IN THE SUPREME COURT OF OHIO

LEAGUE OF WOMEN VOTERS OF OHIO, *et al.*, **Relators,** **v.** **OHIO REDISTRICTING COMMISSION, *et al.*,** **Respondents.**	**Case No. 2021-1449** **Original Action Filed Pursuant to Ohio Const., Art. XIX, Sec. 1(C)(3)**

AFFIDAVIT OF KOSUKE IMAI

Franklin County
/ss
State of Ohio

Now comes affiant Kosuke Imai, having been first duly cautioned and sworn, deposes and states as follows:

1. I am over the age of 18 and fully competent to make this declaration. I have personal knowledge of the statements and facts contained herein.
2. For the purposes of this litigation, I have been asked by counsel for Relators to analyze relevant data and provide my expert opinions.
3. To that end, I have personally prepared the report attached to this affidavit as Exhibit A, and swear to its authenticity and to the faithfulness of the opinions expressed, and, to the best of my knowledge, the accuracy of the factual statements made therein.

FURTHER AFFIANT SAYETH NAUGHT

Executed on 12/09/2021, 2021.

Kosuke Imai
Signed on 2021/12/09 08:01:53 -8:00

Kosuke Imai

Sworn and subscribed before me this 12/09/2021 day of ________, 2021

NOTARY PUBLIC
STATE OF OHIO
Theresa M Sabo
Commission # 2016-RE-619622
Electronic Notary Public
State of Ohio
My Comm Exp. Nov 28, 2026
Notary Stamp 2021/12/09 09:01:53 PST C9A5EA4809AA

Signed on 2021/12/09 08:01:53 -8:00

Notary Public

Notarial act performed by audio-visual communication

Imai Affidavit.pdf

DocVerify ID:	D8970D65-ECBB-4BFE-B639-1BE659039744
Created:	December 09, 2021 07:57:36 -8:00
Pages:	1
Remote Notary:	Yes / State: OH

E-Signature Summary

E-Signature 1: Kosuke Imai (KI)
December 09, 2021 08:01:53 -8:00 [9DB37A030428] [108.26.227.252]
imai@harvard.edu (Principal) (Personally Known)

E-Signature Notary: Theresa M Sabo (TMS)
December 09, 2021 08:01:53 -8:00 [C9A5EA4809AA] [74.142.214.254]
tess.sabo@gmail.com
I, Theresa M Sabo, did witness the participants named above electronically sign this document.

EXHIBIT A

IN THE SUPREME COURT OF OHIO

League of Women Voters of Ohio, *et al.* *Relators*, v. Ohio Redistricting Commission, *et al.* *Respondents*.	Original Action Filed Pursuant to Ohio Const., Art. XIX, Sec. 3(A)

EXPERT REPORT

Kosuke Imai, Ph.D.

December 9, 2021

Table of Contents

EXPERT REPORT

I. INTRODUCTION AND SCOPE OF WORK

1. My name is Kosuke Imai, Ph.D., and I am a Professor in the Department of Government and the Department of Statistics at Harvard University. I specialize in the development of statistical methods for and their applications to social science research. I am also affiliated with Harvard's Institute for Quantitative Social Science.

2. I have been asked by counsel representing the relators in this case to analyze relevant data and provide my expert opinions related to whether Ohio's enacted congressional districting plan (SB 258, which I will refer to as the "enacted plan" in this report) meets the criteria in Article XIX, Section 1(C)(3)(a) of Ohio's Constitution. More specifically, I have been asked to statistically analyze the enacted plan's compliance with Article XIX, Section 1(C)(3)(a)'s requirement that "[t]he general assembly shall not pass a plan that unduly favors or disfavors a political party or its incumbents" by comparing it against other alternative plans that are as or more compliant with other relevant requirements of Article XIX.

II. SUMMARY OF OPINIONS

3. I simulated 5,000 hypothetical plans that are at least as compliant with Article XIX as the enacted plan. The comparison of these simulated plans with the enacted plan yields the following findings:

- The enacted plan unduly favors the Republican Party by giving the Republicans a much greater expected number of seats than in any of my 5,000 simulated plans. Even using the General Assembly's assumptions regarding the appropriate election set and calculation of expected number of seats, the Republican candidates are expected to win 2.8 more seats under the enacted plan than under the average simulated plan.

- The expected number of Republican seats under the enacted plan is a clear statistical outlier. Indeed, any plan that provides for more than 9 expected Republican seats is an outlier. Moreover, the probability of generating the enacted plan's extreme partisan outcome under the non-partisan simulation procedure I used is essentially zero.

- The enacted plan exhibits a significant partisan bias in favor of the Republican Party. Even using the General Assembly's assumptions regarding the appropriate election set and calculation of expected number of seats, the magnitude of bias is much greater under the enacted plan than in any of my 5,000 simulated plans and is a clear statistical outlier, according to several standard metrics used in the academic literature.
- In Hamilton County, the enacted plan cracks Democratic voters to create safe Republican seats, while in Franklin and Cuyahoga counties the enacted plan packs Democratic voters to create additional Republican-leaning districts.

III. QUALIFICATIONS, EXPERIENCE, AND COMPENSATION

4. I am trained as a political scientist (Ph.D. in 2003, Harvard) and a statistician (MA in 2002, Harvard). I have published more than 60 articles in peer reviewed journals, including premier political science journals (e.g., *American Journal of Political Science*, *American Political Science Review*, *Political Science*), statistics journals (e.g., *Biometrika*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society*), and general science journals (e.g., *Lancet*, *Nature Human Behavior*, *Science Advances*). My work has been widely cited across a diverse set of disciplines. For each of the past four years, Clarivate Analytics, which tracks citation counts in academic journals, has named me as a highly cited researcher in the cross-field category for producing "multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science."

5. I started my academic career at Princeton University, where I played a leading role in building interdisciplinary data science communities and programs on campus. I was the founding director of Princeton's Program in Statistics and Machine Learning from 2013 to 2017. In 2018, I moved to Harvard, where I am Professor jointly appointed in the Department of Government and the Department of Statistics, the first such appointment in the history of the university. Outside of universities, between 2017 and 2019, I served as the president of the Society for Political Methodology, a primary academic organization of more than one thousand researchers worldwide who conduct methodological research in political science. My introductory statistics textbook for

social scientists, *Quantitative Social Science: An Introduction* (Princeton University Press, 2017), has been widely adopted at major research universities in the United States and beyond.

6. Computational social science is one of my major research areas. As part of this research agenda, I have developed simulation algorithms for evaluating legislative redistricting since the beginning of this emerging literature. At Harvard, I lead the Algorithm-Assisted Redistricting Methodology (ALARM; https://alarm-redist.github.io/) Project, which studies how algorithms can be used to improve legislative redistricting practice and evaluation.

7. Back in 2014, along with Jonathan Mattingly's team at Duke, my collaborators and I were the first to use Monte Carlo algorithms to generate an ensemble of redistricting plans. Since then, my team has written several methodological articles on redistricting simulation algorithms (Fifield, Higgins, et al. 2020; Fifield, Imai, et al. 2020; McCartan and Imai 2020; Kenny et al. 2021).

8. I have also developed an open-source software package titled `redist` that allows researchers and policy makers to implement the cutting-edge simulation methods developed by us and others (Kenny et al. 2020). This software package can be installed for free on any personal computer with Windows, Mac, or Linux operating system. According to a website that tracks the download statistics of R packages, our software package has been downloaded about 30,000 times since 2016 with an increasing download rate.[1]

9. In addition to redistricting simulation methods, I have also developed the methodology for ecological inference referenced in voting rights cases (Imai, Lu, and Strauss 2008; Imai and Khanna 2016). For example, my methodology for predicting individual's race using voter files and census data was extensively used in a recent decision by the Second Circuit Court of Appeals regarding a redistricting case (Docket No. 20-1668; Clerveaux *et al* v. East Ramapo Central School District).

10. A copy of my curriculum vitae is attached as Exhibit A.

11. I am being compensated at a rate of $450 per hour. My compensation does not

1. https://ipub.com/dev-corner/apps/r-package-downloads/ (accessed on December 6, 2021)

depend in any way on the outcome of the case or on the opinions and testimony that I provide.

IV. METHODOLOGY

12. I conducted simulation analyses to evaluate the enacted plan's compliance with Section 1(C)(3)(a) of Article XIX. Redistricting simulation algorithms generate a representative sample of all possible plans under a specified set of criteria. This allows one to evaluate the properties of a proposed plan by comparing them against those of the simulated plans. If the proposed plan unusually favors one party over another *when compared to* the ensemble of simulated plans, this serves as empirical evidence that the proposed plan is a partisan gerrymander. Furthermore, statistical theory allows us to quantify the degree to which the proposed plan is extreme relative to the ensemble of simulated plans in terms of partisan outcomes.

13. A primary advantage of the simulation-based approach, over the traditional methods, is its ability to account for the political and geographic features that are specific to each state, including spatial distribution of voters and configuration of administrative boundaries. Simulation methods can also incorporate each state's redistricting rules. These state-specific features limit the types of redistricting plans that can be drawn, making comparison across states difficult. The simulation-based approach therefore allows us to compare the enacted plan to a representative set of alternate districting plans subject to Ohio's administrative boundaries, political realities, and constitutional requirements. Appendix A provides a brief introduction to redistricting simulation.

A. Simulation Analysis

14. I have ensured that all my simulated plans are equally or more compliant with Section 2(B) of Article XIX than the enacted plan. My simulation procedure achieves this, in part, by being compliant with the U.S. Constitution and federal law protecting racial minority voting rights, generating contiguous and compact districts, limiting the number of county splits, and respecting the other splitting criteria specified in Section 2(B). I also avoid splitting the counties the enacted plan does not split. Appendix B provides detailed information about this process. For all simulations, I ensure districts fall within a 0.5% deviation from population parity. Although this deviation is greater than the population deviation used in the enacted plan, it only accounts for less

than 4,000 people and hence has no impact on the conclusions of my analysis.

15. Here, I provide a brief overview of the procedure while leaving the details to Appendix B. My simulation proceeds in two steps. First, at the instruction of counsel for the relators, I ensured that every simulated plan has one district in Cuyahoga County with the proportion of black voting age population (BVAP) falling above 42% in order to be compliant with the U.S. Constitution and federal law protecting racial minority voting rights. To do this, I sampled a contiguous and compact district that has an appropriate population size and BVAP proportion within Cuyahoga County. This district always contains the entire city of Cleveland because Section 2(B)(4)(b) prohibits splitting it. Once such a district is generated, I then separately run the simulation algorithm on the rest of the state and generate the remaining 14 districts while making sure that the resulting districts satisfy the requirements specified in Section 2(B). I repeat this procedure 5,000 times to obtain the desired number of simulated plans.

B. Metrics Used to Measure Bias

16. Using the redistricting simulation methodology, I evaluate compliance with Section 1(C)(3)(a) of Article XIX in the set of simulated plans generated by the algorithm as well as the enacted plan. To determine whether the enacted plan unduly favors a particular political party, I compare the expected number of Republican and Democratic seats under the enacted plan against the corresponding number under the simulated plans.

17. I understand that the General Assembly assessed the partisan leanings of the enacted plan using the set of six statewide federal elections from 2012 to 2020 (see Appendix E.1 for the list of these elections). I do not endorse the assumption that using this limited data set can accurately predict the expected number of Republican and Democratic seats under the enacted plan.[2] I nonetheless use this same set of election results data in my analysis so that the differences in conclusions between my analysis and the General Assembly's assessment cannot be attributed to the way in which the partisan leanings of districts are evaluated. Given that these elections

2. I have reviewed the Affidavit of Dr. Christopher Warshaw dated November 30, 2021, which concludes that this set of elections artificially enhances the perception of Democratic Party strength under the enacted plan. I agree with his conclusion in this regard.

enhance the perception of Democratic relative strength, using this assumption effectively gives the enacted plan the benefit of the doubt.

18. I also adopt the General Assembly's approach to computing the expected number of Republican seats under a given redistricting plan. Specifically, I first compute the total number of Republican votes for each district and then sum it across the six statewide federal elections. Dividing this by the total number of two-party votes that are similarly aggregated across these elections yields the Republican two-party vote share for each district. This aggregation method may not be ideal because it gives greater weights to general elections, which tend to have higher turnout than midterm elections. In spite of this potential problem, I follow the General Assembly's approach so that the findings of my analysis can be directly compared to the General Assembly's assessment. I have confirmed that the resulting vote share for each district under the enacted plan is essentially identical to the corresponding district-level vote share presented in the November 16, 2021 statement from Senator Rob McColley. Finally, based on these vote shares, I determine likely winners of all districts based on the vote totals for each statewide election. This gives the total number of expected Republican and Democratic seats for a given plan under the General Assembly's approach.

19. In addition to the expected number of seats, I apply a variety of metrics that are commonly used in the academic literature. These metrics are extensively discussed in Dr. Christopher Warshaw's affidavit, dated November 30, 2021, and the references therein. I have reviewed Dr. Warshaw's articulation of these metrics and they are consistent with my understanding, and appear to be applicable to the facts of this case. Specifically, to measure compliance with Section 1(C)(3)(a), I use the following partisan bias metrics whose definitions are discussed in Dr. Warshaw's affidavit and the references therein.

- Efficiency gap
- Mean-median gap
- Symmetry in the vote-seat curve across parties
- Declination

C. The Determination of Whether the Enacted Plan is a Statistical Outlier Can Provide a Useful Measure of its Partisan Bias

20. Another important benefit of using the redistricting simulation methodology is that it can determine whether or not the enacted plan is a statistical outlier relative to the simulated plans generated under a specified set of criteria. If the enacted plan is a statistical outlier, then the observed difference in partisan outcome between the enacted plan and the simulated plans represents a systematic partisan bias.

21. To determine whether the enacted plan is a statistical outlier, I first estimate the probability of generating a simulated plan that favors a political party at least as much as the enacted plan does. This can be done by simply computing the proportion of the simulated plans that favors a political party equally or more than the enacted plan. If this estimated probability is very small (e.g., less than 0.001), then the enacted plan is a statistical outlier because it is highly unlikely to come from the non-partisan distribution that is used to generate the simulated plans. If the data based on the simulated plans follow the normal distribution, which is a bell-shaped symmetric distribution without skew, then this probability of 0.001, for example, implies that the enacted plan is more than three standard deviations away from the average simulated plans.[3]

22. I also compute the difference in partisan outcome between the enacted plan and the average simulated plan. This allows me to measure the magnitude of partisan bias while accounting for its random variability across the simulated plans. I apply the most commonly used definition of an outlier (Tukey 1977). According to this definition, an outlier represents a data point that is beyond a distance of 1.5 interquartile range (IQR) below the first quartile or above the third quartile. If the data based on the simulated plans were normally distributed, the enacted plan is regarded as an outlier if it is at least 2.70 standard deviations away from the average simulated plan.

D. Description of Redistricting Simulation Software

3. Note that a standard deviation represents the average distance between a data point and the mean.

23. In my analysis, I use the open-source software package for redistricting analysis `redist` (Kenny et al. 2020), which implements a variety of redistricting simulation algorithms as well as other evaluation methods. My collaborators and I have written the code for this software package, so that other researchers and the general public can implement these state-of-the-art methods on their own. I supplement this package with code written primarily to account for the redistricting rules and criteria that are specific to Ohio. All of my analyses are conducted on a laptop. Indeed, all of my analysis code can be run on any personal computer once the required software packages, which are also freely available and open-source, are installed.

