IN THE UNITED STATES DISTRICT COURT
FOR THE MIDDLE DISTRICT OF NORTH CAROLINA

COMMON CAUSE, *et al.*, PLAINTIFFS, v. ROBERT A. RUCHO, in his official capacity as Chairman of the North Carolina Senate Redistricting Committee for the 2016 Extra Session and Co-Chairman of the Joint Select Committee on Congressional Redistricting, *et al.*, DEFENDANTS.	CIVIL ACTION NO. 1:16-CV-1026-WO-JEP THREE-JUDGE COURT
LEAGUE OF WOMEN VOTERS OF NORTH CAROLINA, *et al.*, PLAINTIFFS, v. ROBERT A. RUCHO, in his official capacity as Chairman of the North Carolina Senate Redistricting Committee for the 2016 Extra Session and Co-Chairman of the 2016 Joint Select Committee on Congressional Redistricting, *et al.*, DEFENDANTS.	CIVIL ACTION NO. 1:16-CV-1164-WO-JEP THREE JUDGE PANEL

DECLARATION OF JONATHAN MATTINGLY

Jonathan Mattingly, under penalty of perjury, makes the following declaration:

1. I am over the age of eighteen and competent to testify to the matters set forth herein.

2. I am Professor of Mathematics and Statistical Science at Duke University. My degrees are from the North Carolina School of Science and Math (High School Diploma), Yale University (B.S.), and Princeton University (Ph.D.). I am the Chair of the Mathematics Department at Duke University. A copy of my curriculum vitae is appended as Exhibit 1.

3. As a North Carolinian and mathematician I have developed an interest in quantifying the extent to which the partisan composition of the State's congressional delegation correlates to the statewide political preferences expressed by voters at the polls especially with regard to the 2012 and 2016 congressional elections.

4. Working with my students, I have developed methods to evaluate whether a particular political redistricting produces an election result which falls within or outside the range of typical election results produced by the plethora of plans that reflect established redistricting criteria and that are available for a legislative body to consider and enact. These methods were first developed in 2013 to better understand the extent to which the 2011 congressional redistricting plan reflected the will of the voters as expressed by their votes. Over time we have expanded our analysis to include the 2016 congressional plan, the results of the 2016 elections, as well as the redistricting proposed by the bipartisan redistricting commission of retired North Carolina Judges as part of the "Beyond Gerrymandering" project at Duke University.

5. We concentrate on the congressional districts for the U.S. House of Representatives and use the state of North Carolina and its redistrictings since the 2010

census to explore and showcase our ideas. We begin by observing that for a given set of votes on the state level, the number of representatives elected to a given party can vary drastically depending on the collection of districts used. Yet when randomly drawn "reasonable" redistrictings are used, some results are more common, giving a relatively clear range of expected election outcomes. Furthermore, the structure of the "reasonable" redistrictings gives a clear indication when a particular redistricting is gerrymandered by cracking or packing a particular party. We apply our methods to critique the congressional districts used in the 2012 and 2016 N.C. elections as well as those generated by a bipartisan redistricting commission though never used in practice. We find that the results of the 2012 and 2016 redistrictings are highly inconsistent with the votes cast in those elections. This was already indicated by the startling fact that in the 2012 election, 4 of 13 elected representatives were Democrats, while over 50% of the congressional votes statewide went to the Democratic candidates. On the other hand, our results indicate that the plan produced by a bipartisan redistricting commission from the "Beyond Gerrymandering" project at Duke University is representative of the will of the people.

6. I have been asked by the plaintiffs in Common Cause v. Rucho to share these mathematical analyses with the Court. Our work is described at length in the report appended hereto as Exhibit 2. I will be compensated for my work in the amount of $12,500.

Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury under the laws of the United States that the foregoing statements are true and correct.

This the 6th day of March, 2017.

Jonathan Mattingly

Jonathan Christopher Mattingly
Professor of Mathematics and Statistical Science.

Address:

Department of Mathematics, Duke University, Box 90320, Durham, NC 27708
Email: jonm@math.duke.edu   Web page:www.math.duke.edu/faculty/jonm

Education:

Irwin Avenue Elementary, Charlotte N.C., 1976-1980
Charlotte Country Day Schools, Charlotte N.C., 1980-1986
N.C. School of Science and Mathematics, Durham N.C., High school diploma, 1988
Yale University, B.S. Applied Mathematics, 1992
Ecole Normale de Lyon, Rotary Fellow in Nonlinear and Statistical Physics, 1994
Princeton University, Ph.D. Applied and Computational Mathematics, (Advisor Ya. Sinai) 1998

Principal Appointments:

Department Chair, Duke Mathematics, 2016–2020
Full Professor of Mathematics and Statistical Science, Duke University, 2012 – present
Associate Professor of Mathematics and Statistical Science, Duke University, 2009 – 2012
Associate Professor of Mathematics, Duke University, 2006 – 2009
Saint Flour Summer school in probability Lecturer, Summer 2007
Assistant Professor of Mathematics, Duke University, 2002 – 2005
Member, School of Mathematics, Institute for Advanced Studies, 2002 – 2003
Szegö Assistant Professor of Mathematics, Stanford University, 1998 – 2002
NSF Post-Doctoral Fellow, Stanford University, (mentor George Papanicolaou), 1999 – 2002
Contractor, AT& T Shannon Labs, Summers 1999, 1996

Awards: Fellow of the American Mathematics Society (2014), Fellow of the Institute of Mathematical Statistics (2012).Plenary Speaker SPA Berlin (2010), Speaker National Academy of Science: American-German Frontiers of Science (2009). Plenary Speaker AMS regional meeting (2008). NSF Presidential early career award for scientists and engineers (PECASE), 2006. Alfred P. Sloan Research Fellow, 2005-2007. NSF Career Grant Recipient, 2005-2010. Representative to NAS American-Japanese Frontiers of Science, 2005. NSF Post-Doctoral Fellowship, Stanford University, 1999–2002.

Advisers: Ph.D. Adviser: Yakov G. Sinai, Princeton University. (1994-1998) Post-Doctoral Adviser: George Papanicolaou, Stanford University (1998-2002)

PhD Students: David Anderson 2006 (with Mike Reed), Rachel Thomas 2010 (with Mike Reed), Andrea Watkins 2010. Tiffany Tasky 2012, Prakash Balachandran (with Maggioni) 2011, Sean Lawley (with Mike Reed) 2014, Shishi Luo (with Mike Reed and Katia Koelle) 2014. Current: Brendan Williamson

Postdoctoral fellows: Yuri Bakhtin (now NYU), Scott McKinley (now Univ Florida), Boumediene Hamzi (now London), Matthias Heymann (now Goldman Sachs), Avanti Athreya (now JHU), Oliver Diaz (Trinity College, TX). David Herzog (now Iowa State).

Synergistic Activity:

1. Organizer of: *Seminar on Stochastic Processes* (with Amarjit Budhiraja) ('13), Woman in Probability III ('12) (local orgnaizer with Rick Durrett), Co-organizer of South Eastern Probability Conference, (14'-11'), BRIS workshop for mathematicians and Biologists working on metabolic networks.('10). Oberwolfach Seminar: Ergodic Theory of Markov Processes ('10). SAMSI Year in *Stochastic Dynamics* ('09), AIM Program on *The practice and theory of stochastic simulation* ('07), MSRI program on Dynamical Systems ('07), MSRI

EXHIBIT 1

Mathematical Issues in Stochastic Approaches for Multiscale Modeling ('07), Stochastic dynamics minisymposium, SIAM Snowbird dynamics meeting ('07), Joint Stats meeting satellite conf. on *New Directions in Probability Theory* ('05), AIM Program on *Stochastic and deterministic Navier-Stokes eq.* ('05). Cornell SPDE meeting ('06). Adviser to UNC/Duke Graduate prob. Conf. ('09).
2. Developed new graduate class "Introduction to stochastic differential equations" which is attended by students from many departments. Jointly developed new non-majors class Math Everywhere" (with Ingrid Daubechies and Ezra Miller). Developed new "honors" introduction to probability for mathematically sophisticated undergraduates.
3. Associate Editor: Associate editor for *Annals of probability* (2006-2012), *Comm. in Math. Sci.*, and *SIAM Journal of Mathematical Analysis* (2006-2009), *Stochastic Partial Differential Equations: Analysis and Computations* (2012-2016), *Nonlinearity* (2014-2017)
4. NimBios Scientific Advisory Committee, NSF Math/Bio center,University of Tennessee. (2006-2012)
5. Organized Mentoring of 5-8 N.C. School of Science of Math High school students each year in a year long research project. (2009-present) Served as mentor for two years.

Publications

[1] J. O. B. Tempkin, B. Van Koten, J. C. Mattingly, A. R. Dinner, and J. Weare. Trajectory stratification of stochastic dynamics. *ArXiv e-prints*, October 2016.

[2] M. Hairer and J. Mattingly. The strong Feller property for singular stochastic PDEs. *ArXiv e-prints*, October 2016.

[3] S. Luo and J. C. Mattingly. Scaling limits of a model for selection at two scales. *ArXiv e-prints*, July 2015.

[4] J. C. Mattingly and C. Vaughn. Redistricting and the Will of the People. *ArXiv e-prints*, October 2014.

[5] Nathan Glatt-Holtz, Jonathan C. Mattingly, and Geordie Richards. On Unique Ergodicity in Nonlinear Stochastic Partial Differential Equations. *J. Stat. Phys.*, 166(3-4):618–649, 2017.

[6] Yuri Bakhtin, Tobias Hurth, and Jonathan C. Mattingly. Regularity of invariant densities for 1D systems with random switching. *Nonlinearity*, 28(11):3755–3787, 2015.

[7] David P. Herzog and Jonathan C. Mattingly. Noise-induced stabilization of planar flows II. *Electron. J. Probab.*, 20:no. 113, 37, 2015.

[8] David P. Herzog and Jonathan C. Mattingly. Noise-induced stabilization of planar flows I. *Electron. J. Probab.*, 20:no. 111, 43, 2015.

[9] David P. Herzog and Jonathan C. Mattingly. A practical criterion for positivity of transition densities. *Nonlinearity*, 28(8):2823–2845, 2015.

[10] Sean D. Lawley, Jonathan C. Mattingly, and Michael C. Reed. Stochastic switching in infinite dimensions with applications to random parabolic PDE. *SIAM J. Math. Anal.*, 47(4):3035–3063, 2015.

[11] Stephan Huckemann, Jonathan C. Mattingly, Ezra Miller, and James Nolen. Sticky central limit theorems at isolated hyperbolic planar singularities. *Electron. J. Probab.*, 20:no. 78, 34, 2015.

[12] Elizabeth Munch, Katharine Turner, Paul Bendich, Sayan Mukherjee, Jonathan Mattingly, and John Harer. Probabilistic Fréchet means for time varying persistence diagrams. *Electron. J. Stat.*, 9(1):1173–1204, 2015.

[13] Sean D. Lawley, Jonathan C. Mattingly, and Michael C. Reed. Sensitivity to switching rates in stochastically switched ODEs. *Commun. Math. Sci.*, 12(7):1343–1352, 2014.

[14] Jonathan C. Mattingly and Etienne Pardoux. Invariant measure selection by noise. An example. *Discrete Contin. Dyn. Syst.*, 34(10):4223–4257, 2014.

[15] Thomas Hotz, Stephan Huckemann, Huiling Le, J. S. Marron, Jonathan C. Mattingly, Ezra Miller, James Nolen, Megan Owen, Vic Patrangenaru, and Sean Skwerer. Sticky central limit theorems on open books. *Ann. Appl. Probab.*, 23(6):2238–2258, 2013.

[16] Avanti Athreya, Tiffany Kolba, and Jonathan C. Mattingly. Propagating Lyapunov functions to prove noise-induced stabilization. *Electron. J. Probab.*, 17:no. 96, 38, 2012.

[17] Jonathan C. Mattingly, Natesh S. Pillai, and Andrew M. Stuart. Diffusion limits of the random walk Metropolis algorithm in high dimensions. *Ann. Appl. Probab.*, 22(3):881–930, 2012.

[18] Jonathan C. Mattingly, Scott A. McKinley, and Natesh S. Pillai. Geometric ergodicity of a bead-spring pair with stochastic Stokes forcing. *Stochastic Process. Appl.*, 122(12):3953–3979, 2012.

[19] Martin Hairer and Jonathan C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[20] David F. Anderson and Jonathan C. Mattingly. A weak trapezoidal method for a class of stochastic differential equations. *Commun. Math. Sci.*, 9(1):301–318, 2011.

[21] Martin Hairer and Jonathan C. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electron. J. Probab.*, 16:no. 23, 658–738, 2011.

[22] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.

[23] Jonathan C. Mattingly, Andrew M. Stuart, and M. V. Tretyakov. Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, 48(2):552–577, 2010.

[24] Martin Hairer and Jonathan C. Mattingly. Slow energy dissipation in anharmonic oscillator chains. *Comm. Pure Appl. Math.*, 62(8):999–1032, 2009.

[25] Martin Hairer and Jonathan C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[26] Gautam Iyer and Jonathan Mattingly. A stochastic-Lagrangian particle system for the Navier-Stokes equations. *Nonlinearity*, 21(11):2537–2553, 2008.

[27] Jonathan C. Mattingly and Toufic M. Suidan. Transition measures for the stochastic Burgers equation. In *Integrable systems and random matrices*, volume 458 of *Contemp. Math.*, pages 409–418. Amer. Math. Soc., Providence, RI, 2008.

[28] Jonathan C. Mattingly, Toufic M. Suidan, and Eric Vanden-Eijnden. Anomalous dissipation in a stochastically forced infinite-dimensional system of coupled oscillators. *J. Stat. Phys.*, 128(5):1145–1152, 2007.

[29] Yuri Bakhtin and Jonathan C. Mattingly. Malliavin calculus for infinite-dimensional systems with additive noise. *J. Funct. Anal.*, 249(2):307–353, 2007.

[30] Jonathan C. Mattingly, Toufic Suidan, and Eric Vanden-Eijnden. Simple systems with anomalous dissipation and energy cascade. *Comm. Math. Phys.*, 276(1):189–220, 2007.

[31] H. Lamba, J. C. Mattingly, and A. M. Stuart. An adaptive Euler-Maruyama scheme for SDEs: convergence and stability. *IMA J. Numer. Anal.*, 27(3):479–506, 2007.

[32] David F. Anderson, Jonathan C. Mattingly, H. Frederik Nijhout, and Michael C. Reed. Propagation of fluctuations in biochemical systems. I. Linear SSC networks. *Bull. Math. Biol.*, 69(6):1791–1813, 2007.

[33] Martin Hairer and Jonathan C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)*, 164(3):993–1032, 2006.

[34] Jonathan C. Mattingly and Étienne Pardoux. Malliavin calculus for the stochastic 2D Navier-Stokes equation. *Comm. Pure Appl. Math.*, 59(12):1742–1790, 2006.

[35] Yuri Bakhtin and Jonathan C. Mattingly. Stationary solutions of stochastic differential equations with memory and stochastic partial differential equations. *Commun. Contemp. Math.*, 7(5):553–582, 2005.

[36] Jonathan C. Mattingly and Toufic M. Suidan. The small scales of the stochastic Navier-Stokes equations under rough forcing. *J. Stat. Phys.*, 118(1-2):343–364, 2005.

[37] Martin Hairer and Jonathan C. Mattingly. Ergodic properties of highly degenerate 2D stochastic Navier-Stokes equations. *C. R. Math. Acad. Sci. Paris*, 339(12):879–882, 2004.

[38] Martin Hairer, Jonathan C. Mattingly, and Étienne Pardoux. Malliavin calculus for highly degenerate 2D stochastic Navier-Stokes equations. *C. R. Math. Acad. Sci. Paris*, 339(11):793–796, 2004.

[39] Jonathan C. Mattingly. On recent progress for the stochastic Navier Stokes equations. In *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. XI, 52. Univ. Nantes, Nantes, 2003.

