The Use of Data Visualization in Political Science*

Frederick J. Boehmke[1], Thomas Dainty[1], Samantha J. DeRagon[1], Hyein Ko[2], Taylor Tokos[1], Tianhui Wu[1], and Weidong Zhang[3]

[1]University of Iowa, USA
[2]Case Western Reserve University, USA
[3]University of Texas at El Paso, USA

October 20, 2025

Abstract

We provide an overview of data visualization use in political science. To trace how visualization practices have evolved, we sampled the first 25 articles from every fifth volume of the *American Political Science Review* (*APSR*), beginning with Volume 1 in 1906. This historical sample captures over a century of trends in the frequency, purpose, and form of visualizations. To examine a broader cross-section of the field, we also analyze over 200 articles from recent issues of eight leading political science journals, including both general and subfield-specific ones. For each article, we coded every included visualization using the same procedure as in the *APSR* sample, assigning each a primary purpose and a specific form based on a typology of about 100 distinct types. The resulting dataset—comparing roughly 1,500 visualizations—allows us to assess both long-term changes and the current state of visualization in political science, and to compare usage patterns across journals and subfields.

Keywords: Data Visualization, Research Methods, Scientific Communication.

*We appreciate comments and suggestions from Walter Mebane, Raymond Duch, Emily Anderson, and Thomas Brailey, and audience members at various presentations. This paper began as a class project constructing our cross-sectional sample. That included work from some students that decided not to participate in this article. We thank Danielle Beavers, Abbie Eastman, Samuel Harper, Chase LaSpisa, Carly Millerd, Eric Ngo, Cassidy Oberreuter, and Stephanie Schmitt for their work. All coding was subsequently updated by and the analysis reported here was performed by the current authors.

Word Counts

This section is *not* included in the word count.

```
File: main.tex
Encoding: utf8
Sum count: 6582
Words in text: 5973
Words in headers: 57
Words outside text (captions, etc.): 551
Number of headers: 14
Number of floats/tables/figures: 18
Number of math inlines: 1
Number of math displayed: 0
Subcounts:
  text+headers+captions (#headers/#floats/#inlines/#displayed)
  162+9+178 (2/0/0/0) Subsection
  352+1+39 (1/0/0/0) Section: Introduction
  274+3+56 (1/1/0/0) Section: Sample of Articles
  115+3+0 (1/0/0/0) Section: Visualization Coding Overview
  119+1+0 (1/0/0/0) Subsection: Training
  600+4+6 (1/2/1/0) Subsection: Coding Process and Details
  856+6+160 (1/6/0/0) Section: The Rise of Visualization over Time
  1038+5+14 (1/3/0/0) Section: The Use of Visualization Today
  174+1+0 (1/0/0/0) Section: Conclusion
  167+7+33 (1/0/0/0) Section: Details on Fields, Purposes, and Post-estimation Details
  2116+5+9 (1/0/0/0) Section: List of Visualization Representations Coded} \label{si:viz-list
  0+6+18 (1/2/0/0) Section: Additional Results from our Time-Series Sample} \label{si:ts
  0+6+38 (1/4/0/0) Section: Additional Results from our Cross-Sectional Sample} \label{si:cs


File: output.bbl
Encoding: utf8
Sum count: 2772
Words in text: 2772
Words in headers: 0
Words outside text (captions, etc.): 0
Number of headers: 0
Number of floats/tables/figures: 0
Number of math inlines: 0
Number of math displayed: 0

Total
Sum count: 9354
Words in text: 8745
Words in headers: 57
Words outside text (captions, etc.): 551
Number of headers: 14
Number of floats/tables/figures: 18
Number of math inlines: 1
Number of math displayed: 0
Files: 2
Subcounts:
  text+headers+captions (#headers/#floats/#inlines/#displayed)
  5973+57+551 (14/18/1/0) File(s) total: main.tex
  2772+0+0 (0/0/0/0) File(s) total: output.bbl

(errors:18)
```

1 Introduction

Visualization has long served as valuable tool for exploring features of and communicating meaning from data in various fields. Well-constructed visualizations help readers understand research results. Visualization techniques, best practices, and aesthetics are critical for and used frequently in political science research. But, beyond a couple of studies (Kastellec and Leoni 2007; Traunmüller 2020), relatively little is known about how visualizations are used in the discipline. How has the use of visualization evolved over time? For what purposes? Which types have become most common?