V. EVALUATION OF THE ENACTED PLAN USING THE GENERAL ASSEMBLY'S APPROACH

24. Using the redistricting simulation methodology, I evaluate the enacted plan's compliance with Section 1(C)(3)(a). Appendix E.1 provides the detailed information about data sources. I simulated 5,000 alternative Congressional redistricting plans, using the simulation procedure described in Section IV. As explained in Appendix B, every simulated plan is at least as compliant with Sections 2(B) as the enacted plan. For example, Appendices C and D show that the simulated plans are more compact and have fewer county splits than the enacted plan.

25. I can easily generate additional compliant plans by running the algorithm longer, but for the purpose of my analysis, 5,000 simulated plans will yield statistically precise conclusions. In other words, generating more than 5,000 plans, while possible, will not materially affect the conclusions of my analysis.

26. To evaluate the enacted plan's compliance with Section 1(C)(3)(a), I first compare the expected number of Republican seats under the enacted plan with that under each of my 5,000 simulated plans. Figure 1 shows that under the enacted plan, the Republican Party is expected to win 11 seats.[4] In contrast, under about 80% of the simulated plans, the expected number of Republican seats is only 8, while the Republican Party is expected to win 9 seats under the remaining

4. This prediction of 11 expected seats is based on using the set of six statewide federal elections from 2012 to 2020 that the General Assembly used. Again, I do not endorse the assumption that using this limited data set can accurately predict the expected number of Republican seats.

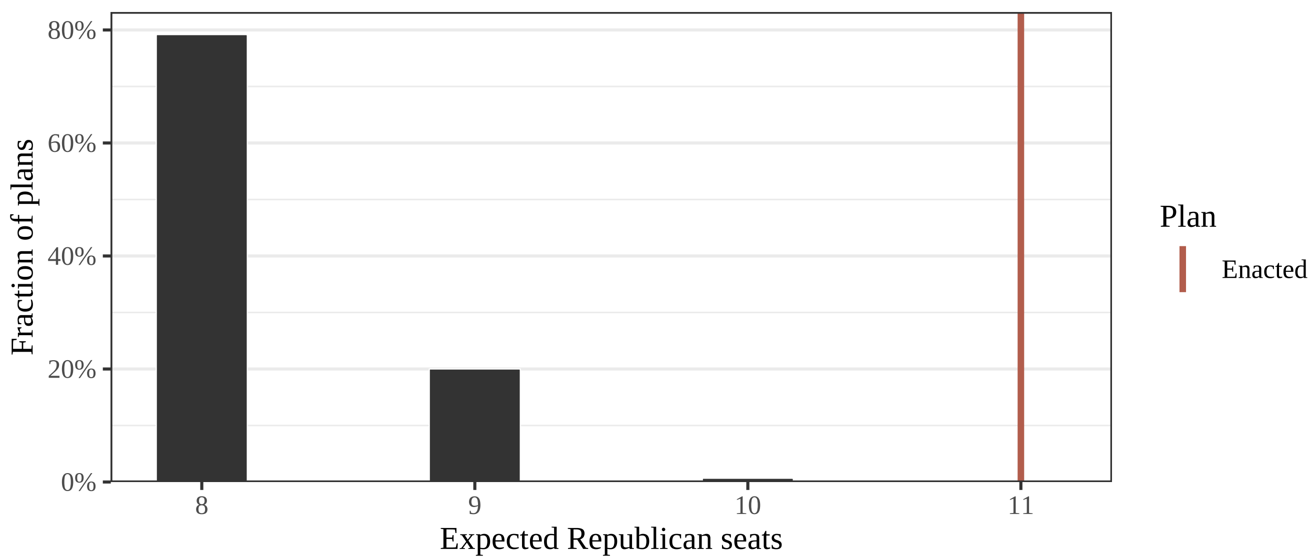

Figure 1: Expected number of Republican seats calculated for the 5,000 simulated plans computed by averaging across the six statewide federal elections from 2012 to 2020. Overlaid is the value for the enacted plan (red).

20% of the simulated plans. In other words, the enacted plan is expected to yield an additional 2.8 Republican seats when compared to the average simulated plan. Indeed, none of my 5,000 simulated plans gives as many Republican seats as the enacted plan. This result implies that the probability of generating the enacted plan's extreme partisan outcome under the non-partisan simulation procedure I used is essentially zero. Thus, any redistricting plan that gives more than 9 seats to the Republican Party, including the enacted plan, is a clear statistical outlier.

27. Under most of the simulated plans, the Republican Party is expected to win 8 seats, which is equivalent to 53% of the Ohio's 15 Congressional seats. This seat proportion is almost identical to the statewide vote share of the Republican Party, which is approximately 52% calculated using the General Assembly's approach and 54% based on the statement made by the Ohio Redistricting Commission in compliance with Section 8(C)(2) of Article XI of the Ohio Constitution. In contrast, under the enacted plan, the expected seat share of the Republican Party is 73%, which is roughly 20 percentage points greater than its expected vote share. As discussed above, this seat share result is a clear statistical outlier. Accordingly, this shows that the enacted plan unduly favors the Republican Party.

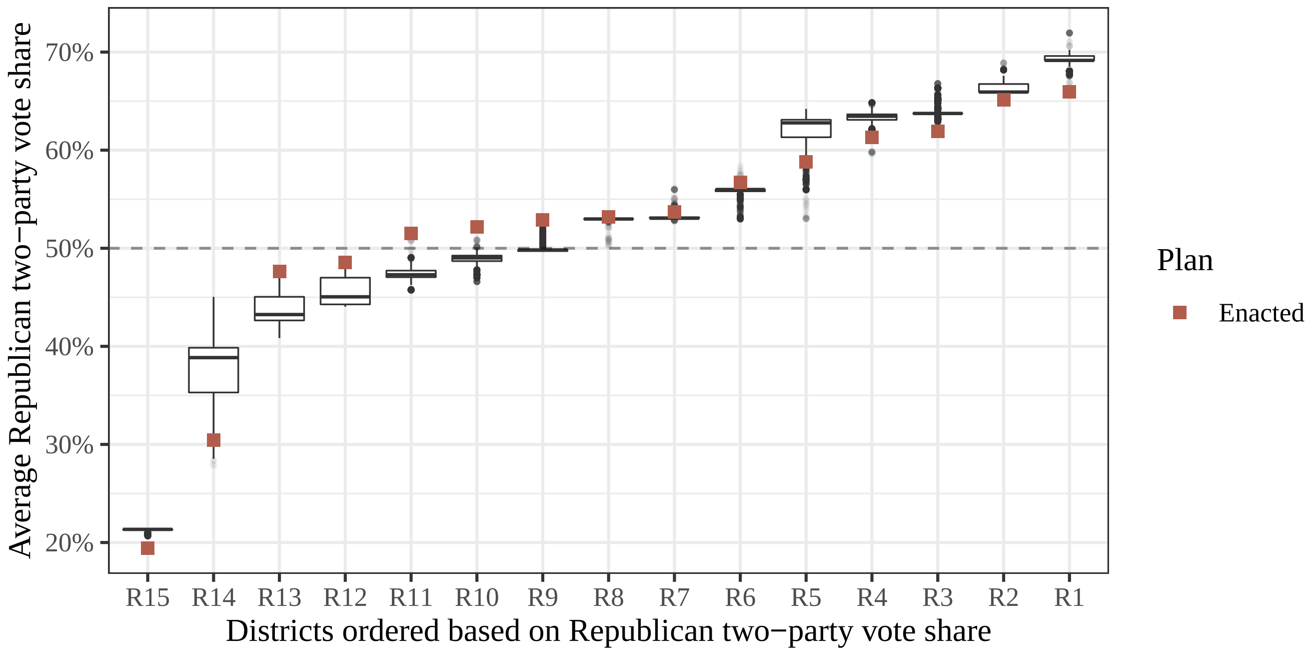

Figure 2: Expected Republican vote share for districts using the six statewide federal elections from 2012 to 2020. For any given plan, the districts are ordered based on their expected Republican vote share. Boxplots represent the distribution of the expected Republican vote share across the simulated plans, whereas the red square correponds to the expected Republican vote share under the enacted plan.

28. Figure 2 further demonstrates the partisan bias of the enacted plan. In this plot, for any given plan (both enacted and simulated), I ordered the districts based on the magnitude of their expected Republican vote share. This means that under any given plan, district R1 yields the highest expected vote share while district R15 is expected to give the least support to the Republican candidate (to be clear, the R1 through R15 district identifiers do not correspond to the Congressional district numbers in the enacted plan). If the expected Republican vote share of each ordered district under the enacted plan (red square) diverges from the corresponding distribution of the simulated plans (boxplot), it constitutes evidence of possible partisan bias. Note that in a boxplot, the "box" contains 50% of the data points (those from 25 percentile to 75 percentile to be exact) with the horizontal line indicating the median value whereas the vertical lines coming out of the box, called "whiskers", indicate the range, which contains most data. Any data points that are beyond these whiskers are considered as outliers according to the second part of the definition

discussed in Section IV.C (paragraph 23).

29. The figure shows clear evidence of the enacted plan's partisan bias. This partisan bias, for the reasons discussed below, further shows that the enacted plan unduly favors the Republican Party. For all of my 5,000 simulated plans, districts R10 and R11 (the 10th and 11th most Republican-leaning districts, respectively) lean toward the Democratic party with the expected median Republican vote share equal to 49.0% and 47.3%, respectively. Indeed, for district R11, none of 5,000 simulated plans are expected to yield as many Republican votes as the enacted plan. Yet under the enacted plan, both of these districts have the expected Republican vote shares above 50%. According to the definition discussed in Section IV.C, these two points associated with the enacted plan are clear statistical outliers, with district R10 and R11 5.2 and 5.8 standard deviations away from the median, respectively.

30. I also find that under the enacted plan, districts R12 and R13 lean much less strongly towards the Democratic party than under all of the simulated plans. Lastly, the enacted plan packs Democratic voters in districts R14 and R15, which are two most Democratic-leaning districts. This is indicated by the fact that these districts have much lower levels of expected Republican vote shares under the enacted plan than under the simulated plans. In contrast, the enacted plan avoids packing Republican voters in the five most Republican districts (districts R1 to R5). Indeed, these districts have much lower levels of expected Republican vote shares under the enacted plan than under the simulated plans. Aside from districts R2 and R5, these points are also statistical outliers. Districts R1 to R5 are 6.8, 1.4, 2.4, 3.7 and 2.0 standard deviations away from the median, respectively.

31. I next use the four partisan bias metrics discussed in Section IV.B to examine the enacted plan's compliance with Section 1(C)(3)(a). I adjusted the sign of each metric so that positive values indicate Republican bias, and values nearer to zero indicate less partisan bias. To summarize the results, as shown in Figure 3, when compared to these simulated plans (black histogram), the enacted plan (red vertical line) is a clear outlier favoring the Republican Party. Indeed, the enacted map is more biased than any of 5,000 simulated plans for all four partisan bias

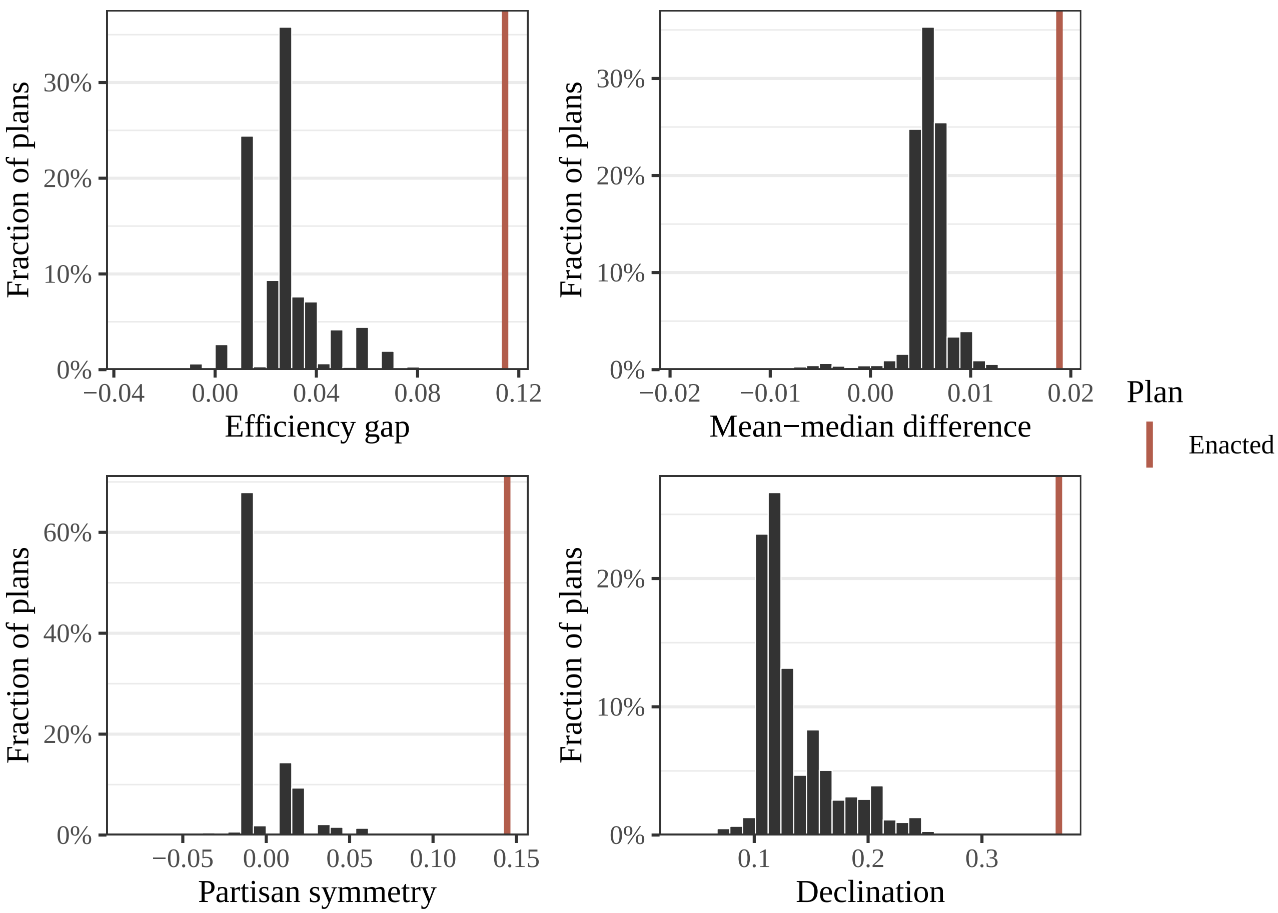

Figure 3: Four partisan bias measures calculated for the 5,000 simulated Congressional redistricting plans computed by averaging across the six federal elections from 2012 to 2020. Overlaid is the value for the enacted plan (red). For each measure, larger values (towards the right) correspond to more Republican-favoring plans.

metrics I considered.

32. The efficiency gap, which captures both cracking and packing, is 15.0% for the enacted map, whereas the average efficiency gap for the simulated plans is only 5.7%. This implies that the enacted plan wastes around 219,000 more Democratic votes on average than the simulated plans, and around 219,000 fewer Republican votes. As shown in the top-left plot of Figure 3, the enacted map is 7.5 standard deviations away from the average simulated plan, and is thus a clear statistical outlier in terms of the efficiency gap metric.

33. The mean-median gap is a measure of asymmetry in the distribution of votes across districts. The existence of packed districts may lead to a large mean-median gap. The top-right plot of the figure shows that the mean-median gap is 0.018 under the enacted plan while the simulated plans score 0.007 on average. Indeed, the enacted plan is 5.7 standard deviations away from the average simulated plan, and is thus a clear statistical outlier in terms of the mean-median gap metric.

