Analysis of Population Deviations in the 2021 Nebraska Legislative Redistricting Adopted and Proposed Maps

Steven R. Dunbar, Ph.D. and Kristin Pfabe, Ph.D.

February 22, 2022

Contents

1 Introduction

A fundamental principle in redistricting is that districts have equal populations; a map that does not adhere to this is called *malapportioned*. In Reynolds v. Sims, the U.S. Supreme Court ruled that state legislative districts must have close to equal populations.

The key argument was the "equal protection under the law" clause of the 14th Amendment. The ideal of "one person, one vote" is now embedded in state constitutions and legislative resolutions.

This analysis reveals malapportionment in the legislative districts adopted in Nebraska's 2021 redistricting process. Although the adopted map stayed

within the allowable deviations of populations of districts, the statistical analysis here shows the deviations have a strong pattern of giving less representational power to urban districts and more representational power to rural districts. Sections 2 and 3 present the conclusions with supporting graphics. Sections 4 and 5 present the supporting work for the conclusions.

2 Conclusions from Analysis of the Adopted Redistricting Map

1. The adopted 2021 Legislative redistricting plan, LEG21-39006, sponsored by Sen. Linehan, has an *evident pattern of over-populating urban Legislative Districts* with more residents, within the allowable 5% population deviation from the ideal district population of 40,031. Correspondingly, rural districts in the adopted Legislative redistricting plan, LEG21-39006, show a pattern of fewer residents, within the allowable 5% population deviation. The allowable district population deviation value of 5% is set by Legislative Resolution 134 to be in accord with other redistricting criteria, e.g. district boundary lines should coincide with boundaries of cities and villages.

   Figure 1 shows the map of 2021 Legislative Districts colored by population deviation, the orange districts showing populations above equal population and the purple districts showing populations below the ideal population. The orange districts cluster in the urban centers of Omaha and Lincoln, while the purple districts cover the rest of the state. (We chose these colors to avoid political connotations and to use best color practices for cartography.) An interactive version of the map allowing zooming and panning, and labeled with exact population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`, in the tab labeled 2021.

2. Statistical analysis of the adopted 2021 Legislative redistricting plan, LEG21-39006, sponsored by Sen. Linehan, shows a *definite correlation of district population deviation with population density.* The statistical analysis also shows the overall trend of district population deviation versus population density is stronger than alternatives. See Figure 2.

3. Advanced statistical analysis and comparison with 30,000 alternative redistricting maps show that the adopted 2021 Legislative redistricting plan, LEG21-39006, is *more extreme* in over-populating urban districts than 99% of alternatives. See Figure 3.

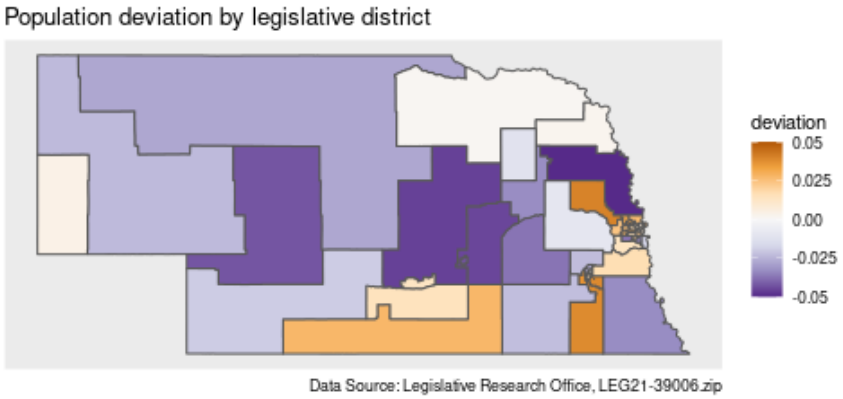

Figure 1: Map of adopted 2021 Legislative Districts, colored by deviation from ideal; purple represents a negative deviation, while orange represents positive deviation.

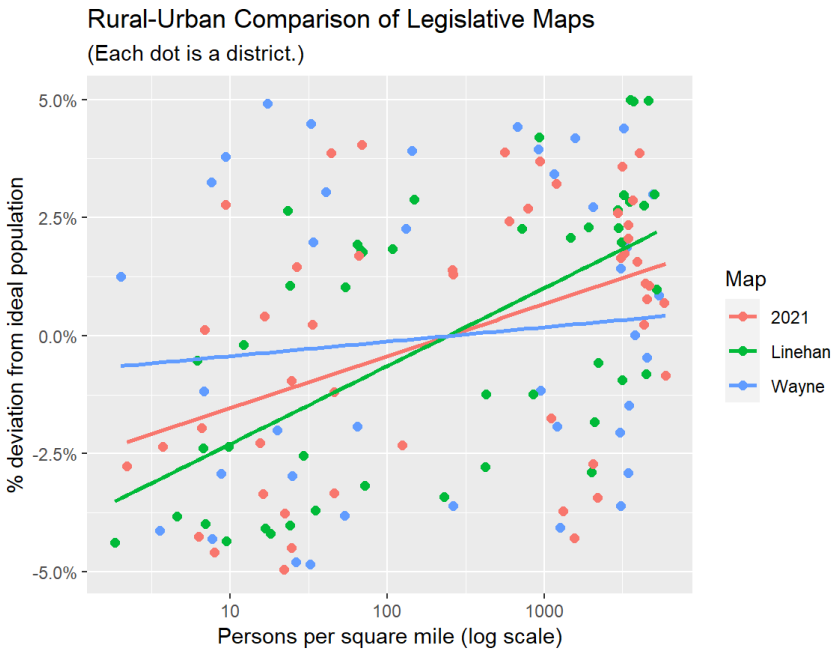

Figure 2: Scatter plot and best-fit lines of population deviation versus population density for proposed and adopted redistricting maps.