[40] J. C. Mattingly. Contractivity and ergodicity of the random map $x \mapsto |x - \theta|$. *Teor. Veroyatnost. i Primenen.*, 47(2):388–397, 2002.

[41] Jonathan C. Mattingly. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Comm. Math. Phys.*, 230(3):421–462, 2002.

[42] Jonathan C. Mattingly. The dissipative scale of the stochastics Navier-Stokes equation: regularization and analyticity. *J. Statist. Phys.*, 108(5-6):1157–1179, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[43] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[44] J. C. Mattingly and A. M. Stuart. Geometric ergodicity of some hypo-elliptic diffusions for particle motions. *Markov Process. Related Fields*, 8(2):199–214, 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).

[45] P. J. Holmes, J. C. Mattingly, and R. W. Wittenberg. Low-dimensional models of turbulence or, The dynamics of coherent structures. In *From finite to infinite dimensional dynamical systems (Cambridge, 1995)*, volume 19 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 177–215. Kluwer Acad. Publ., Dordrecht, 2001.

[46] Weinan E, J. C. Mattingly, and Ya. Sinai. Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation. *Comm. Math. Phys.*, 224(1):83–106, 2001. Dedicated to Joel L. Lebowitz.

[47] Weinan E and Jonathan C. Mattingly. Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation. *Comm. Pure Appl. Math.*, 54(11):1386–1402, 2001.

[48] Jonathan C. Mattingly. Ergodicity of 2D Navier-Stokes equations with random forcing and large viscosity. *Comm. Math. Phys.*, 206(2):273–288, 1999.

[49] J. C. Mattingly and Ya. G. Sinai. An elementary proof of the existence and uniqueness theorem for the Navier-Stokes equations. *Commun. Contemp. Math.*, 1(4):497–516, 1999.

[50] Philip J. Holmes, John L. Lumley, Gal Berkooz, Jonathan C. Mattingly, and Ralf W. Wittenberg. Low-dimensional models of coherent structures in turbulence. *Phys. Rep.*, 287(4):337–384, 1997.

REPORT ON REDISTRICTING: DRAWING THE LINE

JONATHAN C. MATTINGLY

I have been asked by the plaintiffs in Common Cause v Rucho to share with the court the mathematical analysis that I and my students have developed to examine the extent to which the outcome of North Carolina's congressional elections depends on choices the North Carolina General Assembly made in drawing the boundaries of those districts. These methods were first developed in 2013 to better understand the extent to which the 2011 congressional redistricting plan reflected the will of the voters as expressed by their votes. Overtime we have expanded our analysis to include the 2016 congressional plan, the results of the 2016 elections, as well as the redistricting proposed by the simulated bipartisan redistricting commission of retired North Carolina Judges as part of the "Beyond Gerrymandering" project at Duke University.

The method we use is simple and based on standard and accepted mathematical methods. We take the actual votes cast by North Carolinians at the 2012 and 2016 congressional elections and then change the boundaries of the congressional districts to see how the partisan results of the elections change. The results of the report are based on a manuscript by the author and his students which is in preparation (See [2] from references).

In the United States, district representation schemes are used to divide the population into distinct groups, each of which carry a certain amount of representation. This districting, whether by state or within a state, acknowledges that the people's voice is geographically diverse and that we value the expression of that diversity in our government. We take election results to represent the "will of the people", giving the elected officials a mandate to act in the people's name. Hence, it is reasonable to ask if and how this will is affected by the choice of district boundaries. Just how sensitive are election results to our choices for geographic divisions?

The 2011redistricting of North Carolina is a useful example and testing ground for this general line of inquiry. Most would agree that politics had a hand in the North Carolina redistricting process. The motivations were diverse. The twelfth district was drawn to create a majority African-American district. Others seemingly were drawn to split and pack different voting blocks to diminish their political power. The question remains of how large was the effect of the redistricting on the outcome ?

In the 2012 congressional elections, which were based on the 2011 districts, four out of the thirteen congressional seats were filled by Democrats. Yet in seeming contradiction, the majority of votes were cast for Democratic candidates on the statewide level. The election results hinged on the geographic positioning of congressional districts. While this outcome is clearly the result of politically drawn districts, perhaps it is not the result of excessive tampering. Our country has a long history of balancing the rights of urban areas with high population with those of more rural, less populated areas. Our federalist and electoral structures enshrined the idea that majority rule must be balanced with regionalism. It might be that in North Carolina, the subversion of the results of the global vote count would happen in any redistricting which balances the representation of the urban with the rural or the beach with the mountains, and each with the Piedmont. Maybe the vast majority of reasonable districts which one might draw would have these issues due to the geography of the population's distribution. We are left asking the basic question: how much does the outcome depend on the choice of districts? This can be further refined by asking "what are the outcomes for a typical choice of districts ?" or "when should a redistricting be considered outside

EXHIBIT 2

the norm?" These last two refinements require some way of quantifying what the typical outcomes are for a given set of votes. This turns the usual election procedure on its head. Usually one fixes the districts and varies the votes from election to election. We are interested in fixing the votes and then changing the redistricting to observe how the results change. Since we explore these questions in the context of the American political system, we assume that people vote for parties, not people, which is of course not true. However, in these polarized times it is not the worst approximation. We still find the results extremely illuminating.

Once we understand the extent to which election results can vary over a collection of possible redistrictings, we can assign the representativeness of particular redistricting by observing its place in this collection of results. Similarly, with statistics of typical redistricting in hand, we can devise measures of gerrymandering where the effects of packing and cracking blocks of votes can be better identified. We develop indices to measure the representativeness and level of gerrymandering of a given district.

We apply our metrics to analyze and critique the North Carolina U.S. Congressional redistrictings used in the 2012 and 2016 elections as well as the redistrictings developed by a bipartisan group of retired Judges as part of the "Beyond Gerrymandering" project spearheaded by Thomas Ross. We henceforth refer to these redistrictings of interest as NC2012, NC2016, and Judges respectively. See Figures 14–16 in the Appendix for visualizations of these redistrictings. Our analysis uses the actual votes in elections to illuminate the structure and features of a redistricting. In this report, we use the actual votes cast in the 2012 and 2016 N.C. congressional elections.

Using a related methodology, we also assess the degree to which the three redistrictings (NC2012, NC2016, and Judges) are engineered. This is done by seeing how close their properties are to the collection of redistrictings which can be obtained by only small changes. It seems reasonable that the character of an election should not be overly sensitive to small changes in the redistricting if the concept of the "will of the people" is to have any meaning.

No matter what lens is used, the results repeatedly show that the NC2012 and NC2016 redistricting are heavily engineered and produce results which are extremely atypical and at odds with the "will of the people," as illuminated by our analysis. Finer analysis clearly shows that the Democratic voters are clearly packed into a few districts, decreasing their power, while Republican voters are spread more evenly which increases their power. In contrast, election results from the Judges redistricting are quite typical, producing results consistent with what is typically seen. We emphasize that all of these conclusions come from asking what is the typical character and result of an election if we use a "reasonable" redistricting drawn at random without any partisan input, save the possible effect of ensuring a few districts contain a sufficient minority population to comply with the Voting Rights Act (VRA).

1. Main Results: Where do you draw the line?

We emphasize from the start that in contrast to some works (See for example [5] from references), we are not proposing an automated method of creating redistrictings which might be used in practice. Rather, we are proposing a class of ideas for evaluating if a redistricting is truly representative or if it is gerrymandered. We hope this helps draw the line between fair and biased redistricting.

Our analysis begins by generating over 24,000 "reasonable" redistrictings of North Carolina into thirteen U.S. House congressional districts. For each redistricting, we tabulate the votes from a previous election, either 2012 or 2016, to calculate the number of representatives elected from both the Democratic and Republican Parties. We emphasize that we use the actual votes from either the 2012 or 2016 U.S. House of Representative elections. In using these votes, we assume that a vote cast for a Republican or Democrat remains so even when district boundaries are shifted.

By "reasonable," we mean districts which are drawn in a nonpartisan fashion, guided only by the desire to:

- Divide the state population evenly between the thirteen districts.
- Keep the districts geographically connected and compact.
- Refrain from splitting counties as much as possible.
- Ensure that African-American voters are sufficiently concentrated in two districts to give them a reasonable chance to affect the winner.

The precise meaning of "reasonable" is given in Section 2 along with the method we used to generate the over 24,000 "reasonable" redistrictings. We construct our districts by taking Voting Tabulation Districts (VTD) from NC2012 as the fundamental atomic element used as our building blocks. North Carolina is composed of over 2,600 VTDs.

The first criterion above enforces the "one-person-one-vote" doctrine, which dictates that each representative should represent a roughly equal number of people. The second criterion reflects the desire to have districts represent regional interests. The third criterion embodies the idea that districts should not fracture historical political constituencies if possible; counties provide a convenient surrogate for these constituencies. The last criterion, which is dictated by the Voting Rights Act (VRA), asks that two districts have enough African-American voters that they might be reasonably expected to choose the winner in that district. In particular, we emphasize that no voting or registration information is used, nor is any demographic information, except for what is dictated by the VRA as specified in the preceding criterion.

The exact choice of these criteria for our study comes from House Bill 92, which passed the North Carolina House during the 2015 General Assembly legislative session. This bill proposed establishing a bipartisan commission to perform redistricting guided solely by these principles. Since the companion legislation did not pass the North Carolina Senate, the provision never became law. In fact, it is just the latest in a chain of bills which have been introduced over the years with similar criteria and aims.

1.1. **Beyond One-Person-One-vote.** Our results clearly show that there is a large amount of variation in the outcome of an election depending on the districts used. The simple criteria from House Bill 92 are not enough to produce a single preferred outcome of the elections. Rather, there is a distribution of possible outcomes. Our findings in this direction, summarized in Figure 1, clearly show that the results generated by the redistrictings NC2012 and NC2016 are extremely biased towards the Republicans while the Judges redistricting produces acceptably representative results. The NC2012 and NC2016 redistrictings produce results which are highly atypical of the non-partisan redistrictings we randomly drawn according to HB92.

Over 24,000 random, but reasonable, redistrictings were used to generate the probability distributions show in Figure 1. We emphasize that the two plots use the actual votes cast by the electorate in the 2012 and 2016 Congressional elections respectively to determine the outcomes for each redistricting. For the 2012 vote counts, the NC2012 and NC2016 redistrictings both result in four Democratic seats, a result which occurs in less than 0.3% of our collection of over 24,000 redistrictings. The Judges redistricting results in the election of six Democrats, which occurs in over 39% of redistrictings. For the 2016 vote counts, the NC2012 and NC2016 redistrictings results in three Democratic seats, a result which occurs in less than 0.7% of redistrictings. The Judges redistricting results in the election of four Democrats, which occurs in 28% of redistrictings.

1.2. **Measuring Representativeness and Gerrymandering.** While Figure 1 is already quite compelling, it is useful to develop quantitative measures of how representative the results of a given election are. However, gerrymandering goes beyond just affecting the results, but also makes districts so safe that representatives are less responsive to the "will of the people." To measure these effects, we use two indices. The first, which we call the Gerrymandering Index, is based on the plots used to visualize gerrymandering introduced in Section 1.3. It quantifies how packed or depleted the collection of districts is relative to what is expected from the ensemble of "reasonable" redistrictings

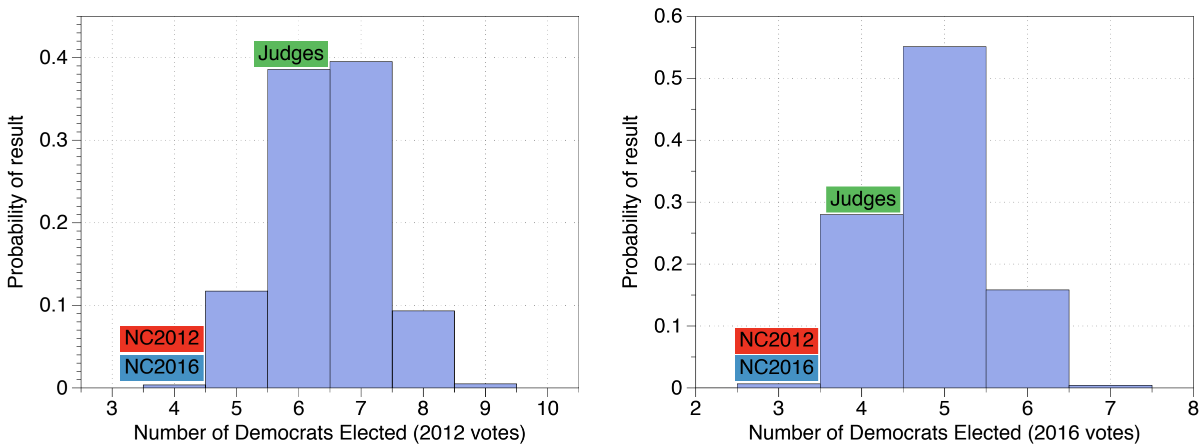


FIGURE 1. Probability of a given number of Democratic wins among the 13 congressional seats using 2012 election (left) and 2016 election (right) votes.

we have created. The second is the measure of how typical are the election results produced by the redistricting in the context of what is seen in the ensemble of "reasonable" redistrictings. In summary, we consider

- **Gerrymandering Index:** Measures the degree to which the percentage of Democratic votes in each district deviates from what is typically seen in our collection of "reasonable" redistrictings. To make the comparison, the districts are ordered from the most Republican to the most Democratic. The squareroot of the sum of the square deviations is the index. Relatively large scores are less balanced than the bulk of the "reasonable" redistrictings in our ensemble. These large indexed redistrictings typically have some districts with many more voters from one party than is normal seen or generally have a higher percentage of one party in many districts than is normal, or have both. How the term "normal" is understood is partially explained in Section 1.3 and completely explained in Section 4.1.
- **Representativeness Index:** Measures how typical the results obtained by a given redistricting are in the context of the collection of "reasonable" redistrictings we have generated. Redistrictings with relatively large values produced an election outcome which is farther from the typical election outcome in the collection of "reasonable" redistrictings. The full details are given in Section 4.2.

As these indices are most useful when values for two different redistricting are compared, we place each redistricting of interest on the plot of the complementary cumulative distribution function for each of the three above measures. This allows us to judge the relative size of each index in the context of our collection of "reasonable" redistrictings.

In a complementary cumulative distribution function, the vertical axis shows the fraction of random redistrictings which have a larger index value than a redistricting with a given index on the horizontal axis. We plot results for the Gerrymandering Index and the Representativeness Index in Figures 2 and 3, respectively. We calculate the probability of each index obtaining a value greater than a given value based on our random redistrictings. We then situate each of our redistrictings of interest (NC2012, NC2016, and Judges) on the plot indicating the fraction of random redistrictings which have a larger index.

We find the probability of constructing a reasonable redistricting plan with index scores equal to or worse than the NC2012 redistricting plan is negligible for both indices; not a single random redistricting has a Gerrymandering Index or a Representative Index that surpasses those of the NC2012 plan. Similarly, we also find that the probability of randomly constructing a redistricting

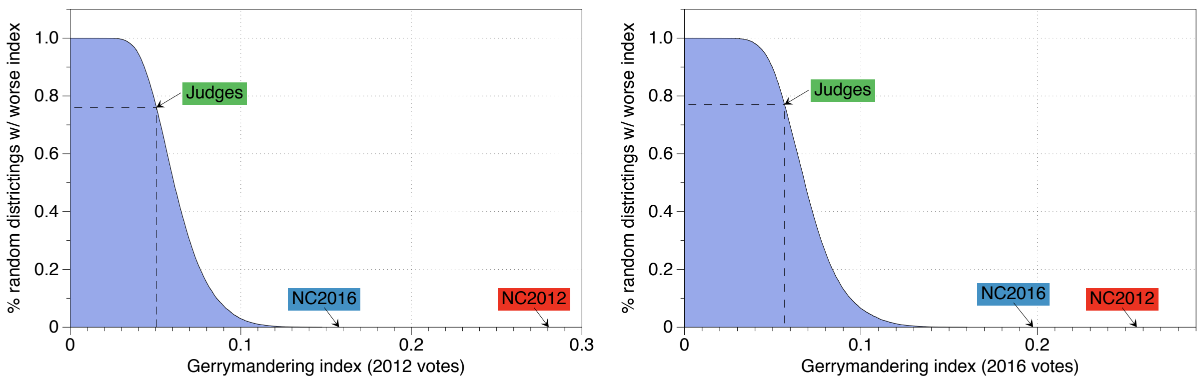


FIGURE 2. Gerrymandering Index for the three districts of interest based on the congressional voting data from 2012 (left) and 2016 (right). No generated redistrictings had a Gerrymandering Index higher than either the NC2012 or the NC2016 redistrictings. The Judges redistricting plan was less gerrymandered than over 75% of the random districts in both sets of voting data, meaning that it is an exceptionally non-gerrymandered redistricting plan.

with indices worse than the NC2016 redistricting plan is vanishingly small. On the contrary, the Judges redistricting plan has a Gerrymandering Index that is better than over 75% of all districts and a Representativeness Index that is better than roughly 75% of all districts. Thus the Judges plan is a very typical plan, and even above average in terms of not being gerrymandered and in being representative of the "will of the people."

