IN THE UNITED STATES DISTRICT COURT FOR THE MIDDLE DISTRICT OF FLORIDA JACKSONVILLE DIVISION

JACKSONVILLE BRANCH OF THE NAACP, et al.,

Plaintiffs,

v.

CITY OF JACKSONVILLE, et al.,

Defendants.

Case No. 3:22-cv-493-MMH-LLL

EXPERT REPORT

Kosuke Imai, Ph.D.

July 21, 2022

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 plaintiffs in this case to analyze relevant data and provide my expert opinions on the role that race played in drawing certain districts in Jacksonville's City Council district plan (hereafter "the enacted plan"). To do so, I simulated 10,000 alternative redistricting plans while adhering to a set of redistricting criteria (hereafter "simulated plans"). I ensured that my simulated plans are generally at least as compliant with the traditional redistricting and other criteria as the enacted plan, on average. These criteria include population equality, compactness, and avoidance of precinct and neighborhood splits. In addition, following the enacted plan, I instructed the simulation algorithm to have no incumbency pairing. I also imposed the avoidance of incumbency pairing for the school board districts, each of which is formed by combining a pair of adjacent City Council districts.

3. Moreover, my simulated plans comply with the Voting Rights Act (VRA). I first conduct a racially polarized voting (RPV) analysis using the official election results. Using the results of this RPV analysis, I then instructed the simulation algorithm to keep a total of at least four VRA-performing districts, in which Black voters would be able to elect the candidate of their choice with a high probability and would cast a majority of votes for such candidate. Thus, my simulation analysis allows me to determine whether and to what extent the enacted plan's inclusion or exclusion of Black voters played a role in drawing relevant district boundaries beyond the purpose of compliance with the VRA and other redistricting criteria.

II. SUMMARY OF OPINIONS

4. My simulation analysis shows that the enacted plan draws its boundary lines such that a disproportionately large number of Black voters are placed into Districts 7, 9, and 10, leading

to unusually high Black voting-age population (BVAP) proportions in these districts in comparison to the simulated plans.[1] As a consequence of this packing of Black voters, Districts 2 and 12 of the enacted plan have much lower BVAP proportions than the simulated plans. Thus, my analysis shows that race played a significant role beyond the purpose of adhering to the traditional and other redistricting criteria including compliance with the VRA.

5. My simulation analysis also shows that the enacted plan dilutes the voting power of Black voters who live in the southwestern parts of Jacksonville and splits the community of Black voters located in the middle of the city. The enacted plan does this by creating four majority-Black districts that are unnecessarily noncompact. In contrast, the simulated plans create at least four VRA-performing districts that are relatively compact while keeping many Black voters together.

III. QUALIFICATIONS, EXPERIENCE, AND COMPENSATION

6. I am trained as a political scientist (Ph.D. in 2003, Harvard) and a statistician (MA in 2002, Harvard). I have published more than 70 articles in peer reviewed journals, including premier political science journals (e.g., *American Journal of Political Science*, *American Political Science Review*, *Political Analysis*), 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."

7. 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 Govern-

1. In this report, I define BVAP as people who are at least 18 years old and either Black alone or Black in combination with any other race per the Census definition.

ment 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 premier 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.

8. 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.

9. 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).

10. 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 a 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.[2]

11. In addition to redistricting simulation methods, I have also developed the methodology for ecological inference, which is commonly used in racially polarized voting analysis of 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

2. https://ipub.com/dev-corner/apps/r-package-downloads/ (accessed on January 17, 2022)

recent decision by the Second Circuit Court of Appeals regarding a redistricting case (*Clerveaux et al. v. East Ramapo Central School District* No. 20-1668).

12. Previously, I have submitted my expert reports, based on redistricting simulation analysis, to the Congressional and General Assembly redistricting cases in Ohio (*League of Women Voters of Ohio et al. v. Ohio Redistricting Commission et al.* The Supreme Court of Ohio, No. 2021–1449; *League of Women Voters of Ohio et al. v. Ohio Redistricting Commission et al.* The Supreme Court of Ohio, No. 2021–1193; *League of Women Voters of Ohio et al. v. Frank LaRose et al.* The Supreme Court of Ohio, No. 2022–0303). In both cases, the Supreme Court of Ohio heavily relied upon my analyses in its decisions (*League of Women Voters of Ohio v. Ohio Redistricting Commission*, Slip Opinion No. 2022-Ohio-65; *Adams v. DeWine*, Slip Opinion No. 2022-Ohio-89).

13. I also submitted expert reports, which utilize redistricting simulation analysis, to the Alabama Congressional redistricting case in the United States District Court for the Northern District of Alabama, Southern Division (*Milligan et al. v. Merrill et al.*, No. 2:21-cv-01530), the Pennsylvania State House redistricting case in the Supreme Court of Pennsylvania (*Benninghoff v. 2021 Legislative Reapportionment Commission*, No. 11 MM 2022), the Kentucky State House and Congressional redistricting cases (*Graham et al. v. Adams et al.*, Commonwealth of Kentucky Franklin Circuit Court Division, No. 22-CI-00047), and the South Carolina Congressional and State House redistricting cases (*The South Carolina State Conference of the NAACP, et al. v. Alexander, et al.* and *The South Carolina State Conference of the NAACP, et al. v. McMaster, et al.*, in the United States District Court for the District of South Carolina, Columbia Division, No. 3:21-cv-03302-JMC-TJH-RMG).

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

15. I am being compensated at a rate of $450 per hour. My compensation does not depend in any way on the outcome of the case or on the opinions and testimony that I provide.

IV. METHODOLOGY

16. I conducted simulation analysis to evaluate whether the enacted plan was drawn using race as a significant factor beyond a set of traditional and other redistricting criteria including compliance with the VRA. Modern redistricting simulation algorithms, including the one I used in my analysis, generate a representative sample of all possible plans that satisfy a specified set of criteria. These criteria may, for example, include requiring a certain degree of population equality, avoiding pairing of incumbents, drawing compact districts, and limiting the number of counties being split. The resulting simulated plans constitute a representative set of alternative plans that comply with these redistricting criteria. One can then evaluate the properties of an enacted plan by comparing it against the simulated plans. If the enacted plan unusually treats particular racial groups in a certain way (e.g. packing and cracking Black voters) *when compared to* the set of simulated plans, this serves as empirical evidence that the enacted plan was likely drawn using race as a significant factor.

17. Furthermore, statistical theory allows us to quantify the degree to which the enacted plan is extreme in terms of racial composition, relative to the ensemble of simulated plans. For example, we can estimate the probability of a simulated plan packing Black people into a district at least as much as the enacted plan does. If this probability is small, then the enacted plan is a statistical outlier because it is highly unlikely to come from the distribution that is used to generate the simulated plans.

18. A primary advantage of the simulation-based approach is its ability to account for the political and geographic features that are specific to each jurisdiction, including spatial distribution of voters and configuration of administrative boundaries. Simulation methods can also incorporate each jurisdiction's redistricting rules, criteria, or guidelines. These jurisdiction-specific features limit the types of redistricting plans that can be drawn, making comparison across jurisdictions and over time difficult. The simulation-based approach, therefore, allows us to compare the enacted plan to a representative set of alternate districting plans subject to Jacksonville's legal requirements. Appendix B provides a brief introduction to redistricting simulation algorithms.

A. Racially Polarized Voting Analysis

19. To ensure that my simulated plans are compliant with the VRA, I conducted a racially polarized voting (RPV) analysis using the official election data from a total of 17 city-wide elections (see the vertical axis of Figure 1 for the list of elections). Specifically, using a standard ecological inference methodology, I estimated the proportions of Black voters who voted for Democratic, Republican, and other candidates in each election. The details of my RPV analysis appears in Appendix C.

20. Figure 1 presents the results of my RPV analysis. In this figure, each dot denotes the estimated proportion of Democratic support among Black voters (solid square) and among White voters (solid circle) for a given election whereas a horizontal line represents its associated 95 percent credible interval.[3] Averaging across all 17 city-wide elections I considered (top row), more than 90% of Black voters (denoted by solid squares) who turned out are estimated to have voted for Democratic candidates. In contrast, White voters (denoted by solid circles) are disproportionately supporting non-Democratic candidates. The average difference in Democratic support between Black and White voters is estimated to be 63 percentage points with the 95% credible interval of $[34, 80]$. The results of my RPV analysis, therefore, imply a clear evidence of racially polarized voting between Black and White voters in Jacksonville.