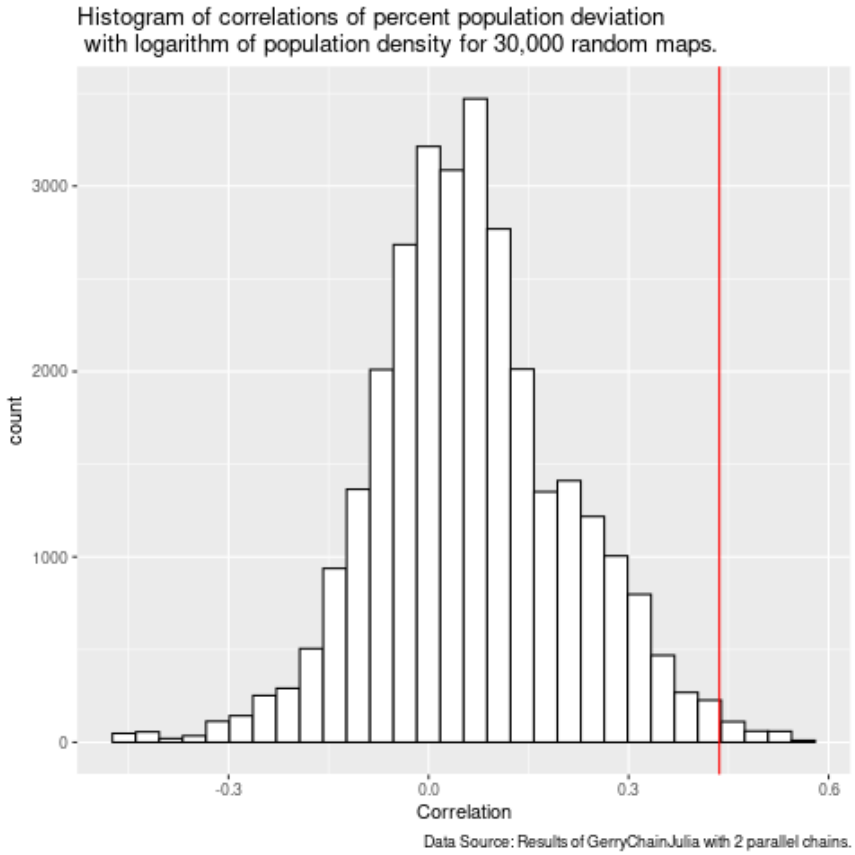

Figure 3: A histogram of correlations of percent population deviation versus logarithm of density for 30,000 alternative maps, compared to the correlation for the adopted 2021 map, indicated with a red line.

3 Comparative Conclusions from Analysis of Proposed Maps

1. Sen. Linehan's first proposed 2021 Legislative redistricting plan, LEG21-39001, which was not adopted but was later modified, shows an *even more pronounced pattern* of over-populating urban Legislative Districts with more residents, within the allowable 5% population deviation. Correspondingly, rural districts in Sen. Linehan's first Legislative redistricting plan LEG21-39001 show a pattern of fewer residents within allowable population deviation. See Figure 4. An interactive version of the map allowing zooming, panning and labeled with population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`. in the tab labeled Linehan.

2. Advanced statistical analysis and comparison with 30,000 alternative maps indicate that the original Legislative redistricting plan, LEG21-39001, sponsored by Sen. Linehan was *even more extreme* in over-populating urban districts that 99.6% of alternatives.

3. In contrast, Sen. Wayne's proposed, but not adopted, redistricting plan, LEG21-13001, shows *no pattern* of greater or lesser population deviation from ideal for urban or rural districts. See Figure 5. An interactive version of the map allowing zooming, panning and labeled with exact population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`. in the tab labeled Wayne.

4. Advanced statistical analysis and comparison with 30,000 alternative maps indicate that the defeated Legislative redistricting plan (LB13-001 sponsored by Sen. Wayne) was *neither extreme nor unusual* in over-populating urban districts.

5. Advanced statistical analysis of the 2011 Legislative Districts using 2010 population data show that in comparison with 10,000 alternative maps that the adopted 2011 plan was *neither extreme nor unusual* in over-populating urban districts. See Figure 6 .

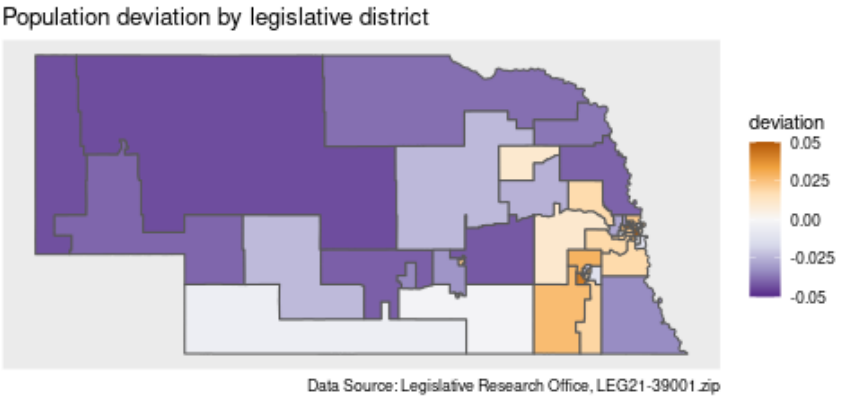

Figure 4: Map of Linehan's proposed Legislative Districts, colored by deviation from ideal; purple represents a negative deviation, while orange represents positive deviation.