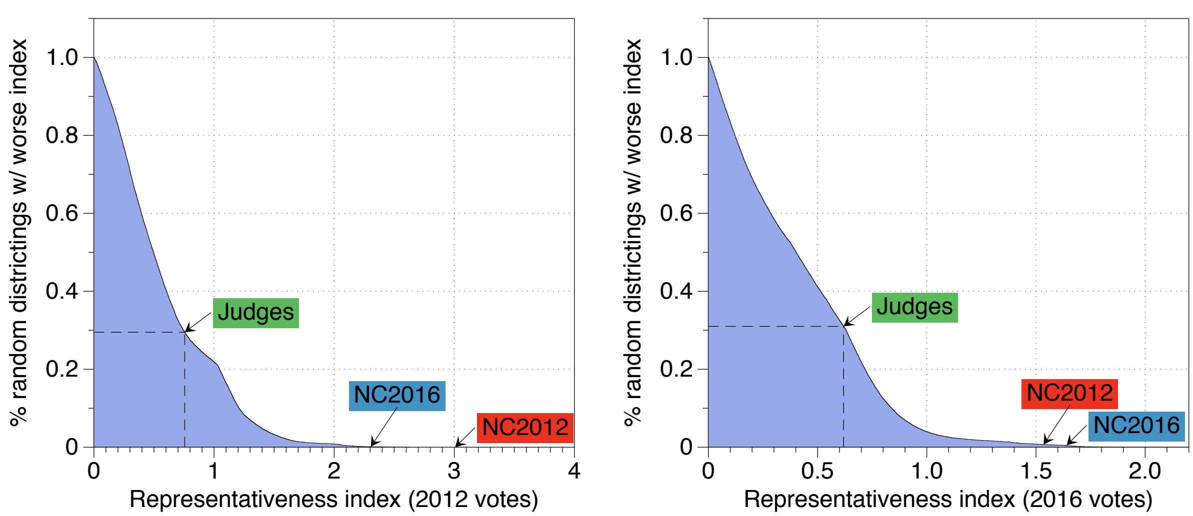


FIGURE 3. Representativeness Index for the three districts of interest using congressional voting data from 2012 (left) and 2016 (right). No redistrictings was less representative than the NC2012 nor NC2016 redistricting plans. Roughly 30% of redistricting plans were less representative than the Judges redistricting plan in both sets of voting data, meaning that the Judges plan was reasonably representative.

1.3. **Visualizing Gerrymandering.** While the reductive power of a single number can be quite compelling, we have also developed a simple graphical representation to summarize the properties of a given redistricting relative to the collection of reference redistrictings. The goal was to create a graphical representation which would make visible when a particular redistricting packed or fractured voters from a particular party to reduce its political power.

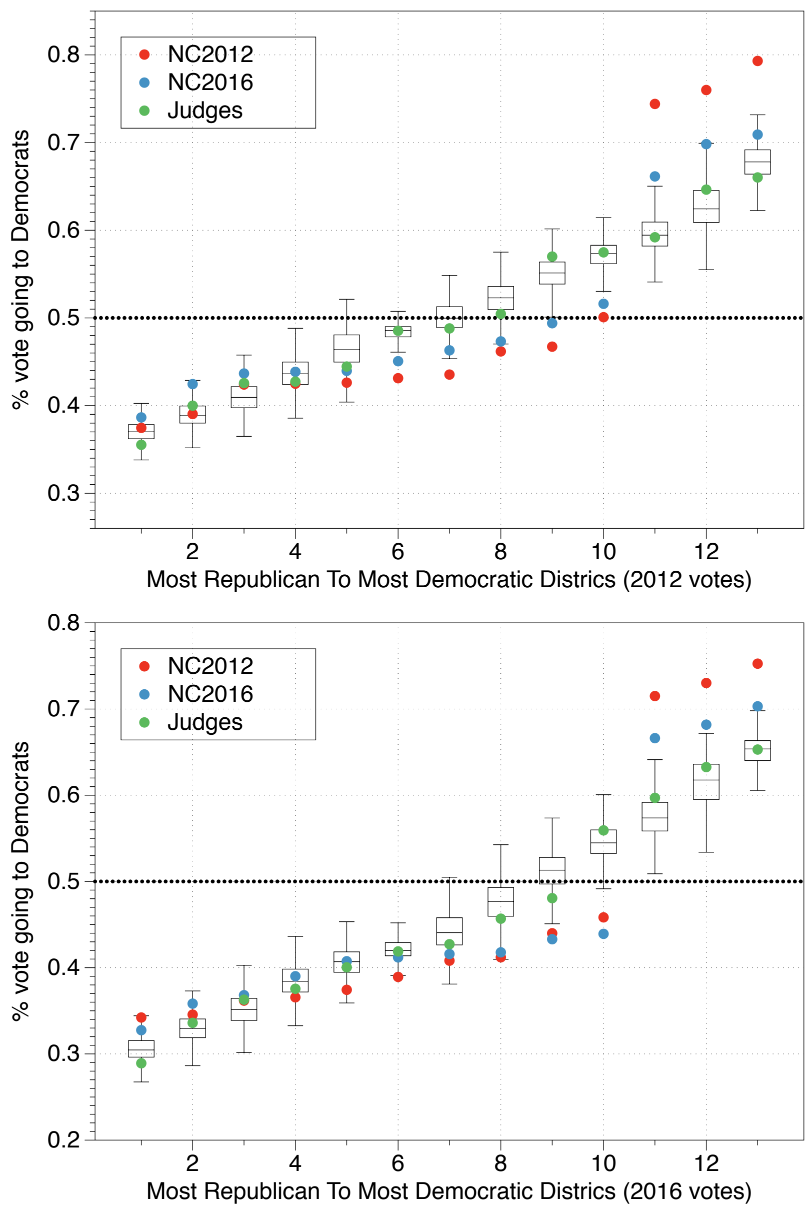


FIGURE 4. After ordering districts from most Republican to most Democrat, these box-plots summarize the distribution of the percentage of Democratic votes for the district in each ranking position for the votes from 2012 (top) and 2016 (bottom). We compare our statistical results with the three redistricting plans of interest. The Judges plan seems to be typical while NC2012 and NC2016 have more Democrats than typical in the most Democratic districts and fewer in districts which are often switching between Democrat and Republican, depending on the votes.

One first needs to begin by discovering the natural structure of the geographical distribution of votes in the state when viewed through the lens of varying over "reasonable" redistrictings. We begin by ordering the thirteen congressional districts which make up a redistricting from lowest to highest based on the percentage of Democratic votes in each district. Since there are essentially only two parties, nothing would change if we instead considered the percentage of Republican votes.

We are interested in the random distribution of this thirteen dimensional vector. Since it is difficult to visualize such a high dimensional distribution, we summarize the distribution by considering the marginal distribution of each position in this vector and summarize it in a classical box-plot for each component of the thirteen dimensional vector in Figure 4. That is to say, we examine the distribution of votes that make up the percentage of Democratic votes in the most Republican district. Then we repeat the process for the second most Republican district. Continuing for each of the rankings, we obtain thirteen box-plots which we arrange horizontally on the same plot.

The box-plots are standard plots, meaning that within each box-plot the central line gives the median percentage, the ends of the box give the location of the upper quartile and the lower quartile (25% of the results exist below and above these lines). The outer bracketing line defines an interval containing either the maximum and minimum values, or three halves the distance of the quantiles from the mean, which ever is smaller. In the interest of visual clarity, we have not plotted any outliers. On top of these box-plots, we have overlaid the percentages for the NC2012, NC2016, and the Judges redistricting. In Figure 5, we also include plots that displays histograms rather than box-plots. These plots are richer in detail. Yet, the detail makes it harder to estimate confidence intervals.

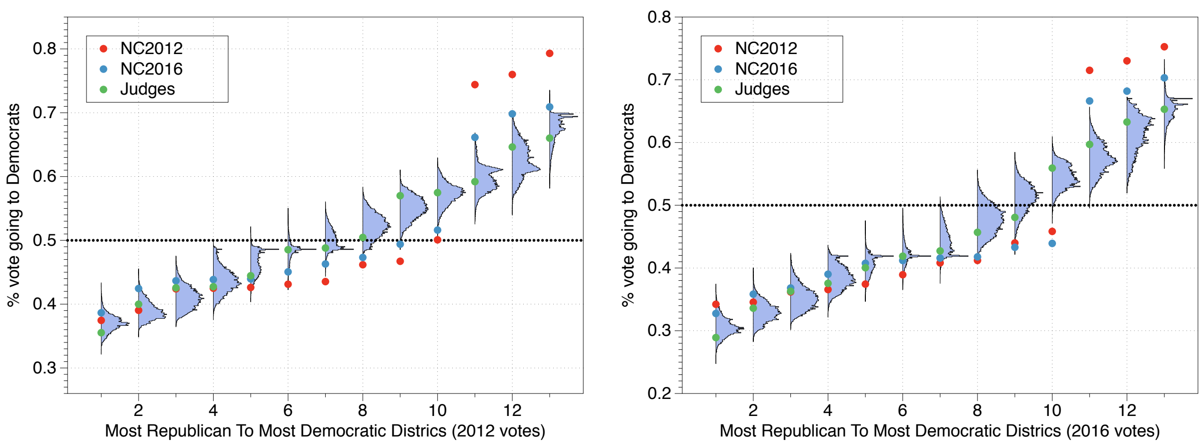


Figure 5. We present the same data as in Figure 4, but, with histograms replacing the box-plots. Note that the distribution of the sixth most Republican district (district with label 6 on the plots) is quite peaked in both the 2012 and 2016 votes, the Judges results are centered directly on this peak while the NC2012 and NC2016 lie well outside the extent of the distribution.

The structure of these plots highlights the typical structure of the redistrictings in our ensemble and by extension, the spatial-political structure of the voters in North Carolina. It can then be used to reveal the structure of our three redistrictings of interested. Observe that for both the 2012 and 2016 votes, the centers of the box-plots form a relatively straight, gradually increasing line from the most Republican district (labeled 1) to the most Democratic (labeled 13). The Judges districts mirror this structure. Furthermore, most of the percentages from the Judges districts fall inside the box on the box-plot which marks the central 50% of the distribution. The NC2012 and NC2016 have a different structure. There is a large jump between the tenth and eleventh most Republican district (those with labels 10 and 11, respectively). In the NC2012 redistricting, the fifth through tenth most Republican districts have more Republicans than one would typically see in our ensemble of "reasonable" redistrictings. In the NC2016 redistricting, the shifting starts with the sixth most Republican district and runs through the tenth most Republican district (labeled 6-10). In both cases, the votes removed from the central districts have largely been added to the three most Democratic districts (labeled 11-13). In the 2012 votes, this moved three to four districts that typically would have been above the 50% line to below the 50% line, meaning that these districts elected the Republican rather than the Democrat. With the 2016 votes, the changes in structure only moved the tenth most Republican district across the 50% line.

Forgetting about the election outcomes, the structure has implications for the competitiveness of districts and likely political polarization. Rather than a gradual increase at a constant rate from left to right as the Judges redistricting and the ensemble of box-plots, the NC2012 and NC2016 redistrictings have significantly more Democrats in the three most Democratic districts and fairly safe Republican majorities in the first eight most Republican districts. It is not hard to argue that this leads to a more polarized legislative delegation with fewer centrist representatives being elected on both ends of the political spectrum.

Figure 4 can be used to motivate and explain the Gerrymandering Index for our redistrictings of interest. For example, to calculate the Gerrymandering Index for NC2012, one sums the square of the distance from the red dots to the mean in each distribution from 1 to 13. The Gerrymandering Index is the square root of this sum. To aid with visualization, recall that the line though the center of each box is the median which, in these cases, is close to the mean. Clearly, this index captures some of the features of Figure 4 discussed in the previous paragraph.

It is remarkable how stable the structures in Figure 4 are across the 2012 and 2016 votes. The 2016 plot is largely a downward shift of the 2012 plot. This stability largely explains why the two plots in Figure 2 of the Gerrymandering Index look so similar. It also speaks to the stability of the Representativeness Index in Figure 3.

1.4. **Stability of Election Results.** We also explored the degree to which the NC2012, NC2016 and Judges redistrictings are representative of the nearby redistrictings, where we interpret nearby to mean that roughly 10% of the VTDs are swapped between districts. (See the next paragraph for more precise description.) By switching nearby VTDs among districts we are able to assess whether small changes impact the characteristics of the districts or not. We found that the districts within the NC2012 and NC2016 redistricting plans had a Gerrymandering Index which was significantly larger than the nearby redistrictings while the the Judges plan had a Gerrymandering Index which was in the middle of the range produced by nearby redistrictings. In other words, switching nearby districts made the NC2012 and NC 2016 redistrictings less partisan but did not change the characteristics of the Judges redistricting. This suggests that the NC 2012 and NC2016 redistricting, in contrast to the Judges redistricting, were precisely engineered and tuned to achieve a partisan goal and that the components of the NC2012 and NC 2016 redistrictings redistricting were not randomly chosen.

More precisely, we randomly sample "resonable" redistrictings which are near the NC2012, NC2016, and Judges redistrictings in the sense that no single district differs by more VTDs than a set threshold from the redistricting understudy. To set the threshold, we observe that among the over 24,000 redistrictings we generated the average district size is around 210 VTDs. In a particular typical redistricting from our ensemble, the sizes varied from 152 to 278 VTDs. With these numbers in mind, we set our threshold to be 40 VTDs. Since every VTD switched is counted twice, once for the district it is leaving and once for the district it is entering, this amounts to a total of around 10% of the VTDs switching districts.

Figure 6 shows the results of these analyses applied to the NC2012, NC2016, and Judges redistrictings. The redistrictings sampled around NC2012 have markedly better Gerrymandering Indices than NC2012 itself. The results are less dramatic for NC2016, but telling nonetheless. This

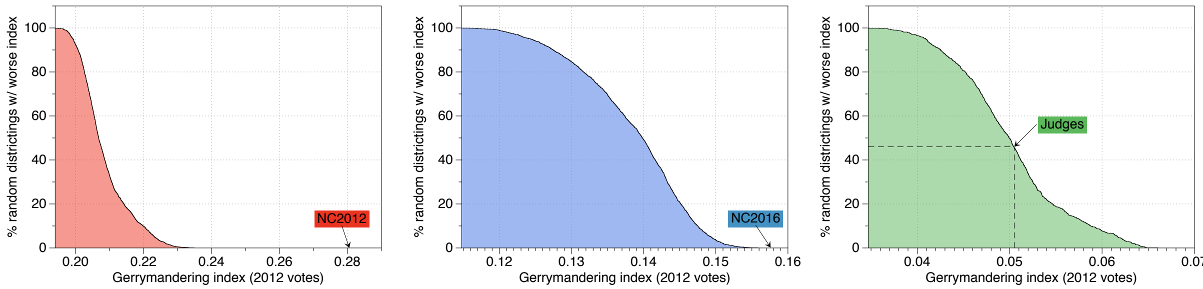


FIGURE 6. Gerrymandering Index based on random samples drawn from nearby the three redistrictings of interest: NC2012 (left), NC2016 (center), and Judges (right). Only for the Judges are the other points in the neighborhood similar to the redistricting of interest. All plots use the 2012 votes.

shows that a randomly chosen redistricting near NC2012 (or respectively NC2016) has very different properties than NC2012 (or respectively NC2016). This is convincing evidence that the NC2012 and NC2016 redistrictings were deliberately constructed to have unusual properties. It would have been unlikely to choose such a singularly unusual redistricting by chance. In contrast, the Judges redistricting has a Gerrymandering Index which is quite typical of its nearby redistrictings. It is worse than around 50% of those nearby it and hence better than 50% of those nearby it. Thus, it is very representative of its nearby redistrictings.