B. Simulation Setup

21. I conducted a simulation analysis by generating a total of 10,000 alternative redistricting plans with the following properties:

- All simulated districts are geographically contiguous.[4]
- All simulated plans keep Districts 5, 6, and 11 of the enacted plan as they are.[5]

3. The point estimates in this figure are based on posterior mean. They are not necessarily centered in the 95% confidence intervals because of the asymmetry of posterior distributions for each election.

4. Following the enacted plan, I limit water crossings for the St. Johns River to the area of Mill Cove, particularly on the connections of Dames Point Bridge.

5. These districts are of little relevance to my simulation analysis because none of them borders the challenged districts. They are also relatively compact.

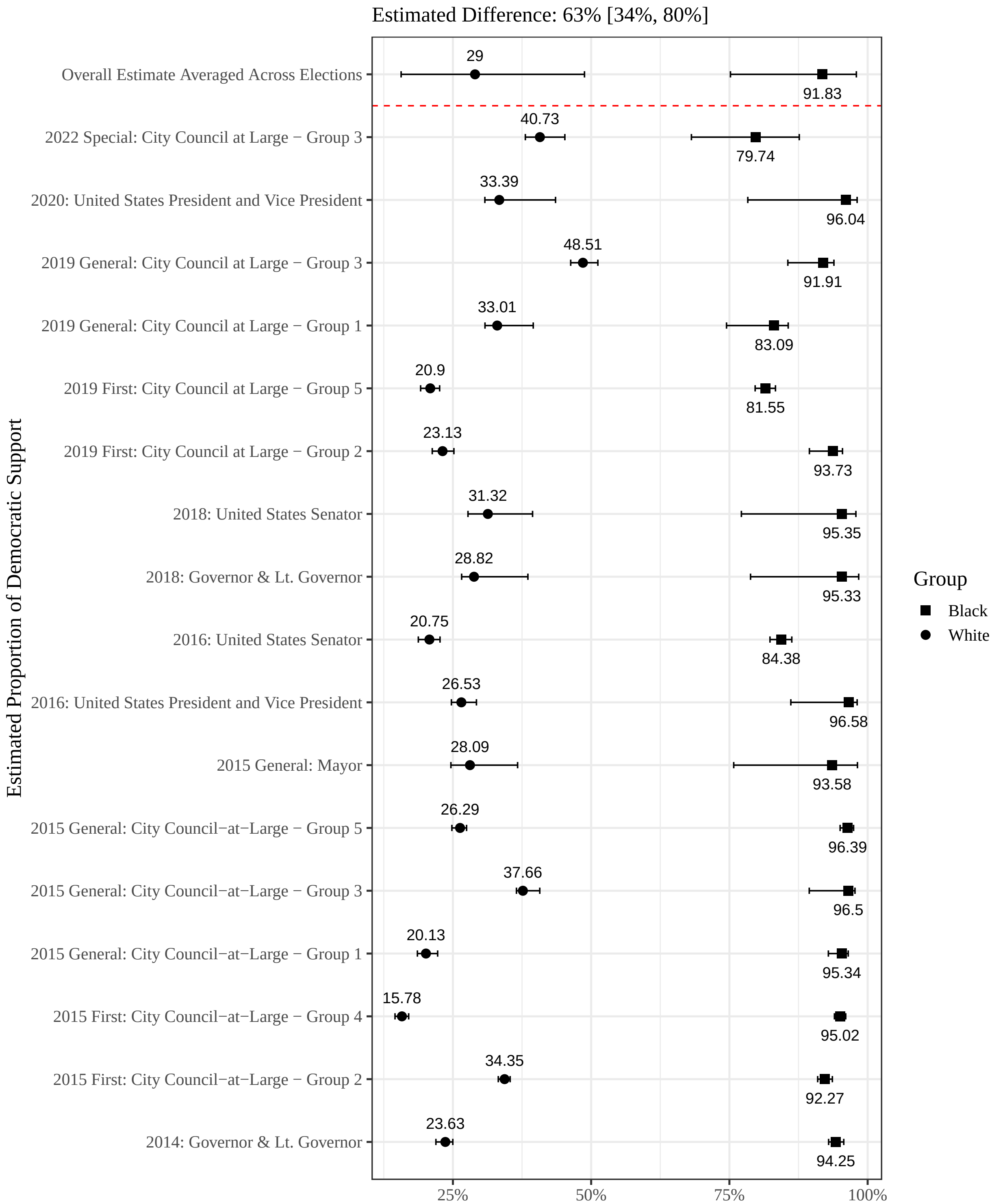

Figure 1: Estimated proportions of Democratic support among Black and White voters for a given election. Each dot denotes a point estimate whereas a horizontal line represents its associated 95 percent credible interval. The results imply a high degree of racially polarized voting in Jacksonville with Black voters (solid squares) overwhelmingly supporting Democratic candidates and White voters (solid circles) disproportionately supporting non-Democratic candidates.

- All simulated districts do not exceed an overall population deviation of $\pm$ 4.8%, which is the maximum population deviation under the enacted plan (see Figure 6 in Appendix E).
- Districts under the simulated plans are similar to or more compact than those under the enacted plan, on average (Figures 7, 8, and 9 in Appendix F).
- All simulated plans have fewer split precincts than the enacted plan (see Figure 10 in Appendix G).
- Almost all simulated plans have fewer split neighborhoods than the enacted plan (see Figure 11 Appendix H).
- All simulated plans place incumbents in separate districts
  - incumbents for City Council districts must be in separate City Council districts
  - incumbents for School Board districts must be in separate School Board districts, each of which is formed by combining two adjacent City Council districts.

22. In addition, I instructed the simulation algorithm such that all simulated plans are compliant with the VRA. This is done by instructing the simulation algorithm to generate a total of at least four VRA-performing districts. A VRA-performing district is defined as a district where the candidate of choice for Black voters is predicted to win at least two thirds of the time and the votes cast by Black voters are likely to form a majority of the votes received by such candidate. Figure 12 of Appendix I shows that like the enacted plan, all simulated plans have a total of at least four VRA-performing districts.

23. My 10,000 simulated plans were generated by only considering the above criteria, using the merge-split type simulation algorithm with the enacted plan as a starting plan (E. A. Autry et al. 2021; Carter et al. 2019; briefly described in Appendix D). I provide detailed information about my simulation procedure in Appendix D. This simulation analysis enables me to examine whether and to what extent race was used as a significant factor in determining the district boundaries of the enacted plan beyond the purpose of adhering to the above redistricting criteria including compliance with the VRA.

24. For my simulation analysis, I can easily generate additional plans by running the

algorithm longer, but for the purpose of my analysis, 10,000 simulated plans for each county will yield statistically precise conclusions. In other words, generating more than 10,000 plans, while possible, will not materially affect the conclusions of my analysis. Figure 5 of Appendix A shows the enacted plan (left) and an example simulated plan (right) with precinct-level BVAP proportions.

C. Description of Redistricting Simulation Software

25. In my analysis, I used the two open-source software packages for redistricting analysis, `redist` (Kenny et al. 2020) and `redistmetrics` (Kenny et al. 2022), which implement a variety of redistricting simulation algorithms as well as other evaluation methods and metrics. My collaborators and I have developed these software packages, so that other researchers and the general public can implement these state-of-the-art methods on their own. I supplemented these packages with code written primarily to account for the redistricting rules, criteria, and guidelines that are specific to Jacksonville. All of my analyses were conducted on a personal computer. Indeed, all of my analysis code can be replicated by running my code 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

26. Using the 10,000 simulated plans, I evaluate how race played a role in determining district boundaries of the enacted plan beyond the purpose of satisfying the redistricting criteria described above, including compliance with the VRA. Using these simulated plans, I conduct two types of analyses: outlier and dislocation analyses. Outlier analysis compares the enacted plan with the simulated plans in terms of district-level BVAP proportions. If the enacted plan significantly departs from the simulated plans with respect to how Black voters are distributed across different districts, it constitutes evidence that race was a significant factor in the design of the districts, beyond what was necessary to comply with the VRA. Dislocation analysis identifies the areas of Jacksonville where Black voters are affected most by the enacted plan in comparison with the simulated plans. Dislocation analysis can be used to present evidence of packing and cracking, which are common forms of racial gerrymandering.