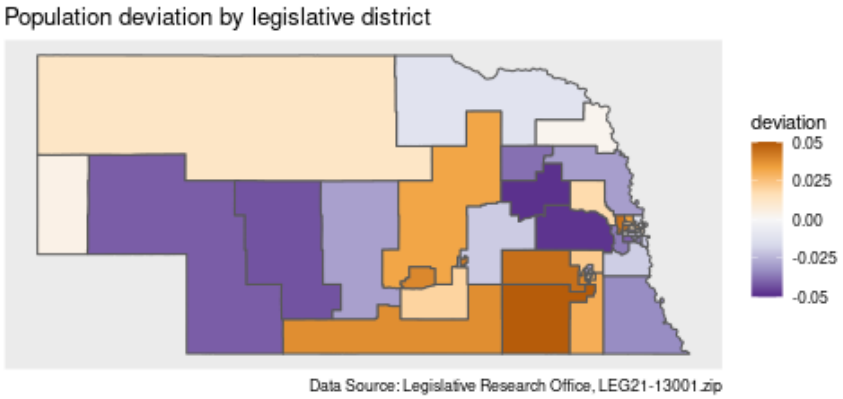

Figure 5: Map of Wayne's proposed Legislative Districts, colored by deviation from ideal; purple represents a negative deviation, while orange represents positive deviation.

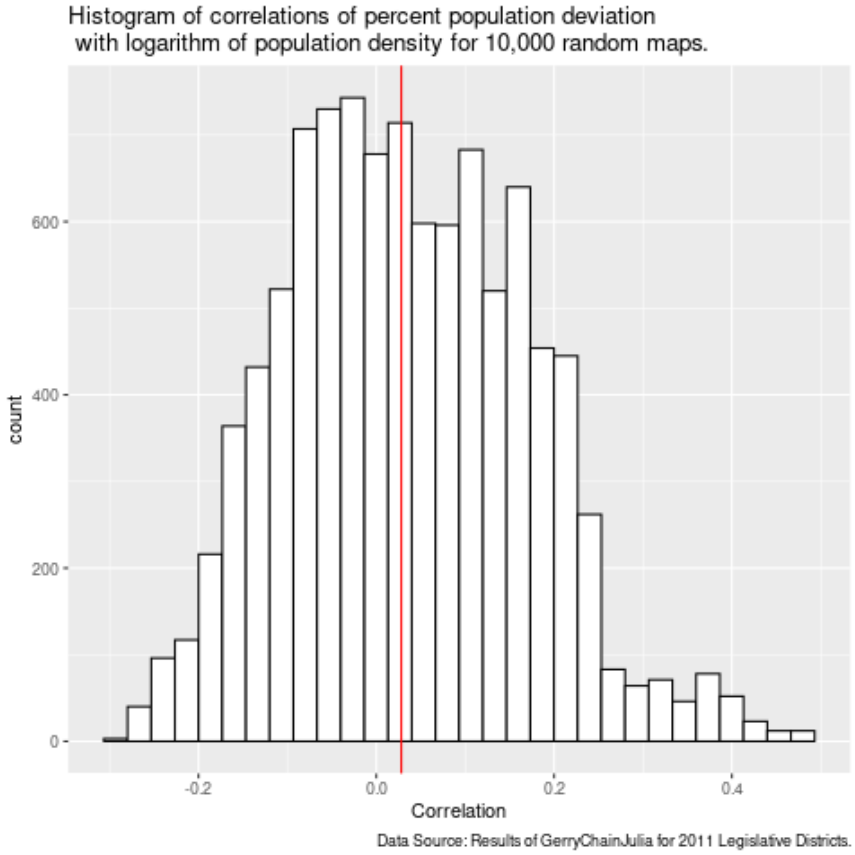

Figure 6: Correlation of deviation with density for the 2011 map compared to 10,000 alternative maps.

4 Analysis of the Adopted Redistricting Map

1. The analysis of the adopted 2021 legislative redistricting map uses the shapefile `LEG21-39006.zip` provided by the Legislative Research Office, see `https://drive.google.com/file/d/1hH4Y_g2CkPJz-PbusfNEhXO4_uI-x820/view?usp=sharing`. Using this shapefile with the code in the Code Section creates Figure 1. The legislative districts are colored by the percentage of population deviation from the ideal population of 40,031 residents per district. Purple indicates the district is under ideal population and orange indicates over ideal population. The intensity of the color indicates the magnitude of deviation, with a greater deviation having a more intense color. This legislative district map shows most of the orange districts intersect the major urban areas of Omaha and Lincoln. The purple districts are in the rural parts of the state. An evident pattern of consistently over-populating most of the urban Legislative Districts with more residents, within the allowable 5% population deviation, is apparent in Figure 1. An interactive version of the map allowing zooming and panning, and labeled with population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`. in the tab labeled 2021.

   A revealing statistical evaluation of the pattern is the following. Calculate the population density, in persons per square mile, of each district. Create a scatterplot of percent deviation from ideal versus population density. For the 49 Legislative districts, population density ranges from about 6 persons per square mile to over 6000 persons per square mile. Because the density ranges over 3 orders of magnitude, using this direct density is inconvenient, so instead use the base-10 logarithm of the population density. The logarithm has a range approximately 0 to 4.