1.5. **Summary of Main Results.** By sampling over 24,000 reasonable redistrictings, we explore the distribution of different election outcomes by estimating the probabilities of the numbers of Democrats elected from North Carolina to the U.S. House of Representatives. Our sampling of reasonable redistrictings also allows us to estimate the distribution of winning margins in each district as well as the value of two indices representing gerrymandering and representativeness. In every one of our tests, we have found that the NC2012 and NC2016 redistricting plans are extraordinarily anomalous, suggesting that (i) these districts are heavily gerrymandered, (ii) they do not represent the "will of the people" and (iii) they dilute the votes of one party. We have also uncovered evidence that these two redistricting plans employ packing and cracking. On the contrary, the redistricting plan produced by a bipartisan redistricting commission of retired judges from the Beyond Gerrymandering project produced results which were highly typical among our 24,000 reasonable redistrictings. The Judges plan was exceptionally non-gerrymandered, was a typical representation of the "will of the people," and does not seem to pack or crack either party.

We also explored the degree to which the NC2012, NC2016 and Judges redistrictings were representative of the nearby redistrictings, where we interpret nearby to mean that roughly 10% of the VTDs are switched out of any given district. We found that the NC2012 and NC2016 redistrictings were significantly more gerrymandered then those around them while the Judges redistricting was similar to those nearby. This seems to imply that the NC2012 and NC2016 redistrictings were carefully engineered and tuned, and not randomly chosen among those with a certain basic structure.

The remainder of the paper is organized as follows. In the remainder of this section we describe the Beyond Gerrymandering project. In Section 2, we describe how we construct our distribution of "reasonable" redistrictings and sample it using Markov chain Monte Carlo. In Section 2.4, we describe how we further sub-select our samples based on a series of thresholds to better reflect the proposed bipartisan redistricting legislation HB92. In Section 2.5, we discuss how the parameters of our distribution are chosen to produce "reasonable" redistrictings. In Section 3, we explore the effect of the Voting Rights Act provision in HB92 on the outcome of the elections. In Section 4, we give the missing details in the construction of the Representativeness and Gerrymandering Indices. In Section 5, we show that our results are relatively insensitive to our choice of parameters. We also provide evidence that our Markov chain Monte Carlo is running sufficiently long to produce results from the desired distribution. In Section 6, we provide some details about the data used and some of the more technical choices made in the preceding analysis. Finally in Section 7, we make some concluding remarks and discuss future directions. The Appendix of the paper gives some sample maps drawn by our algorithm as well as for the NC2012, NC2016, and Judges redistrictings.

1.6. **The Beyond Gerrymandering Project.** The Beyond Gerrymandering[1] project was a collaboration between UNC system President Emeritus and Davidson College President Emeritus Thomas W. Ross, Common Cause, and the POLIS center at the Sanford School at Duke University. The project's goal is to educate the public on how an independent, impartial redistricting process would work. The project formed an independent redistricting commission made up of ten retired jurists, with each political party represented by an equal number. The commission used strong, clear criteria to create a new North Carolina congressional map based on NC House Bill 92 from the 2015 legislative season. All federal rules related to the Voting Rights Act were followed but no political data, election results or incumbents addresses were considered when creating new districts. The commission met twice over the summer of 2016 to deliberate and draw maps. The maps resulting from this simulated redistricting commission were released in August 2016. The Judges agreed on a redistricting at the level of Voting Tabulation Districts (VTD). This coarser redistricting was refined at the level of census blocks to achieve districts with less than 0.1% population deviation. The original VTD based maps are used in our study.

2. Random Sampling of Reasonable Redistrictings

Central to our analysis is the ability to generate a large number of different "reasonable" redistrictings. This is accomplished by sampling a probability distribution on possible redistrictings of North Carolina. The distribution is constructed so that it is concentrated on "reasonable" redistrictings. We then filter the randomly drawn redistrictings, using only those which satisfy our criteria for being "reasonable."

As already mentioned, we take our definition of "reasonable" redistrictings from the unratified House Bill 92 (HB92) from the 2015 Session of the North Carolina General Assembly [2] which stated that a bipartisan commission should draw up redistrictings while observing the following principles:

- §120-4.52(f): Districts must be contiguous; areas that meet only at points are not considered to be contiguous.
- §120-4.52(c): Districts should have close to equal populations, with deviations from the ideal population division within 0.1%.
- §120-4.52(g): Districts should be reasonably compact, with (1) the maximum length and width of any given district being as close to equal as possible and (2) the total perimeter of all districts being as small as possible.
- §120-4.52(e): Counties will be split as infrequently as possible and into as few districts as possible. The division of Voting Tabulation Districts (VTDs) will also be minimized.
- §120-4.52(d): Redistrictings should comply with pre-existing federal and North Carolina state law, such as the Voting Rights Act (VRA) of 1965.
- §120-4.52(h): Districts shall not be drawn with the use of (1) political affiliations of registered voters, (2) previous election results, or (3) demographic information other than

[1] For more information see `https://sites.duke.edu/polis/projects/beyond-gerrymandering/`

[2] Nonpartisan Redistricting Commission. House Bill 92. General Assembly of North Carolina Session 2015. House DRH10039-ST-12 (02/05)

population. An exception may be made only when adhering to federal law (such as the VRA).

We restrict our probability distribution to redistrictings which have connected districts. The remaining principles are encoded in a *score function* which is minimized by redistrictings that are most successful at satisfying the remaining design principles. We introduce some mathematical formalisms in order to describe the score function.

We represent the state of North Carolina as a graph G with edges E and vertices V. Each vertex represents a Voting Tabulation District (VTD) and an edge between two vertices exists if the two VTDs are adjacent on the map. This graph representing the North Carolina voting landscape has over 2500 vertices and over 8000 edges.

Since North Carolina has thirteen seats in the U.S. House of Representatives, we define a redistricting plan to be a function from the set of vertices to the integers between one and thirteen. More formally, recalling that V was the set of vertices, we represent a redistricting plan by a function $\xi : V \to \{1, 2, \ldots, 13\}$. If a VTD is represented be a vertex $v \in V$, then $\xi(v) = i$ means that the VTD in question belongs to district i. Similarly for $i \in \{1, 2, \ldots, 13\}$ and a plan ξ, the i-th district, which we denote by $D_i(\xi)$, is given by the set $\{v \in V : \xi(v) = i\}$. We wish to only consider redistricting plans ξ such that each district $D_i(\xi)$ is a single connected component. We will denote the collection of all redistricting plans with connected districts by $\mathcal{R}$.

2.1. **The Score Function.** We now wish to define a function J that assigns a nonnegative number $J(\xi)$ to every redistricting $\xi \in \mathcal{R}$. To do this, we employ functions J_p, J_I, J_c, and J_m that measure how well a given redistricting satisfies the individual principles outlined in HB92. The *population score* $J_p(\xi)$ measures how well the redistricting ξ partitions the population of North Carolina into 13 equal parts. The *isoperimetric score* $J_I(\xi)$ measures how compact the districts are by returning the sum of the isoperimetric constants for each district, a quantity which is minimized by a circle. The *county score* $J_c(\xi)$ measures the number of counties split between multiple districts; the minimum is achieved when there are no split counties. Lastly, the *minority score* $J_m(\xi)$ measures the extent to which the districts with the largest percentage of African-Americans achieve stipulated target percentages. With these, we then define our score function J to be a weighted sum of J_p, J_I, J_c, and J_m; we use a weighted combination so as to not give one of the above scores undue influence since all of the score functions do not necessarily change on the same scale. Specifically, we define:

$$J(\xi) = w_p J_p(\xi) + w_I J_I(\xi) + w_c J_c(\xi) + w_m J_m(\xi), \tag{1}$$

where w_p, w_I, w_c, and w_m are a collection of positive weights.

To describe the individual score functions, we attach to our graph $G = (V, E)$ some data which gives relevant features of each VTD. We define the positive functions $\text{pop}(v)$, $\text{area}(v)$, and $\text{AA}(v)$ for a vertex $v \in V$ as respectively the total population, geographic area, and African-American population of the VTD associated with the vertex v. We extend these functions to a collection of vertices $B \subset V$ by

$$\text{pop}(B) = \sum_{v \in B} \text{pop}(v), \quad \text{area}(B) = \sum_{v \in B} \text{area}(v), \quad \text{AA}(B) = \sum_{v \in B} \text{AA}(v)\,. \tag{2}$$

We think of the boundary of a district $D_i(\xi)$ as the subset of the edges E which connect vertices inside of $D_i(\xi)$ to vertices outside of $D_i(\xi)$. We write $\partial D_i(\xi)$ for the boundary of the district $D_i(\xi)$. Since we want to include the exterior boundary of each district (the section bordering an adjacent state or the ocean), we add to V the vertex o which represents the "outside" and connect it with an edge to each vertex representing a VTD which is on the boundary of the state. We always assume that any redistricting ξ always satisfies $\xi(v) = 0$ if and only if $v = o$. Since ξ always satisfies $\xi(o) = 0$ and hence $o \notin D_i(\xi)$ for $i \geq 1$, it does not matter that we have not defined $\text{area}(o)$ or $\text{pop}(o)$, as o is never included in the districts.

Given an edge $e \in E$ which connects the two vertices $v, \tilde{v} \in V$, we define boundary(e) to be the length of common border of the VTDs associated with the vertex v and $\tilde{v}$. As before, we extend the definition to the boundary of a set of edges $B \subset E$ by

$$\text{boundary}(B) = \sum_{e \in B} \text{boundary}(e)\,. \tag{3}$$

With these preliminaries out of the way, we turn to defining the first three score functions used to assess the goodness of a redistricting.

2.1.1. *The population score function.* We define the population score by

$$J_p(\xi) = \sum_{i=1}^{13} \Big(\frac{\text{pop}(D_i(\xi))}{\text{pop}_{\text{Ideal}}} - 1\Big)^2, \quad \text{pop}_{\text{Ideal}} = \frac{N_{pop}}{13}$$

where N_{pop} is the total population of North Carolina, $\text{pop}(D_i(\xi))$ is the population of the district $D_i(\xi)$ as defined in (2), and $\text{pop}_{\text{Ideal}}$ is the population that each district should have according to the 'one person one vote' standard; namely, $\text{pop}_{\text{Ideal}}$ is equal to one-thirteenth of the total state population.

2.1.2. *The Isoperimetric score function.* The Isoperimetric score J_I, which measure the compactness of a district, is the ratio of the perimeter to the total area of each district. The Isoperimetric score is minimized for a circle, which is the most compact shape. Hence we define

$$J_I(\xi) = \sum_{i=1}^{13} \frac{\big[\text{boundary}(\partial D_i(\xi))\big]^2}{\text{area}(D_i(\xi))}\,.$$

where $\partial D_i(\xi)$ is the set of edges which define the boundary, boundary$(\partial D_i(\xi))$ is the length of the boundary of district D_i and area$(D_i(\xi))$ is its area.

This compactness measure is one of two measures often used in the legal literature where it is referred to as *the parameter score* (See [11, 12] from references). This second measure, usually referred to as *the dispersion score*, is more sensitive to overly elongated districts, though the parameter score also penalizes them. The dispersion score does not penalize undulating boundaries while the parameter score (our J_I) does.

2.1.3. *The county score function.* The county score function $J_c(\xi)$ penalizes redistrictings which contain single counties contained in two or more districts. We refer to these counties as split counties. The score consists of the number of counties split over two different districts times a factor $W_2(\xi)$ plus a large constant M_C times the number of counties split over three of more different districts times a second factor $W_3(\xi)$. Specifically, we define:

$$\begin{aligned} J_c(\xi) =& \{\#\text{ counties split between 2 districts}\} \cdot W_2(\xi) \\ &+ M_C \cdot \{\#\text{ counties split between} \geq 3\text{ districts}\} \cdot W_3(\xi) \end{aligned}$$

where M_C is a large constant and the weights $W_2(\xi)$ and $W_3(\xi)$ are defined by

$$W_2(\xi) = \sum_{\substack{\text{counties} \\ \text{split between} \\ \text{2 districts}}} \begin{pmatrix}\text{Fraction of county VTDs in 2nd largest} \\ \text{intersection of a district with the county}\end{pmatrix}^{\frac{1}{2}}$$

$$W_3(\xi) = \sum_{\substack{\text{counties} \\ \text{split between} \\ \geq 3\text{ districts}}} \begin{pmatrix}\text{Fraction of county VTDs not in 1st or 2nd} \\ \text{largest intersection of a district with the county}\end{pmatrix}^{\frac{1}{2}}$$

The factors $W_2(\xi)$ and $W_3(\xi)$ make the score function vary in a more continuous fashion, which encourages reduction of the smaller fraction of a split county.

2.1.4. *The Voting Rights Act or minority score function.* It is less clear what it means for a redistricting to comply with the VRA. African-American voters make up approximately 20% of the eligible voters in North Carolina. Since 0.2 is between $\frac{2}{13}$ and $\frac{3}{13}$, the current judicial interpretation of the VRA stipulates that at least two districts should have enough African-American representation so that this demographic may elect a candidate of their choice. However, the NC2012 redistricting plan was ruled unconstitutional because two districts, each containing over 50% African-Americans, were ruled to have been packed too heavily with African-Americans, diluting their influence in other districts. The NC2016 redistricting was accepted based on racial considerations of the VRA and contained districts that held 44.48% African-Americans, and 36.20% African-Americans. The amount of deviation constitutionally allowed from these numbers is unclear.

Based on these considerations, we chose a VRA score function which awards lower scores to redistrictings which had one district with at least 44.48% African-Americans and a second district with at least 36.20% African-Americans. We write

$$J_m(\xi) = \sqrt{H(44.48\% - m_1)} + \sqrt{H(36.20\% - m_2)}, \tag{4}$$

where m_1 and m_2 represent the percentage of African-Americans in the districts with the highest and second highest percent of African-Americans, respectively. H is the function defined by $H(x) = 0$ for $x \leq 0$ and $H(x) = x$ for $x \geq 0$. We chose this function to make the transition smoother, and we utilize the square root function to encourage districts that are just above the threshold to be less probable than when no square root is included. Notice that whenever $m_1 \geq 4.484\%$ and $m_2 \geq 36.20\%$ we have that $J_m = 0$.

2.2. **The Probability Distributions on Redistrictings.** We now use the score function $J(\xi)$ to assign a probability to each redistricting $\xi \in \mathcal{R}$ that makes redistrictings with lower scores more likely. Fixing a $\beta > 0$, we define the probability of ξ, denoted by $\mathcal{P}_\beta(\xi)$, by

$$\mathcal{P}_\beta(\xi) = \frac{e^{-\beta J(\xi)}}{\mathcal{Z}_\beta} \tag{5}$$

where $\mathcal{Z}_\beta$ is the normalization constant defined so that $\mathcal{P}_\beta(\mathcal{R}) = 1$. Specifically,

$$\mathcal{Z}_\beta = \sum_{\xi \in \mathcal{R}} e^{-\beta J(\xi)}\,.$$

The positive constant β is often called the "inverse temperature" in analogy with statistical mechanics and gas dynamics. When β is very small (the high temperature regime), different elements of $\mathcal{R}$ have close to equal probability. As β increases ("the temperature decreases"), the measure concentrates the probability around the redistrictings $\xi \in \mathcal{R}$ which minimize $J(\xi)$.