A. Outlier Analysis

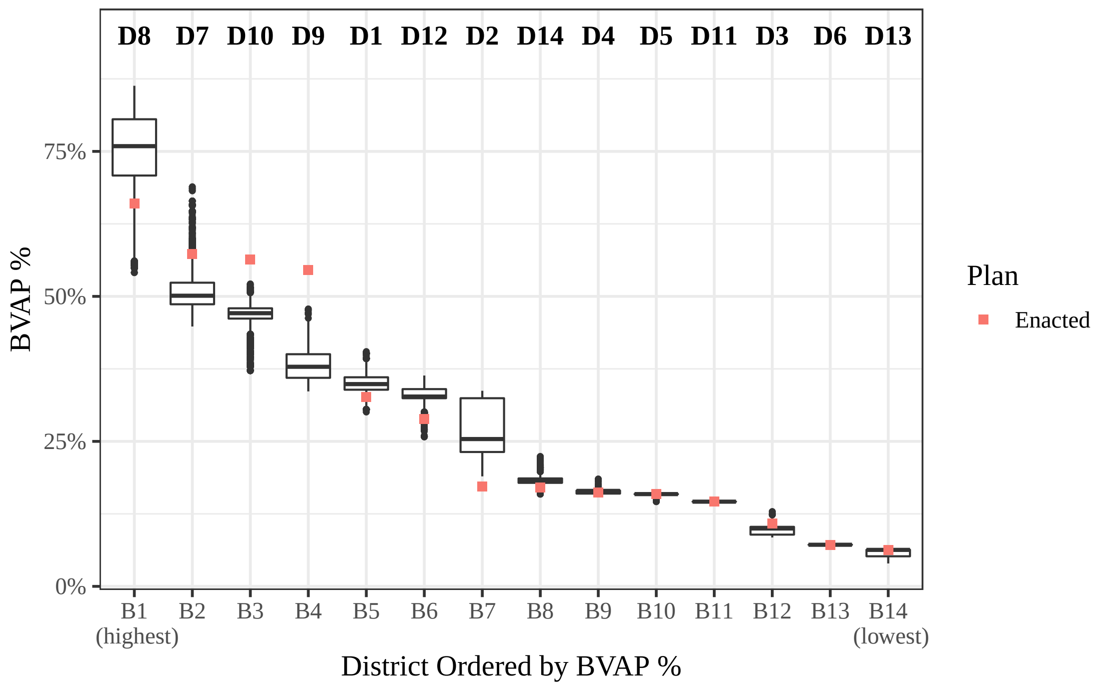

Figure 2: District-level ordered Black voting age population (BVAP) proportions under the simulated and enacted plans. For any given plan, the districts are ordered based on their BVAP proportion, ranging from the district with the highest BVAP proportion ("B1") to that with the lowest BVAP proportion ("B14"). Boxplots represent the distribution of the BVAP proportion for each ordered district under the simulated plans, whereas the red square correponds to the BVAP proportion under the enacted plan with its district label given at the top (e.g., "D8").

27. Figure 2 shows the results of my outlier analysis. In this plot, for any given plan (both enacted and simulated), I ordered the districts based on the magnitude of their BVAP proportion. This means that under any given plan, District B1 has the highest BVAP proportion while District B14 has the lowest BVAP proportion (to be clear, the B1 through B14 district identifiers do not correspond to the City Council district numbers in the enacted plan, which is labeled with "D" at the top of the plot). If the expected BVAP proportion of each ordered district under the enacted plan (red square) diverges from the corresponding distribution of the simulated plans (boxplot), it constitutes evidence that race was used in drawing district boundaries under the enacted plan beyond the purpose of satisfying the set of redistricting criteria specified in Section IV.B including

compliance with the VRA.

28. 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 a common definition of outlier (Tukey 1977).

29. The figure shows that the three ordered districts with the second to fourth highest BVAP proportions (i.e., ordered districts B2, B3, and B4) have much higher BVAP proportion under the enacted plan than the simulated plans. These ordered districts correspond to Districts 7, 10, and 9 (i.e., D7, D10, and D9), respectively, under the enacted plan. In fact, none of my 10,000 simulated plans have higher BVAP proportions for ordered districts B3 and B4 than the corresponding districts. For ordered district B2, only 360 simulated plans (or 3.6%) have higher BVAP proportions than the corresponding district under the enacted plan. Thus, Districts 9 and 10 of the enacted plan are clear outliers in terms of their BVAP proportions while District 7 is also unusual with comparison to the simulated plans. In other words, a disproportionately large number of Black voters are unnecessarily packed into these districts in comparison with the simulated plans.

30. Under the enacted plan, this packing of Black voters into Districts 9 and 10 leads to unusually low BVAP proportions of ordered districts B6 and B7, corresponding Districts 12 and 2 (i.e., D12 and D2), respectively. In fact, none of my 10,000 simulated plans have lower BVAP proportions for ordered district B7 than the corresponding ordered district under the enacted plan. For ordered district B6, only 66 of my 10,000 simulated plans (or 0.66%) have lower BVAP proportions than the corresponding ordered districts under the enacted plan, respectively.

31. Critically, there is a wide gap in BVAP proportion between ordered districts B4 and B5 and between B6 and B7 under the enacted plan whereas no such gap exists under the simulated plans with BVAP proportion more smoothly changing across ordered districts. These results present clear evidence that race played a significant role in determining these relevant district

boundaries of the enacted plan beyond the purpose of satisfying the redistricting criteria specified in Section IV.B, including compliance with the VRA. Appendix J further demonstrates that this conclusion of my outlier analysis is robust to a different definition of VRA performance.

B. Dislocation Analysis

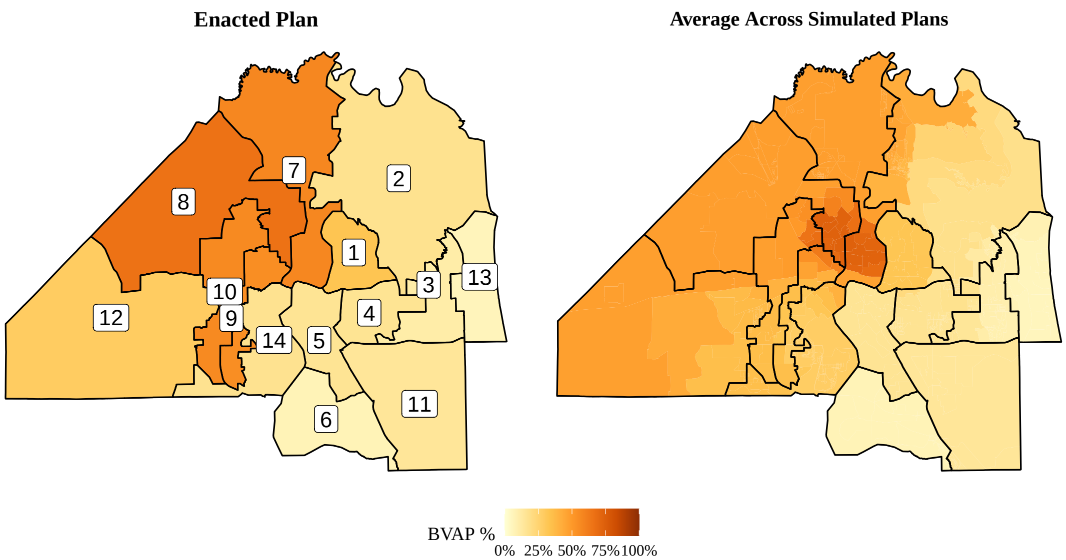

Figure 3: Comparison of the district-level Black Voting-Age Population (BVAP) proportions under the enacted plan (left) with those averaged across the simulation plans. The left map presents the BVAP proportion of each district under the enacted plan, while the right map shows, for each precinct, the average BVAP proportion of the district to which the precinct is assigned across the simulated plans. The enacted district boundaries are shown with thick black lines.

32. I next conduct the dislocation analysis. Figure 3 compares the district-level BVAP proportions under the enacted plan (left) with those averaged across the simulated plans (right). Specifically, the left map presents the BVAP proportion of each district under the enacted plan, while the right map shows, for each precinct, the average BVAP proportion of the district to which the precinct is assigned across the simulated plans. The difference between these two maps appears in Figure 4. This figure shows, for any given precinct, the extent to which the enacted plan yields a higher (orange) or lower (purple) district-level BVAP proportion, on average, when compared to

the simulated plans.