   Each Legislative district has an ideal equal population of 40,031 persons. Because of the difficulty of drawing legislative districts with exactly this number of persons, the law allows up to a 5% allowable deviation from ideal, [1], [2]. The deviation of population is a percentage deviation from ideal for each district. Over the 49 districts, this is in the allowable range of $\pm 5\%$. This is comparable in magnitude to the base-10 logarithm of the population density. Both the logarithm of density and percent deviation are at a convenient scale for graphing and analysis.

A scatter plot of the percent deviation (vertical scale) versus the logarithm of the population density (horizontal scale) is in Figure 7. The broad Census Bureau definition of "urban" is census units with a population density of at least 1000 persons per square mile, [3]. Hence, this analysis classifies districts with a population density greater than 1000 residents per square mile (logarithm is greater than 3) as urban. The analysis classifies remaining districts as rural. Points representing urban districts are orange, points representing rural districts are purple. The dividing line in population density is a vertical dashed line at 3, urban to the right, rural to the left. The dividing line in population deviation is a horizontal dashed line at 0. The best fit regression line, or trend line, is the thick black line.

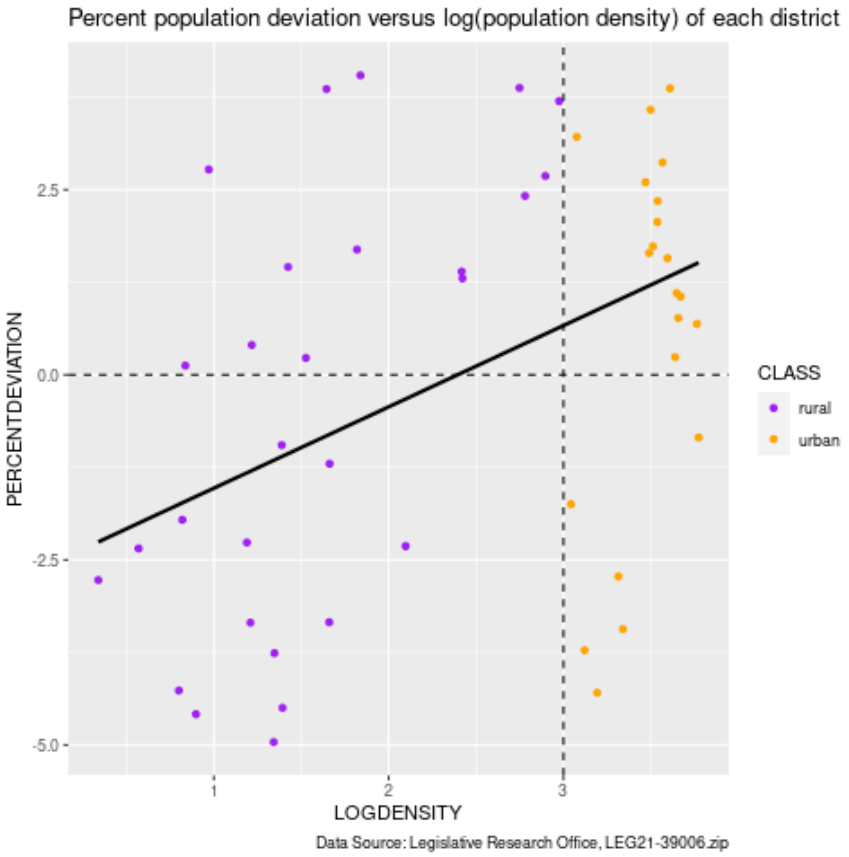

Figure 7: Scatter plot of deviation versus density for all districts for the adopted 2021 LEG21-39006 maps.

The clustering of points and the trend line on the scatter plot shows relationships among the points. At the simplest level, 15 of 21 districts classified as urban have positive percent deviation, that is, have greater than ideal population size. Those 15 districts are 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 18, 20, 26, 29, and 31. See Table 1 for more detail about those districts. All of the legislative districts in Douglas County have positive deviation but some districts are not classified as urban by this definition because they contain large tracts of agricultural, industrial or infrastructure (e.g. Eppley Airport). An example is District 13.

Table 1: Table of over-populated urban districts with general description, county, deviation and Senator

District	General Area	County	Percent Deviation	Senator
4	West Omaha	Douglas	+2.60	Hilkemann
5	South Omaha	Douglas	+2.35	McDonnell
6	Northwest Omaha	Douglas	+1.74	M. Cavanaugh
7	East Omaha	Douglas	+1.06	Vargas
8	Northwest Omaha	Douglas	+1.10	Hunt
9	Central Omaha	Douglas	+0.69	J. Cavanuagh
10	Northwest Omaha	Douglas	+3.87	DeBoer
11	North Omaha	Douglas	+0.77	McKinney
12	Southwest Omaha	Douglas	+1.64	Lathrop
14	Papillion	Sarpy	+2.06	Arch
18	Bennignton	Douglas	+3.21	Lindstrom
20	Southwest Omaha	Douglas	+3.58	McCollister
26	North Lincoln	Lancaster	+0.24	Hansen
29	South Lincoln	Lancaster	+1.57	Bostar
31	Southwest Omaha	Douglas	+2.87	Pahls

On the other hand, half of the 28 districts classified as rural are below ideal population, and half are above ideal population. To quantify this trend we use two statistical measures. The first is the Pearson correlation coefficient, or just the correlation for short. The correlation for deviation-versus-density for the final 2021 districts is $r = 0.315$. The correlation between the transformed data of percent deviation versus logarithm of density $r = 0.436$. These are positive, indicating increasing deviation with increasing population density. (The transformation to the logarithm of population density affects the value of the correla-

tion. However, scaling the 2 sets of data to the same magnitudes does not affect the positive association.) The size of the positive correlation is better understood in comparison to other proposed redistricting maps. The comparison appears later in this report. However, the correlation already shows that increasing population density is associated with greater deviation from ideal. That is, the plan tends to pack urban districts with more residents.