2.3. **Generating Random Redistrictings.** If we neglect the fact that the individual districts in a redistricting need to be connected, then there are more than $13^{2500} \approx 7.2 \times 10^{2784}$ different redistrictings, larger than both the current estimate for the number of atoms in the universe (between 10^{78} and 10^{82}) and the estimated number of seconds since the Big Bang (4.3×10^{17}). While there are significantly fewer redistrictings in $\mathcal{R}$ (the set of simply connected redistrictings), it is not practical to enumerate all redistrictings to find those with the lowest values of J (i.e. the most probable ones).

The standard, very effective way to escape this curse of dimensionality is to use a Markov chain Monte Carlo (MCMC) algorithm to sample from the probability distribution $\mathcal{P}_\beta$. The basic idea is to define a random walk on $\mathcal{R}$ which has $\mathcal{P}_\beta$ as its unique, attracting stationary measure. We do this using the standard Metropolis-Hastings algorithm.

The Metropolis-Hastings algorithm is designed to use one Markov transition kernel Q (the proposal chain) to sample from another Markov transition kernel that has a unique stationary distribution μ (the target distribution). $Q(\xi, \xi')$ gives the probability of moving from the redistricting ξ to the redistricting ξ' in the proposal Markov chain and is readily computable. We use Q to draw a sample distributed according to μ. The algorithm proceeds as follows:

(1) Choose some initial state $\xi \in \mathcal{R}$.
(2) Propose a new state ξ' with transition probabilities given by $Q(\xi, \xi')$.
(3) Accept the proposed state with probability $p = \min\left(1, \frac{\mu(\xi')q(\xi',\xi)}{\mu(\xi)Q(\xi,\xi')}\right)$.
(4) Repeat steps 2 and 3.

The stationary distribution of this Markov chain matches the stationary measure μ. Thus, the states can be treated as samples from the desired distribution. The stationary measure we would like to sample is $\mathcal{P}_\beta$. We sample from three possible initial states: NC2012, NC2016, and Judges redistrictings. Since this algorithm is designed to converge to a unique stationary measure $\mathcal{P}_\beta$, any results should be independent of the initial starting point. However, this assumes the parameters have been chosen so that the time to equilibrate is short enough to happen during our runs. We show that the results are relatively independent of the initial condition in Section 5.2, which lends credence to the assertion that the algorithm is equilibrating.

We define the proposal chain Q used for proposing new redistrictings in the following way:

(1) Uniformly pick a conflicted edge at random. An edge, $e = (u, v)$ is a conflicted edge if $\xi(u) \neq \xi(v)$, $\xi(u) \neq 0$, $\xi(v) \neq 0$.
(2) For the chosen edge $e = (u, v)$, with probability $\frac{1}{2}$, either:

$$\xi'(w) = \begin{cases} \xi(w) & w \neq u \\ \xi(v) & u \end{cases} \qquad \text{or} \qquad \xi'(w) = \begin{cases} \xi(w) & w \neq v \\ \xi(u) & v \end{cases}$$

Let conflicted(ξ) be the number of conflicted edges for redistricting ξ. Then we have $Q(\xi, \xi') = \frac{1}{2}\frac{1}{\text{conflicted}(\xi)}$. The acceptance probability is given by:

$$p = \min\left(1, \frac{\text{conflicted}(\xi)}{\text{conflicted}(\xi')} e^{-\beta(J(\xi') - J(\xi))}\right)$$

If a redistricting ξ' is not connected, then we refuse the step, which is equivalent to setting $J(\xi') = \infty$.

Given a fixed set of weights (w_p, w_i, w_c, w_m), one still needs to determine an appropriate β so that typical samples from the distribution are "reasonable" redistrictings. If β is chosen to be too large, the algorithm will seek out a local minimum and leave this minimum with very low probability, meaning that it may require a large amount of steps to switch between high quality redistrictings. If β is chosen to be too low, then the algorithm will never find the locally good districts as it will choose redistrictings indiscriminately.

There are several well established ideas in the literature to overcome these challenges, including simulated annealing (e.g. [13]), parallel tempering (e.g. [6]) and simulated tempering (e.g. [5]). In the present work, we examine simulated annealing, in which β is set to be small at first until a certain number of steps are accepted (in the sense of step (3) from the algorithm in Section 2.3). This allows the system to explore the space of redistrictings more freely. Next, β is increased linearly to a maximum value over the course of a defined number of steps. This slowly "cools" the systems, relaxing it into a redistricting ξ which has a relatively low score $J(\xi)$. Finally, β is kept at this fixed maximum value for a defined number of steps so that the algorithm locally samples the measure $\mathcal{P}_\beta$ sufficiently long enough to produce a good redistricting.

The principal results quoted in Section 1 use the low β to be zero over 40,000 steps, linearly increase β to one over 60,000 steps, and fix β to be one for 20,000 steps before taking a sample.

This process is repeated for each sample redistricting. One potential critique with using simulated annealing is that the results may depend on the number of steps chosen above. We make a standard test to confirm that we have taken an appropriate number of steps by doubling each number of steps and repeating our analysis. The results of this test, which are found in Section 5.2, show that doubling the lengths has little effect on the results.

2.4. **Thresholding the sampled redistrictings.** Although drawing a redistricting with a bad score is unlikely when using the MCMC algorithm from Section 2.3 combined with the probability distribution given in (5), it is still possible. Additionally, the use of simulated annealing also increases the chance that we become stuck in a local minimum with a less than desirable score function, as such local minimum may take longer than the time we spend at high β to escape. These local trapping events can often lead to samples with less than perfect score functions. Lastly, our score functions do not perfectly encapsulate our redistricting design aesthetic. For example, since the isoperimetric score function is the sum of the individual isoperimetric scores of each district, it is still possible to have one bad district if the rest have exceptionally small isoperimetric scores.

Since we want to maximize the degree of compliance with HB92, we only use samples which pass an additional set of thresholds, one for each of the selection criteria. This additional layer of rejection sampling was also used in reference [5], though the authors of reference [5] chose to reweigh the samples to produce the uniform distribution over the set redistrictings that satisfy the thresholds. We prefer to continue to bias our sampling according to the score function so better redistrictings are given higher weights. We now detail our thresholding requirements.

It is our experience from the Beyond Gerrymandering project that redistrictings which use VTDs as their building blocks and have less that 1% population deviation can readily be driven to 0.1% population deviation by breaking the VTDs into census tracts and performing minimal alterations to the overall redistricting plan. We thus only accept redistrictings that have no districts above 1% population deviation. Many of our samples have deviations considerably below this value. It is important to emphasize that we require this of every district in the redistricting. In Section 5.1, we show that the results are quantitatively extremely similar, and qualitatively identical, when the population threshold is decreased from 1% to 0.75 % and then to 0.5%.

We have found that districts with isoperimetric scores under 60 are almost always reasonably compact. Thus, we choose to accept a redistricting only if each district in the plan has an isoperimetric ratio less than 60. The Judges redistricting plan would be accepted under this threshold as its least compact district has an isoperimetric score of 53.5. Neither NC2012 nor NC2016 would be accepted with this thresholding as the least compact districts of each plan have isoperimetric scores of 434.65 and 80.1, respectively. We also note that only two of the thirteen districts for the NC2012 plan meet our isoperimetric score threshold, whereas eight of the thirteen districts of NC2016 fall below the threshold.

Though redistrictings which split a single county in three are infrequent, they do occur among our samples. Since these are undesirable, we only accept redistrictings for which no counties are split across three or more districts. Note that, in order to satisfy population requirements, we must allow counties to be split into two districts because of the large populations of Wake and Mecklenburg Counties which each contain a population larger than a single Congressional district's ideal population. We do not explicitly threshold based on number of split counties, though redistrictings with more split counties have a higher scores, and hence are less favored.

To build a threshold based on minority requirements of the VRA, we note that the NC2016 redistricting was deemed to satisfy the VRA by the courts. The districts in this plan with the two highest proportion of African-Americans to total population are composed of 44.5% and 36.2%

African-Americans. With this in mind, we only accept redistrictings if the districts with the two highest percentages of African-American population have at least 40% and 33%, respectively.

The effect of all of these thresholds was to select around 16% of the samples initially produced by our MCMC runs. Though this leads to unused samples, it ensures that all of the redistrictings used meet certain minimal standards. This in turn allowed us to better adhere to the spirit of HB92. The reported 24,000 samples used in our study refer to those left after thresholding. The full data set of samples was in excess of 150,000. That being said, we show in Section 5.1 that results without thresholding were quantitatively very close and qualitatively identical. As already mentioned, we also show that decreasing the population threshold from 1% to 0.75% and then to 0.5% also has little effect on the quantitative results and no effect on the qualitative conclusions.

2.5. **Determining the weight parameters.** As we have mentioned above, we have four independent weights (w_p, w_I, w_c, w_m) used in balancing the effect of the different scores in the total score $J(\xi)$. In addition to these parameters, we also have the low and high temperatures corresponding respectively to the max and min β used in the simulated annealing. Since β multiplies the weights, one of these degrees of freedom is redundant and can be set arbitrarily. We chose to fix the low temperature (high value of β) to be one.

To select appropriate parameters, we employ the following tuning method:

(1) Set all weights to zero.
(2) Find the smallest w_p such that a fraction of the results are within a desired threshold (for the current work we ensured that at least 25% of the redistrictings were below 0.5% population deviation, however we typically did much better than this).
(3) Using the w_p from the previous step, find the smallest w_I such that a fraction of the redistrictings have all districts below a given isoperimetric ratio (we ensured that at least 10% of the results were below this threshold; we chose a threshold of 60 (see section 2.4)).
(4) If above criteria for population is no longer met, repeat steps 2 through 4 until both conditions are satisfied
(5) Using the w_p and w_I from the previous steps, find the smallest w_m such that at least 50% of all redistrictings have at least one district with more than 40% African-Americans and a second district has at least 33% African-Americans.
(6) If the thresholds for population were overwhelmed by increasing w_m, repeat steps 2 through 6. If the thresholds for compactness were overwhelmed, repeat steps 3 through 6.
(7) Using the w_p, w_I, and w_m from the previous steps, find the smallest w_c such that we nearly always only have two county splits, and the number of two county splits are, on average, below 25 two county splits.
(8) If the thresholds for population are no longer satisfied, repeat steps 2 through 8. If the criteria for the compactness is no longer met, repeat steps 3 through 8. If the criteria for the minority populations is not satisfied, repeat steps 5 through 8. Otherwise, finish with a good set of parameters.

With this process, we settle on parameters $w_p = 3000$, $w_I = 2.5$, $w_c = 0.4$, and $w_m = 800$ and have used these parameters for all of the results presented in the main results above (section 1). In Section 5.3, we show reasonable variations of these choices have little qualitative effect on the results.

3. The effect of the VRA

The NC2012 districts were labeled unconstitutional as a racial gerrymander. We investigate the effect of the VRA on election outcomes by considering samples taken from simulations that do not take the VRA into account, which is to say that we set $w_m = 0$. We examine the distribution of elected Democrats along with the histogram box-plots in Figure 7. We find that the VRA, even with the more modest thresholds of 40% and 33% required African-Americans, favors the Republican party. Without the VRA, there is roughly a 65% chance that 7 or more Democrats will be elected, with a 20% chance that 8 Democrats will be elected; in contrast, with the VRA considered, there is a 50% chance that 7 or more Democrats are elected, with a 10% chance that 8 Democrats are elected. Even with this bias towards the Republican party caused by the VRA, the results produced by the NC2012 elections are still very atypical.

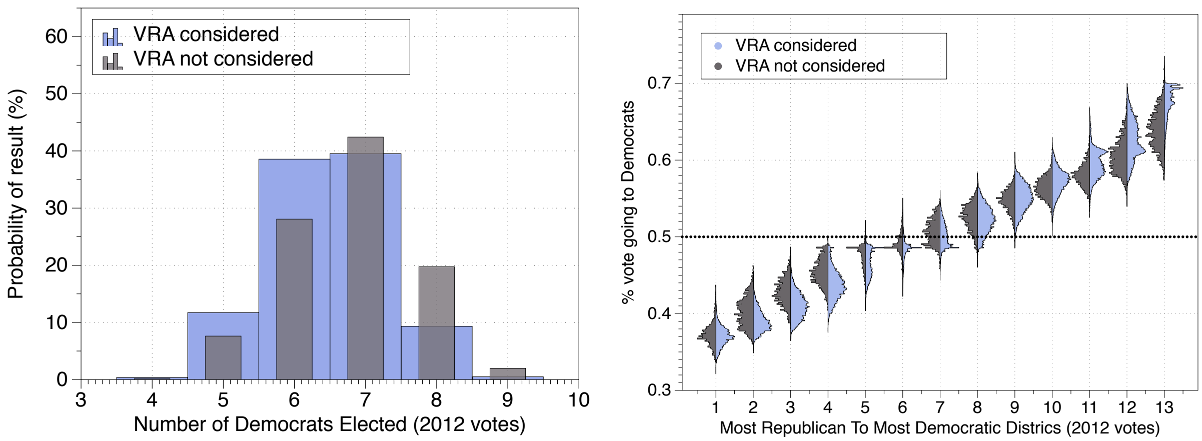


FIGURE 7. We display changes of the distribution of election results when the VRA is not taken into consideration (left). The histogram formed from the distribution of our main results overlays this image with the gray shaded histogram. We display changes to the histogram of the box-plot when comparing the results when VRA is considered or not (right).

4. Details of the Indices

We begin by expounding and clarifying how we compute the Gerrymandering Index and the Representativeness Index.

4.1. **Details of Gerrymandering Index.** To compute the Gerrymandering Index, we examine the mean percentage of Democratic votes in each of the thirteen districts when the districts are ordered from most to least Republican (see Figure 4). To calculate the Gerrymandering Index for any given redistricting plan, we take the Democratic votes for each district when the districts are again ordered from most to least Republican. The differences between the mean and the observed democratic percentage are taken for each district using a given set of votes. These differences are then each squared and summed over the 13 districts. The square root of this sum of squares is our Gerrymandering Index.

The Gerrymandering Index is smallest when all of the ordered Democratic vote percentages are precisely the mean values. However, this is likely not possible as the percentages in the different districts are highly correlated. To understand the range of possible values, we plot the complementary cumulative distribution function of the Gerrymandering Index of our ensemble of randomly generated reasonable redistrictings (see Figure 2). This gives a context in which to interpret any one score.

The mean percentages for the collection of redistricting we generated is

$$(0.37, 0.39, 0.41, 0.44, 0.46, 0.48, 0.50, 0.52, 0.55, 0.57, 0.60, 0.63, 0.67)\,.$$

If a given redistricting is associated with the sorted winning Democratic percentages

$$(0.36, 0.38, 0.39, 0.40, 0.41, 0.42, 0.43, 0.44, 0.49, 0.52, 0.64, 0.66, 0.7)\,.$$

then the Gerrymandering Index for the redistricting is the square root of

$$
\begin{aligned}
(0.37-0.36)^2 &+ (0.39-0.38)^2 + (0.41-0.39)^2 \\
&+ (0.44-0.40)^2 + (0.46-0.41)^2 + (0.48-0.42)^2 + (0.50-0.43)^2 \\
&+ (0.52-0.44)^2 + (0.55-0.49)^2 + (0.57-0.52)^2 + (0.60-0.64)^2 \\
&+ (0.63-0.66)^2 + (0.67-0.7)^2 = 0.0291
\end{aligned}
$$

In summary, in this example the Gerrymandering Index is $\sqrt{0.0291} = 0.17$.

4.2. **Details of Representativeness Index.** To calculate the Representativeness Index, we first construct a modified histogram of election results that captures how close an election was to swapping results. To do this for a given redistricting plan, we examine the least Republican district in which a Republican won, and the least Democratic district in which a Democrat won. We then linearly interpolate between these districts and find where the interpolated line intersects with the 50% line. For example, in the 2012 election, the 9th most Republican district elected a Republican with 53.3% of the vote, and the fourth most Democratic district won their district with 50.1% of the vote. We would then calculate where these two vote counts cross the 50% line, which will be

$$\frac{50-(100-50.1)}{53.3-(100-50.1)} \approx 0.03, \tag{6}$$

and add this to the number of Democratic seats won to arrive at the continuous value of 4.03. This index allows us to construct a continuous variable that contains information on the number of Democrats elected, and also demonstrates how much safety there is in the victory.