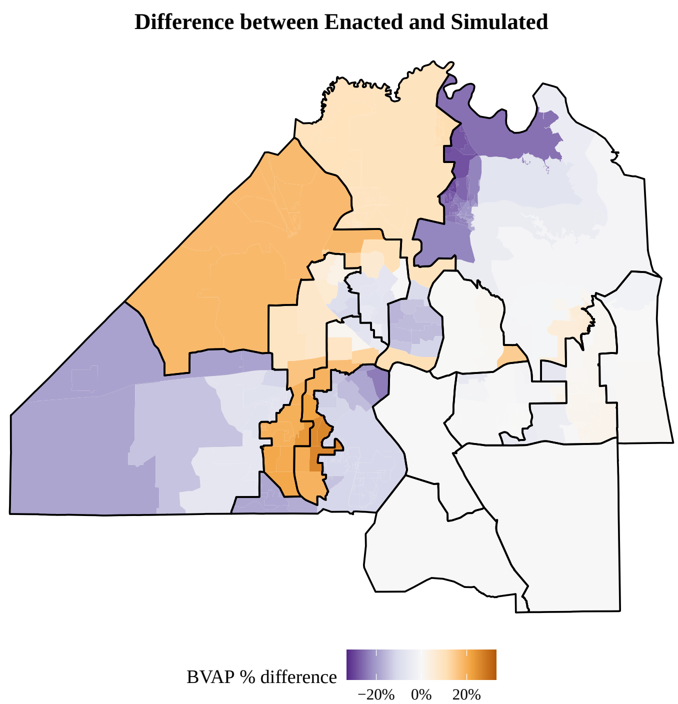

Figure 4: Average difference in district-level Black Voting-Age Population (BVAP) proportions between the enacted and simulated plans. The map presents the difference between the enacted plan as shown in the left map of Figure 3 and the average simulated plans as shown in the right map of the same figure.

33. The two figures show several differences between the enacted and simulated plans in terms of how Black voters are assigned to different districts. First, the voters who live in District 12 under the enacted plan are likely to belong to a district with a much greater BVAP proportion under the simulated plans. Under the enacted plan, the presence of District 10, which is a highly noncompact majority-Black district (BVAP 56.4%), prevents District 12 from having a higher BVAP proportion. Indeed, the BVAP proportion of District 12 under the enacted plan is only 28.9%. In contrast, the simulated plans tend to combine the area in the middle of District 12, which has a relatively high BVAP proportion, with the area of District 10 bordering District 8, where a large number of Black voters live. The resulting simulated district has a much higher BVAP proportion than District 12 and yet is much more compact than District 10.

34. Second, under the enacted plan, District 8 has the highest BVAP proportion, extend-

ing eastward from the western city border with a hook-shaped area that contains many precincts with a large number of Black voters (see the left map of Figure 5 in Appendix A). In addition, District 7 has the second highest BVAP proportion, reaching into the same area where many Black voters live, and Districts 9 and 10 do the same, extending from the southern part of the city. Thus, the enacted plan divides the communities of Black voters into Districts 7, 8, 9, and 10, some of which are highly noncompact.

35. In contrast, the simulated plans tend to create compact districts with the highest BVAP proportions at the center of Jacksonville where the largest number of Black voters live. These simulated districts, which are highlighted by the area with darkest brown color in the right map of Figure 3, tend to keep communities of Black voters together.

36. Furthermore, under the enacted plan, District 2 has a low BVAP proportion of 17.2%. Indeed, although this district has the seventh highest BVAP proportion under the enacted plan (i.e., B7), this proportion is unusually low in comparison to the BVAP proportion of the corresponding simulated district (see Figure 2). According to my dislocation analysis, the reason is that the western part of this district tends to belong to a district with a much higher BVAP proportion. This is because these precincts are often included in a simulated district that contains many precincts located in District 7 of the enacted plan where a relatively large number of Black voters live (see e.g., the right map of Figure 5 in Appendix A).

37. Lastly, the precincts, which are part of District 14 under the enacted plan, tend to belong to a district with a higher BVAP proportion under the simulated plans. In addition, under the simulated plans, the precincts, which form the southern parts of Districts 9 and 10 of the enacted plan, belong to a district with much lower BVAP proportion. This is in part because Districts 9 and 10, which are highly noncompact, pack Black voters, lowering the BVAP proportion of District 14 under the enacted plan. The results support the conclusion that packing a disproportionately large number of Black voters into several districts under the enacted plan dilute their voting power by lowering the BVAP proportions of other districts.

Pursuant to 28 U.S.C. § 1746, I hereby declare under penalty of perjury that the forgoing is true and correct:

Executed, this day, July 20, 2022, in Cambridge, Massachusetts.

____

Kosuke Imai, Ph.D.

VI. APPENDIX

A. Enacted Plan and An Example Simulated Plan

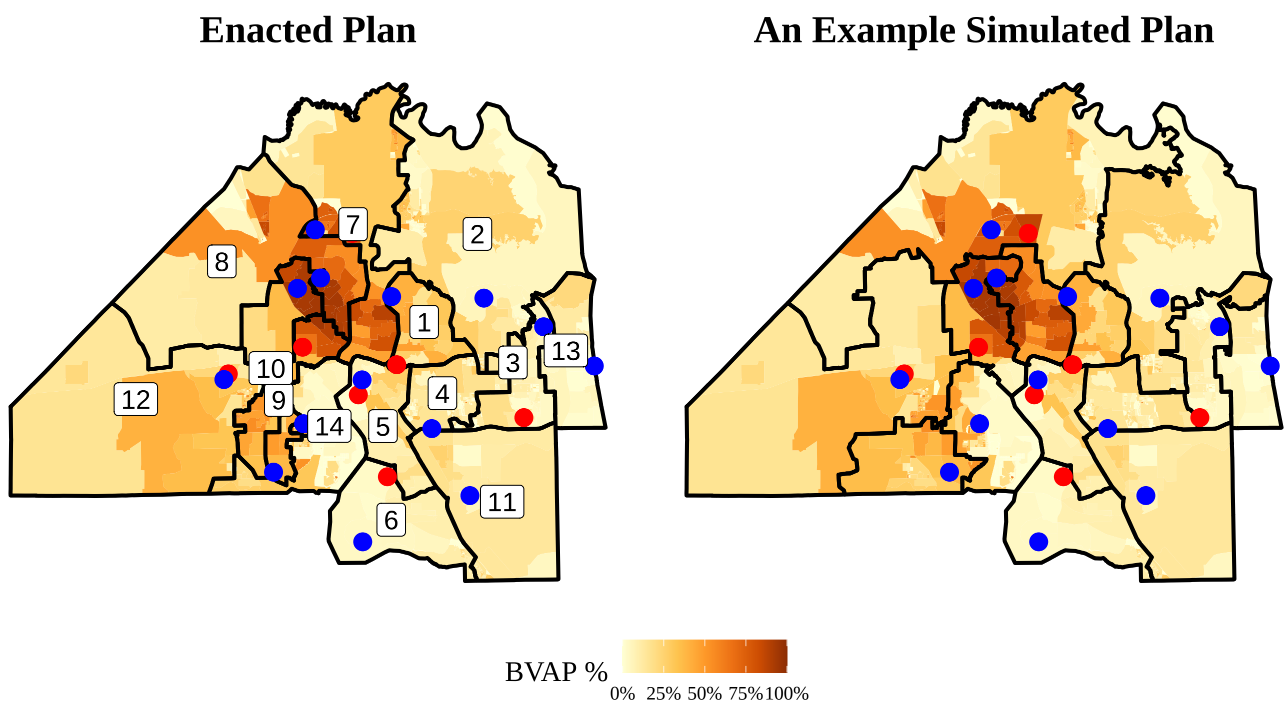

Figure 5: The enacted plan (left) and example simulated plan (right) with precinct-level Black Voting-Age Population (BVAP) proportion. The black solid lines represent the district boundaries, the residence locations of City Council incumbents are shown as blue solid circles, while those of School Board incumbents are indicated by red solid circles. Each precinct is colored according to its BVAP proportion.