The second measure of relationship is the slope of the best-fit line through the scatter plot. The slope for the best-fit line through the final 2021 deviation-versus-density scatter plot is 4.733×10^{-6}. This slope is very small in magnitude because the rise in deviations only changes by about 0.1 while the run, the change in magnitude of the density, is about 6000. Nevertheless, the slope is positive. The slope for the best-fit line through the final 2021 percent deviation versus logarithm of population density scatter plot is 1.098 which is easier to interpret, justifying the transformation of variables. The important aspect of both is that the slope of the trend line is positive. Here again, the size of the positive slope is better understood in comparison to other proposed redistricting maps, the comparison appears below.

Even before comparison, a fundamental observation is: The correlation and the slope are positive, indicating the trend that districts with greater population density also have over-population deviation from ideal. Correspondingly, districts with lesser population density also have under population deviation from ideal.

2. Advanced statistical analysis using `GerryChainJulia`.

   In her opinion for Rucho v. Common Cause [4] Supreme Court Justice Kagan says "today's mapmakers can generate thousands of possibilities at the touch of a key and then choose the one giving their party maximum advantage" but "the same technologies and data that ... facilitate ... gerrymanders also enable courts to discover them". Mathematicians and computer scientists have developed scientific tools to detect and measure gerrymandering.

   `GerryChainJulia` is an open-source program specifically designed to statistically sample and evaluate alternative redistricting maps. The article [5] documents and evaluates the algorithms implemented in `GerryChainJulia` in more general terms. The code repository [6] documents the code for `GerryChainJulia`, written in the very fast and modern mathematical programming language Julia developed at MIT.

Briefly, `GerryChainJulia` makes a sequence of random changes, called a Markov Chain Monte Carlo method, to a given redistricting plan. The changes create a new set of valid districts at each step. The `GerryChainJulia` program evaluates characteristics of the individual districts and of the entire redistricting plan at each step. In our use of `GerryChainJulia`, the characteristics we evaluated were the population density of each generated district and the population deviation from ideal of each generated district. The program calculates correlation of the population deviation and the population density for each alternative redistricting plan created at each step. The result is a sequence of correlation values for alternative redistrictings that can be statistically compared to the correlation value for the adopted 2021 LEG21-39006.

The larger in magnitude a correlation between two variables is, the stronger is the relationship between the variables. To evaluate how unusual a strong correlation value of the original redistricting is with respect to randomly generated comparative values, we rank the correlation among all the alternatives. If the ranking of the particular redistricting correlation, is in say, the 99th percentile among all alternatives we call it an *outlier*. There is some chance an evaluation of a correlation as an outlier can occur because of the random nature of the Markov chain Monte Carlo method. A mathematical theorem in [7] shows that under conditions satisfied by the Markov Chain Monte Carlo method, the probability that an unusual redistricting occurring at the 99th percentile has about a 2% chance of occurring. In other words, observing by chance a redistricting which is unusual in this degree of correlation is a rare event. It is not expected by chance alone to see such a degree of correlation among a group of valid redistrictings. Therefore, it is reasonable to believe the unusual nature of the correlation occurs for some reason other than chance.

We ran the `GerryChainJulia` program with 2 parallel chains of 15,000 steps for a combined total of 30,000 alternative redistrictings to evaluate for correlation. Recall that the final 2021 LEG21-39006 had a positive correlation value of 0.315. Ranking this positive correlation for the adopted 2021 LEG21-39006 against 30,000 alternative maps, the correlation of the adopted map was at the 99.64 percentile. The mathematical theorem cited above predicts that among all similar runs, a percentile this large would occur about only 0.7% of the time. That is, the degree of correlation of the final 2021 redistricting LEG21-39006

compared to 30,000 randomly created alternatives makes it reasonable to believe the unusual nature of the correlation occurs for some reason other than chance. See Figure 3.

Similar results occurred in shorter testing runs done earlier:

- Comparing the adopted 2021 LEG21-39006 correlation to 10,000 correlations from two parallel chains of 5,000 steps each resulted in the adopted map correlation at the 99.4 percentile with a probability of about 1.3% of occurring.

- Comparing the adopted 2021 LEG21-39006 correlation to 20,000 correlations from two parallel chains of 10,000 steps each resulted in the adopted map correlation at the 99.1 percentile with a probability of about 1.8% of occurring.

This tells us that `GerryChainJulia` worked consistently.

Summarizing the results of these experiments in redistricting, a redistricting map having a correlation value at the 99th percentile or greater has less than a 2% probability of occurring solely by chance.

3. To determine if this degree of correlation is unusual among other redistricting plans, in the next section we make analogous comparisons with the 2011 maps and with other proposed redistricting plans considered in the September 2021 Special Session of the Nebraska Legislature.