Political scientists have long recognized the value of data visualization. None other than Edward Tufte, for example, was trained as a Political Scientist. More recently, King, Tomz and Wittenberg (2000) argues that statistical results could be presented in more reader-friendly ways, by reporting quantities such as predicted values, expected values, and first differences. In 2007, Kastellec and Leoni analyzed visualizations in three leading political science journals—the *American Political Science Review (APSR)*, *American Journal of Political Science (AJPS)*, and *Political Analysis (PA)*—from which they identified three key trends: (i) tables far outnumbered figures, (ii) tables primarily summarized data and regression results, and (iii) figures were used for post-estimation tasks, such as illustrating predicted probabilities, with no figures used to present regression results directly. In 2020, Traunmüller studied the *AJPS* from 2003 to 2018 and found a steady increase in the number of figures over time, an increase in the use of figures to interpret model results, and a modest effect of the number of figures on citations to an article.

Building on this earlier work, we provide a substantially more detailed overview of visualization in political science by surveying nearly one thousand journal articles.[1] We trace the usage of data visualization over more than a century and then assess its current state. To do so, we collect a

[1]Our primary focus is to analyze how articles in leading political science journals use graphical displays, rather than to provide suggestions or *best practices* on data visualization. About visualization strategies, see Tufte (1983) and Cleveland (1985) for classics. For more recent publications, see Cairo (2016); Few (2009); Franconeri et al. (2021); Schwabish (2016).

long time series of articles from one general journal and a current snapshot of articles from multiple journals. Our results show that prior to the 1960s, visualizations were used sparingly. By the late twentieth century their use had increased to one per ten pages and now sits at nearly one per three pages. The majority of figures seek to interpret estimates from statistical models.

2 Sample of Articles

To explore long-term trends in visualization use in political science we focused on articles published in the *APSR*. Starting with Volume 1, Issue 1 in 1906, we collected the first twenty-five articles from each volume at five-year intervals. Earlier volumes differed in size and content, which we addressed in three ways. First, if an issue lacked sufficient articles, we moved to the next issue until we had enough. Second, earlier issues often included different content types such as commentary pieces, forums, note from the editors, or presidential addresses. We annotated all types but excluded editor notes and presidential addresses from twenty-five sampled articles. Third, we excluded book reviews or News and Notes pieces, as they are typically non-research content.

Our sample of cross-sectional visualizations come from eight political science journals: the *AJPS*, *APSR*, *International Studies Quarterly* (*ISQ*), *Journal of Conflict Resolution* (*JCR*), *Journal of Peace Research* (*JPR*), *Journal of Politics* (*JOP*), *PA*, and *Public Opinion Quarterly* (*POQ*).[2] This list was created from a class assignment by fourteen students in Spring 2023 focusing on prominent general journals or top subfield journals. We chose sufficient issues to include at least ten articles per journal, drawing from recent years while avoiding special issues.

In total, we have 600 articles in our time-series sample and 242 articles in our cross-sectional sample.[3] We reviewed each article, identified figures, and captured screenshots including titles and notes, if present.[4] Figure 1 shows article counts for our cross-sectional sample by journal issue and

[2] Although *APSR* appears in both datasets, the specific articles do not overlap.

[3] Articles were saved and named journal-volume*XXX*-issue*XX*-article*XX*, with leading zeros (e.g., ajps-vol066-iss04-art01.pdf).

[4] Figures were named journal-volume*XXX*-issue*XX*-article*XX*-pageXXXX-figureXX, with leading zeros (e.g., ajps-vol066-iss04-art01-p0799-fig01.png).

volume, organized by article type. Of the 242 articles, regular articles were most common, while shorter formats, such as letters, are present in journals that include those types of manuscripts.