38. The left map of Figure 5 shows the enacted plan with color shades indicating the precinct-level BVAP proportion while the right map shows an example simulated plan. Table 1 presents the summary of the enacted plan in terms of district-level compactness and the number of split neighborhoods.

Table 1: Compactness and Neighborhood Splits of Enacted Districts.

District	Polsby-Popper	Reock	Convex Hull	Neighborhood Splits
1	0.532	0.502	0.821	4
2	0.284	0.489	0.753	12
3	0.203	0.311	0.671	6
4	0.498	0.495	0.810	5
5	0.457	0.367	0.752	4
6	0.501	0.454	0.793	2
7	0.179	0.299	0.624	6
8	0.310	0.489	0.733	11
9	0.148	0.189	0.543	13
10	0.192	0.202	0.608	17
11	0.671	0.508	0.943	1
12	0.563	0.546	0.904	9
13	0.441	0.351	0.791	2
14	0.222	0.316	0.643	8

B. Introduction to Redistricting Simulation

39. 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 Alabama, Michigan, North Carolina, Ohio, and Pennsylvania.[6]

40. 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 (E.

6. Examples include 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). Expert Report of Kosuke Imai, League of Women Voters of Ohio et al. v. Ohio Redistricting Commission et al. (2021). Expert Report of Kosuke Imai, Milligan et al. v. Merrill et al. (2021).

Autry et al. 2020; E. A. Autry et al. 2021; 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 generate a representative sample of plans.

41. These modern 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 algorithms may 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 from a target population. In addition, the generated plans reach a certain level of geographic compactness, on average. Other criteria, based on the jurisdiction 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.

42. 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.

43. 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.

44. The MCMC algorithms (E. Autry et al. 2020; E. A. Autry et al. 2021; 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.

45. 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.

C. Racially Polarized Voting Analysis Details

46. My Racially Polarized Voting (RPV) analysis consists of the following several steps. The same procedure is applied separately to each election. First, I estimate the turnout probability for each racial category — Black, White, and others — by dividing the number of those who turned out by the number of registered voters. This estimate is based on the official election turnout data obtained from Duval County. Second, I conduct an 3×3 Ecological Inference analysis to estimate the vote choice probabilities, conditional on turning out, for each racial group, where the three vote choice categories are Democrat, Republican, and others. When there are multiple candidates from the same party in an election, I combine their votes into Democrat, Republican, and others. For this, I use the Multinomial Dirichlet model as implemented in Lau, Moore, and Kellermann 2007. Third, I combine these estimates with the voting-age population for each racial category from the 2020 Census data to estimate the Democratic and Republican support rate for each precinct. Lastly, the precinct-level support estimates are aggregated to districts under a given redistricting plan.

47. I conduct my RPV analysis separately for each of the 17 elections across 2014-2020

(see the vertical axis of Figure 1 for the list of these elections). Consistent with best practices, first I tune the parameters used to fit the aforementioned Multinomial Dirichlet model. Specifically, for each of the 17 elections I run an independent tuning algorithm 10 times with 25,000 iterations each. Then, I use those tuned parameters to initialize the Markov chain Monte Carlo algorithm for fitting the Multinomial Dirichlet model. I run 17 independent chains of 125,000 draws and a burn-in of 1,000 and thin to every 5th draw. Finally, to calculate my aggregate estimate of racial support averaged across all the elections, I take the weighted average of estimates across elections where the weight is proportional to the number of voters of a given race who turned out in that election.

D. Simulation Analysis Details

48. I used the merge-split type MCMC algorithm (E. Autry et al. 2020; E. A. Autry et al. 2021; Carter et al. 2019) as implemented in Kenny et al. 2020. To match decisions made by the City Council, I restrict water crossings over the St. John's River between districts to the area of the Dames Point Bridge. This is done by manually removing the edges between the relevant precincts in the adjacency graph. I freeze three irrelevant districts in the southeast (Districts 5, 6, and 11) that are neither challenged nor bordering any of the challenged districts, and simulate the remaining 11 districts. When evaluating districts, I combine the 11 simulated districts with the 3 frozen districts for full, valid plans of 14 districts. I use the following set of constraints to ensure that the simulated plans have a set of properties specified in Section IV.B: population equality (maximum deviation less than or equal to the value of the enacted plan, approximately 4.8%), compactness ($\rho = 1.25$), avoid pairing incumbents in city council and school board elections (weight 50), and minimizing neighborhood splits (weight 0.5).

49. In addition, I also include a constraint (of weight 10) that encourages at least top four districts by BVAP proportion to perform: (1) the candidate of choice by Black voters is predicted to win at least 12 out of the 17 elections, and (2) Black voters on average would have been the majority of the electorate for that candidate.

50. Using the above specification, I run 4 independent chains with 40,000 iterations with 5,000 burn-in draws for each chain. I then thin these chains to every 5th draw, resulting in the final set of 28,000 draws. Following the simulations, I subset my sample to the plans that do not pair incumbents for City Council or School Board districts. In addition, I remove 3950 plans that fail to generate School Board districts without incumbency pairing (note that a School Board district must also consist of two adjacent City Council districts). This is done by using the enumeration algorithm to generate all possible ways to create valid School Board districts using given each simulated plan (Fifield, Imai, et al. 2020). I further remove 41 plans, in which at least one of the top four BVAP districts does not perform according to the aforementioned VRA criteria. Finally, I randomly subset my remaining 24009 valid simulated plans down to a final sample of 10,000 plans, which is sufficiently many to give statistical precision needed for my analysis.

E. Maximum Population Deviation of the Simulated Districts

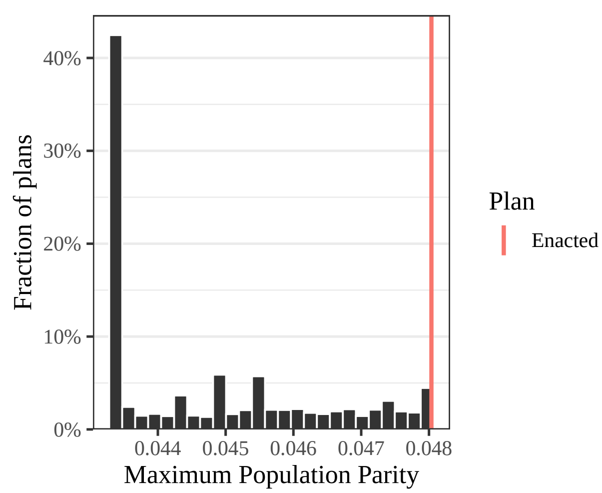

Figure 6: Maximum population deviation under the simulated plans. The histogram shows the maximum population deviation under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. All simulated plans have a smaller maximum population deviation than the enacted plan.

F. Compactness of the Simulated Districts

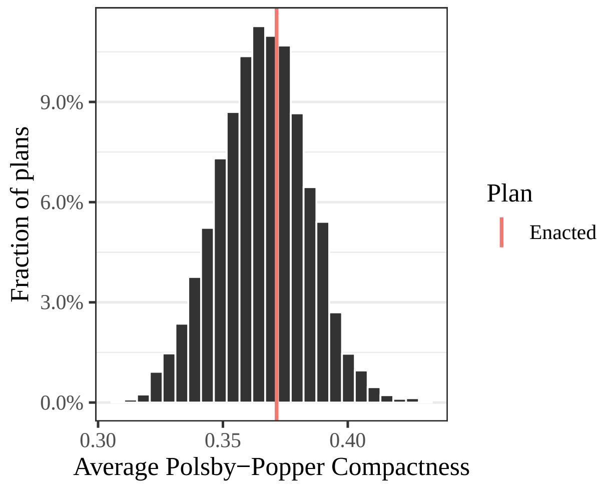

Figure 7: Average compactness based on the Polsby-Popper score under the simulated plans. The histogram shows the Posby-Popper compactness score under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. A greater value indicates a more compact district. On average, the simulated plans tend to have an average compactness score similar to that of the enacted plan.