Fractional parts close to zero suggest that the most competitive Democratic race is less likely to go Democratic than the most competitive Republican race is to go Republican. On the other hand, fractional parts close to one suggest that the most competitive Republican race is less likely to go Republican than the most competitive Democratic race is to go Democratic. Instead of simply creating a histogram of the number of seats won by the Democrats, in Figure 8 we construct a histogram of our new interpolated value. We define the representativeness as the distance from the interpolated value to the mean value of this histogram (shown in the dashed line). These are the values we report in Figure 3. For the 2012 vote data, we find that the mean interpolated Democratic seats won is 7.01, and the Judges plan yields a value of 6.28, giving a Representative Index of $|7.01 - 6.28| = 0.73$. The NC2012 and NC2016 plans both have representative indices greater than two.

5. Testing the Sensitivity of Results

We wish to ensure that our algorithm has sampled the space of redistrictings in a robust way. We use this section to carefully study the effect of changing the set of threshold values, changing the weights in our distribution, and changing simulated annealing parameters on election results. We also verify that the choice of the initial district does not influence our results and that this information is lost as the algorithm updates the redistrictings.

5.1. **Varying thresholds.** Achieving a 0.1% population deviation is the only statute of HB92 that we violate. Although we have noted above that the Judges original redistrictings in the 'Beyond Gerrymandering' project were all slightly over 1% population deviation, and splitting VTDs to fall below this threshold had little impact on the election results. We test this for our own redistrictings by changing the population threshold to 0.75% and 0.5%. The results are shown in Figure 9, for which we have used the 2012 vote data. We find that tightening the population threshold has negligible impact on the number of Democrats elected, and that the variation in the histogram box-plots is barely perceptible. In the 0.5% population deviation threshold plots, we have discarded

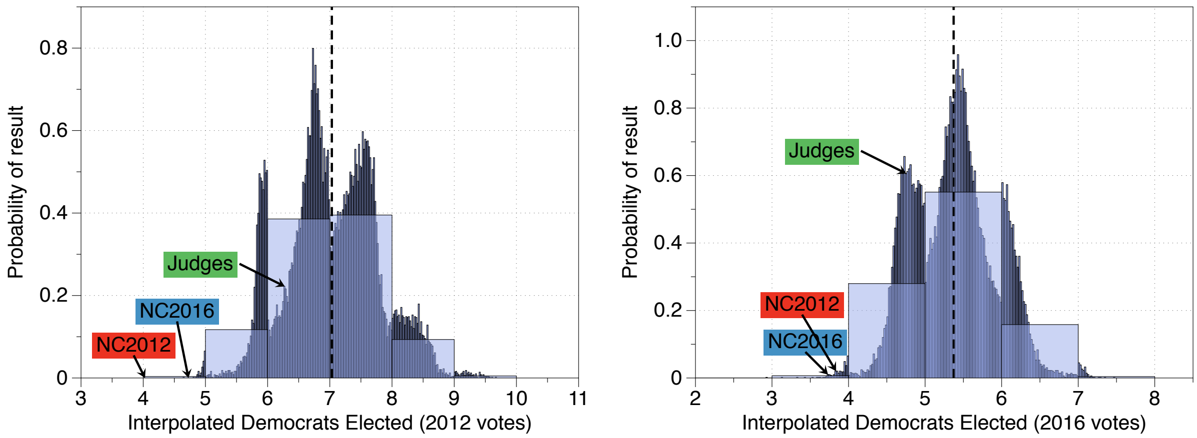


FIGURE 8. For the 2012 votes (left) and the 2016 votes (right), we plot the interpolated winning margins, which give the number of seats won by the Democrats in finer detail. We determine the mean of this new histogram and display it with the dashed line. The Representativeness Index is defined to be the distance from this mean value. The histogram presented in Figure 1 is overlaid on this plot for reference.

over half of our results and we still do not see any significant changes. These results support our claim that splitting VTDs to achieve a less than 0.1% deviation will have a negligible effect on our conclusions.

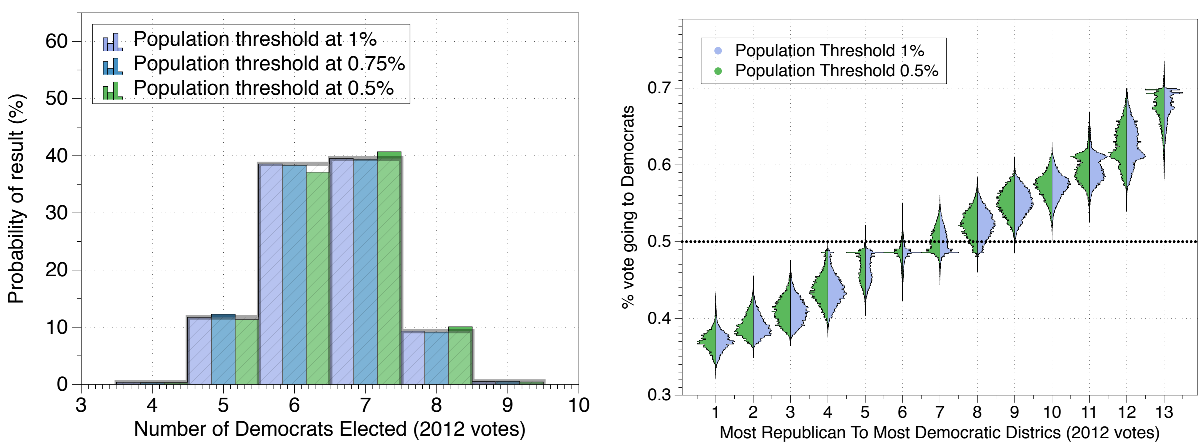
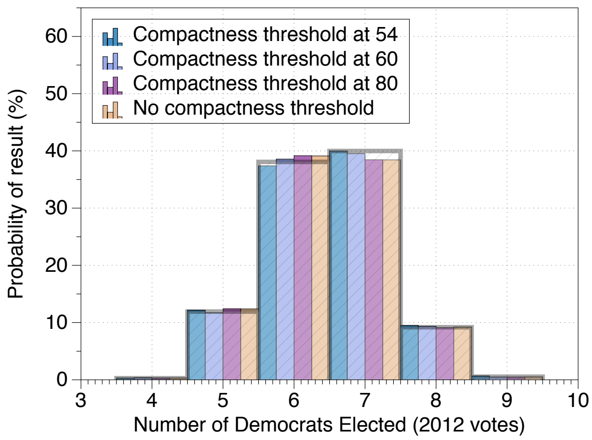


FIGURE 9. We display changes of the distribution of election results with changes to the population threshold (left). The histogram formed with 1% population deviation overlays this image with the gray shaded histogram. We display changes to the histogram of the box-plot when comparing 1% population deviation threshold with 0.5% (right).

Next, we note that there is no corresponding law to dicate an choice of compactness threshold. The NC2016 districts have a maximum isoperimetric ratio of around 80, and the NC2012 districts have a maximum of over 400. The Judges redistricting has a district with maximum isoperimetric ratio of around 54. To test the effect of setting different compactness thresholds, we repeat our analysis by choosing 54, 80 and no threshold for the maximum isoperimetric ratio of all districts within a redistricting. We find that relaxing the compactness threshold minimally changes the election results as demonstrated in Figure 10. We note that having no threshold does not mean that we have arbitrarily large compactness values. This is because of the cooling process in the simulated annealing algorithm and the fact that we continue to penalize large compactness scores. We find that we have an average maximum isoperimetric ratio of around 75 and that we rarely see redistrictings with maximal ratio larger than 120.



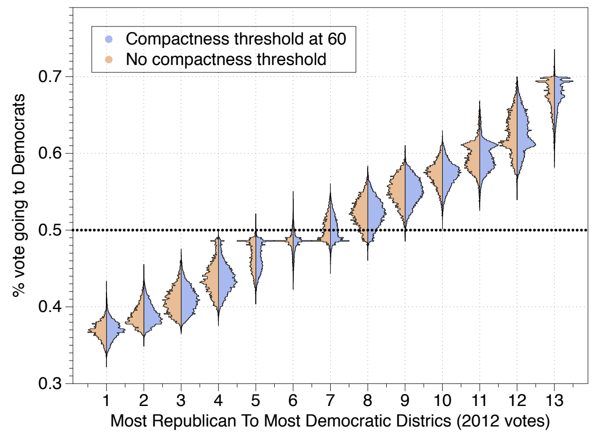


FIGURE 10. We display changes of the distribution of election results with changes to the compactness threshold (left). The histogram formed with a maximum of 60 for the isoperimetric ratio overlays this image with the gray shaded histogram. We display changes to the histogram of the box-plot when comparing a maximum of 60 in the isoperimetric ratio without any thresholding on compactness (right).

5.2. **Independence of initial conditions and simulated annealing parameters.** There is a possible pitfall of using simulated annealing: we may become trapped in local regions, leaving us unable to explore the entire space of redistrictings. This may be because we have cooled the system down too quickly, keeping it trapped in a local region, or it may be because the likelihood of finding a path out of one local region of redistrictings and into another is small. We note that we have animated our algorithm and have found that districts may travel from one end of the state to another; such motion suggests that many types of redistrictings are sampled, and it is reasonable to hypothesize that as districts exchange locations, they lose information on past configurations. To more fully vet this idea, we examine the effect of (i) choosing a different initial redistricting in our algorithm, and (ii) doubling the simulated annealing parameters, thus cooling the system down twice as slowly. To clarify the point (ii), instead of remaining hot ($\beta = 0$) for 40K steps, cooling linearly for 60K steps, and remaining cold ($\beta = 1$) for 20K steps, we instead remain hot for 80K steps, cool linearly for 120K steps, and remain cold for 40K steps. We then check to see if the election results are altered by changing these conditions and display our results in Figure 11.

We find that the changes with respect to both initial conditions and the slowdown of the annealing process have little effect on the election results. There are slight effects; for example, the initial condition for the NC2012 redistricting has a 15% chance of electing five Democrats rather than the 12% chance we have seen before. We note that these are exploratory runs, so we have less than 1000 accepted districtings for the NC2012 and NC2016 initial conditions (each has close to 1000) and less than 2500 runs for the increased cooling times. These sample sizes are robust enough to provide a general trend but are subject to statistical variations. Hence the small sample sizes are a possible and likely culprit of these variations.

5.3. **Varying weights.** We have proposed a methodology for determining the weights in the score function that is primarily concerned with obtaining a high percent of redistrictings below our chosen

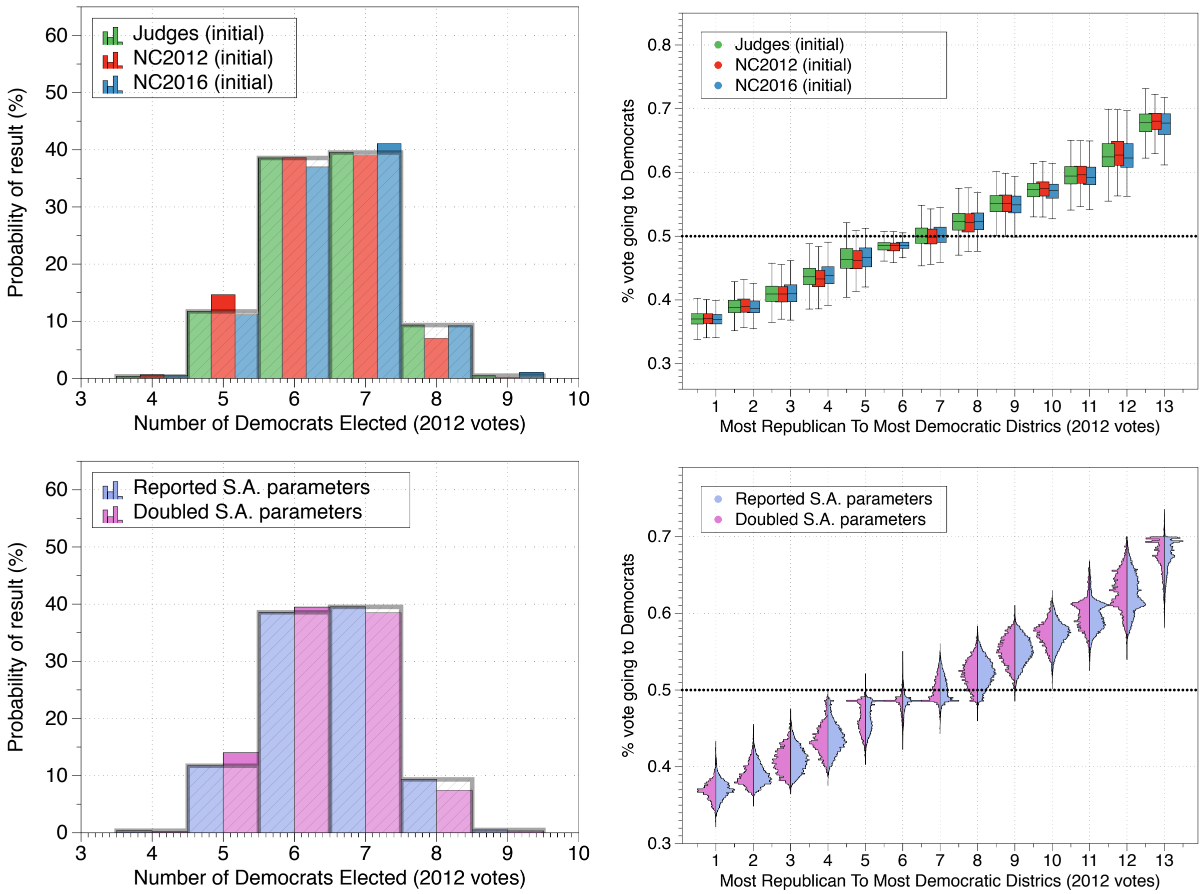


FIGURE 11. We display the probability distribution of elected Democrats with respect to initial conditions (top left) and the original versed doubled simulated annealing parameters (bottom left). The histogram formed with the Judges as an initial condition and the previously reported simulated annealing parameters overlays this image with the gray shaded histogram. We display traditional box-plots for the three initial conditions as we need to compare three results rather than two (top right) along with the histogram box-plots to compare the effect of changing the simulated annealing parameters (bottom right).

threshold values (see section 2.5). We note that other parameters may be chosen, and here we test whether making a different choice will affect the statistics on the election outcomes. We are in a four dimensional space, meaning that the parameter space is very large. Exploring this space thoroughly would come at an extraordinary computational cost. We instead perform a simple sensitivity test on our current location in the parameter space by exploring the four dimensional space in four linearly independent directions. We explore over three directions by significantly increasing and decreasing w_p, w_I, and w_m. For the fourth direction, we note that we could simply increase or decrease w_c; however, we thought it might be interesting to increase and decrease β instead. Because changing β is equivalent to changing all parameters, this forms a fourth linearly independent search direction, and provides us with information similar to changing w_c. This leads us to examine eight different parameter sets, which still requires a large number of runs. To cut down on the computational cost, we take advantage of the result presented in section 5.1 above, where we conclude that ignoring the compactness threshold has a minimal effect on our results.

The compactness threshold is by far the most restrictive, so omitting it will allow us to sample more redistrictings with fewer runs.

We present our results in Figure 12, and find that the results are very robust in all examined directions of changing parameters. We note, however, that the percentage of redistrictings that falls below our compactness acceptance threshold does change with varying parameters. Based on our result that election results are robust with respect to large changes in the compactness threshold, we conclude that significant changes in the parameters will have little effect on the statistical results of the election data.

