.18>.