5 Analysis of Proposed Maps

1. Comparative Analysis of Sen. Linehan's First Proposed 2021 Legislative Redistricting plan (LEG21-39001),

   (a) Sen. Linehan's first proposed 2021 legislative redistricting plan LEG21-39001 was not adopted, the Legislature amended it into the final plan. This first proposed redistricting plan LEG21-39001 shows an *even more pronounced pattern* of over-populating urban Legislative Districts, within the allowable 5% population deviation.

   The analysis of the adopted 2021 map uses the shapefile provided by the Legislative Research Office, originally downloaded from the Redistricting Committee's website. The shapefile is no

longer available on the Redistricting Committee website. However, the shapefile accompanies the source material for this report. Using this information creates Figure 4. An interactive version of the map allowing zooming, panning and labeled with population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`. in the tab labeled Linehan. This map shows the even more pronounced pattern of consistently over-populating most of the urban Legislative Districts with more residents. Most of the orange colored districts intersect the major urban areas of Omaha and Lincoln. The purple colored districts are in the rural parts of the state.

The more numerical evaluation of the pattern follows the previous analysis. A scatter plot of the per-cent deviation (vertical scale) versus the (logarithm of) the density (horizontal scale) is in Figure 8. Note that of 20 districts classified as urban based 15 have positive deviation indicating greater than ideal population. On the other hand of 29 districts classified as rural, 19 have negative deviation and population less than ideal, while only 10 rural districts have greater than ideal population.

The Pearson correlation coefficient for deviation-versus-density for the adopted 2021 districts is $r = 0.550$ indicating an *even stronger* relationship between density and deviation than before. In terms of the transformed variable of percent deviation versus logarithm of population density the correlation coefficient is 0.626 compared to the previous value of 0.436.

The second measure of relationship is the slope of the best-fit line through the scatter plot. The slope for the best-fit line through Linehan's proposed 2021 raw deviation versus density scatter plot is 8.86×10^{-6} which is greater than the slope of the adopted plan, indicating an even stronger trend for denser districts to be over-populated. In terms of the transformed variable of percent deviation versus logarithm of population density the best-fit line slope is 1.651, greater than the previous value of 1.06. This comparison can be seen easily in Figure 7.

(b) We also did the advanced statistical analysis on Sen. Linehan's first proposed map, again using the open source program `GerryChainJulia`. Once again the program makes a sequence of random changes to the starting redistricting plan LEG21-39001, creating a new valid set of districts at each step. The program evaluates characteristics

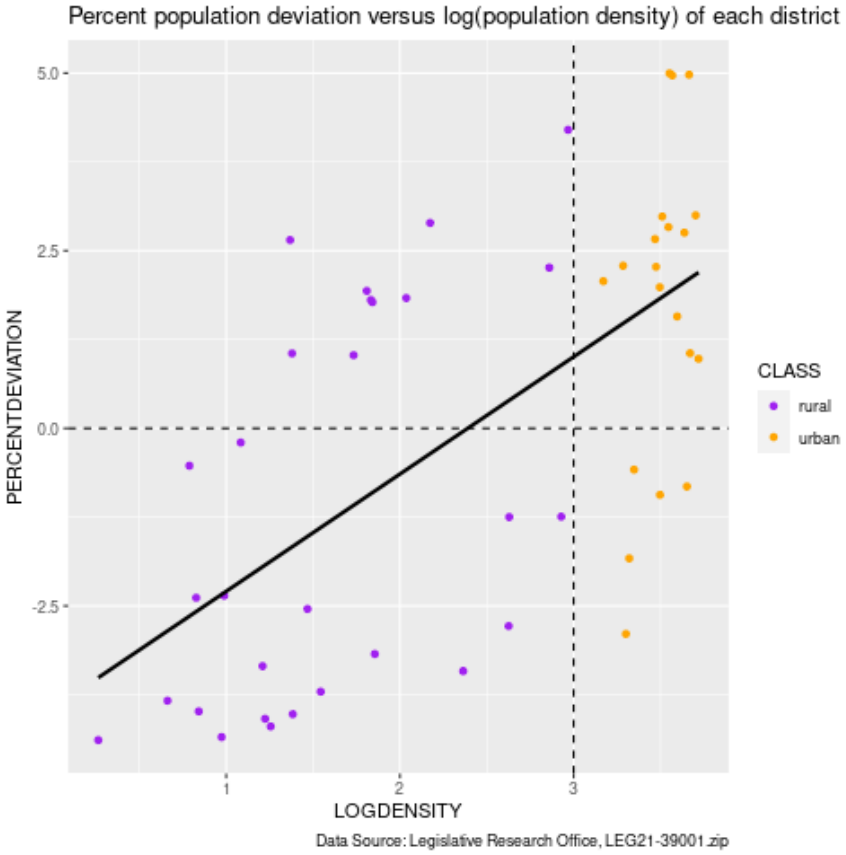

Figure 8: Scatter plot of deviation versus density for all districts for the proposed 2021 LEG21-39001 map.

of each district and of the entire redistricting plan at each step. As before, the characteristics we evaluated were the population density of each district and the population deviation from ideal of each district. The correlation of the population deviation and the population density can then be calculated for each created alternative redistricting. The result is a sequence of correlation values for alternative redistrictings that to compare to the correlation value for the adopted 2021 LEG21-39006.

We ran the `GerryChainJulia` program for two parallel chains with 2 parallel chains of 15,000 steps for a combined total of 30,000 alternative redistricting to evaluate. Recall that the first proposed 2021 LEG21-39001 had a positive correlation value of 0.550. Ranking this positive correlation for the first proposed 2021 LEG21-39001 against 30,000 alternative maps, the correlation of the adopted map was at the 99.64 percentile. The mathematical theorem predicts that among all similar runs percentile this large has only about 0.7% probability of occurring solely by chance. This demonstrates the Sen. Linehan's first proposed map had an even greater association of higher density districts with overpopulation which is unreasonable to attribute to chance.