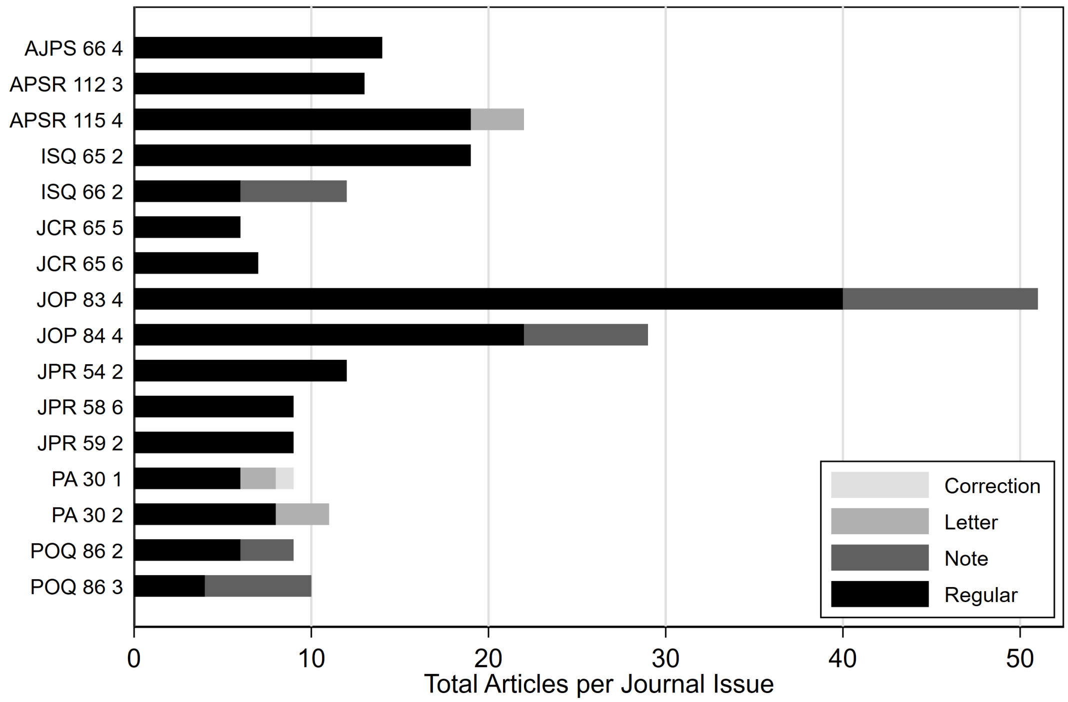

Figure 1: Number of Articles in Sample by Journal Volume and Issue, by Article Type

3 Visualization Coding Overview

For each article we sought to collect all included visualizations for which we would identify two primary items: the form(s) of visualization used and the figure's primary purpose. Our approach to visualization forms was broad. We began with a pre-compiled listing of nearly 100 types drawn from various sources, including Few (2009) and Schwabish (2016). This list includes common forms like bar charts and scatter plots, but also less common ones such as beeswarm or waterfall plots. We expanded this list when new forms were encountered. To capture the purpose of each visualization, we created a nine-item typology covering uses such as theory representation, data description, or post-estimation interpretation. We conducted extensive training and calibration before full coding.

3.1 Training

To prepare for our final coding, which was done by the current team of authors, we first completed extensive training and calibration using a stratified sample of 123 visualizations from the initial cross-sectional sample collected from the original class project. The training sample included up to three examples of each visual representation type to ensure comprehensive coverage. We then reviewed these figures, coding up to four representation types and two purposes per figure. After training, we reconciled differences and created coding rules. We adjusted our list of forms through consolidation and incorporation of additional visual representations and purposes as needed. We then revised the coding scheme and added instructions to help distinguish among the nearly 100 different possible forms.

3.2 Coding Process and Details

The clarifications from training resulted in the final coding procedure for both cross-sectional and time series data. Each coder was assigned figures grouped by journal issue and volume. Then, the coder ensured that all article entries, with or without figures, contained the following basic information: journal name, volume, issue, year, article number (sequentially starting at one for all articles in the issue), article type, article field, article title, author(s), and DOI number. For articles with figures, coders entered page number, figure and subfigure identifiers, and the number of subfigures sharing the same purpose and representation type. All of this information was entered into the dataset (see Supplemental Information, Section A).

For each figure-subfigure entry, coders recorded the representation forms used (see Supplemental Information B for details). Many figures use multiple representations to convey extensive amounts of information. For example, a figure might include a scatterplot along with a best-fit line and a shaded area plot to convey a confidence interval for the best-fit line. In such cases, all representations were labeled; purely referential elements to facilitate interpretation (e.g., a vertical line at $x = 1$) were excluded.

Coders then assigned a purpose to each figure-subfigure type by examining the figure and the text of the article, using the list in Supplemental Information A Table 3. A figure could have multiple purposes. We found that all but a few figures could be assigned a single purpose. The main restriction for specifying the purpose of figures is that any figure coded as “descriptive” should only be labeled as description, as most figures could be considered descriptive to some extent. By limiting the “descriptive” label to instances where it is the primary purpose, we prevent coders from overusing this category in coding process. Figures coded with post-estimation as their purpose had an additional column of post-estimation details, specifying the type of post-estimation the figure represents.