34. Partisan symmetry is based on the idea that each party should receive half of the seats if they each receive 50% of votes. The bottom-left plot of Figure 3 shows that the enacted plan scores 14.1% on this metric while the simulated plans score 1.8%, on average. This suggests that under the enacted plan, the Republican Party would gain roughly 2.1 more seats than the Democrats, for a hypothetical tied election. In contrast, the simulated plans would give only 0.3 more seats to the Republican Party than the Democrats in the same situation. The enacted plan is 7.4 standard deviations away from the average simulated plan, and is thus a clear statistical outlier in terms of the partisan symmetry metric.

35. Lastly, the declination metric represents another measure of asymmetry in the vote distribution. As shown in the bottom-right plot of the figure, the enacted plan also scores worse on this metric than any of the 5,000 simulated plans. Specifically, the enacted plan scores 0.42 whereas the simulated plans earn 0.21 on average. The enacted plan is 9.3 standard deviations away from the average simulated plan, and is thus a clear statistical outlier in terms of the declination metric.

36. Thus, all of the partisan bias metrics show that the enacted plan is a clear statistical

outlier, favoring the Republican Party, when compared to the simulated plans. Indeed, the enacted plan has a worse partisan bias than any of my 5,000 simulated plans.

VI. LOCAL ANALYSIS OF SELECTED COUNTIES

37. Partisan bias in the enacted plan is apparent not just in statewide summary statistics, as shown above, but also at the local level. To illustrate this, I performed a detailed analysis of the Congressional districts in Hamilton, Franklin, and Cuyahoga counties. My analysis of these cities shows that the enacted plan packs a disproportionately large number of Democratic voters into some districts while cracking Democratic voters in other districts to create Republican-leaning seats.

38. My analysis of each county proceeds as follows. For each precinct, I first compute the expected two-party vote share of the district to which the precinct is assigned under the enacted plan. I then perform the same calculation under each simulated plan and average these expected vote shares across all of the simulated plans. Comparison of these two numbers reveals whether the enacted plan assigns a precinct to a district whose political leaning is different from what would be expected under the simulated plans. As in Section V, the results shown below are based on the General Assembly's approach that uses the statewide federal elections from 2012-2020.

A. Hamilton County

39. I begin by illustrating the above calculation through an example. Precinct 061031BEZ of Cincinnati lies within District 1 of the enacted map, which has an expected Republican two-party vote share of 51.53%. However, the same precinct belongs to different districts in most of the simulated maps, each with their own Republican vote share. The average Republican vote share for the districts to which this precinct is assigned across all of the simulated plans is 44.85%, which is 6.68 percentage points lower than under the enacted plan. So, based on the representative set of simulated plans that have less partisan bias, precinct 061031BEZ is assigned to a more Republican-leaning district under the enacted plan than under the average simulation plan.

40. The left map of Figure 4 presents the expected vote shares of districts under the

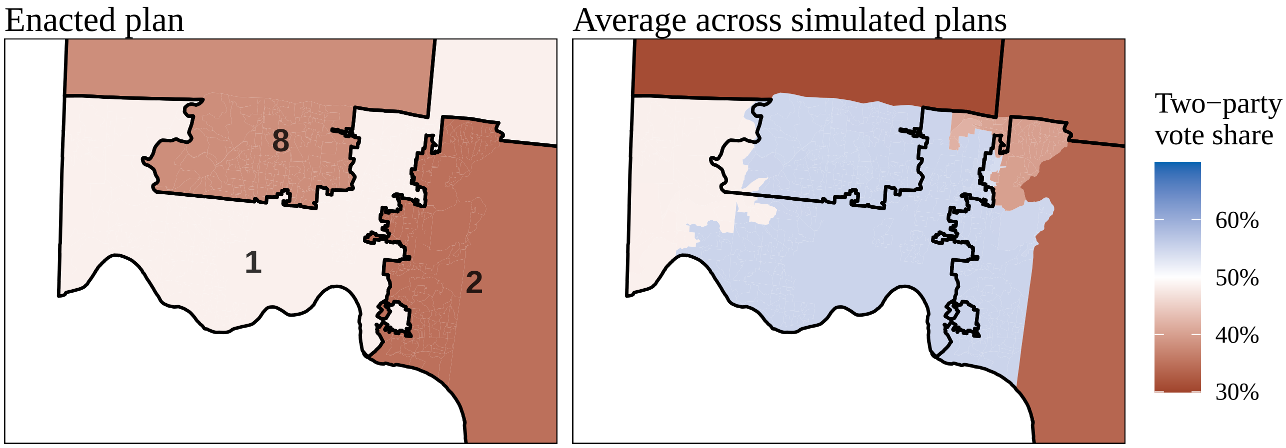

Figure 4: Congressional districts in Hamilton County. The left map presents the expected two-party vote shares of districts under the enacted plan, while the right map shows, for each precinct, the average expected two-party vote share of districts to which the precinct is assigned across the simulated plans. The enacted district boundaries are shown with thick black lines. While under the simulated plans, Cincinnati and its environs are expected to belong to a Democratic-leaning district, the enacted plan cracks Democratic voters, leading to solely Republican districts.

enacted plan, while the right map shows, for each precinct, the average expected two-party vote share of districts to which the precinct is assigned across the simulated plans. Under the enacted plan, Democratic areas are cracked to yield three Republican-leaning districts, despite a significant concentration of Democratic voters in and around Cincinnati. This is especially apparent with the two unusual protrusions of Districts 2 and 8 into Hamilton County, which split the county twice. The simulated plans, in comparison, are expected to only split Hamilton County once. As the right figure indicates, the area covered by these protrusions would normally be expected to belong to a Democratic district, but as a result of being lumped with adjacent districts in the enacted plan, instead belongs to safely Republican districts.

41. As a result of these manipulations and additional splits of Hamilton County, the enacted plan has no Democratic seats under the average statewide federal contest, whereas the simulated plans are expected to yield a Democratic seat. So in Hamilton County alone, cracking of Democratic voters nets Republicans an entire seat.

B. Franklin County

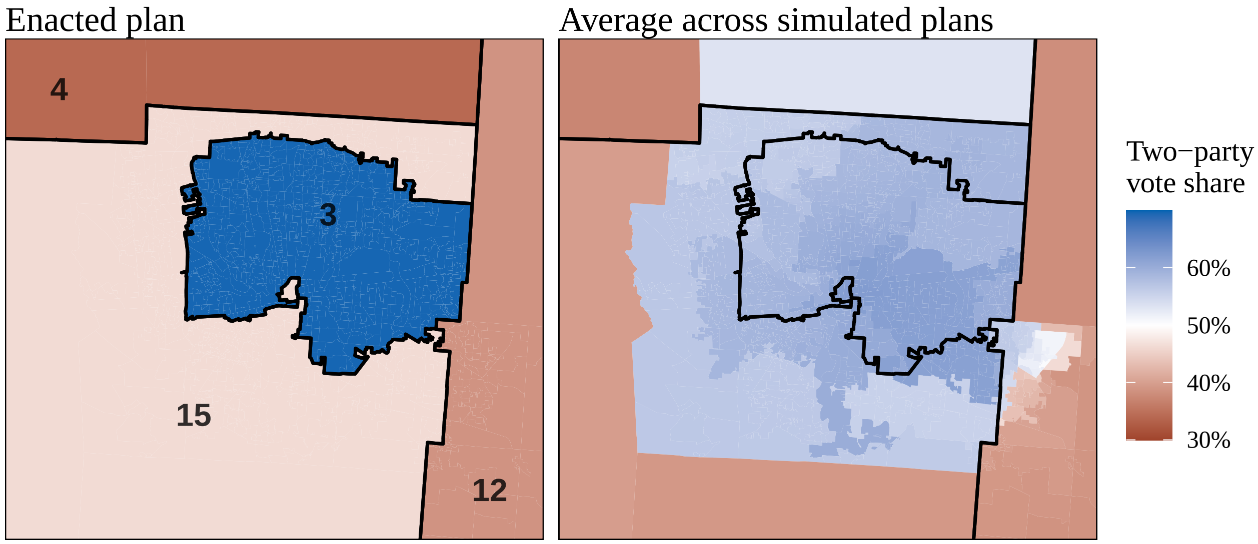

Figure 5: Congressional districts in Franklin County. The left map presents the expected two-party vote shares of districts under the enacted plan, while the right map shows, for each precinct, the average expected two-party vote share of districts to which the precinct is assigned across the simulated plans. The enacted district boundaries are shown with thick black lines. While under the simulated plans, all of Franklin County are expected to belong to a Democratic district, the enacted plan packs Democratic voters, leaving much of the city of Columbus in a Republican district stretching most of the way to Cincinnati.

42. Analogous to Figure 4, Figure 5 compares the enacted plan with the simulated plans in Franklin County. Unlike in Hamilton County, the enacted plan packs Democratic voters into a single, heavily Democratic, District 3, leaving Districts 4, 12, and 15 to be safely Republican. Much of the area inside Franklin County belongs to a safe Republican district under the enacted plan. In contrast, under the simulated plans, the entire area of Franklin County is expected to belong to a Democratic-leaning district, as is Delaware County and part of Fairfield County.

43. By confining Democratic voters to a single district containing part of Columbus, the enacted plan deprives Democratic voters in the rest of the county of a reasonable opportunity to elect a Democratic candidate. In doing so, the enacted plan yields around one additional seat for Republicans, on average, when compared to the simulated plans.

C. Cuyahoga County

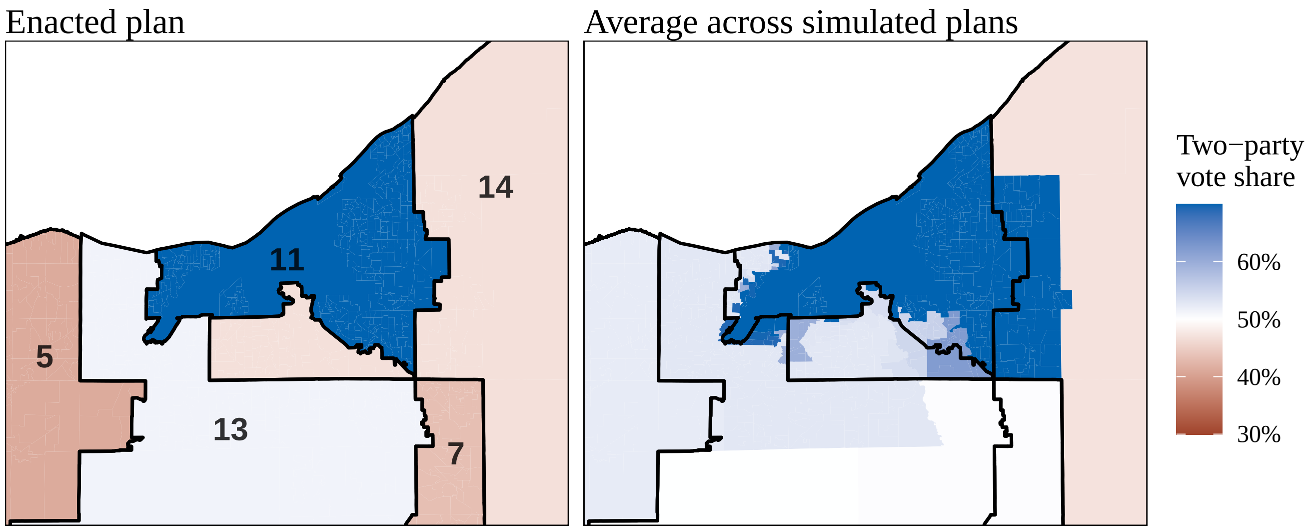

Figure 6: Congressional districts in Cuyahoga County. The left map presents the expected two-party vote shares of districts under the enacted plan, while the right map shows, for each precinct, the average expected two-party vote share of districts to which the precinct is assigned across the simulated plans. The enacted district boundaries are shown with thick black lines. While under the simulated plans, the suburbs of Cleveland are expected to belong to either Democratic districts or highly competitive districts, the enacted plan packs urban Democratic voters, leaving the remainder of Cuyahoga County and nearby areas in Republican districts.

44. Figure 6 is constructed just like Figures 4 and 5. Districts in Cuyahoga County are more constrained than in Franklin County, based on the need to avoid splitting the city of Cleveland, as well as Voting Rights Act considerations. Even so, the enacted plan differs in key ways from the average simulated plan. First, it overly packs Democratic voters in District 11, as indicated by Figure 2 where District 11 corresponds to the least Republican-leaning district (R15). More importantly, Districts 5, 7, 13, and 14 in the enacted plan are drawn to crack the remaining Democratic voters outside of Cleveland and in the cities of Lorain and Akron. The result of this is to create three Republican-leaning districts and only one competitive district. In contrast, under the simulated plans, all of the areas south and west of Cleveland are generally expected to belong to competitive or Democratic-leaning districts.

VII. APPENDIX

A. Introduction to Redistricting Simulation

1. In recent years, redistricting simulation algorithms have played an increasingly important role in court cases involving redistricting plans. Simulation evidence has been presented to courts in many states, including Michigan, North Carolina, Ohio, and Pennsylvania.[5]

2. Over the past several years, researchers have made major scientific advances to improve the theoretical properties and empirical performance of redistricting simulation algorithms. All of the state-of-the-art redistricting simulation algorithms belong to the family of Monte Carlo methods. They are based on random generation of spanning trees, which are mathematical objects in graph theory (DeFord, Duchin, and Solomon 2021). The use of these random spanning trees allows these state-of-the-art algorithms to efficiently sample a representative set of plans (Autry et al. 2020; Carter et al. 2019; McCartan and Imai 2020; Kenny et al. 2021). Algorithms developed earlier, which do not use random spanning trees and instead rely on incremental changes to district boundaries, are often not able to do so.

3. These algorithms are designed to sample plans from a specific probability distribution, which means that every legal redistricting plan has certain odds of being generated. The algorithms put as few restrictions as possible on these odds, except to ensure that, on average, the generated plans meet certain criteria. For example, the probabilities are set so that the generated plans reach a certain level of geographic compactness, on average. Other criteria, based on the state in question, may be fed into the algorithm by the researcher. In other words, this target distribution is based on the weakest assumption about the data under the specified constraints.

4. In addition, the algorithms ensure that all of the sampled plans (a) are geographically contiguous, and (b) have a population which deviates by no more than a specified amount

5. Declaration of Dr. Jonathan C. Mattingly, Common Cause v. Lewis (2019); Testimony of Dr. Jowei Chen, Common Cause v. Lewis (2019); Testimony of Dr. Pegden, Common Cause v. Lewis (2019); Expert Report of Jonathan Mattingly on the North Carolina State Legislature, Rucho v. Common Cause (2019); Expert Report of Jowei Chen, Rucho v. Common Cause (2019); Amicus Brief of Mathematicians, Law Professors, and Students in Support of Appellees and Affirmance, Rucho v. Common Cause (2019); Brief of Amici Curaiae Professors Wesley Pegden, Jonathan Rodden, and Samuel S.-H. Wang in Support of Appellees, Rucho v. Common Cause (2019); Intervenor's Memo, Ohio A. Philip Randolph Inst. et al. v. Larry Householder (2019); Expert Report of Jowei Chen, League of Women Voters of Michigan v. Benson (2019).

from a target population.

5. There are two types of general Monte Carlo algorithms which generate redistricting plans with these guarantees and other properties: sequential Monte Carlo (SMC; Doucet, Freitas, and Gordon 2001) and Markov chain Monte Carlo (MCMC; Gilks, Richardson, and Spiegelhalter 1996) algorithms.