2. We used Sen. Wayne's proposed map LEG21-13001 to compare with Sen. Linehan's first proposed 2021 LEG21-3001 and final legislative redistricting plans LEG21-39006. The analysis of the map uses the shapefile provided by the Legislative Research Office. The shapefile is no longer available on the Redistricting Committee website. Using this information creates Figure 5. This map shows no apparent pattern of orange or purple districts appearing in either rural or urban areas. The distribution appears to be random. An interactive version of the map allowing zooming, panning and labeled with population deviation values is at `https://kristie3.github.io/#rural-and-urban-comparison-legislative-maps`. in the tab labeled Wayne.

   Statistical evaluation of the pattern follows the previous analyses. A scatter plot of the percent deviation (vertical scale) versus the logarithm of the density (horizontal scale) is in Figure **??**. Of 22 districts classified as urban, 12 have positive deviation and 10 have negative deviation. Of 27 districts classified as rural, 14 have positive deviation and 13 have negative deviation. The Pearson correlation coefficient for the raw deviation-versus-density for the final 2021 districts is

$r = 0.112$ indicating a weak relationship between density and deviation. The Pearson correlation coefficient for percent deviation versus logarithm of density for the Wayne district proposal is $r = 0.067$ indicating virtually no relationship between density and deviation. The second measure of relationship is the slope of the best-fit line through the scatter plot. The slope for the best-fit line through the Wayne's proposed 2021 percent deviation versus logarithm of density scatter plot is 0.3058 indicating essentially only a weak trend between density and deviation. For deviation versus density, the best-fit slope is 0 to significant digits.

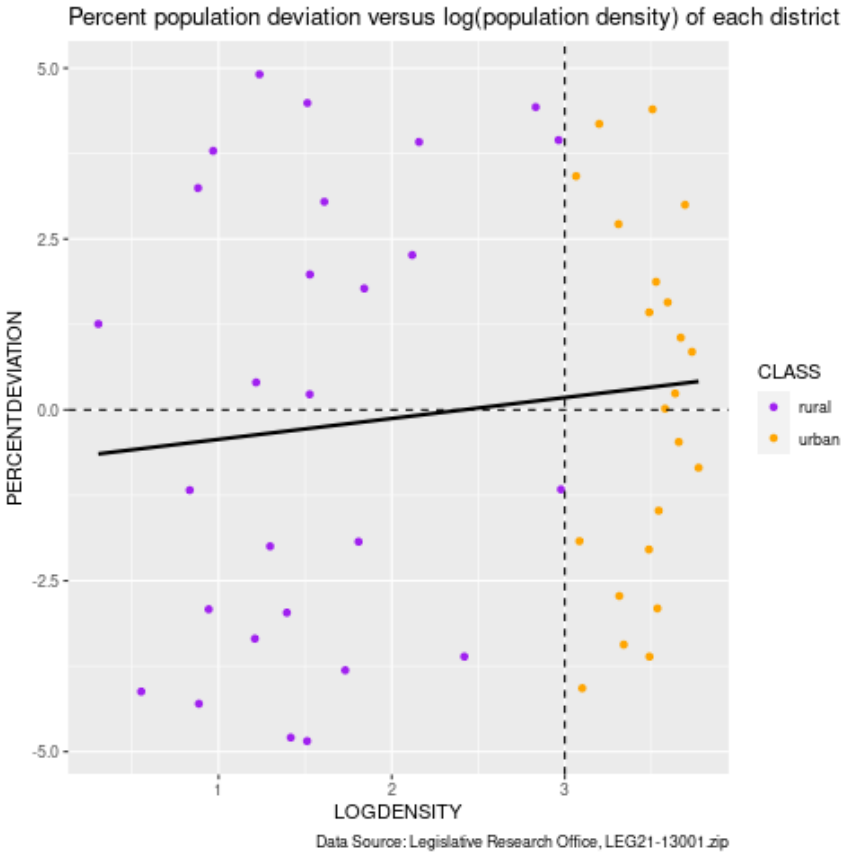

Figure 9: Scatter plot of deviation versus density for all districts for the proposed 2021 LEG21-13001 maps.

(a) We also did the advanced statistical analysis on Sen. Wayne's

proposed redistricting. Again we ran the `GerryChainJulia` program for two parallel chains with 2 parallel chains of 15,000 steps for a combined total of 30,000 alternative redistricting to evaluate. Ranking the correlation value for Sen. Wayne's proposed map against 30,000 alternative maps, the correlation of the proposed map was at the 85th percentile. The mathematical theorem predicts that among all similar runs, a percentile this large would occur about 31% of the time. These ranking indicate that the correlation in Sen. Wayne's map is *neither extreme nor unusual* with respect to these measures.

3. Advanced statistical analysis of the 2011 Legislative Districts.

   The correlation of deviation with density for the 2011 redistricting plan was 0.0454. Using the scaled values of percent population deviation versus logarithm of population density, the correlation was 0.028. This is essentially no correlation, in contrast to the correlations in the 2021 LEG21-39006 plan.