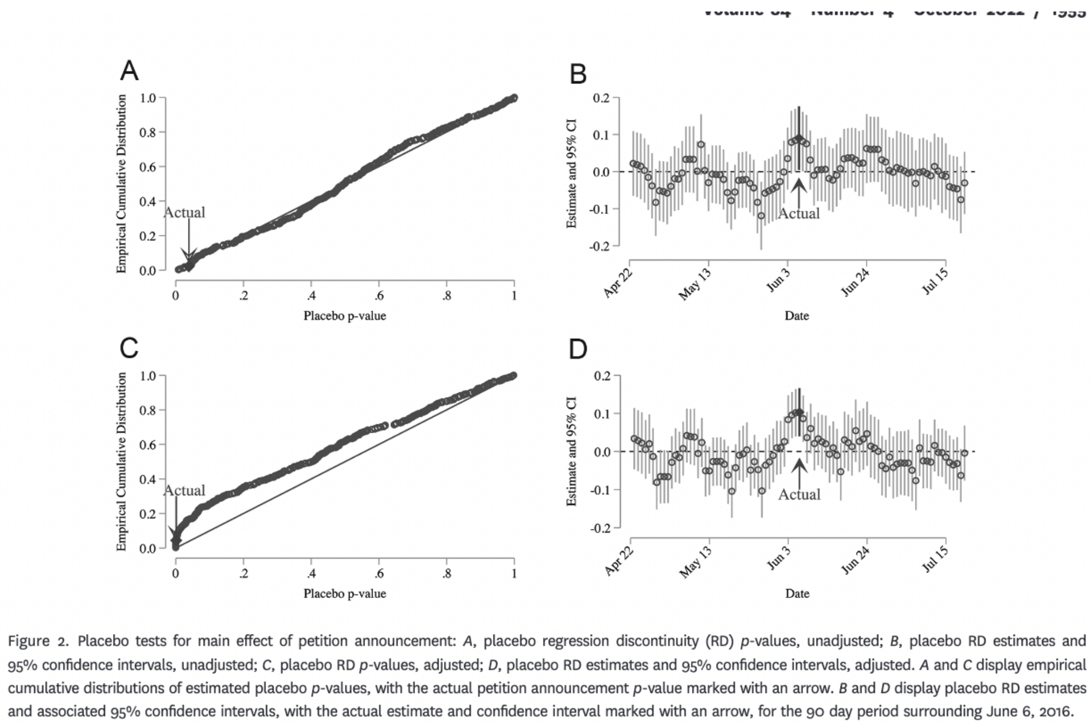

Figure 2: Sample Figure from Gordon and Yntiso (2022)

Examples illustrate our coding. Figure 2 from Gordon and Yntiso (2022) is one of the more complicated cases, with multiple representations types and repeated subfigure formats. Subfigures A and C appear to be line graphs. However, the darker lines are actually a series of scattered points. Because the points are the primary element, these subfigures were coded as scatterplots; the line, included only for reference, was not coded as a separate representation. Subfigures B and D are scatterplots with spike plots. The primary purpose of these visualizations is to show the point estimates, making the points themselves the main visual element. Confidence intervals are presented as thin lines whose heights represent values; thus, these figures were coded as spike plots as the second visual representation. The horizontal line in subfigures B and D was excluded from the final coding because it serves only as a reference.

All four subfigures in figure 2 were coded as post-estimation, specifically "Predicted Values/Probabilities" because A and C present p-values, while B and D show predicted values. Since subfigures A and C share the same visual representation and purpose, and B and D do the same, we entered information of figure 2 into two rows, i.e., two figure-subfigure types. We then assigned a value of two of the column representing the number of subfigures that shared the same visual purpose and representation type. Even for articles without explicit subfigure identifications, we applied this coding rule consistently.

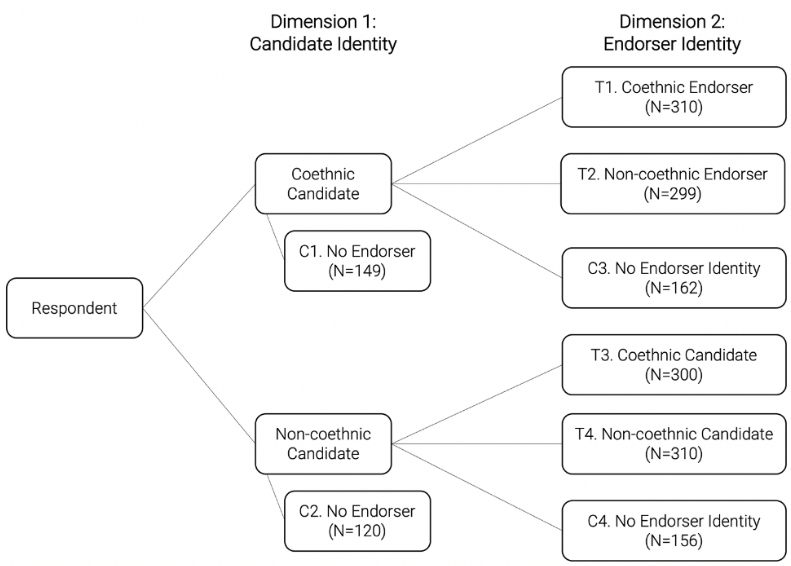

Figure 1. Experimental design: subjects assigned to main treatment and control conditions

Figure 3: Sample Figure from Arriola, Choi and Gichohi (2022)

Not all figures in our dataset were data-driven. Figure 3, for example, depicts the group assignment of each respondent in an experimental design (Arriola, Choi and Gichohi 2022). This figure shows the paths of subject assignments, coded as a flowchart with the purpose of research design. This example illustrates how visualization can simplify descriptions of a research design that may be over complicated in text.