6. The SMC algorithm (McCartan and Imai 2020; Kenny et al. 2021) samples many redistricting plans in parallel, starting from a blank map. First, the algorithm draws a random spanning tree and removes an edge from it, creating a "split" in the map, which forms a new district. This process is repeated until the algorithm generates enough plans with just one district drawn. The algorithm calculates a weight for each plan in a specific way so that the algorithm yields a representative sample from the target probability distribution. Next, the algorithm selects one of the drawn plans at random. Plans with greater weights are more likely to be selected. The algorithm then draws another district using the same splitting procedure and calculates a new weight for each updated plan that comports with the target probability distribution. The whole process of random selection and drawing is repeated again and again, each time drawing one additional district on each plan. Once all districts are drawn, the algorithm yields a sample of maps representative of the target probability distribution.

7. The MCMC algorithms (Autry et al. 2020; Carter et al. 2019) also form districts by drawing a random spanning tree and splitting it. Unlike the SMC algorithm, however, these algorithms do not draw redistricting plans from scratch. Instead, the MCMC algorithms start with an existing plan and modify it, merging a random pair of districts and then splitting them a new way.

8. Diagnostic measures exist for both these algorithms which allow users to make sure the algorithms are functioning correctly and accurately. The original papers for these algorithms referenced above provide more detail on the algorithm specifics, empirical validation of their performance, and the appropriateness of the chosen target distribution.

B. Implementation Details

9. In my analysis, I use the SMC algorithm for several reasons. First, unlike the MCMC algorithms, the SMC algorithm generates nearly independent samples, leading to a diverse set of redistricting plans that satisfy the specified constraints. Second, the SMC algorithm avoids splitting political subdivision boundaries where possible, an important consideration in the case of Ohio. Third, Sections 2(B)(2) and 2(B)(3) require districts to be compact and contiguous, respectively. The SMC algorithm automatically satisfy both of these requirements. Appendix C shows that most of simulated plans generate more compact districts than the enacted plan according to the Polsby-Popper measure, which is a common metric of compactness used in the academic literature.

10. My simulation proceeds in two steps. First, I sample a district in Cuyahoga County using a Voting Rights Act (VRA) constraint to be compliant with Section 2(B)(1). At the instruction of counsel for the relators, I sample one district within Cuyahoga County such that its BVAP proportion falls above 42%. This is done by using the constraint of the form $\sqrt{\max\left(x_b - B(x_b), 0\right)}$, where x_b is the share of a district's VAP that is Black, and $B(x_b)$ returns the target BVAP percentages closest to x_b from the set $\{0.02, 0.08, 0.42\}$. This is a common way to formulate the VRA constraint (Herschlag et al. 2020). Note that I also instructed the algorithm to never split the City of Cleveland, in accordance with Section 2(B)(4)(b), and not to split Cuyahoga County three times or more, in accordance with Sections 2(B)(4)(a) and 2(B)(5).

11. Once a district is sampled within Cuyahoga, I generate the remaining 14 districts within the rest of the state without the VRA constraint. In this second step, I incorporate several split constraints. According to Section 2(B)(4)(b), municipalities with population between 100,000 people and the Congressional ratio of represetation, that reside in a county with population greater than the Congressional ratio of representation, should not be split. In addition to the City of Cleveland, this provision also applies to the City of Cincinnati. I instruct the SMC algorithm to never split either of these municipalities.

12. Section 2(B)(5) requires that of Ohio's 88 counties, at least 65 counties should not

be split; no more than 18 counties can be split no more than once; no more than 5 counties can be split no more than twice. I made sure that all of my simulated plans satisfy this requirement by not splitting the counties the enacted plan does not split and imposing a constraint that discourages the algorithm from splitting a county. This is accomplished in two pieces. First, the SMC algorithm, by design, can be instructed to attempt to follow county boundaries where possible by drawing spanning trees within counties and then between them; I use this feature. Additionally, I penalize a district which splits a county twice with a score of 3, and I penalize a district which splits a county three or more times with a score of 100. A penalty of 100 is so severe that any such district is effectively discarded. These parameter values are chosen such that the diversity of the simulated plans is reasonable while minimizing the number of county splits.

13. As shown in Appendix D, all of my simulation plans have fewer county splits than the enacted plan. In addition, while the enacted plan splits Hamilton and Cuyahoga counties twice, only 8 of my 5,000 simulated plans split two counties twice. 35.9% of the simulated plans split only Franklin County twice whereas the remaining simulated plans split no counties twice.

14. Section 2(B)(4)(a) applies to single municipality or township that exceeds the Congressional ratio of representation. The only municipality or township that satisfies this criteria is the City of Columbus. The provision states that the map drawers "shall attempt to include a significant portion of that municipal corporation or township in a single district and may include in that district other municipal corporations or townships that are located in that county and whose residents have similar interests as the residents of the municipal corporation or township that contains a population that exceeds the congressional ratio of representation." To satisfy this requirement, I impose a penalty of 0.5 for each additional district that encompasses any part of the city. This has the effect of ensuring that the city is not split into many different districts. Again, this parameter value is chosen such that the diversity of the simulated plans is reasonable while appropriately discouraging Columbus splits. Like the enacted plan, all of my simulated plans split Columbus into two districts but in different ways.

15. According to Section 2(B)(6), for counties that are split by a congressional district,

the part of the district that falls within county lines must be geographically contiguous within the county. This requirement is mathematically guaranteed by the properties of the SMC algorithm; by drawing spanning trees hierarchically, within and then across counties, it is impossible to split off a district which has two discontiguous pieces inside one county.

16. Section 2(B)(7) requires that two congressional districts can share at most the territory of a single county, excepting counties with population greater than 400,000, where another county can be shared. Like Section 2(B)(6), this requirement is guaranteed by the SMC algorithm: each new district will split at most one county, whereas a 2(B)(7) violation would require two districts to each split the same two counties.

17. Section 2(B)(8) states, "The authority drawing the districts shall attempt to include at least one whole county in each congressional district." This provision does not apply when a district is contained entirely within a county or when in conflict with federal law. This requirement is guaranteed by the enacted plans' choice of counties to split: with the exception of Cuyahoga and Franklin counties, which are each large enough to have a district contained entirely within them, every other split county is surrounded by counties which are not split. Since I do not permit the algorithm to split these surrounding counties, every other district is either contained within a single county or includes the entirety of one of these surrounding counties.

C. Compactness of the Simulated Plans

18. I now show that the simulated plans are more compliant with Section 2(B)(2), which requires districts to be compact, than the enacted plan. I use the Polsby–Popper (Polsby and Popper 1991) and edge-removal (DeFord, Duchin, and Solomon 2021; McCartan and Imai 2020) scores, two commonly-used quantitative measures of district compactness. For the edge-removal compactness, I present the fraction of edge kept so that like the Polsby–Popper score, a greater value implies a higher level of compactness. Figure 7 shows that a vast majority of the simulated plans are more compact than the enacted plan according to the Polsby–Popper score. If I instead use the edge-removal compactness score, all of the simulated plans have superior compactness when compared to the enacted plan. The result clearly implies that it is possible to be compliant

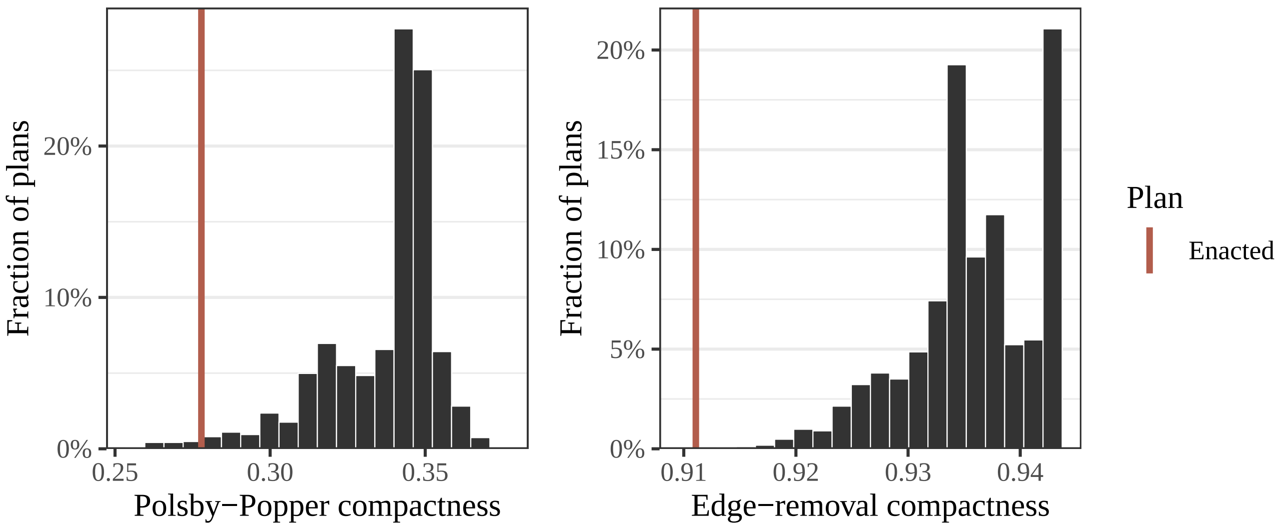

Figure 7: Polsby–Popper and edge–removal compactness scores for the simulated redistricting plans. Overlaid are scores for the enacted plan (red). For both measures, larger values indicate more compact districts.

with Section 1(C)(3)(a) without sacrificing the compliance with Section 2(B)(2).

D. County Splits of the Simulated Plans

19. Similar to compactness, it is possible to be compliant with Section 1(C)(3)(a) without splitting counties more than the enacted plan. The left plot of Figure 8 shows that the number of counties split once is much less under any of the simulated plans than under the enacted plan. The same finding applies to the number of counties that are split twice. As a result, the total number of counties split under the enacted plan is much greater than that under any of the simulated plans.

E. References and Materials Considered

E.1. Data Sources

Data Aquisition

- I analyze a total of 13 statewide elections: US President (2012, 2016, 2020), US Senate (2012, 2016, 2018), Secretary of State (2014, 2018), Governor (2014, 2018), Attorney General (2018), Treasurer (2018), Auditor (2018)

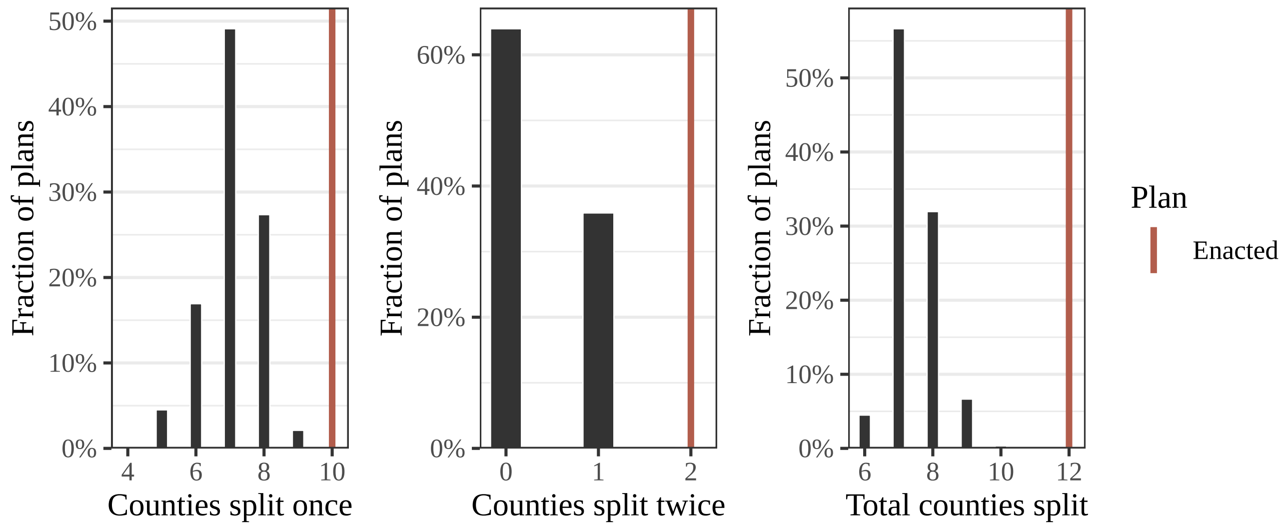

Figure 8: The number of county splits for the simulated redistricting plans. Overlaid are the scores for the enacted plan (red). The left plot shows the number of counties that are split once under each plan, whereas the middle plot presents the number of counties that are split twice under each plan. The right plot shows the number of counties that are split either once or twice. No county is split more than twice under both the enacted plan and any of the simulated plans.

- The six statewide federal elections I use to implement the General Assembly's approach: US President (2012, 2016, 2020), US Senate (2012, 2016, 2018)

- The 2016, 2018, and 2020 precinct-level shapefiles were acquired from the Voting and Election Science Team at the University of Florida and Wichita State University. This data is publicly available on the Harvard Dataverse, an online repository of social science data. Those shapefiles were joined to precinct-level election returns from the Ohio Secretary of State's office, which had been processed and cleaned by OpenElections.

- The 2012 and 2014 election returns pro-rated to the 2010 VTD level were acquired from Bill Cooper. Counsel has informed that Bill Cooper provided the following description of the data: The 2012 results are disaggregated to the block level (based on block centroids) from the statewide 2012 precinct file. The 2014 results are based on a geocoding of about 3.15 million voters who cast ballots in Nov. 2014. These addresses were matched to census blocks and the blocks were aggregated to the precinct level. These virtual precincts were next matched to the 2014 election results and then disaggregated back to the block level, with block-level matches. When aggregated to the congressional level, the differences are measured in the tenths of a percent for House contests. As a final step, these datasets were aggregated from the block-level to the 2010 VTD level. Finally, it is important to note that there is a 2% to 3% undercount statewide for all votes cast in the 2014 election.

- Given the missing votes for the 2014 contests in Lorain County, the VTD-level totals in that county were approximated using the official precinct 2014 returns. First, after identifying the township, city, or village of each 2014 precinct, the official precinct-level returns were aggregated up to that level. Those municipality-level returns were then disaggregated for each candidate down to the VTDs in each municipality, proportionally to the vote counts for the candidate running for the same office and party in the 2018 midterm cycle.
- The 2020 Census Block shapefiles, total population by race and ethnicity, and voting age population by race and ethnicity were obtained directly from the Census FTP portal.
- The 2020 Census place block assignment files (for city and village boundaries and VTD block assignment files) were obtained from the Census website.
- The 2020 Census county subdivision shapefiles (for Ohio township boundaries) were obtained from the Census website.
- The enacted plan data were gathered from the text of SB258, and cleaned into a block equivalency file.
- Geolocated congressional incumbent names and addresses, which were gathered by Carl Klarner, were acquired through Redistricting Data Hub. For new incumbents who came into office following the 2021 general election (Shontel Brown, Mike Carey), their addresses and geolocated locations were given to me by counsel for the plaintiffs.

Data Processing

- The datasets that were on the 2020 census block level (total population, voting age population, Census place assignment, VTD assignment, enacted plan) were joined to the 2020 Census block shapefile.

- The datasets that were not on the level of the census block (2016, 2018, and 2020 election returns – precinct; 2012 and 2014 election returns – 2010 VTD) were disaggregated down to the 2020 census block level. Then, the resulting data were joined to the 2020 Census block shapefile.

- For the 2020 Census county subdivision shapefile, each 2020 Census block was assigned to its corresponding county subdivision assignment by overlaying the county subdivision shapefile onto the 2020 Census blocks.

- Given that some of Ohio's voting districts are geographically discontiguous, the separate discontiguous pieces of each voting district were identified.