   We did the advanced statistical analysis with `GerryChainJulia` as we did for the final 2021 plan but with a slight difference. The 2010 Census data shapefile has 193,352 Census Blocks compared to the 2020 Census data files with 119,103 Census Blocks. (It is not known to us why the number of 2010 Census Blocks is 62% larger than the number of 2020 Census Blocks. However, many of the 2010 Census Blocks had zero population so perhaps the Census Bureau consolidated blocks so that most have non-zero population.) The larger number of Census Blocks prevented running two parallel chains in the memory on available computers, so we ran the single chain test [8] with 10,000 steps using 2010 population data. This gave a comparison of the correlation of percent deviation versus logarithm of population density with 10,000 alternative maps. The result is that the correlation of percent deviation with logarithm of population density for the 2011 redistricting was at the 50.4 percentile with a chance of about 99.56% of occurring at that percentile or greater among comparable runs of the chain. The conclusion is that the adopted 2011 plan was *neither extreme nor unusual* in packing population into urban districts.

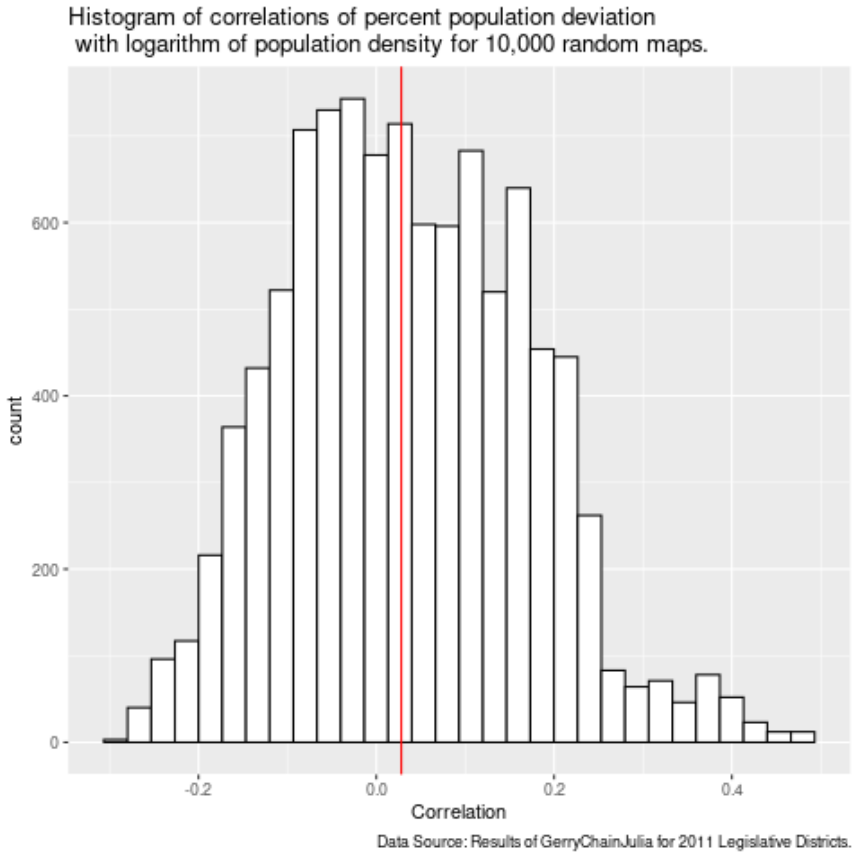

Figure 10: A histogram of correlations of percent population deviation versus logarithm of density for 10,000 alternative maps, compared to the correlation for the adopted 2011 map, indicated with a red line.

6 Acknowledgments

This analysis was undertaken on our own initiative, motivated by our interest in the application of mathematics, data science, statistics, and information technology to public policy and good government. This research and analysis was not sponsored or funded by any organization or individual. All of the research was carried out on personally owned computers using open-source software.

We wish to thank the Metric Geometry and Gerrymandering Group for sponsoring a series of conferences on mathematical analysis of gerrymandering that we individually attended at our own expense. We also wish to thank the Metric Geometry and Gerrymandering Group for creating the `GerryChainJulia` software.

Steven R. Dunbar, Ph.D., is Professor Emeritus of Mathematics at the University of Nebraska-Lincoln. Kristin Pfabe, Ph.D, is Professor of Mathematics at Nebraska Wesleyan University.

7 References

References

[1] Nebraska Legislature. Nebraska consitution. https://nebraskalegislature.gov/laws/articles.php?article=III-5, May 1999.

[2] Nebraska Legislature. Legislative resolution 134. https://nebraskalegislature.gov/FloorDocs/107/PDF/Enroll/LR134.pdf, May 2021.

[3] Bureau of the Census. Geographic areas reference manual, November 1994.

[4] Supreme Court of the United States. Rucho v. common cause. `https://www.supremecourt.gov/opinions/18pdf/18-422_9ol1.pdf`, May 2018.

[5] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of markov chains for redistricting. *Harvard Data Science Review*, 3 2021. https://hdsr.mitpress.mit.edu/pub/1ds8ptxu.

[6] Parker Rule, Matthew Sun, and Bhushan Suwal. mggg/gerrychainjulia: Minor fixes + save as hdf5, March 2021.

[7] Maria Chikina, Alan Frieze, Jonathan C. Mattingly, and Wesley Pegden. Separating effect from significance in markov chain tests. *Statistics and Public Policy*, 0(0):1–14, 2020.

[8] Maria Chikina, Alan Frieze, and Wesley Pegden. Assessing significance in a markov chain without mixing. *Proceedings of the National Academy of Sciences*, 114(11):2860–2864, 2017.