4 The Rise of Visualization over Time

We first demonstrate the growth of data visualization over time in the *APSR* using 600 articles spanning 115 years. When the *APSR* began in 1906, the art and practice of data visualization was relatively mature: the late nineteenth century was the “Golden Age” of statistical graphics (Friendly and Wainer 2021). Yet this period was also coming to a close as the rise of quantitative statistical analysis led to the “Modern Dark Ages” of data visualization (Friendly and Wainer 2021).

Visualizations were not common early on in the *APSR*, with the first appearing in 1926. We identified the earliest table and figure published in *APSR* (see figure 4).[5] The first table appeared in Story (1909), presenting a cross-tabulation of occupational categories by sex from the Philippine census. In 1918, the first figure was published in Sauer (1918)–a map of Missouri depicting racial demographics and congressional districts. While Sauer (1918) included several visualizations, they were all maps intended to support spatial arguments. These visualizations were primarily descriptive, indicating a potentially limited role of visualization in the journal’s early decades.

TABLE I.

This table shows the occupations of the Chinese in the Philippines as classified in the Census of 1905.

	Male.	Female.	Total.	Per cent of total number in same class.
Agriculture	601	1	602	0.048
Professional work	86	1	87	0.34
Domestic and personal	9,696	107	9,803	1.70
Trade and transportation	23,330	34	23,364	10.26
Manufacturing and mechanical	6,670	40	6,710	0.69
Unknown or unproductive	688	843	1,531	0.04

Cf. Philippine Census, vol. ii, pp. 894–5.

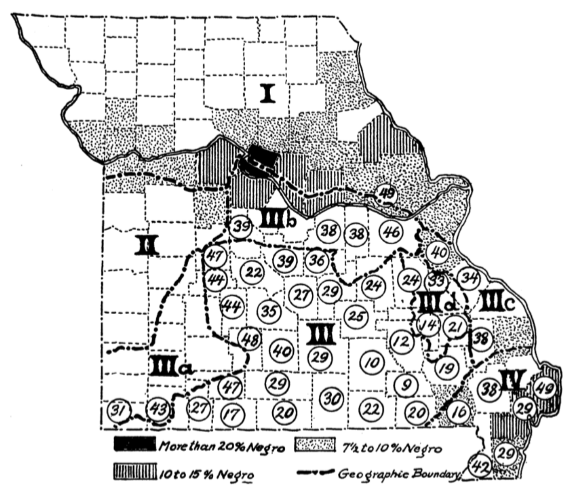

FIG. 1. GEOGRAPHIC DIVISIONS OF MISSOURI

Figure 4: First Table and Figure Published in the APSR

[5] To identify the first instances of tables and figures, we considered only those with explicit labels such as ’Table 1’, or ’Figure 2’ by the author(s).

Figure 5 shows the average number of figures (blue line) and tables (purple line) per page from 1906 to 2021. Both were rare before the 1960s, indicating little use of data visualization or statistical analysis (at least as reported in tables). The use of tables surges first in the 1960s, with a huge jump from one table for every forty pages in 1956 to one table every three pages by 1966. This spike ends in 1986 with a drop to one table every six pages, after which it remains steady through 2021. The use of figures shows a steadier increase beginning in 1961, with about one figure for every ten pages for the next fifty years before jumping to almost one figure for every five pages in 2011 and then to one for every three pages in 2021. These trends mirror the broader resurgence of visualization with the rise of exploratory data analysis (Tukey 1977) and computer-generated graphics (Friendly 2008). Comparing the trends in figures and tables suggests that the use of figures and tables both accelerates in the middle of the twentieth century, likely corresponding with an increase in the use of quantitative methods. Furthermore, the drop in tables in the 1980s does not signal a shift to using figures in their place — use of figures begins an upward trend around that time, but starts slowly and takes a quarter century to pick up steam.