Data Aggregation

- The full block-level dataset was aggregated up to the level of the 2020 voting districts, taking into account (a) discontiguous voting districts and (b) splits of voting districts by the enacted plan.

- The final municipality ID was constructed on the aggregated dataset. Where a VTD belonged to a village or a city, the municipality ID took the value of that village or city. Otherwise, it took the value of the county subdivision of the VTD. Then, discontiguous municipalities or townships were identified, and assigned to unique identifiers. The final municipality ID concatenates the original municipality ID, the identifier for each discontiguous piece, and a county identifier, so that it identifies a unique contiguous piece of a municipality within a given county.

E.2. References

Autry, Eric, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. 2020. "Multi-scale merge-split Markov chain Monte Carlo for Redistricting." *arXiv preprint arXiv:2008.08054.*

Carter, Daniel, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. 2019. "A Merge-Split Proposal for Reversible Monte Carlo Markov Chain Sampling of Redistricting Plans." *arXiv preprint arXiv:1911.01503.*

DeFord, Daryl, Moon Duchin, and Justin Solomon. 2021. "Recombination: A Family of Markov Chains for Redistricting." Https://hdsr.mitpress.mit.edu/pub/1ds8ptxu, *Harvard Data Science Review* (March 31, 2021). https://doi.org/10.1162/99608f92.eb30390f. https://hdsr.mitpress.mit.edu/pub/1ds8ptxu.

Doucet, Arnaud, Nando de Freitas, and Neil Gordon. 2001. *Sequential Monte Carlo methods in practice.* New York: Springer.

Fifield, Benjamin, Michael Higgins, Kosuke Imai, and Alexander Tarr. 2020. "Automated Redistricting Simulation Using Markov Chain Monte Carlo." *Journal of Computational and Graphical Statistics* 29 (4): 715–728.

Fifield, Benjamin, Kosuke Imai, Jun Kawahara, and Christopher T Kenny. 2020. "The essential role of empirical validation in legislative redistricting simulation." *Statistics and Public Policy* 7 (1): 52–68.

Gilks, Walter R., Sylvia Richardson, and David J. Spiegelhalter. 1996. *Markov chain Monte Carlo in Practice.* London: Chapman & Hall.

Herschlag, Gregory, Han Sung Kang, Justin Luo, Christy Vaughn Graves, Sachet Bangia, Robert Ravier, and Jonathan C Mattingly. 2020. "Quantifying gerrymandering in North Carolina: Supplemental Appendix." *Statistics and Public Policy* 7 (1): 30–38.

Imai, Kosuke, and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Record." *Political Analysis* 24 (2): 263–272.

Imai, Kosuke, Ying Lu, and Aaron Strauss. 2008. "Bayesian and Likelihood Inference for 2×2 Ecological Tables: An Incomplete Data Approach." *Political Analysis* 16 (1): 41–69.

Kenny, Christopher T., Shiro Kuriwaki, Cory McCartan, Evan Rosenman, Tyler Simko, and Kosuke Imai. 2021. "The Use of Differential Privacy for Census Data and its Impact on Redistricting: The Case of the 2020 U.S. Census." *Science Advances* 7, no. 41 (October): 1–17.

Kenny, Christopher T., Cory McCartan, Benjamin Fifield, and Kosuke Imai. 2020. *redist: Computational Algorithms for Redistricting Simulation.* https://CRAN.R-project.org/package=redist.

McCartan, Cory, and Kosuke Imai. 2020. "Sequential Monte Carlo for sampling balanced and compact redistricting plans." *arXiv preprint arXiv:2008.06131.*

Polsby, Daniel D, and Robert D Popper. 1991. "The third criterion: Compactness as a procedural safeguard against partisan gerrymandering." *Yale Law & Policy Review* 9 (2): 301–353.

Tukey, John W. 1977. *Exploratory Data Analysis.* Pearson.

EXHIBIT A

Curriculum Vitae

Kosuke Imai

Curriculum Vitae

November 2021

Contact Information

1737 Cambridge Street
Institute for Quantitative Social Science
Harvard University
Cambridge, MA 02138

Phone: 617-384-6778
Email: Imai@Harvard.Edu
URL: https://imai.fas.harvard.edu

Education

Ph.D. in Political Science, Harvard University (1999–2003)

A.M. in Statistics, Harvard University (2000–2002)

B.A. in Liberal Arts, The University of Tokyo (1994–1998)

Positions

Professor, Department of Government and Department of Statistics, Harvard University (2018 – present)

Professor, Department of Politics and Center for Statistics and Machine Learning, Princeton University (2013 – 2018)

- Founding Director, Program in Statistics and Machine Learning (2013 – 2017)

Professor of Visiting Status, Graduate Schools of Law and Politics, The University of Tokyo (2016 – present)

Associate Professor, Department of Politics, Princeton University (2012 – 2013)

Assistant Professor, Department of Politics, Princeton University (2004 – 2012)

Visiting Researcher, Faculty of Economics, The University of Tokyo (August, 2006)

Instructor, Department of Politics, Princeton University (2003 – 2004)

Honors and Awards

1. Invited to read "Experimental Evaluation of Computer-Assisted Human Decision-Making: Application to Pretrial Risk Assessment Instrument." before the Royal Statistical Society Research Section, London (2021).

2. *Excellence in Mentoring Award*, awarded by the Society for Political Methodology (2021).

3. *Statistical Software Award* for developing statistical software that makes a significant research contribution, for "fastLink: Fast Probabilistic Record Linkage," awarded by the Society for Political Methodology (2021).

4. *Highly Cited Researcher* (cross-field category) for "production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science," awarded by Clarivate Analytics (2018, 2019, 2020).

5. *President*, The Society for Political Methodology (2017–2019). *Vice President and President-elect* (2015–2017).

6. *Elected Fellow*, The Society for Political Methodology (2017).

7. *The Nils Petter Gleditsch Article of the Year Award* (2017), awarded by *Journal of Peace Research*.

8. *Statistical Software Award* for developing statistical software that makes a significant research contribution, for "mediation: R Package for Causal Mediation Analysis," awarded by the Society for Political Methodology (2015).

9. *Outstanding Reviewer Award* for Journal of Educational and Behavioral Statistics, given by the American Educational Research Association (2014).

10. *The Stanley Kelley, Jr. Teaching Award*, given by the Department of Politics, Princeton University (2013).

11. *Pi Sigma Alpha Award* for the best paper presented at the 2012 Midwest Political Science Association annual meeting, for "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan," awarded by the Midwest Political Science Association (2013).

12. Invited to read "Experimental Designs for Identifying Causal Mechanisms" before the Royal Statistical Society Research Section, London (2012).

13. Inaugural recipient of the *Emerging Scholar Award* for a young scholar making exceptional contributions to political methodology who is within ten years of their terminal degree, awarded by the Society for Political Methodology (2011).

14. *Political Analysis Editors' Choice Award* for an article providing an especially significant contribution to political methodology, for "Estimation of Heterogeneous Treatment Effects from Randomized Experiments, with Application to the Optimal Planning of the Get-out-the-vote Campaign," awarded by the Society for Political Methodology and Oxford University Press (2011).

15. *Tom Ten Have Memorial Award* for the best poster presented at the 2011 Atlantic Causal Inference Conference, for "Identifying Treatment Effect Heterogeneity through Optimal Classification and Variable Selection," awarded by the Departments of Biostatistics and Statistics, University of Michigan (2011).

16. Nominated for the *Graduate Mentoring Award*, The McGraw Center for Teaching and Learning, Princeton University (2010, 2011).

17. *New Hot Paper*, for the most-cited paper in the field of Economics & Business in the last two months among papers published in the last year, for "Misunderstandings among Experimentalists and Observationalists about Causal Inference," named by Thomson Reuters' ScienceWatch (2009).

18. *Warren Miller Prize* for the best article published in *Political Analysis*, for "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," awarded by the Society for Political Methodology and Oxford University Press (2008).

19. *Fast Breaking Paper* for the article with the largest percentage increase in citations among those in the top 1% of total citations across the social sciences in the last two years, for "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," named by Thomson Reuters' ScienceWatch (2008).

20. *Pharmacoepidemiology and Drug Safety Outstanding Reviewer Recognition* (2008).

21. *Miyake Award* for the best political science article published in 2005, for "Do Get-Out-The-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments," awarded by the Japanese Political Science Association (2006).

22. *Toppan Prize* for the best dissertation in political science, for *Essays on Political Methodology*, awarded by Harvard University (2004). Also, nominated for American Political Science Association E.E. Schattschneider Award for the best doctoral dissertation in the field of American government and politics.

Publications in English

Book

Imai, Kosuke. (2017). *Quantitative Social Science: An Introduction.* Princeton University Press. Translated into Japanese (2018), Chinese (2020), and Korean (2021).

Stata version (2021) with Lori D. Bougher.

Tidyverse version (forthcoming) with Nora Webb Williams

Refereed Journal Articles

1. Fan, Jianqing, Kosuke Imai, Inbeom Lee, Han Liu, Yang Ning, and Xiaolin Yang. "Optimal Covariate Balancing Conditions in Propensity Score Estimation." *Journal of Business & Economic Statistics*, Forthcoming.

2. Imai, Kosuke, Zhichao Jiang, D. James Greiner, Ryan Halen, and Sooahn Shin. "Experimental Evaluation of Computer-Assisted Human Decision-Making: Application to Pretrial Risk Assessment Instrument." (with discussion) *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Forthcoming. To be read before the Royal Statistical Society.

3. Imai, Kosuke, In Song Kim, and Erik Wang. "Matching Methods for Causal Inference with Time-Series Cross-Sectional Data." *American Journal of Political Science*, Forthcoming.

4. Imai, Kosuke and Michael Lingzhi Li. "Experimental Evaluation of Individualized Treatment Rules." *Journal of the American Statistical Association*, Forthcoming.

5. de la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. "Experimental Design and Statistical Inference for Conjoint Analysis: The Essential Role of Population Distribution." *Political Analysis*, Forthcoming.

6. Kenny, Christopher T., Shiro Kuriwaki, Cory McCartan, Evan Rosenman, Tyler Simko, and Kosuke Imai. (2021). "The Use of Differential Privacy for Census Data and its Impact on Redistricting: The Case of the 2020 U.S. Census." *Science Advances*, Vol. 7, No. 7 (October), pp. 1-17.

7. Imai, Kosuke and James Lo. (2021). " Robustness of Empirical Evidence for the Democratic Peace: A Nonparametric Sensitivity Analysis." *International Organization*, Vol. 75, No. 3 (Summer), pp. 901–919.

8. Imai, Kosuke, Zhichao Jiang, and Anup Malani. (2021). "Causal Inference with Interference and Noncompliance in the Two-Stage Randomized Experiments." *Journal of the American Statistical Association*, Vol. 116, No. 534, pp. 632-644.

9. Imai, Kosuke, and In Song Kim. (2021). "On the Use of Two-way Fixed Effects Regression Models for Causal Inference with Panel Data." *Political Analysis*, Vol. 29, No. 3 (July), pp. 405–415.

10. Imai, Kosuke and Zhichao Jiang. (2020). "Identification and Sensitivity Analysis of Contagion Effects with Randomized Placebo-Controlled Trials." *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol. 183, No. 4 (October), pp. 1637–1657.

11. Fifield, Benjamin, Michael Higgins, Kosuke Imai, and Alexander Tarr. (2020). "Automated Redistricting Simulation Using Markov Chain Monte Carlo." *Journal of Computational and Graphical Statistics*, Vol. 29, No. 4, pp. 715–728.

12. Fifield, Benjamin, Kosuke Imai, Jun Kawahara, and Christopher T. Kenny. (2020). "The Essential Role of Empirical Validation in Legislative Redistricting Simulation." *Statistics and Public Policy*, Vol. 7, No 1, pp. 52–68.

13. Ning, Yang, Sida Peng, and Kosuke Imai. (2020). "Robust Estimation of Causal Effects via High-Dimensional Covariate Balancing Propensity Score." *Biometrika*, Vol. 107, No. 3 (September), pp. 533—554.

14. Chou, Winston, Kosuke Imai, and Bryn Rosenfeld. (2020). "Sensitive Survey Questions with Auxiliary Information." *Sociological Methods & Research*, Vol. 49, No. 2 (May), pp. 418–454.

15. Imai, Kosuke, Gary King, and Carlos Velasco Rivera. (2020). "Do Nonpartisan Programmatic Policies Have Partisan Electoral Effects? Evidence from Two Large Scale Randomized Experiments." *Journal of Politics*, Vol. 82, No. 2 (April), pp. 714–730.

16. Zhao, Shandong, David A. van Dyk, and Kosuke Imai. (2020). "Propensity-Score Based Methods for Causal Inference in Observational Studies with Non-Binary Treatments." *Statistical Methods in Medical Research*, Vol. 29, No. 3 (March), pp. 709–727.

17. Lyall, Jason, Yang-Yang Zhou, and Kosuke Imai. (2020). "Can Economic Assistance Shape Combatant Support in Wartime? Experimental Evidence from Afghanistan." *American Political Science Review*, Vol. 114, No. 1 (February), pp. 126–143.

18. Kim, In Song, Steven Liao, and Kosuke Imai. (2020). "Measuring Trade Profile with Granular Product-level Trade Data." *American Journal of Political Science*, Vol. 64, No. 1 (January), pp. 102-117.

19. Enamorado, Ted and Kosuke Imai. (2019). "Validating Self-reported Turnout by Linking Public Opinion Surveys with Administrative Records." *Public Opinion Quarterly*, Vol. 83, No. 4 (Winter), pp. 723—748.

20. Blair, Graeme, Winston Chou, and Kosuke Imai. (2019). "List Experiments with Measurement Error." *Political Analysis*, Vol. 27, No. 4 (October), pp. 455–480.

21. Egami, Naoki, and Kosuke Imai. "Causal Interaction in Factorial Experiments: Application to Conjoint Analysis." *Journal of the American Statistical Association*, Vol. 114, No. 526 (June), pp. 529-540.

22. Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. (2019). "Using a Probabilistic Model to Assist Merging of Large-scale Administrative Records." *American Political Science Review*, Vol. 113, No. 2 (May), pp. 353–371.

23. Imai, Kosuke and In Song Kim. (2019) "When Should We Use Linear Fixed Effects Regression Models for Causal Inference with Longitudinal Data?." *American Journal of Political Science*, Vol. 63, No. 2 (April), pp. 467–490.

24. Imai, Kosuke, and Zhichao Jiang. (2018). "A Sensitivity Analysis for Missing Outcomes Due to Truncation-by-Death under the Matched-Pairs Design." *Statistics in Medicine*, Vol. 37, No. 20 (September), pp. 2907–2922.

25. Fong, Christian, Chad Hazlett, and Kosuke Imai. (2018). "Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Efficacy of Political Advertisements." *Annals of Applied Statistics*, Vol. 12, No. 1, pp. 156–177.

26. Hirose, Kentaro, Kosuke Imai, and Jason Lyall. (2017). "Can Civilian Attitudes Predict Insurgent Violence?: Ideology and Insurgent Tactical Choice in Civil War" *Journal of Peace Research*, Vol. 51, No. 1 (January), pp. 47–63.

27. Imai, Kosuke, James Lo, and Jonathan Olmsted. (2016). "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review*, Vol. 110, No. 4 (December), pp. 631–656.

28. Rosenfeld, Bryn, Kosuke Imai, and Jacob Shapiro. (2016). "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions." *American Journal of Political Science*, Vol. 60, No. 3 (July), pp. 783–802.

29. Imai, Kosuke and Kabir Khanna. (2016). "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Record." *Political Analysis*, Vol. 24, No. 2 (Spring), pp. 263–272.