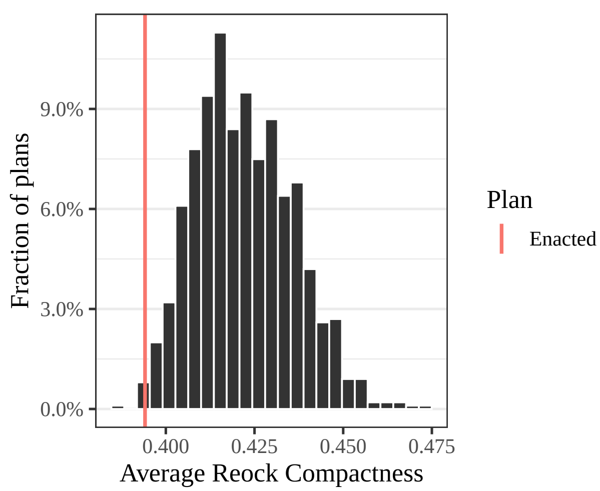

Figure 8: Average compactness based on the Reock score under the simulated plans. The histogram shows the Reock compactness score under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. A greater value indicates a more compact district. The simulated plans tend to have a greater average compactness score than the enacted plan.

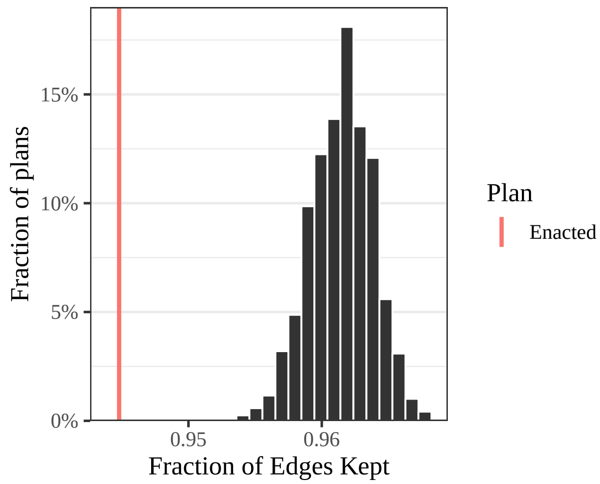

Figure 9: Compactness based on the fraction of edges kept score under the simulated plans. The histogram shows the fraction of edges kept compactness score under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. A greater value indicates a more compact district. All simulated plans have a greater average compactness score than the enacted plan.

G. Precinct Splits of the Simulated Districts

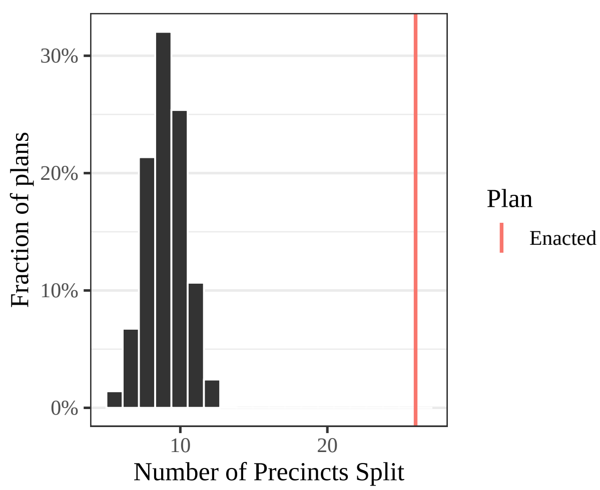

Figure 10: Number of split precincts under the simulated plans. The histogram shows the number of split precincts under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. All simulated plans have fewer split precincts than the enacted plan.

H. Neighborhood Splits of the Simulated Districts

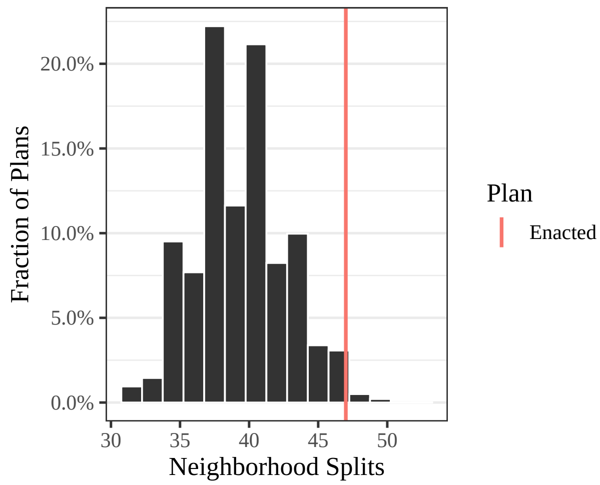

Figure 11: Number of split neighborhoods under the simulated plans. The histogram shows the number of split neighborhoods under the simulated plans whereas the red vertical line represents the corresponding number under the enacted plan. A vast majority of the simulated plans have fewer split neighborhoods than the enacted plan.

I. Voting Rights Act Compliance of the Simulated Plans

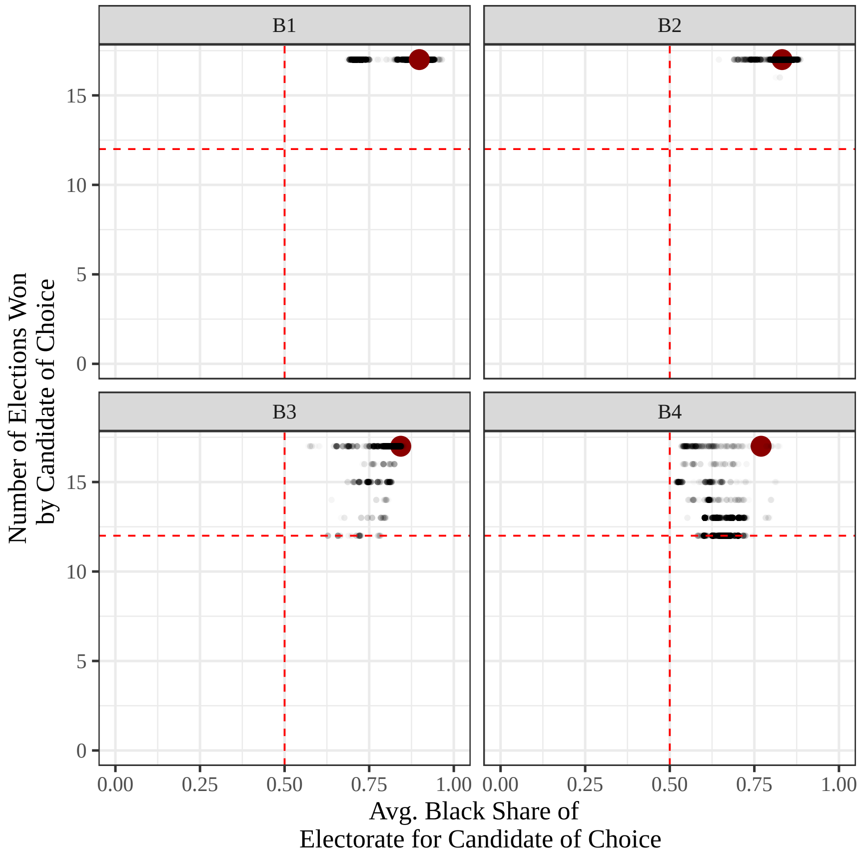

Figure 12: Estimated number of elections won by the candidate of choice for Black voters (x-axis) and the proportion of such candidate's votes cast by Black voters. Each plot represents one of the four VRA-performing districts under the enacted and simulated plans, ordered by the BVAP proportion ("B1" in the top left indicates the district with the highest BVAP proportion, "B2" in the top right represents the one with the second highest district, and so on). In each plot, a red solid circle represents the corresponding district under the enacted plan whereas a black solid circle indicates the corresponding district for each simulated plan. A dot in the north east corner indicates the district is performing under the corresponding plan. Like the enacted plan, all simulated plans have a total of at least four VRA-perfoming districts.