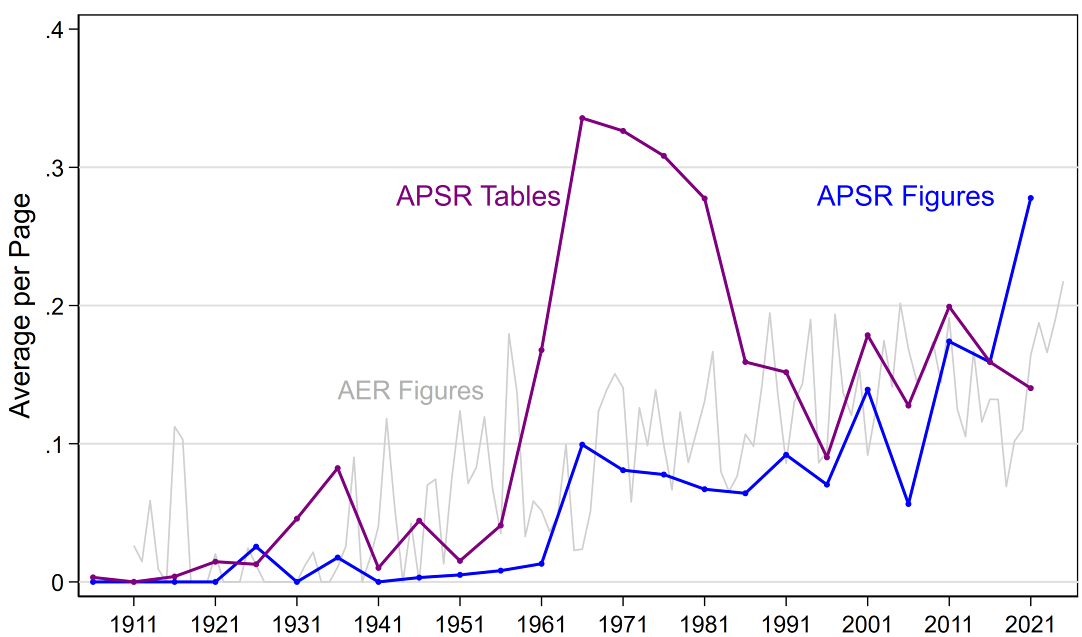

Figure 5: Total Number of Figures and Tables per Year

Note: Data on AER figures through 2017 from Schwabish (2022); extended by authors through 2023.

We compare the use of visualization in political science to other fields. Figure 5 depicts the number of figures per page in articles appearing in the first issue of every volume of the *American Economic Review* from 1911-2017 as collected by Schwabish (2022) with a light gray line.[6] The *AER* data show an earlier initial rise in visualization among economists around 1940. Political science matches this jump two decades later after which it uses visualizations at a slightly lower rate until the increase in 2016 pushes it into the lead.

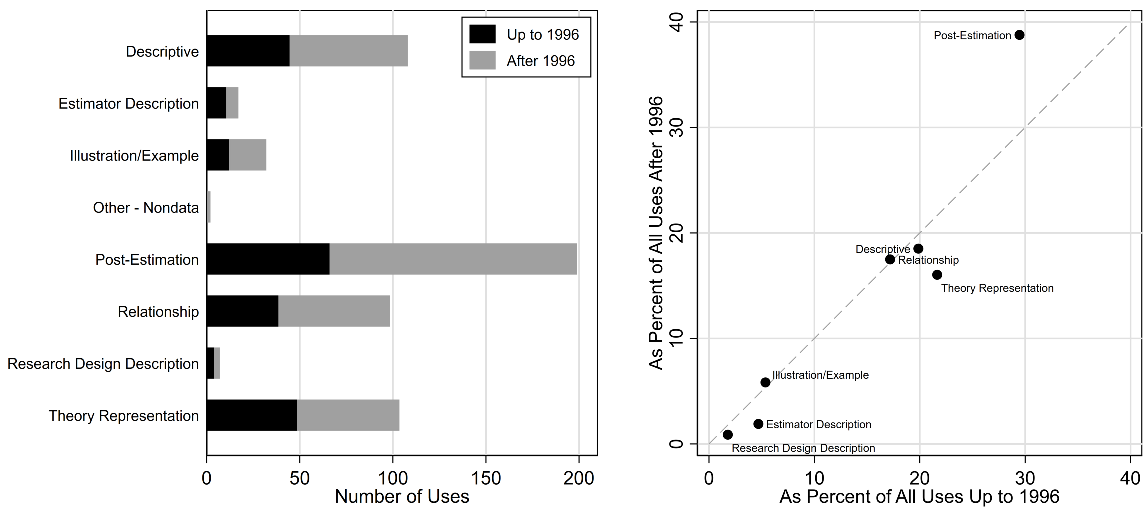

Figure 6: Figure Purposes Stay Consistent Over Time

Note: Purposes coded at the subfigure level and inversely weighted by number of purposes per subfigure.

The purpose of political science visualizations has changed little over time. Splitting the data in 1996, which corresponds to the recent rise in visualization, nearly halves the number of figures (40% of figures appear before 1996). Figure 6 shows the frequency of purposes in both time periods.[7] The stacked bar chart on the left shows the count of purposes per figure while the scatter plot

[6]We downloaded the associated data from the author's website (https://policyviz.com/2018/08/22/working-paper-graphs-in-the-american-economic-review/ on June 13, 2025. We processed the data to enhance comparability to our data, for example by excluding book reviews.