30. Blair, Graeme, Kosuke Imai, and Yang-Yang Zhou. (2015). "Design and Analysis of the Randomized Response Technique." *Journal of the American Statistical Association*, Vol. 110, No. 511 (September), pp. 1304–1319.

31. Imai, Kosuke and Marc Ratkovic. (2015). "Robust Estimation of Inverse Probability Weights for Marginal Structural Models." *Journal of the American Statistical Association*, Vol. 110, No. 511 (September), pp. 1013–1023. (lead article)

32. Lyall, Jason, Yuki Shiraito, and Kosuke Imai. (2015). "Coethnic Bias and Wartime Informing." *Journal of Politics*, Vol. 77, No. 3 (July), pp. 833–848.

33. Imai, Kosuke, Bethany Park, and Kenneth Greene. (2015). "Using the Predicted Responses from List Experiments as Explanatory Variables in Regression Models." *Political Analysis*, Vol. 23, No. 2 (Spring), pp. 180–196. Translated in Portuguese and Reprinted in *Revista Debates* Vol. 9, No 1.

34. Blair, Graeme, Kosuke Imai, and Jason Lyall. (2014). "Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan." *American Journal of Political Science*, Vol. 58, No. 4 (October), pp. 1043–1063.

35. Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. (2014). "mediation: R Package for Causal Mediation Analysis." *Journal of Statistical Software*, Vol. 59, No. 5 (August), pp. 1–38.

36. Imai, Kosuke and Marc Ratkovic. (2014). "Covariate Balancing Propensity Score." *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, Vol. 76, No. 1 (January), pp. 243–263.

37. Lyall, Jason, Graeme Blair, and Kosuke Imai. (2013). "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review*, Vol. 107, No. 4 (November), pp. 679-705. Winner of the Pi Sigma Alpha Award.

38. Imai, Kosuke and Teppei Yamamoto. (2013). "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments." *Political Analysis*, Vol. 21, No. 2 (Spring), pp. 141–171. (lead article).

39. Imai, Kosuke and Marc Ratkovic. (2013). "Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation." *Annals of Applied Statistics*, Vol. 7, No. 1 (March), pp. 443–470. Winner of the Tom Ten Have Memorial Award. Reprinted in *Advances in Political Methodology*, R. Franzese, Jr. ed., Edward Elger, 2017.

40. Imai, Kosuke, Dustin Tingley, and Teppei Yamamoto. (2013). "Experimental Designs for Identifying Causal Mechanisms."(with discussions) *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol. 176, No. 1 (January), pp. 5–51. (lead article) Read before the Royal Statistical Society, March 2012.

41. Imai, Kosuke, and Dustin Tingley. (2012). "A Statistical Method for Empirical Testing of Competing Theories." *American Journal of Political Science*, Vol. 56, No. 1 (January), pp. 218–236.

42. Blair, Graeme, and Kosuke Imai. (2012). "Statistical Analysis of List Experiments." *Political Analysis*, Vol. 20, No. 1 (Winter), pp. 47–77.

43. Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. (2011). "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review*, Vol. 105, No. 4 (November), pp. 765–789. Reprinted in *Advances in Political Methodology*, R. Franzese, Jr. ed., Edward Elger, 2017.

44. Bullock, Will, Kosuke Imai, and Jacob N. Shapiro. (2011). "Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan." *Political Analysis*, Vol. 19, No. 4 (Autumn), pp. 363–384. (lead article)

45. Imai, Kosuke. (2011). "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association*, Vol. 106, No. 494 (June), pp. 407–416. (featured article)

46. Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth Stuart. (2011). "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." *Journal of Statistical Software*, Vol. 42 (Special Volume on Political Methodology), No. 8 (June), pp. 1–28.

47. Imai, Kosuke, Ying Lu, and Aaron Strauss. (2011). "eco: R Package for Ecological Inference in 2×2 Tables." *Journal of Statistical Software*, Vol. 42 (Special Volume on Political Methodology), No. 5 (June), pp. 1–23.

48. Imai, Kosuke and Aaron Strauss. (2011). "Estimation of Heterogeneous Treatment Effects from Randomized Experiments, with Application to the Optimal Planning of the Get-out-the-vote Campaign." *Political Analysis*, Vol. 19, No. 1 (Winter), pp. 1–19. (lead article) Winner of the Political Analysis Editors' Choice Award.

49. Imai, Kosuke, Luke Keele, and Dustin Tingley. (2010). "A General Approach to Causal Mediation Analysis." *Psychological Methods*, Vol. 15, No. 4 (December), pp. 309–334. (lead article)

50. Imai, Kosuke and Teppei Yamamoto. (2010). "Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analysis." *American Journal of Political Science*, Vol. 54, No. 2 (April), pp. 543–560.

51. Imai, Kosuke, Luke Keele, and Teppei Yamamoto. (2010). "Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science*, Vol. 25, No. 1 (February), pp. 51–71.

52. King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, and Héctor Hernández Llamas. (2009). "Public Policy for the Poor? A Randomized Ten-Month Evaluation of the Mexican Universal Health Insurance Program." (with a comment) *The Lancet*, Vol. 373, No. 9673 (April), pp. 1447–1454.

53. Imai, Kosuke, Gary King, and Clayton Nall. (2009). "The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation." (with discussions) *Statistical Science*, Vol. 24, No. 1 (February), pp. 29–53.

54. Imai, Kosuke. (2009). "Statistical Analysis of Randomized Experiments with Nonignorable Missing Binary Outcomes: An Application to a Voting Experiment." *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, Vol. 58, No. 1 (February), pp. 83–104.

55. Imai, Kosuke, Gary King, and Olivia Lau. (2008). "Toward A Common Framework of Statistical Analysis and Development." *Journal of Computational and Graphical Statistics*, Vol. 17, No. 4 (December), pp. 892–913.

56. Imai, Kosuke. (2008). "Variance Identification and Efficiency Analysis in Experiments under the Matched-Pair Design." *Statistics in Medicine*, Vol. 27, No. 4 (October), pp. 4857–4873.

57. Ho, Daniel E., and Kosuke Imai. (2008). "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: California Alphabet Lottery, 1978–2002." *Public Opinion Quarterly*, Vol. 72, No. 2 (Summer), pp. 216–240.

58. Imai, Kosuke, Gary King, and Elizabeth A. Stuart. (2008). "Misunderstandings among Experimentalists and Observationalists: Balance Test Fallacies in Causal Inference." *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol. 171, No. 2 (April), pp. 481–502. Reprinted in *Field Experiments and their Critics*, D. Teele ed., New Haven: Yale University Press, 2013.

59. Imai, Kosuke, Ying Lu, and Aaron Strauss. (2008). "Bayesian and Likelihood Ecological Inference for 2×2 Tables: An Incomplete Data Approach." *Political Analysis*, Vol. 16, No. 1 (Winter), pp. 41–69.

60. Imai, Kosuke. (2008). "Sharp Bounds on the Causal Effects in Randomized Experiments with "Truncation-by-Death"." *Statistics & Probability Letters*, Vol. 78, No. 2 (February), pp. 144–149.

61. Imai, Kosuke and Samir Soneji. (2007). "On the Estimation of Disability-Free Life Expectancy: Sullivan's Method and Its Extension." *Journal of the American Statistical Association*, Vol. 102, No. 480 (December), pp. 1199–1211.

62. Horiuchi, Yusaku, Kosuke Imai, and Naoko Taniguchi. (2007). "Designing and Analyzing Randomized Experiments: Application to a Japanese Election Survey Experiment." *American Journal of Political Science*, Vol. 51, No. 3 (July), pp. 669–687.

63. Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. (2007). "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis*, Vol. 15, No. 3 (Summer), pp. 199–236. (lead article) Winner of the Warren Miller Prize.

64. Ho, Daniel E., and Kosuke Imai. (2006). "Randomization Inference with Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association*, Vol. 101, No. 475 (September), pp. 888–900.

65. Imai, Kosuke, and David A. van Dyk. (2005). "MNP: R Package for Fitting the Multinomial Probit Model." *Journal of Statistical Software*, Vol. 14, No. 3 (May), pp. 1–32. abstract reprinted in *Journal of Computational and Graphical Statistics* (2005) Vol. 14, No. 3 (September), p. 747.

66. Imai, Kosuke. (2005). "Do Get-Out-The-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments." *American Political Science Review*, Vol. 99, No. 2 (May), pp. 283–300.

67. Imai, Kosuke, and David A. van Dyk. (2005). "A Bayesian Analysis of the Multinomial Probit Model Using Marginal Data Augmentation." *Journal of Econometrics*, Vol. 124, No. 2 (February), pp. 311–334.

68. Imai, Kosuke, and David A. van Dyk. (2004). "Causal Inference With General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association*, Vol. 99, No. 467 (September), pp. 854–866.

69. Imai, Kosuke, and Gary King. (2004). "Did Illegal Overseas Absentee Ballots Decide the 2000 U.S. Presidential Election?" *Perspectives on Politics*, Vol. 2, No. 3 (September), pp. 537–549. Our analysis is a part of *The New York Times* article, "How Bush Took Florida: Mining the Overseas Absentee Vote" By David Barstow and Don van Natta Jr. July 15, 2001, Page 1, Column 1.

Invited Contributions

1. Imai, Kosuke, and Zhichao Jiang. (2019). "Comment: The Challenges of Multiple Causes." *Journal of the American Statistical Association*, Vol. 114, No. 528, pp. 1605—1610.

2. Benjamin, Daniel J., *et al.* (2018). "Redefine Statistical Significance." *Nature Human Behaviour*, Vol. 2, No. 1, pp. 6–10.

3. de la Cuesta, Brandon and Kosuke Imai. (2016). "Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections." *Annual Review of Political Science*, Vol. 19, pp. 375–396.

4. Imai, Kosuke (2016). "Book Review of *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* by Guido W. Imbens and Donald B. Rubin." *Journal of the American Statistical Association*, Vol. 111, No. 515, pp. 1365–1366.

5. Imai, Kosuke, Bethany Park, and Kenneth F. Greene. (2015). "Usando as respostas previsíveis da abordagem list-experiments como variaveis explicativás em modelos de regressão." *Revista Debates*, Vol. 9, No. 1, pp. 121–151. First printed in *Political Analysis*, Vol. 23, No. 2 (Spring).

6. Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. (2014). "Comment on Pearl: Practical Implications of Theoretical Results for Causal Mediation Analysis." *Psychological Methods*, Vol. 19, No. 4 (December), pp. 482–487.

7. Imai, Kosuke, Gary King, and Elizabeth A. Stuart. (2014). "Misunderstandings among Experimentalists and Observationalists: Balance Test Fallacies in Causal Inference." in *Field Experiments and their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences*, D. L. Teele ed., New Haven: Yale University Press, pp. 196–227. First printed in *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol. 171, No. 2 (April).

8. Imai, Kosuke, Dustin Tingley, and Teppei Yamamoto. (2013). "Reply to Discussions of "Experimental Designs for Identifying Causal Mechanisms"." *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, Vol. 173, No. 1 (January), pp. 46–49.

9. Imai, Kosuke. (2012). "Comments: Improving Weighting Methods for Causal Mediation Analysis." *Journal of Research on Educational Effectiveness*, Vol. 5, No. 3, pp. 293–295.

10. Imai, Kosuke. (2011). "Introduction to the Virtual Issue: Past and Future Research Agenda on Causal Inference." *Political Analysis*, Virtual Issue: Causal Inference and Political Methodology.

11. Imai, Kosuke, Booil Jo, and Elizabeth A. Stuart. (2011). "Commentary: Using Potential Outcomes to Understand Causal Mediation Analysis." *Multivariate Behavioral Research*, Vol. 46, No. 5, pp. 842–854.

12. Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. (2010). "Causal Mediation Analysis Using R," in *Advances in Social Science Research Using R*, H. D. Vinod (ed.), New York: Springer (Lecture Notes in Statistics), pp. 129–154.

13. Imai, Kosuke, Gary King, and Clayton Nall. (2009). "Rejoinder: Matched Pairs and the Future of Cluster-Randomized Experiments." *Statistical Science*, Vol. 24, No. 1 (February), pp. 65–72.

14. Imai, Kosuke. (2003). "Review of Jeff Gill's *Bayesian Methods: A Social and Behavioral Sciences Approach*," *The Political Methodologist*, Vol. 11 No. 1, 9–10.

Refereed Conference Proceedings

1. Svyatkovskiy, Alexey, Kosuke Imai, Mary Kroeger, and Yuki Shiraito. (2016). "Large-scale text processing pipeline with Apache Spark," *IEEE International Conference on Big Data*, Washington, DC, pp. 3928-3935.

Other Publications and Manuscripts

1. Goldstein, Daniel, Kosuke Imai, Anja S. Göritz, and Peter M. Gollwitzer. (2008). "Nudging Turnout: Mere Measurement and Implementation Planning of Intentions to Vote."

2. Ho, Daniel E. and Kosuke Imai. (2004). " The Impact of Partisan Electoral Regulation: Ballot Effects from the California Alphabet Lottery, 1978–2002." Princeton Law & Public Affairs Paper No. 04-001; Harvard Public Law Working Paper No. 89.

3. Imai, Kosuke. (2003). "Essays on Political Methodology," *Ph.D. Thesis.* Department of Government, Harvard University.

4. Imai, Kosuke, and Jeremy M. Weinstein. (2000). "Measuring the Economic Impact of Civil War," Working Paper Series No. 51, Center for International Development, Harvard University.

Selected Manuscripts

1. McCartan, Cory, Jacob Brown, and Kosuke Imai. "Measuring and Modeling Neighborhoods."

2. Ben-Michael, Eli, D. James Greiner, Kosuke Imai, and Zhichao Jiang. "Safe Policy Learning through Extrapolation: Application to Pre-trial Risk Assessment."

3. Tarr, Alexander and Kosuke Imai. "Estimating Average Treatment Effects with Support Vector Machines."

4. McCartan, Cory and Kosuke Imai. "Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans."

5. Imai, Kosuke and Zhichao Jiang. "Principal Fairness for Human and Algorithmic Decision-Making."

6. Papadogeorgou, Georgia, Kosuke Imai, Jason Lyall, and Fan Li. "Causal Inference with Spatio-temporal Data: Estimating the Effects of Airstrikes on Insurgent Violence in Iraq."

7. Eshima, Shusei, Kosuke Imai, and Tomoya Sasaki. "Keyword Assisted Topic Models."

8. Tarr, Alexander, June Hwang, and Kosuke Imai. "Automated Coding of Political Campaign Advertisement Videos: An Empirical Validation Study."

9. Olivella, Santiago, Tyler Pratt, and Kosuke Imai. "Dynamic Stochastic Blockmodel Regression for Social Networks: Application to International Conflicts."

10. Chan, K.C.G, K. Imai, S.C.P. Yam, Z. Zhang. "Efficient Nonparametric Estimation of Causal Mediation Effects."

11. Barber, Michael and Kosuke Imai. "Estimating Neighborhood Effects on Turnout from Geocoded Voter Registration Records."

12. Hirano, Shigeo, Kosuke Imai, Yuki Shiraito, and Masaki Taniguchi. "Policy Positions in Mixed Member Electoral Systems: Evidence from Japan."

Publications in Japanese

1. Imai, Kosuke. (2007). "Keiryō Seijigaku niokeru Ingateki Suiron (Causal Inference in Quantitative Political Science)." *Leviathan*, Vol. 40, Spring, pp. 224–233.