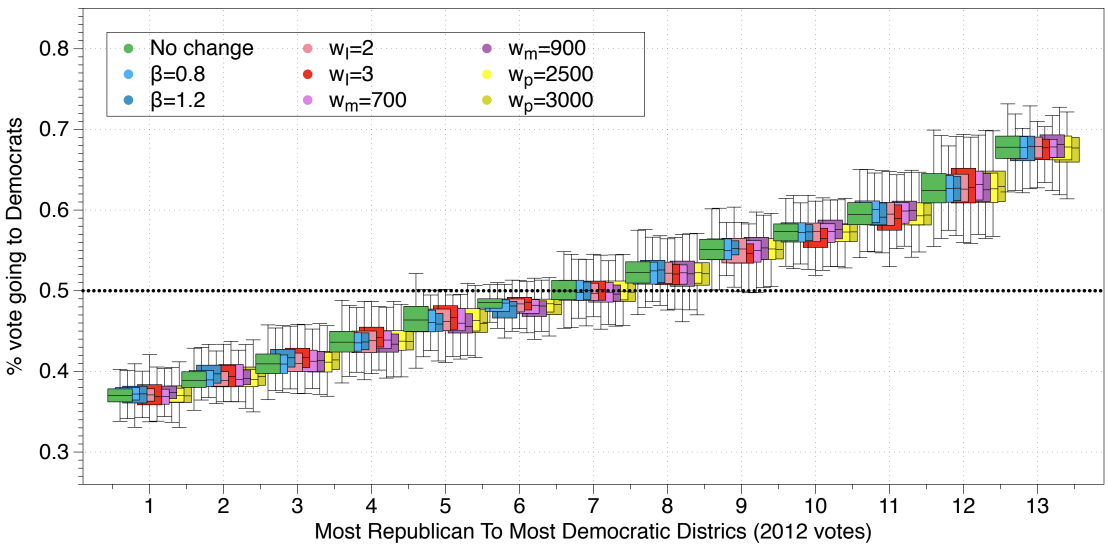


FIGURE 12. We display standard box-plots and demonstrate how the election results change with respect to changing the values of the weights.

6. TECHNICAL DISCUSSIONS

6.1. **Data sources and extraction.** The VTD geographic data were taken from the NCGA website (see [7] from references) and the United States Census Bureau website (see [3] from references), which provide for each VTD its area, population count of the 2010 census, the county in which the VTD lies, its shape and location. Perimeter lengths shared by VTDs were extracted in ArcMap from this data. Minority voting age population was found on the NCGA website using 2010 census data (see [8] from references). Data for the vote counts in each VTD for the 2012 House elections was taken from Harvard's Election Data Archive Dataverse (see [4] from references). Vote count data for the 2016 house elections was provided by NCSBE Public Data (See [10] from references). We note that for the 2016 election, VTD data was not reported for all VTDs, but rather for each precinct; 2447 of the precincts are VTDs, meaning that we have data for the majority of the 2692 VTDs. However 172 precincts contain multiple VTDs, 66 VTDs were reported with split data, and 7 VTDs were reported with complex relationships. To extrapolate VTD data on those contained in the 172 precincts containing multiple VTDs, we split the votes for a precinct among the VTDs it contained proportional to the population of each VTD. For the split VTDs, those containing multiple precincts, we simply added up the votes among the precincts it contained. There was no extrapolation for these VTDs. For the VTDs with complex relationships, we divided up the votes using estimates based on the geography and population of the VTDs. We note that roughly 10% of the population lies in the VTDs with imperfect data, and that we do not expect significant deviation in our results based on the above approximations.

6.2. **Examining nearby redistrictings within a distance.** The random sampling of the nearby districts is accomplished by running the same MCMC algorithm described in Section 2.3 with the small modification that if a proposed step ever tries to increase the deviation between any of the districts from the original redistricting in question (either NC2012, NC2016, or the Judges) above 40 VTDs, then the step is rejected and the chain does not move on that round. Alternatively, one can think of $J(\xi) = \infty$ for any $\xi \in \mathcal{R}$ which has a district that differs from the original redistricting by more that 40 VTDs. As before, we then threshold the results for NC2016 and the Judges on the Population Score, the County Score, and the Minority Score as described in Section 2.4. We do not threshold on the Isoperimetric Score as keeping the redistricting near the original is likely sufficient. We do not threshold the NC2012 at all since most of the redistrictings close to NC2012 would fail the Population threshold.

We examine the difference in the local complementary cumulative distribution function (thresholded and not) to get a sense of how accurate the NC2012 local complementary cumulative distribution function is without thresholding.

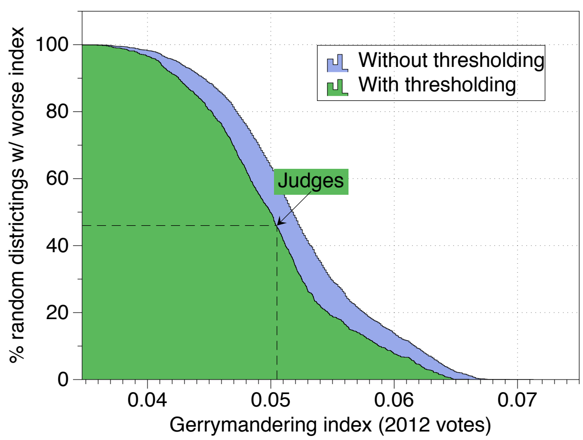


FIGURE 13. There is not a large difference between the thresholded and unthresholded results.

7. DISCUSSION

The most basic critique of this work is that we have assumed that the candidate does not matter, that a vote for the Democrat or Republican will not change, even after the districts are rearranged. Furthermore, as districts become more polarized and many elections results become a forgone conclusion, voter turn out is likely suppressed. While we could try to correct for these effects, we find the simplicity and power of using the actual votes very compelling. The results are so striking that we feel they are still very illuminating. In using 2012 and 2016 data, we have only used presidential election year data. Unfortunately, the 2014 U.S. congressional election in North Carolina contained an unopposed race which prevents the support for both parties being expressed in the VTDs contained in that district. In reference [1], the missing votes were replaced with votes from the Senate race. However, since we had two full elections, namely 2012 and 2016, which needed little to no alterations, we chose not to include the 2014 votes in this study.

REFERENCES

[1] Sachet Bangia, Bridget Dou, Jonathan C. Mattingly, Sophie Guo, and Christy Vaughn. Quantifying gerrymandering. `https://services.math.duke.edu/projects/gerrymandering/`, 2015.

[2] Sachet Bangia, Christy Vaughn Graves, Gregory Herschlag, Han Sung Kang, Justin Luo, Jonathan Mattingly, and Robert Ravier. Redistricting: Drawing the line. In Preperation, March 2017.
[3] U.S. Census Borough. Shapefiles for congressional districts. `http://www.census.gov/geo/maps-data/data/tiger-line.html`, 2010.
[4] Harvard's Election Data Archive Dataverse. 2012 house elections votes. `\\https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/21919`, 2012.
[5] Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. A new automated redistricting simulator using markov chain monte carlo. *Work. Pap., Princeton Univ., Princeton, NJ*, 2015.
[6] Ulrich HE Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *Chemical Physics Letters*, 281(1):140–150, 1997.
[7] North Carolina Legislature. 2010 VTD geographic data. http://www.ncleg.net/representation/Content/BaseData/BD2016.aspx, 2010.
[8] North Carolina Legislature. 2010 VTD minority data. `\\http://www.ncleg.net/representation/Content/BaseData/BD2011.aspx`, 2010.
[9] J. C. Mattingly and C. Vaughn. Redistricting and the Will of the People. *ArXiv e-prints*, October 2014.
[10] North Carolina Board of Elections. 2016 general election for US house by precinct. `http://dl.ncsbe.gov/index.html?prefix=ENRS/2016_11_08`, 2016.
[11] Richard H. Pildes and Richard G. Niemi. Expressive harms, "bizarre districts," and voting rights: Evaluating election-district appearances after Shaw v. Reno. *Michigan Law Review*, 92(3):483–587, Dec 1993.
[12] American Bar Association Section of Administrative Law Practice, Regulatory, and J. Gerald Hebert. *The realist's guide to redistricting: avoiding the legal pitfalls*. American Bar Association, Section of Administrative Law and Regulatory Practice,, 2nd ed. edition, 2010.
[13] Peter JM Van Laarhoven and Emile HL Aarts. Simulated annealing. In *Simulated Annealing: Theory and Applications*, pages 7–15. Springer, 1987.

Jonathan Mattingly, Departments of Mathematics and Statistical Science, Duke University, Durham NC 27708
E-mail address: `jonm@math.duke.edu`

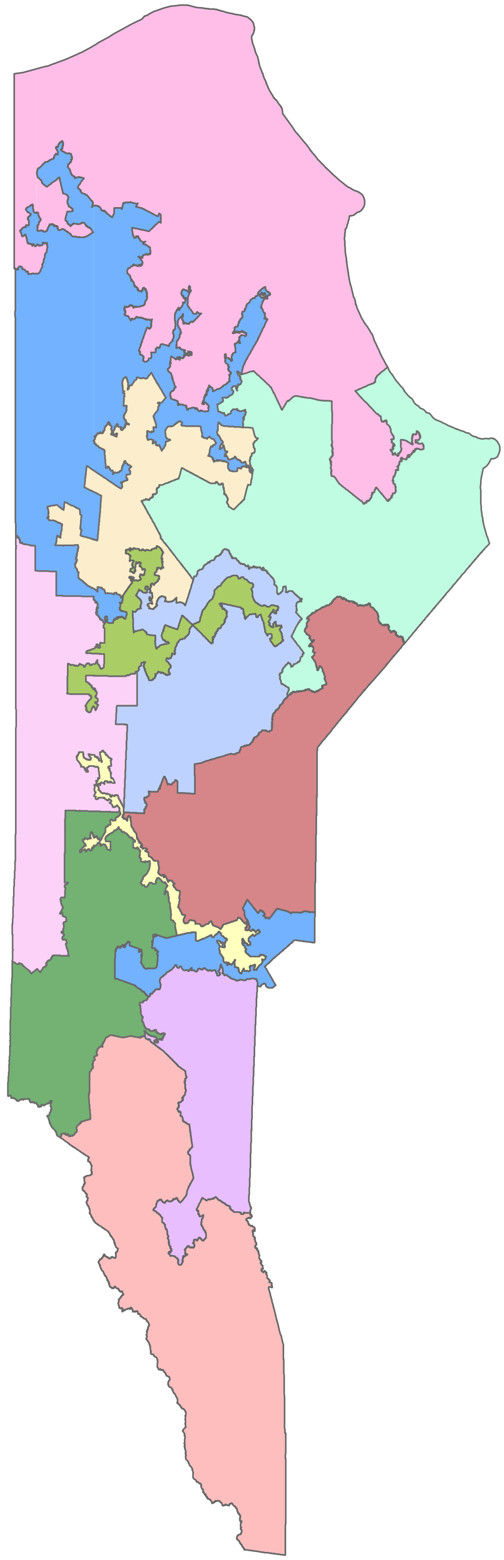


FIGURE 14. Map for NC2012

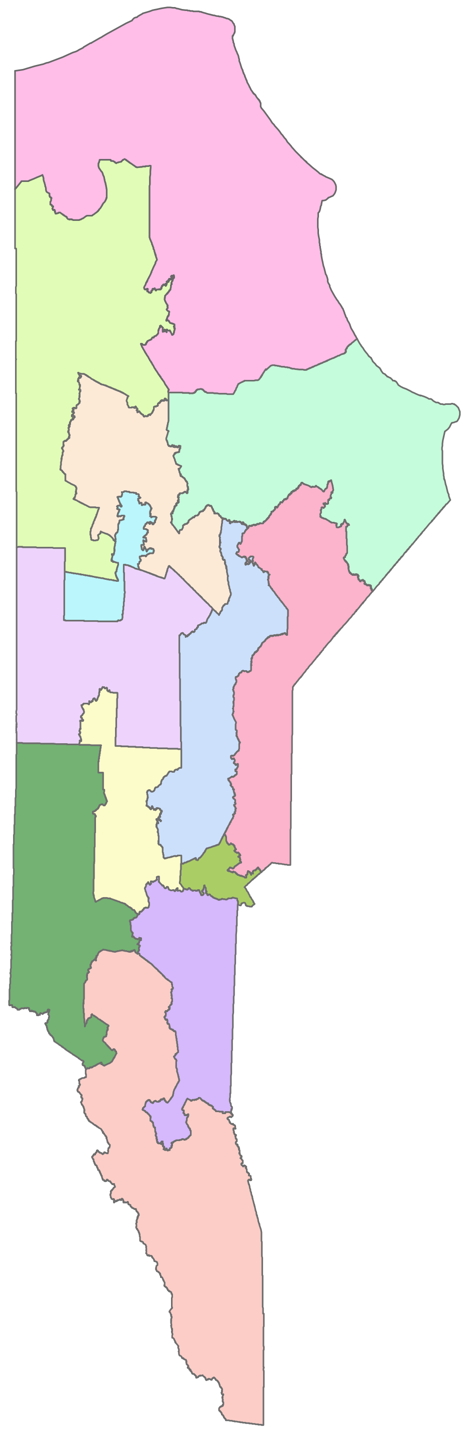


Figure 15. Map NC2016

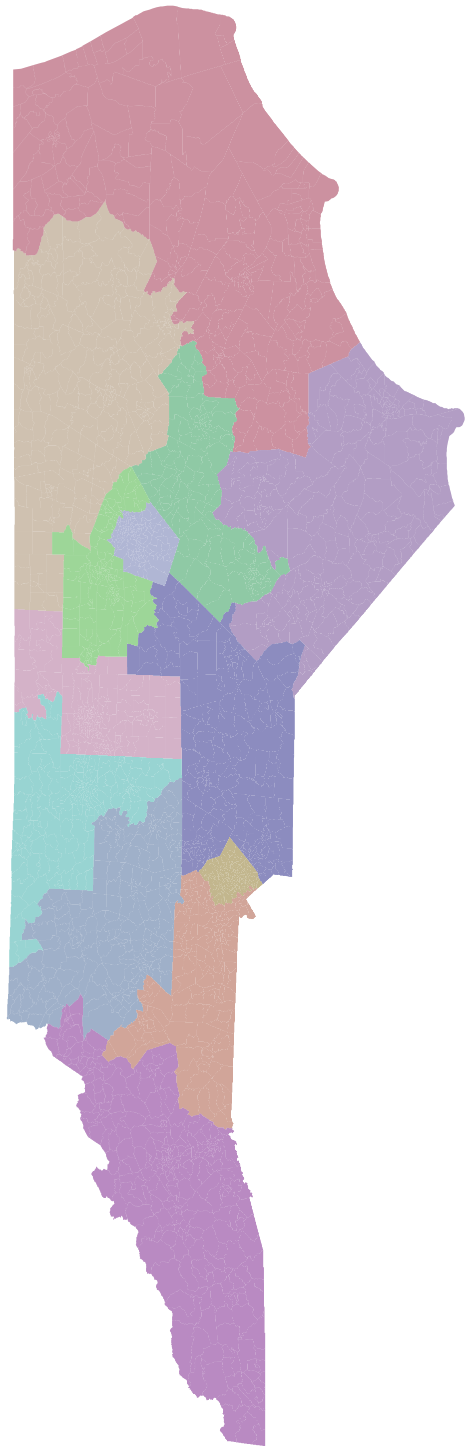


FIGURE 16. Map produce by bipartisan redistricting commission of retired Judges from Beyond Gerrymandering project