[7]This analysis uses the data at the subfigure-type level and weights purposes/representations by the total number that occur in each subfigure. For example, a figure with two subfigures that contain a scatterplot containing a rug plot and a separate line plot would have two representations in the first figure and one in the second. Each therefore has a weight of one-third.

on the right shows the use of each within each time period as a proportion to account for the fact that the latter time period includes more total figures. Most purposes fall close to the dashed line representing equal relative use. The top four uses—descriptive, relationship, theory representation, and post-estimation—remain consistent in both eras, with only theory representation moving by more than one spot. In both periods, the most common use is to present results after estimation, which has grown from thirty to forty percent, while figures representing theories or describing estimators have declined.

The increase in the use of visualization since 1996 also corresponds to a substantial increase in the number of distinct visual representation forms used. Figure 7 shows that traditional forms, such as line plots, scatterplots, and flowcharts were used regularly in both periods. Several graphical types have shown a substantial increase, including dot plots, bar charts, maps of various forms, and spike plots. The latter, now widely used to depict confidence intervals in post-estimation figures, illustrates how new forms have accompanied shifts in analytical practice.

The widening variety of visualization forms parallels other signs of growing sophistication. Figures now carry more information, as shown in figure 8, which tracks the average number of subfigures per figure and representations per subfigure. Once the use of figures becomes common in the 1960s, the number of subfigures bounces around fairly consistently between 1.25 and 1.75 until 2021, when it jumps from its previous peak of 1.8 in 2016 to nearly 2.8. The number of representations increases more steadily starting in 1981 from a baseline near one to 1.2 in the 1990s and 2000s before passing 1.4 in 2016 on its way to 1.75 in 2021. This shifts show that as visualization has expanded, political scientists have reported more information in their figures through multiple subfigures and more distinct representation forms.

5 The Use of Visualization Today

Our cross-sectional sample consists of 242 articles containing 888 figures, which yields 1,731 subfigures and 958 distinct figure-subfigure types (i.e., unique purpose/representation combinations). On average, articles feature 3.9 figures and 7.5 subfigures. These numbers align with Traunmüller (2020), who reported 3.6 figures per article in 2018 from his examination of the *AJPS*