2. Horiuchi, Yusaku, Kosuke Imai, and Naoko Taniguchi. (2005). "Seisaku Jyōhō to Tōhyō Sanka: Field Jikken ni yoru Kensyō (Policy Information and Voter Participation: A Field Experiment)." *Nenpō Seijigaku (The Annals of the Japanese Political Science Association)*, 2005–I, pp. 161–180.

3. Taniguchi, Naoko, Yusaku Horiuchi, and Kosuke Imai. (2004). "Seitō Saito no Etsuran ha Tohyō Kōdō ni Eikyō Suruka? (Does Visiting Political Party Websites Influence Voting Behavior?)" *Nikkei Research Report*, Vol. IV, pp. 16–19.

Statistical Software

1. Eshima, Shusei, Kosuke Imai, and Tomoya Sasaki. "Keyword Assisted Topic Models." The Comprehensive R Archive Network and GitHub. 2020.

2. Li, Michael Lingzhi and Kosuke Imai. "evalITR: Evaluating Individualized Treatment Rules." available through The Comprehensive R Archive Network and GitHub. 2020.

3. Egami, Naoki, Brandon de la Cuesta, and Kosuke Imai. "factorEx: Design and Analysis for Factorial Experiments." available through The Comprehensive R Archive Network and GitHub. 2019.

4. Kim, In Song, Erik Wang, Adam Rauh, and Kosuke Imai. "PanelMatch: Matching Methods for Causal Inference with Time-Series Cross-Section Data." available through GitHub. 2018.

5. Olivella, Santiago, Adeline Lo, Tyler Pratt, and Kosuke Imai. "NetMix: Mixed-membership Regression Stochastic Blockmodel for Networks." available through CRAN and Github. 2019.

6. Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. "fastLink: Fast Probabilistic Record Linkage." available through The Comprehensive R Archive Network and GitHub. Winner of the Statistical Software Award. 2017.

7. Khanna, Kabir, and Kosuke Imai. "wru: Who Are You? Bayesian Predictions of Racial Category Using Surname and Geolocation." available through The Comprehensive R Archive Network and GitHub. 2015.

8. Fifield, Benjamin, Christopher T. Kenny, Cory McCartan, and Kosuke Imai. "redist: Markov Chain Monte Carlo Methods for Redistricting Simulation." available through The Comprehensive R Archive Network and GitHub. 2015.

9. Imai, Kosuke, James Lo, and Jonathan Olmsted. "emIRT: EM Algorithms for Estimating Item Response Theory Models." available through The Comprehensive R Archive Network. 2015.

10. Blair, Graeme, Yang-Yang Zhou, and Kosuke Imai. "rr: Statistical Methods for the Randomized Response Technique." available through The Comprehensive R Archive Network and GitHub. 2015.

11. Fong, Christian, Marc Ratkovic, and Kosuke Imai. "CBPS: R Package for Covariate Balancing Propensity Score." available through The Comprehensive R Archive Network and GitHub. 2012.

12. Egami, Naoki, Marc Ratkovic, and Kosuke Imai. "FindIt: R Package for Finding Heterogeneous Treatment Effects." available through The Comprehensive R Archive Network and GitHub. 2012.

13. Kim, In Song, and Kosuke Imai. “wfe: Weighted Linear Fixed Effects Regression Models for Causal Inference.” available through The Comprehensive R Archive Network. 2011.

14. Shiraito, Yuki, and Kosuke Imai. “endorse: R Package for Analyzing Endorsement Experiments.” available through The Comprehensive R Archive Network and GitHub. 2012.

15. Blair, Graeme, and Kosuke Imai. “list: Statistical Methods for the Item Count Technique and List Experiments.” available through The Comprehensive R Archive Network and GitHub. 2011.

16. Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. “mediation: R Package for Causal Mediation Analysis.” available through The Comprehensive R Archive Network and GitHub. 2009. Winner of the Statistical Software Award. Reviewed in *Journal of Educational and Behavioral Statistics*.

17. Imai, Kosuke. “experiment: R Package for Designing and Analyzing Randomized Experiments.” available through The Comprehensive R Archive Network. 2007.

18. Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth Stuart. “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference.” available through The Comprehensive R Archive Network and GitHub. 2005.

19. Imai, Kosuke, Ying Lu, and Aaron Strauss. “eco: Ecological Inference in 2×2 Tables.” available through The Comprehensive R Archive Network and GitHub. 2004.

20. Imai, Kosuke, and David A. van Dyk. “MNP: R Package for Fitting the Multinomial Probit Model.” available through The Comprehensive R Archive Network and GitHub. 2004.

21. Imai, Kosuke, Gary King, and Olivia Lau. “Zelig: Everyone’s Statistical Software.” available through The Comprehensive R Archive Network. 2004.

External Research Grants

Principal Investigator

1. National Science Foundation (2021–2024). “Collaborative Research: Causal Inference with Spatio-Temporal Data on Human Dynamics in Conflict Settings.” (Algorithm for Threat Detection Program; DMS-2124463). Principal Investigator (with Georgia Papadogeorgou and Jason Lyall) $485,340.

2. National Science Foundation (2021–2023). “Evaluating the Impacts of Machine Learning Algorithms on Human Decisions.” (Methodology, Measurement, and Statistics Program; SES-2051196). Principal Investigator (with D. James Greiner and Zhichao Jiang) $330,000.

3. Cisco Systems, Inc. (2020–2022). “Evaluating the Impacts of Algorithmic Recommendations on the Fairness of Human Decisions.” (Ethics in AI; CG# 2370386) Principal Investigator (with D. James Greiner and Zhichao Jiang) $110,085.

4. The Alfred P. Sloan Foundation (2020–2022). “Causal Inference with Complex Treatment Regimes: Design, Identification, Estimation, and Heterogeneity.” (Economics Program;

2020–13946) Co-Principal Investigator (with Francesca Dominici and Jose Zubizarreta) $996,299

5. Facebook Research Grant (2018). $25,000.
6. National Science Foundation (2016–2021). "Collaborative Conference Proposal: Support for Conferences and Mentoring of Women and Underrepresented Groups in Political Methodology." (Methodology, Measurement and Statistics and Political Science Programs; SES–1628102) Principal Investigator (with Jeffrey Lewis) $312,322. Supplement (SES–1831370) $60,000.
7. The United States Agency for International Development (2015–2017). "Unemployment and Insurgent Violence in Afghanistan: Evidence from the Community Development Program." (AID–OAA–A–12–00096) Principal Investigator (with Jason Lyall) $188,037
8. The United States Institute of Peace (2015–2016). "Assessing the Links between Economic Interventions and Stability: An impact evaluation of vocational and skills training in Kandahar, Afghanistan," Principal Investigator (with David Haines, Jon Kurtz, and Jason Lyall) $144,494.
9. Amazon Web Services in Education Research Grant (2014). Principal Investigator (with Graeme Blair and Carlos Velasco Rivera) $3,000.
10. Development Bank of Latin America (CAF) (2013). "The Origins of Citizen Support for Narcos: An Empirical Investigation," Principal Investigator (with Graeme Blair, Fabiana Machado, and Carlos Velasco Rivera). $15,000.
11. The International Growth Centre (2011–2013). "Poverty, Militancy, and Citizen Demands in Natural Resource-Rich Regions: Randomized Evaluation of the Oil Profits Dividend Plan for the Niger Delta" (RA–2010–12–013). Principal Investigator (with Graeme Blair). $117,116.
12. National Science Foundation, (2009–2012). "Statistical Analysis of Causal Mechanisms: Identification, Inference, and Sensitivity Analysis," (Methodology, Measurement, and Statistics Program and Political Science Program; SES–0918968). Principal Investigator. $97,574.
13. National Science Foundation, (2009–2011). "Collaborative Research: The Measurement and Identification of Media Priming Effects in Political Science," (Methodology, Measurement, and Statistics Program and Political Science Program; SES–0849715). Principal Investigator (with Nicholas Valentino). $317,126.
14. National Science Foundation, (2008–2009). "New Statistical Methods for Randomized Experiments in Political Science and Public Policy," (Political Science Program; SES–0752050). Principal Investigator. $52,565.
15. National Science Foundation, (2006–2009). "Collaborative Research: Generalized Propensity Score Methods," (Methodology, Measurement and Statistics Program; SES–0550873). Principal Investigator (with Donald B. Rubin and David A. van Dyk). $460,000.
16. The Telecommunications Advancement Foundation, (2004). "Analyzing the Effects of Party Webpages on Political Opinions and Voting Behavior," Principal Investigator (with Naoko Taniguchi and Yusaku Horiuchi). $12,000.

Adviser and Statistical Consultant

1. National Science Foundation (2016–2017). "Doctoral Dissertation Research: Crossing Africa's Arbitrary Borders: How Refugees Shape National Boundaries by Challenging Them." (Political Science Program, SES–1560636). Principal Investigator and Adviser for Co-PI Yang-Yang Zhou's Dissertation Research. $18,900.

2. Institute of Education Sciences (2012–2014). "Academic and Behavioral Consequences of Visible Security Measures in Schools" (R305A120181). Statistical Consultant (Emily Tanner-Smith, Principal Investigator). $351,228.

3. National Science Foundation (2013–2014). "Doctoral Dissertation Research: Open Trade for Sale: Lobbying by Productive Exporting Firm" (Political Science Program, SES–1264090). Principal Investigator and Adviser for Co-PI In Song Kim's Dissertation Research. $22,540.

4. National Science Foundation (2012–2013). "Doctoral Dissertation Research: The Politics of Location in Resource Rent Distribution and the Projection of Power in Africa" (Political Science Program, SES–1260754). Principal Investigator and Adviser for Co-PI Graeme Blair's Dissertation Research. $17,640.

Invited Short Courses and Outreach Lectures

1. Short Course on Causal Inference and Statistics – Department of Political Science, Rice University, 2009; Institute of Political Science, Academia Sinica, 2014.

2. Short Course on Causal Inference and Identification, The Empirical Implications of Theoretical Models (EITM) Summer Institute – Harris School of Public Policy, University of Chicago, 2011; Department of Politics, Princeton University, 2012.

3. Short Course on Causal Mediation Analysis – Summer Graduate Seminar, Institute of Statistical Mathematics, Tokyo Japan, 2010; Society for Research on Educational Effectiveness Conference, Washington DC, Fall 2011, Spring 2012, Spring 2015; Inter-American Development Bank, 2012; Center for Education Research, University of Wisconsin, Madison, 2012; Bobst Center for Peace and Justice, Princeton University, 2014; Graduate School of Education, University of Pennsylvania, 2014; EITM Summer Institute, Duke University, 2014; Center for Lifespan Psychology, Max Planck Institute for Human Development, 2015; School of Communication Research, University of Amsterdam, 2015; Uppsala University, 2016

4. Short Course on Covariate Balancing Propensity Score – Society for Research on Educational Effectiveness Conference, Washington DC, Spring 2013; Uppsala University, 2016

5. Short Course on Matching Methods for Causal Inference – Institute of Behavioral Science, University of Colorado, Boulder, 2009; Department of Political Science, Duke University, 2013.

6. Lecture on Statistics and Social Sciences – New Jersey Japanese School, 2011, 2016; Kaisei Academy, 2012, 2014; Princeton University Wilson College, 2012; University of Tokyo, 2014

Selected Presentations

1. Distinguished speaker, Harvard College Summer Program for Undergraduates in Data Science, 2021.
2. Keynote speaker, Kansas-Western Missouri Chapter of the American Statistical Association, 2021.
3. Invited plenary panelist, Association for Computing Machinery Conference on Fairness, Accountability, and Transparency (ACM FAccT) 2021.
4. Keynote speaker, Taiwan Political Science Association, 2020.
5. Keynote speaker, Boston Japanese Researchers Forum, Massachusetts Institute of Technology, 2020.
6. Keynote speaker, Causal Mediation Analysis Training Workshop, Mailman School of Public Health, Columbia University, 2020.
7. Keynote speaker, Special Workshop on Evidence-based Policy Making. World Economic Forum, Centre for the Fourth Industrial Revolution, Japan, 2020.
8. Distinguished speaker, Institute for Data, Systems, and Society. Massachusetts Institute of Technology, 2019.
9. Keynote speaker, The Harvard Experimental Political Science Graduate Student Conference, Harvard University, 2019.
10. Invited speaker, Beyond Curve Fitting: Causation, Counterfactuals, and Imagination-based AI. Association for the Advancement of Artificial Intelligence, Spring Symposium, Stanford University, 2019.
11. Inaugural speaker, Causal Inference Seminar, Departments of Biostatistics and Statistics, Boston University, 2019.
12. Keynote speaker, The Second Latin American Political Methodology Meeting, Universidad de los Andes (Department of Political Science), 2018.
13. Keynote speaker, The First Latin American Political Methodology Meeting, Pontifical Catholic University of Chile (Department of Political Science), 2017.
14. Keynote speaker, Workshop on Uncovering Causal Mechanisms, University of Munich (Department of Economics), 2016.
15. Keynote speaker, The National Quality Registry Research Conference, Stockholm, 2016.
16. Keynote speaker, The UK-Causal Inference Meeting, University of Bristol (School of Mathematics), 2015.
17. Keynote speaker, The UP-STAT Conference, the Upstate Chapters of the American Statistical Association, 2015.
18. Keynote speaker, The Winter Conference in Statistics, Swedish Statistical Society and Umeå University (Department of Mathematics and Mathematical Statistics), 2015.

19. Inaugural invited speaker, The International Methods Colloquium, Rice University, 2015.
20. Invited speaker, The International Meeting on Experimental and Behavioral Social Sciences, University of Oxford (Nuffield College), 2014.
21. Keynote speaker, The Annual Conference of Australian Society for Quantitative Political Science, University of Sydney, 2013.
22. Keynote speaker, The Graduate Student Conference on Experiments in Interactive Decision Making, Princeton University. 2008.