J. Robustness Analysis

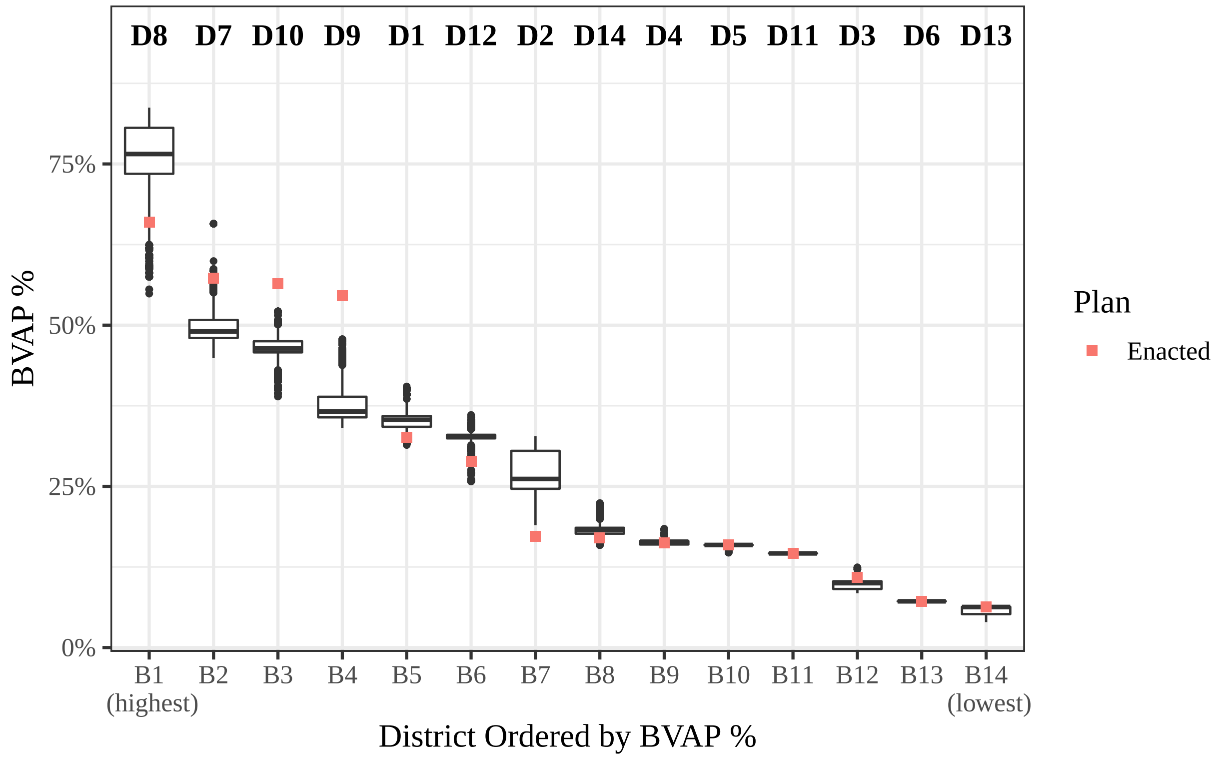

Figure 13: Outlier analysis based on the simulated plans that have at least four VRA-performing districts where a candidate of choice for Black voters is predicted to win 14 or more out of 17 elections. The figure compares an enacted district (red square) with its corresponding simulated districts (boxplot) using the simulated plans with this alternative definition of VRA performance.

51. I examine the robustness of my outlier analysis to a different way in which a district is determined to perform according to the VRA. Specifically, I define a district to be VRA-performing if a candidate of choice for Black voters is predicted to win 14 or more (instead of 12 or more as in the original analysis) out of 17 elections and a majority of votes received by such a candidate are cast by Black voters. Out of my initial 24009 valid simulated plans, a total of 7248 simulated plans have at least four VRA-performing districts based on this new definition. Figure 13 shows the results that are very similar to those of my original outlier analysis (see Figure 2). Thus, the conclusions of my outlier analysis in Section V.A are robust to this different definition of VRA-performance.

K. Data Sources

52. 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 Bureau's Decennial Census API. These were disaggregated proportionally by voting age population down to the 2020 Census block shapefile.

53. The block assignment file for the enacted plan was provided by counsel. The precinct boundaries for the 2020 Duval precincts were provided by counsel. Note that population overlap between 2014 and 2020 precincts was calculated to be 96.2%. Thus, for simplicity, I use the 2020 precincts throughout my analysis. The correspondence between the Census block shapefile and precinct shapes was identified by assigning each block to the precinct with which it had the most area overlap. The 2014, 2015, 2016, 2018, and 2020 precinct-level election results were provided by counsel. These were disaggregated proportionally by voting age population down to the 2020 Census block shapefile.

54. The above datasets were then joined together to form the block level data. The block level data were then aggregated to geographically contiguous components of precincts within each district of the enacted plan. A small number of precincts are split to ensure contiguity and accurate representation of the enacted plan boundaries.

55. Data on neighborhood boundaries are the City of Jacksonville Official Neighborhoods file and were provided by counsel.

56. To conduct the ecological inference analysis, precinct demographic summaries from the Duval Supervisor of Elections website for each election were used. These reports provide voter demographics, including party and race, by precinct. These were converted from PDFs to machine readable format. They were then joined with election results at the precinct level. This allows the data for the ecological inference analysis to be separately from the redistricting simulation analysis to avoid introducing assumptions about precinct splits.

L. 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.*

Autry, Eric A., Daniel Carter, Gregory J. Herschlag, Zach Hunter, and Jonathan C. Mattingly. 2021. "Metropolized Multiscale Forest Recombination for Redistricting." *Multiscale Modeling & Simulation* 19 (4): 1885–1914.

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.

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.

———. 2022. *redistmetrics: Redistricting Metrics.* https://CRAN.R-project.org/package=redistmetrics.

Lau, Olivia, Ryan T Moore, and Michael Kellermann. 2007. "eiPack: R× C ecological inference and higher-dimension data management." *New Functions for Multivariate Analysis* 7 (1): 43.

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

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

EXHIBIT A

Curriculum Vitae

Kosuke Imai

Curriculum Vitae

June 2022

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. James Francis Hannan Lectureship. Department of Statistics and Probability, Michigan State University (2022; declined).
2. 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 (2022).
3. *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, 2021).
4. *Excellence in Mentoring Award*, awarded by the Society for Political Methodology (2021).
5. *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).
6. *President*, The Society for Political Methodology (2017–2019). *Vice President and President-elect* (2015–2017).
7. *Elected Fellow*, The Society for Political Methodology (2017).
8. *The Nils Petter Gleditsch Article of the Year Award* (2017), awarded by *Journal of Peace Research*.
9. *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).
10. *Outstanding Reviewer Award* for Journal of Educational and Behavioral Statistics, given by the American Educational Research Association (2014).
11. *The Stanley Kelley, Jr. Teaching Award*, given by the Department of Politics, Princeton University (2013).
12. *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).
13. Invited to read "Experimental Designs for Identifying Causal Mechanisms" before the Royal Statistical Society Research Section, London (2012).
14. 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).
15. *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).

16. *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).

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

18. *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).

19. *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).

20. *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).

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

22. *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).

23. *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

Books

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 (2022) with Nora Webb Williams

Llaudet, Elena, and Kosuke Imai. (2023). *Data Analysis for Social Science: A Friendly and Practical Introduction.* Princeton University Press.

Refereed Journal Articles

1. 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." *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, Forthcoming.

2. Olivella, Santiago, Tyler Pratt, and Kosuke Imai. "Dynamic Stochastic Blockmodel Regression for Social Networks: Application to International Conflicts." *Journal of the American Statistical Association*, Forthcoming.

3. 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.

4. 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.

5. 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.

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

7. de la Cuesta, Brandon, Naoki Egami, and Kosuke Imai. (2022). "Experimental Design and Statistical Inference for Conjoint Analysis: The Essential Role of Population Distribution." *Political Analysis*, Vol. 30, No. 1 (January), pp. 19–45.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

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

23. 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.

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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)

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

35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. 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).

41. 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.

42. 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.

43. 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.

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

45. 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.

46. 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)

47. 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)

48. 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.

49. 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.

50. 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.

51. 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)

52. 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.

53. 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.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. 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.

65. 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.

66. 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.

67. 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.

68. 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.

69. 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.

70. 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.

71. 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. (2022). "Causal Diagrams and Social Science Research." *Probabilistic and Causal Inference: The Works of Judea Pearl*. Geffner, Hector and Dechter, Rina and Halpern, Joseph Y. (eds). Association for Computing Machinery and Morgan & Claypool, pp. 647–654.

2. 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.

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

4. 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.

5. 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.

6. 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).

7. 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.

8. 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).

9. 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.

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

11. 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.

12. 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.

13. 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.

14. 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.

15. 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. Ben-Michael, Eli, Kosuke Imai, and Zhichao Jiang. "Policy Learning with Asymmetric Utilities."