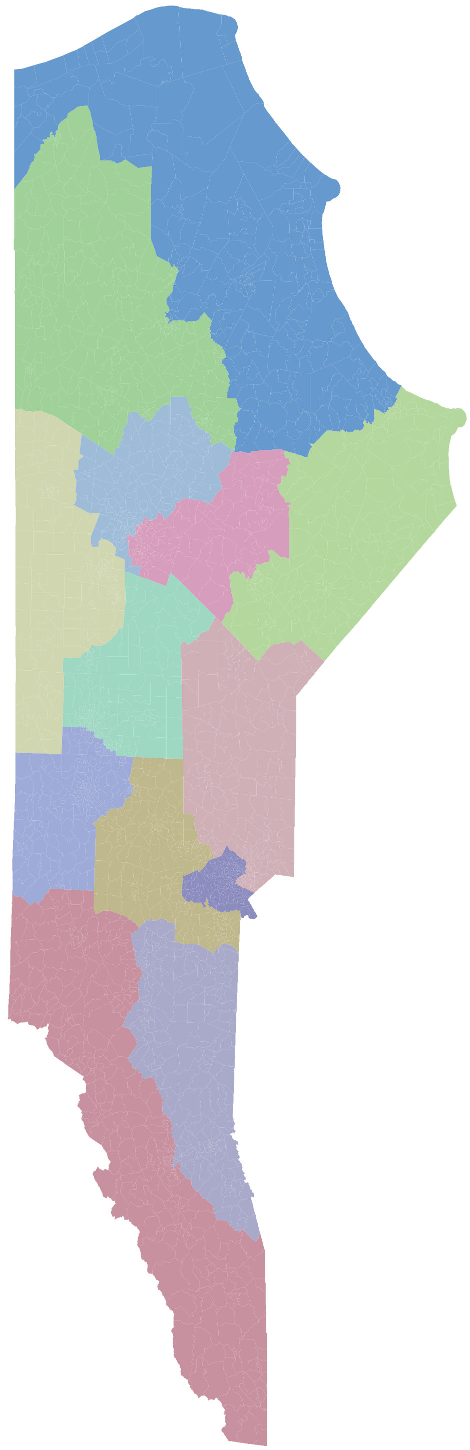


Figure 17. First sample redistricting generated by MCMC

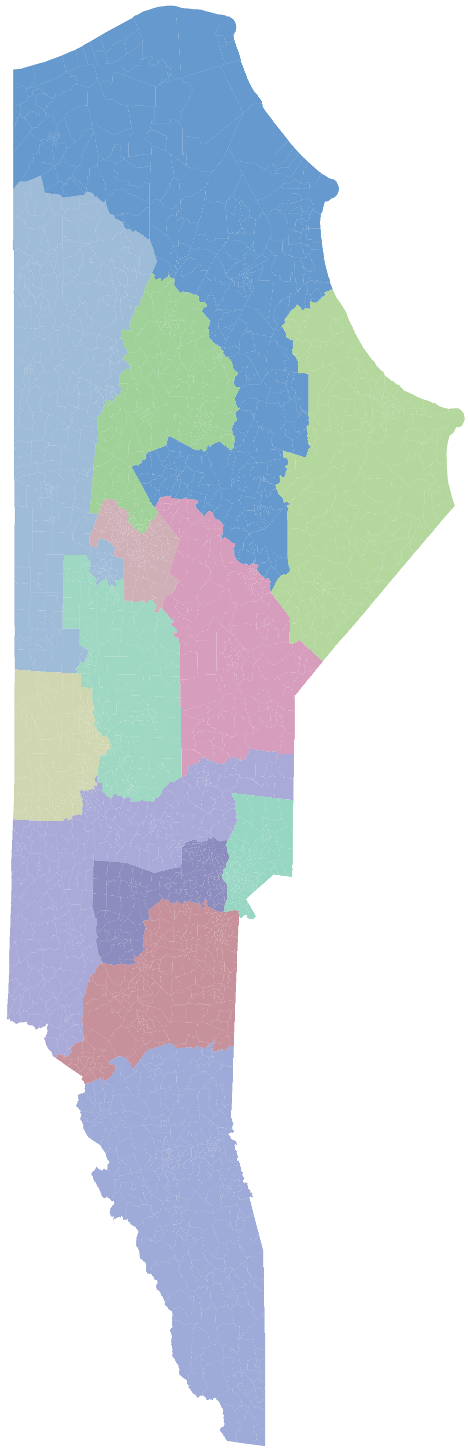


FIGURE 18. Second sample redistricting generated by MCMC

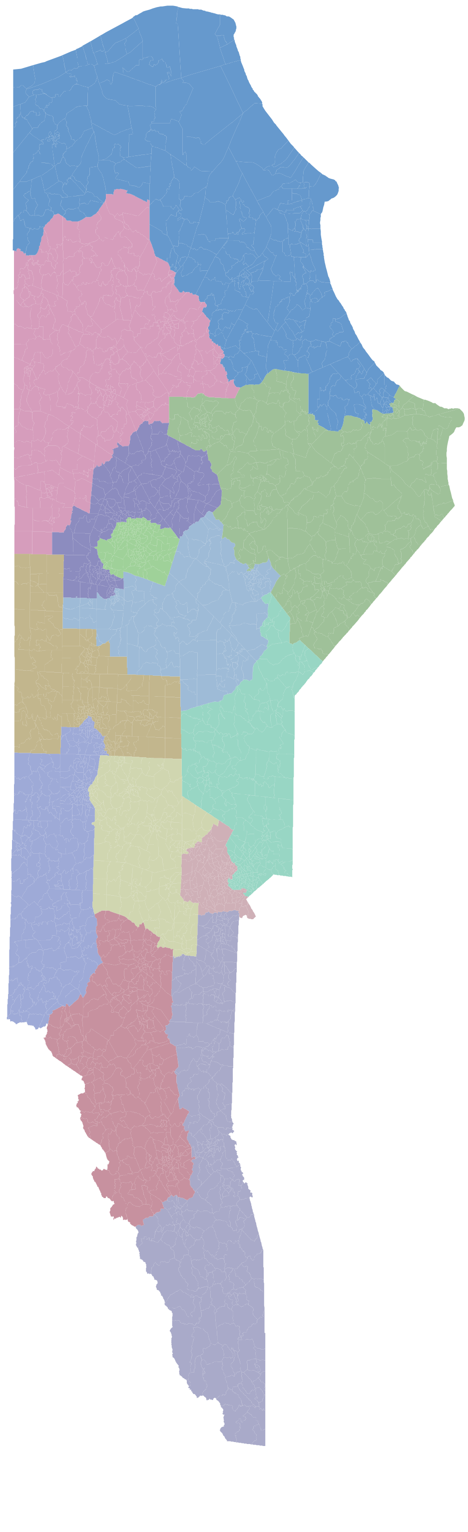


FIGURE 19. Third sample redistricting generated by MCMC

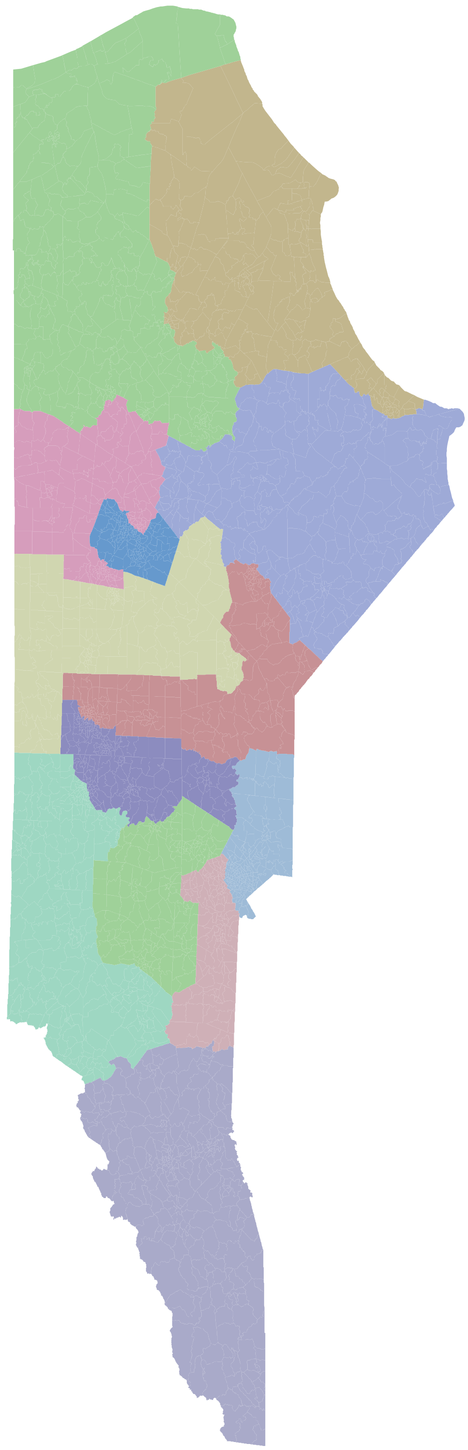


Figure 20. Fourth sample redistricting generated by MCMC

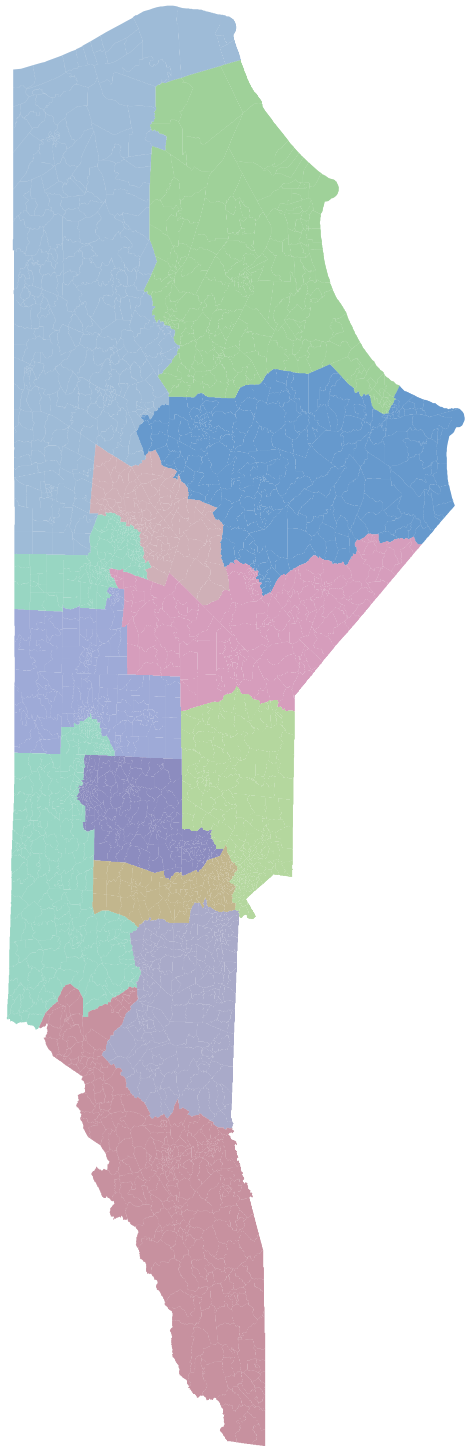


FIGURE 21. Fifth sample redistricting generated by MCMC

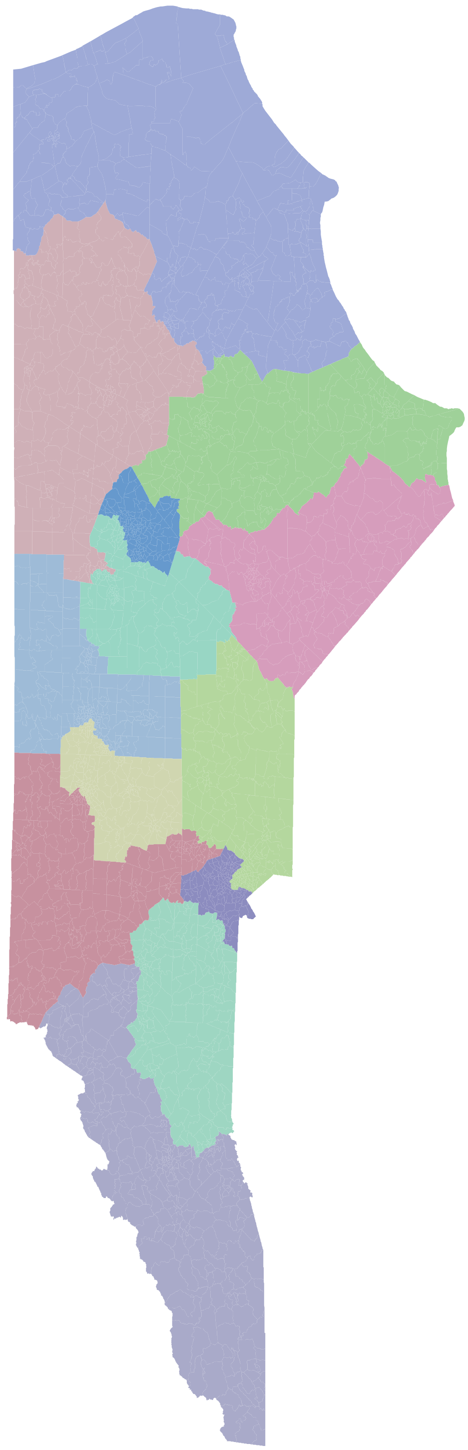


FIGURE 22. Sixth sample redistricting generated by MCMC

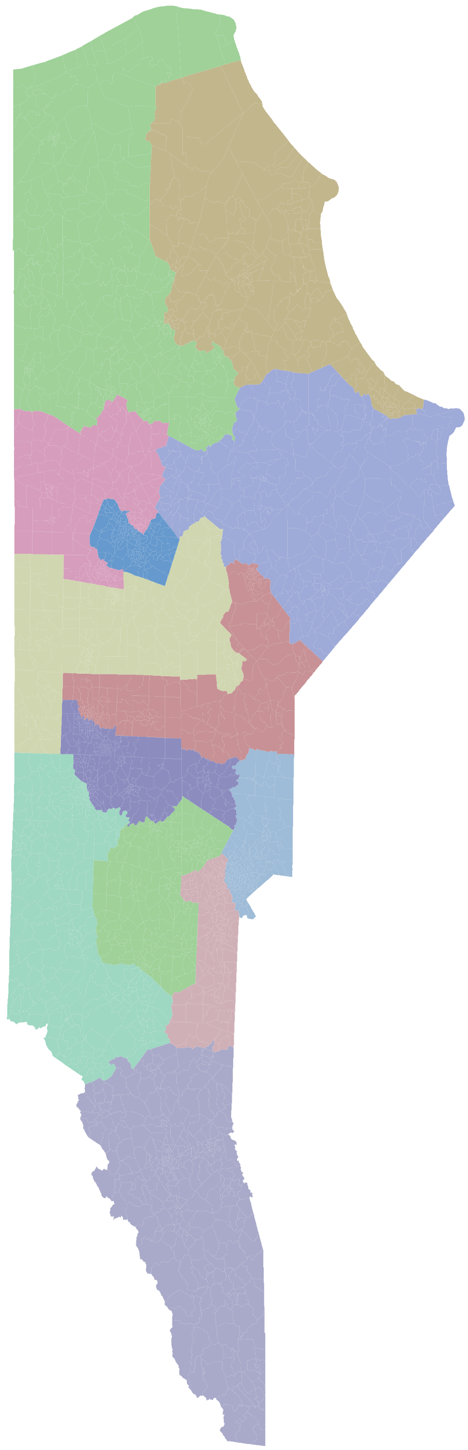


Figure 23. Seventh sample redistricting generated by MCMC

CERTIFICATE OF SERVICE

I hereby certify that I have this day served a copy of the foregoing Expert Disclosure of Jonathan Mattingly via email and by depositing a copy thereof in an envelope bearing sufficient postage in the United States mail, addressed to the following persons at the following addresses, which are the last addresses known to me:

Alexander McClure Peters
James Bernier, Jr.
N.C. Dept. of Justice
P.O. Box 629
Raleigh, NC 27602
jbernier@ncdoj.gov
apeters@ncdoj.gov

Thomas A. Farr
Michael D. McKnight
Phillip J. Strach
Ogletree Deakins Nash Smoak & Stewart, P.C.
4208 Six Forks Road, Suite 1100
Raleigh, NC 27609
michael.mcknight@odnss.com
phil.strach@ogletreedeakins.com
thomas.farr@ogletreedeakins.com

This the 6th day of March, 2017.

Caroline P. Mackie

Caroline P. Mackie