Conferences Organized

1. The Asian Political Methodology Meetings (January 2014, 2015, 2016, 2017, 2018; co-organizer)
2. The Experimental Research Workshop (September 2012; co-organizer)
3. The 12th World Meeting of the International Society for Bayesian Analysis (June 2012; a member of the organizing committee)
4. Conference on Causal Inference and the Study of Conflict and State Building (May 2012; organizer)
5. The 28th Annual Society for Political Methodology Summer Meeting (July 2011; host)
6. Conference on New Methodologies and their Applications in Comparative Politics and International Relations (February 2011; co-organizer)

Teaching

Courses Taught at Harvard

1. Stat 286/Gov 2003 Causal Inference (formally Stat 186/Gov 2002): introduction to causal inference
2. Gov 2003 Topics in Quantitative Methodology: causal inference, applied Bayesian statistics, machine learning

Courses Taught at Princeton

1. POL 245 Visualizing Data: exploratory data analysis, graphical statistics, data visualization
2. POL 345 Quantitative Analysis and Politics: a first course in quantitative social science
3. POL 451 Statistical Methods in Political Science: basic probability and statistical theory, their applications in the social sciences
4. POL 502 Mathematics for Political Science: real analysis, linear algebra, calculus
5. POL 571 Quantitative Analysis I: probability theory, statistical theory, linear models
6. POL 572 Quantitative Analysis II: intermediate applied statistics

7. POL 573 Quantitative Analysis III: advanced applied statistics
8. POL 574 Quantitative Analysis IV: advanced applied statistics with various topics including Bayesian statistics and causal inference
9. Reading Courses: basic mathematical probability and statistics, applied bayesian statistics, spatial statistics

Advising

Current Students

1. Soubhik Barari (Government)
2. Adam Breuer (Computer Science and Government). To be Assistant Professor, Department of Government and Department of Computer Science, Dartmouth College
3. Jacob Brown (Government)
4. Ambarish Chattopadhyay (Statistics)
5. Shusei Eshima (Government)
6. Georgina Evans (Government)
7. Dae Woong Ham (Statistics)
8. Christopher T. Kenny (Government)
9. Michael Lingzhe Li (MIT, Operations Research Center)
10. Jialu Li (Government)
11. Cory McCartan (Statistics)
12. Sayumi Miyano (Princeton, Politics)
13. Sun Young Park (Government)
14. Casey Petroff (Political Economy and Government)
15. Averell Schmidt (Kennedy School)
16. Sooahn Shin (Government)
17. Tyler Simko (Government)
18. Soichiro Yamauchi (Government)
19. Yi Zhang (Statistics)

Current Postdocs

1. Eli Ben-Michael
2. Evan Rosenman

Former Students

1. Alexander Tarr (Ph.D. in 2021, Department of Electrical and Computer Engineering, Princeton University; Dissertation Committee Chair)
2. Connor Jerzak (Ph.D. in 2021, Department of Government, Harvard University). Postdoctoral Fellow, Linkoping University. To be Assistant Professor, Department of Government, University of Texas, Austin
3. Shiro Kuriwaki (Ph.D. in 2021, Department of Government, Harvard University). Postdoctoral Fellow, Stanford University. To be Assistant Professor, Department of Political Science, Yale University
4. Erik Wang (Ph.D. in 2020, Department of Politics, Princeton University). Assistant Professor, Department of Political and Social Change, Australian National University
5. Diana Stanescu (Ph.D. in 2020, Department of Politics, Princeton University). Postdoctoral Fellow, Stanford University
6. Nicole Pashley (Ph.D. in 2020, Department of Statistics, Harvard University). Assistant Professor, Department of Statistics, Rutgers University
7. Asya Magazinnik (Ph.D. in 2020, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, Massachusetts Institute of Technology
8. Max Goplerud (Ph.D. in 2020, Department of Government, Harvard University). Assistant Professor, Department of Political Science, University of Pittsburgh
9. Naoki Egami (Ph.D. in 2020, Department of Politics, Princeton University; Dissertation Committee Chair). Assistant Professor, Department of Political Science, Columbia University
10. Brandon de la Cuesta (Ph.D. in 2019, Department of Politics, Princeton University). Postdoctoral Fellow, Center on Global Poverty and Development, Stanford University
11. Yang-Yang Zhou (Ph.D. in 2019, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, University of British Columbia
12. Winston Chou (Ph.D. in 2019, Department of Politics, Princeton University). Senior Data Scientist at Apple
13. Ted Enamorado (Ph.D. in 2019, Department of Politics, Princeton University; Dissertation Committee Chair). Assistant Professor, Department of Political Science, Washington University in St. Louis
14. Benjamin Fifield (Ph.D. in 2018, Department of Politics, Princeton University; Dissertation Committee Chair). Data Scientist, American Civil Liberties Union
15. Tyler Pratt. (Ph.D. in 2018, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, Yale University
16. Romain Ferrali (Ph.D. in 2018, Department of Politics, Princeton University). Assistant Professor, Aix-Marseille School of Economics

17. Julia Morse (Ph.D. in 2017, Woodrow Wilson School, Princeton University). Assistant Professor, Department of Political Science, University of California, Santa Barbara
18. Yuki Shiraito (Ph.D. in 2017, Department of Politics, Princeton University; Dissertation Committee Chair). Assistant Professor, Department of Political Science, University of Michigan
19. Carlos Velasco Rivera (Ph.D. in 2016, Department of Politics, Princeton University). Research Scientist, Facebook
20. Gabriel Lopez Moctezuma (Ph.D. in 2016, Department of Politics, Princeton University). Assistant Professor, Division of the Humanities and Social Sciences, California Institute of Technology
21. Graeme Blair (Ph.D. in 2016, Department of Politics, Princeton University). Assistant Professor, University of California, Los Angeles
22. Jaquilyn R. Waddell Boie (Ph.D. in 2015, Department of Politics, Princeton University). Private consultant
23. Scott Abramson (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, University of Rochester
24. Michael Barber (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, Brigham Young University
25. In Song Kim (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, Massachusetts Institute of Technology
26. Alex Ruder (Ph.D. in 2014, Department of Politics, Princeton University). Senior Community Economic Development Advisor, Federal Reserve Bank of Atlanta
27. Meredith Wilf (Ph.D. in 2014, Department of Politics, Princeton University). Senior Director, Capital Rx
28. Will Bullock. (Ph.D. candidate, Department of Politics, Princeton University). Senior Researcher, Facebook
29. Teppei Yamamoto (Ph.D. in 2011, Department of Politics, Princeton University; Dissertation Committee Chair). Associate Professor, Department of Political Science, Massachusetts Institute of Technology
30. Dustin Tingley (Ph.D. in 2010, Department of Politics, Princeton University). Professor, Department of Government, Harvard University
31. Aaron Strauss (Ph.D. in 2009, Department of Politics, Princeton University). Former Executive Director, Analyst Institute
32. Samir Soneji (Ph.D. in 2008, Office of Population Research, Princeton University; Dissertation Committee Chair). Associate Professor, Department of Health Behavior at the Gillings School of Global Public Health, University of North Carolina, Chapel Hill
33. Ying Lu (Ph.D. in 2005, Woodrow Wilson School, Princeton University; Dissertation Committee Chair). Associate Professor, Steinhardt School of Culture, Education, and Human Development, New York University

Former Predocs and Postdocs

1. Zhichao Jiang (Postdoctoral Fellow, 2016–2019). Assistant Professor, Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst
2. Adeline Lo (Postdoctoral Fellow, 2016–2019). Assistant Professor, Department of Political Science, University of Wisconsin, Madison
3. Yunkyu Sohn (Postdoctoral Fellow, 2016–2018). Assistant Professor, School of Political Science and Economics, Waseda University
4. Xiaolin Yang (Postdoctoral Fellow, 2015–2017). Research Scientist, Amazon
5. Santiago Olivella (Postdoctoral Fellow, 2015–2016). Associate Professor, Department of Political Science, University of North Carolina
6. Drew Dimmery (Predoctoral Fellow, 2015–2016). Research Scientist, Facebook
7. James Lo (Postdoctoral Fellow, 2014–2016). Assistant Professor, Department of Political Science, University of Southern California
8. Steven Liao (Predoctoral Fellow, 2014–2015). Assistant Professor, Department of Political Science, University of California, Riverside
9. Michael Higgins (Postdoctoral Fellow, 2013–2015). Associate Professor, Department of Statistics, Kansas State University
10. Kentaro Hirose (Postdoctoral Fellow, 2012–2015). Assistant Professor, Waseda Institute for Advanced Studies
11. Chad Hazlett (Predoctoral Fellow, 2013–2014). Associate Professor, Departments of Political Science and Statistics, University of California, Los Angeles
12. Florian Hollenbach (Predoctoral Fellow, 2013–2014). Associate Professor, Department of International Economics, Government and Business at the Copenhagen Business School
13. Marc Ratkovic (Predoctoral and Postdoctoral Fellow, 2010–2012). Assistant Professor, Department of Politics, Princeton University

Editorial and Referee Service

Co-editor for *Journal of Causal Inference* (2014 – present)

Associate editor for *American Journal of Political Science* (2014 – 2019), *Journal of Business & Economic Statistics* (2015 – 2024), *Journal of Causal Inference* (2011 – 2014), *Journal of Experimental Political Science* (2013 – 2017), *Observational Studies* (2014 – present), *Political Analysis* (2014 – 2017).

Editorial board member for *Asian Journal of Comparative Politics* (2014 – present), *Journal of Educational and Behavioral Statistics* (2011 – present), *Journal of Politics* (2007 – 2008, 2019–2020), *Journal of Research on Educational Effectiveness* (2014 – 2016), *Political Analysis* (2010 – 2013), *Political Science Research and Methods* (2019 – present).

Guest editor for *Political Analysis* virtual issue on causal inference (2011).

Referee for *ACM Computing Surveys*, *American Economic Journal: Applied Economics*, *American Economic Review: Insights*, *American Journal of Epidemiology*, *American Journal of Evaluation*, *American Journal of Political Science*, *American Political Science Review*, *American Politics Research*, *American Sociological Review*, *Annals of Applied Statistics*, *Annals of Statistics*, *Annals of the Institute of Statistical Mathematics*, *Biometrics*, *Biometrika*, *Biostatistics*, *BMC Medical Research Methodology*, *British Journal of Mathematical and Statistical Psychology*, *British Journal of Political Science*, *Canadian Journal of Statistics*, *Chapman & Hall/CRC Press*, *Child Development*, *Communications for Statistical Applications and Methods*, *Computational Statistics and Data Analysis*, *Electoral Studies*, *Econometrica*, *Econometrics*, *Empirical Economics*, *Environmental Management*, *Epidemiology*, *European Union Politics*, *IEEE Transactions on Information Theory*, *International Journal of Biostatistics*, *International Journal of Epidemiology*, *International Journal of Public Opinion Research*, *International Migration Review*, John Wiley & Sons, *Journal of Applied Econometrics*, *Journal of Applied Statistics*, *Journal of Biopharmaceutical Statistics*, *Journal of Business and Economic Statistics*, *Journal of Causal Inference*, *Journal of Computational and Graphical Statistics*, *Journal of Conflict Resolution*, *Journal of Consulting and Clinical Psychology*, *Journal of Econometrics*, *Journal of Educational and Behavioral Statistics*, *Journal of Empirical Legal Studies*, *Journal of Multivariate Analysis*, *Journal of Official Statistics*, *Journal of Peace Research*, *Journal of Politics*, *Journal of Research on Educational Effectiveness*,*Journal of Statistical Planning and Inference*, *Journal of Statistical Software*, *Journal of Survey Statistics and Methodology*, *Journal of the American Statistical Association (Case Studies and Applications; Theory and Methods)*, *Journal of the Japanese and International Economies*, *Journal of the Japan Statistical Society*, *Journal of the Royal Statistical Society (Series A; Series B; Series C)*, *Law & Social Inquiry*, *Legislative Studies Quarterly*, *Management Science*, Multivariate Behavioral Research, National Science Foundation (Economics; Methodology, Measurement, and Statistics; Political Science), Natural Sciences and Engineering Research Council of Canada, *Nature Machine Intelligence*, *NeuroImage*, *Osteoporosis International*, *Oxford Bulletin of Economics and Statistics*, *Pharmaceutical Statistics*, *Pharmacoepidemiology and Drug Safety*, *PLOS One*, *Policy and Internet*, *Political Analysis*, *Political Behavior*, *Political Communication*, *Political Research Quarterly*, *Political Science Research and Methods*, *Population Health Metrics*, *Population Studies*, *Prevention Science*, *Proceedings of the National Academy of Sciences*, Princeton University Press, *Psychological Methods*, *Psychometrika*, *Public Opinion Quarterly*, *Quarterly Journal of Economics*, *Quarterly Journal of Political Science*, *Review of Economics and Statistics*, Routledge, Sage Publications, *Scandinavian Journal of Statistics*, *Science*, Sloan Foundation, Springer, *Sociological Methodology*, *Sociological Methods & Research*, *Statistical Methodology*, *Statistical Methods and Applications*, *Statistical Methods in Medical Research*, *Statistical Science*, *Statistica Sinica*, *Statistics & Probability Letters*, *Statistics in Medicine*, *Systems Biology*, U.S.-Israel Binational Science Foundation, *Value in Health*, *World Politics*.

University and Departmental Committees

Harvard University

Department of Government

Member, Curriculum and Educational Policy Committee (2020–2021)

Member, Second-year Progress Committee (2019–2020)

Member, Graduate Placement Committee (2019–2020)

Member, Graduate Admissions Committee (2018–2019)

Member, Graduate Poster Session Committee (2018–2019)

Department of Statistics

Chair, Senior Faculty Search Committee (2021–2022)

Member, Junior Faculty Search Committee (2018–2019)

Member, Second-year Progress Committee (2018–2019, 2020–2021)

Princeton University

University

Executive Committee Member, Program in Statistics and Machine Learning (2013–2018)

Executive Committee Member, Committee for Statistical Studies (2011-2018)

Member, Organizing Committee, Retreat on Data and Information Science at Princeton (2016)

Member, Council of the Princeton University Community (2015)

Member, Search Committee for the Dean of College (2015)

Member, Committee on the Library and Computing (2013–2016)

Member, Committee on the Fund for Experimental Social Science (2013–2018)

Member, Personally Identifiable Research Data Group (2012–2018)

Member, Research Computing Advisory Group (2013–2018)

Member, Task Force on Statistics and Machine Learning (2014–2015)

Department of Politics

Chair, Department Committee on Research and Computing (2012–2018)

Chair, Formal and Quantitative Methods Junior Search Committee (2012–2013, 2014–2015, 2016–2017)

Chair, Reappointment Committee (2015–2016)

Member, Diversity Initiative Committee (2014–2015)

Member, American Politics Junior Search Committee (2012–2014)

Member, Department Chair's Advisory Committee (2010–2013, 2015–2016)

Member, Department Priority Committee (2012–2013, 2014–2015, 2016–2017)

Member, Formal and Quantitative Methods Curriculum Committee (2005–2006)

Member, Formal and Quantitative Methods Junior Search Committee (2009–2010, 2015–2016)

Member, Formal and Quantitative Methods Postdoc Search Committee (2009–2018)

Member, Graduate Admissions Committee (2012–2013)
Member, Reappointment Committee (2014–2016)
Member, Space Committee (2014–2016)
Member, Undergraduate Curriculum Committee (2014–2015)
Member, Undergraduate Exam Committee (2007–2008)
Member, Undergraduate Thesis Prize Committee (2005–2006, 2008–2011)

Center for Statistics and Machine Learning

Executive Committee Member (2016–2018)
Member, Search Committee (2015–2017)

Services to the Profession

National Academies of Sciences, Engineering, and Medicine

Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Panel on the Review and Evaluation of the 2014 Survey of Income and Program Participation Content and Design (2014–2017)

National Science Foundation

Proposal Review Panel (2020)

The Society for Political Methodology

President (2017–2019)
Vice President and President Elect (2015–2017)
Annual Meeting Committee, Chair (2011)
Career Award Committee (2015–2017)
Program Committee for Annual Meeting (2012), Chair (2011)
Graduate Student Selection Committee for the Annual Meeting (2005), Chair (2011)
Miller Prize Selection Committee (2010–2011)
Statistical Software Award Committee (2009–2010)
Emerging Scholar Award Committee (2013)

American Statistical Association

Journal of Educational and Behavioral Statistics Management Committee (2016 – present)

Others

External Expert, Department of Methodology, London School of Economics and Political Science (2017)

Memberships

American Political Science Association; American Statistical Association; Midwest Political Science Association; The Society for Political Methodology.

CERTIFICATE OF SERVICE

I, Freda J. Levenson, hereby certify that on this 10th day of December, 2021, I caused a true and correct copy of the foregoing document to be served by email upon the counsel listed below:

Bridget C. Coontz, bridget.coontz@ohioago.gov
Julie M. Pfeiffer, julie.pfeiffer@ohioago.gov
Michael Walton, michael.walton@ohioago.gov

Counsel for Respondent Ohio Secretary of State LaRose

Phillip J. Strach, phil.strach@nelsonmullins.com
Thomas A. Farr, tom.farr@nelsonmullins.com
John E. Branch, III, john.branch@nelsonmullins.com
Alyssa M. Riggins, alyssa.riggins@nelsonmullins.com

Counsel for Respondents House Speaker Robert R. Cupp and Senate President Matt Huffman

/s/ Freda J. Levenson
Freda J. Levenson (0045916)
Counsel for Relators