2. Imai, Kosuke, Evan Rosenman, and Santiago Olivella. "Addressing Census data problems in race imputation via fully Bayesian Improved Surname Geocoding and name supplements."

3. Imai, Kosuke and Michael Lingzhi Li. "Statistical Inference for Heterogeneous Treatment Effects Discovered by Generic Machine Learning in Randomized Experiments."

4. Ham, Dae Woong, Kosuke Imai, and Lucas Janson. "Using Machine Learning to Test Causal Hypotheses in Conjoint Analysis."

5. Goplerud, Max, Kosuke Imai, Nicole E. Pashley. "Estimating Heterogeneous Causal Effects of High-Dimensional Treatments: Application to Conjoint Analysis."

6. Malani, Anup, Phoebe Holtzman, Kosuke Imai, Cynthia Kinnan, Morgen Miller, Shailender Swaminathan, Alessandra Voena, Bartosz Woda, and Gabriella Conti. "Effect of Health Insurance in India: A Randomized Controlled Trial."

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

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

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

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

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

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

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

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

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

16. 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 and Co-Principal Investigator

1. National Science Foundation (2022-2025). "Collaborative Research: Understanding the Evolution of Political Campaign Advertisements over the Last Century." (Accountable Institutions and Behavior Program, SES–2148928). Principal Investigator (with Michael Crespin and Bryce Dietrich) $538,484.

2. 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.

3. 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.

4. 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.

5. 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

6. Facebook Research Grant (2018). $25,000.

7. 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.

8. 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

9. 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.

10. Amazon Web Services in Education Research Grant (2014). Principal Investigator (with Graeme Blair and Carlos Velasco Rivera) $3,000.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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. Shusei Eshima (Government)
4. Georgina Evans (Government). To be Research Scientist, Google Brain
5. Dae Woong Ham (Statistics)
6. Zeyang Jia (Statistics)
7. Christopher T. Kenny (Government)
8. Jialu Li (Government)
9. Cory McCartan (Statistics)
10. Sayumi Miyano (Princeton, Politics)
11. Sun Young Park (Government)
12. Casey Petroff (Political Economy and Government)
13. Averell Schmidt (Kennedy School)
14. Sooahn Shin (Government)
15. Tyler Simko (Government)
16. Dom Valentino (Government)
17. Soichiro Yamauchi (Government) To be Data Scientist at Google
18. Yi Zhang (Statistics)

Current Postdocs

1. Eli Ben-Michael. To be Assistant Professor, Department of Statistics and Data Science and Heinz College of Informations Systems and Public Policy
2. Evan Rosenman

Former Students

1. Ambarish Chattopadhyay (Ph.D. in 2022, Department of Statistics, Harvard University). To be Postdoctoral Fellow, Stanford University

2. Jacob Brown (Ph.D. in 2022, Department of Government, Harvard University). To be Postdoctoral Fellow, Princeton Unviersity, followed by Assistant Professor, Department of Political Science, Boston University

3. Michael Lingzhe Li (Ph.D. in 2021, Operations Research, MIT). Postdoctoral Fellow, MIT. To be Assistant Professor, Technology and Operations Management Unit, Harvard Business School

4. Alexander Tarr (Ph.D. in 2021, Department of Electrical and Computer Engineering, Princeton University; Dissertation Committee Chair)

5. 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

6. 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

7. Erik Wang (Ph.D. in 2020, Department of Politics, Princeton University). Assistant Professor, Department of Political and Social Change, Australian National University

8. Diana Stanescu (Ph.D. in 2020, Department of Politics, Princeton University). Postdoctoral Fellow, Stanford University

9. Nicole Pashley (Ph.D. in 2020, Department of Statistics, Harvard University). Assistant Professor, Department of Statistics, Rutgers University

10. Asya Magazinnik (Ph.D. in 2020, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, Massachusetts Institute of Technology

11. Max Goplerud (Ph.D. in 2020, Department of Government, Harvard University). Assistant Professor, Department of Political Science, University of Pittsburgh

12. Naoki Egami (Ph.D. in 2020, Department of Politics, Princeton University; Dissertation Committee Chair). Assistant Professor, Department of Political Science, Columbia University

13. Brandon de la Cuesta (Ph.D. in 2019, Department of Politics, Princeton University). Postdoctoral Fellow, Center on Global Poverty and Development, Stanford University

14. Yang-Yang Zhou (Ph.D. in 2019, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, University of British Columbia

15. Winston Chou (Ph.D. in 2019, Department of Politics, Princeton University). Senior Data Scientist at Apple

16. 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

17. Benjamin Fifield (Ph.D. in 2018, Department of Politics, Princeton University; Dissertation Committee Chair). Data Scientist, American Civil Liberties Union

18. Tyler Pratt. (Ph.D. in 2018, Department of Politics, Princeton University). Assistant Professor, Department of Political Science, Yale University

19. Romain Ferrali (Ph.D. in 2018, Department of Politics, Princeton University). Assistant Professor, Aix-Marseille School of Economics

20. Julia Morse (Ph.D. in 2017, Woodrow Wilson School, Princeton University). Assistant Professor, Department of Political Science, University of California, Santa Barbara

21. Yuki Shiraito (Ph.D. in 2017, Department of Politics, Princeton University; Dissertation Committee Chair). Assistant Professor, Department of Political Science, University of Michigan

22. Carlos Velasco Rivera (Ph.D. in 2016, Department of Politics, Princeton University). Research Scientist, Facebook

23. 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

24. Graeme Blair (Ph.D. in 2016, Department of Politics, Princeton University). Assistant Professor, University of California, Los Angeles

25. Jaquilyn R. Waddell Boie (Ph.D. in 2015, Department of Politics, Princeton University). Private consultant

26. Scott Abramson (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, University of Rochester

27. Michael Barber (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, Brigham Young University

28. In Song Kim (Ph.D. in 2014, Department of Politics, Princeton University). Associate Professor, Department of Political Science, Massachusetts Institute of Technology

29. Alex Ruder (Ph.D. in 2014, Department of Politics, Princeton University). Principal Advisor, Federal Reserve Bank of Atlanta

30. Meredith Wilf (Ph.D. in 2014, Department of Politics, Princeton University). Senior Director, Capital Rx

31. Will Bullock. (Ph.D. candidate, Department of Politics, Princeton University). Senior Researcher, Facebook

32. Teppei Yamamoto (Ph.D. in 2011, Department of Politics, Princeton University; Dissertation Committee Chair). Associate Professor, Department of Political Science, Massachusetts Institute of Technology

33. Dustin Tingley (Ph.D. in 2010, Department of Politics, Princeton University). Professor, Department of Government, Harvard University
34. Aaron Strauss (Ph.D. in 2009, Department of Politics, Princeton University). Former Executive Director, Analyst Institute
35. 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
36. 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

- Mmeber, Senior Lecturer Search Committee (2022–2023)
- Member, Curriculum and Educational Policy Committee (2020–2021, 2022–2023)
- 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 Review Committee member, Department of Political Science, University of Rochester (2022)

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.

Expert Reports

1.

2. Graham *et al.* *v.* Adams *et al.* Commonwealth of Kentucky Franklin Circuit Court Division, Case No. 22-CI-00047

3. League of Women Voters of Ohio *et al.* *v.* Frank LaRose *et al.* The Supreme Court of Ohio, Case No. 2022–0303

4. Meryl Neiman, *et al.* *v.* Secretary of State Frank LaRose, *et al.* The Supreme Court of Ohio, Case No. 2022–0298

5. Benninghoff *v.* 2021 Legislative Reapportionment Commission. The Supreme Court of Pennsylvania, Case No. 11 MM 2022

6. The Pennsylvania Legislative Reapportionment Commission, January 2022.

7. The South Carolina State Conference of the NAACP, *et al.* *v.* McMaster, *et al.* United States District Court for the District of South Carolina Columbia Division, Case No. 3-21-cv-03302-JMC-TJH-RMG

8. Milligan *et al.* *v.* Merrill *et al.* United States District Court for the Northern District of Alabama, Case No. 2:2021cv01530

9. League of Women Voters of Ohio *et al.* *v.* Ohio Redistricting Commission *et al.* The Supreme Court of Ohio, Case No. 2021–1193

10. League of Women Voters of Ohio *et al.* *v.* Ohio Redistricting Commission *et al.* The Supreme Court of Ohio, Case No. 2021–1449