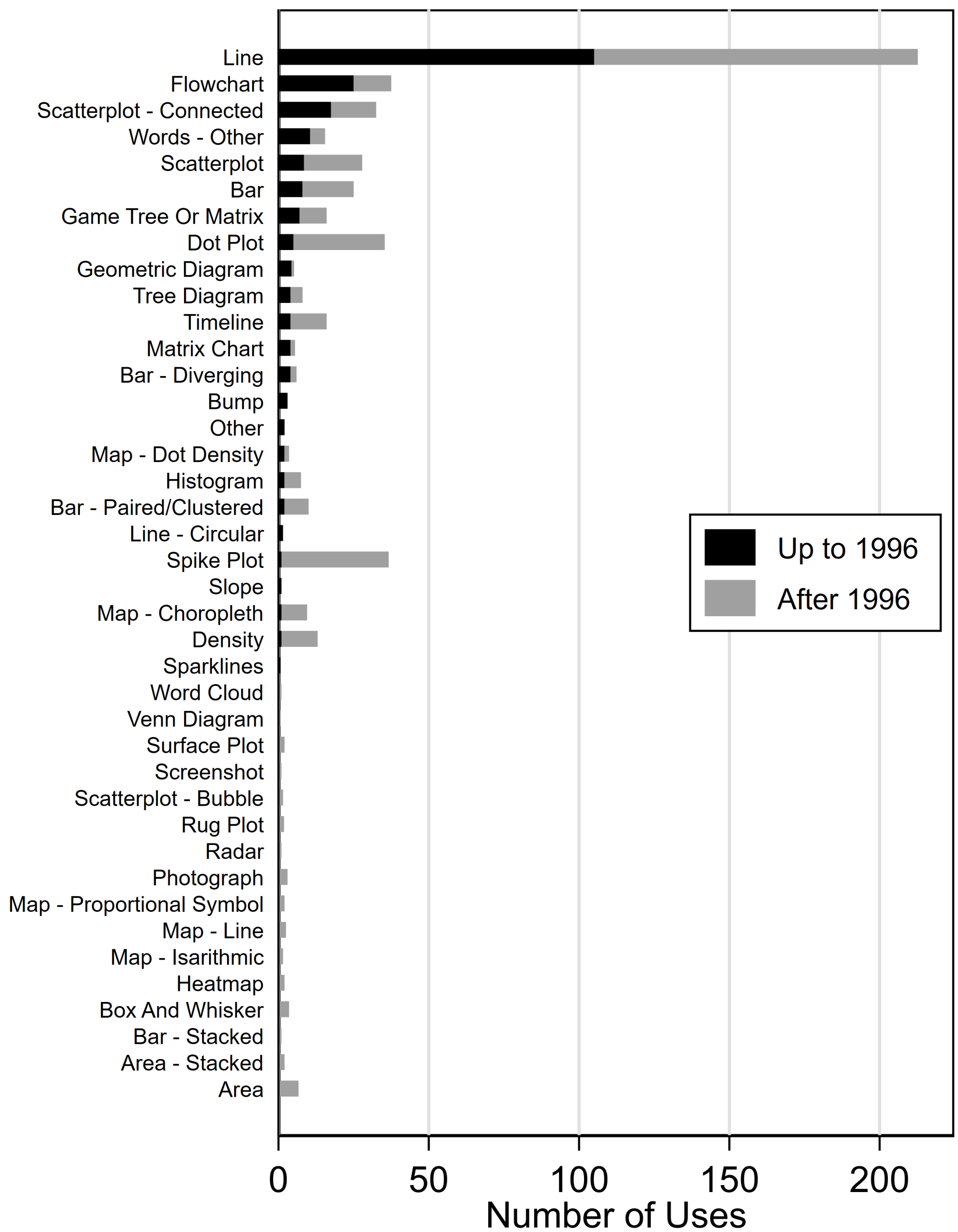

Figure 7: Representation Types Diversify Over Time

Note: Representations coded at the subfigure level and inversely weighted by the number of representations contained in all subfigures for each figure.

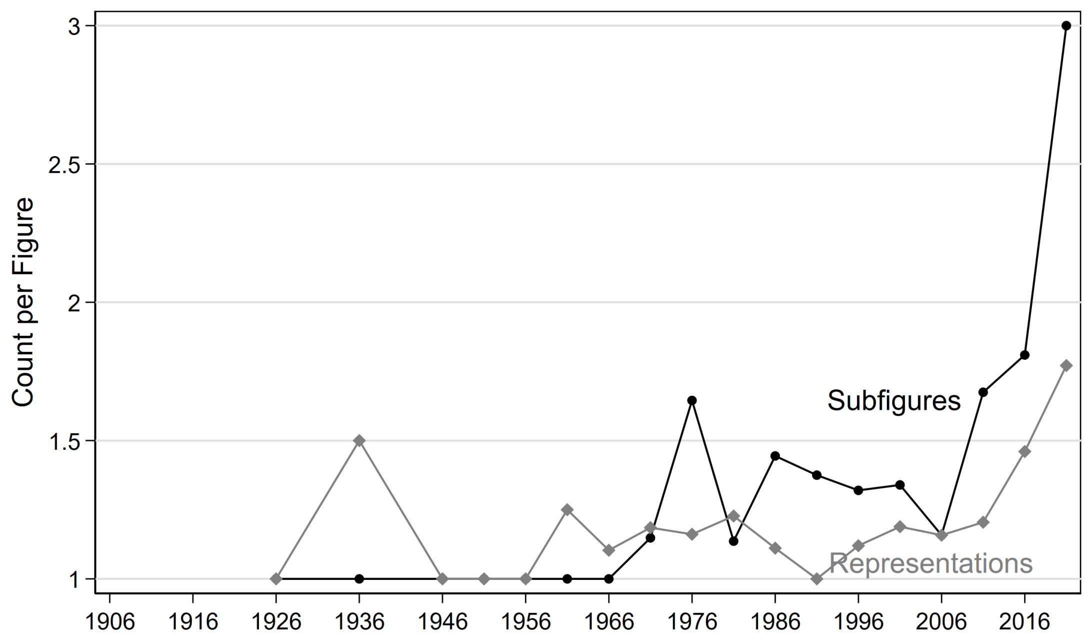

Figure 8: Figures Became more Complex Over Time

Note: Number of representations per figure counts the number of distinct representation forms that appear within a single figure (if one subfigure) or that appear in any of its subfigures (if more than one).

between February 2003 and March 2018 and appears to continue the trend he identified from just one figure per article in 2003. Kastellec and Leoni (2007) found an average of 1.8 figures per article in a sample of 52 articles in 2006. In our data, figures with a single subfigure are somewhat more common (63%) than in Traunmüller (2020) (54%), though slightly less so in our articles from the *AJPS* (61%).

We also analyze the purpose of figures and subfigures using our nine-category coding scheme. We report results at the figure level by considering all purposes assigned to the different subfigure types within each figure. In almost all cases, each figure-subfigure type was coded with a single purpose, with only 12 out of 958 figure-subfigures coded with more than one. In describing these results, we do not weight by the frequency of each subfigure type appearing within a figure. Instead, we weight each purpose by the number of purposes assigned to a given subfigure-figure type. Thus, if a figure was coded with two purposes, each purpose was assigned half the weight of a figure coded with only one.

Figure 9 illustrates that interpreting estimated results is the dominant purpose. Nearly sixty percent of all cases fall into this category, suggesting that King, Tomz and Wittenberg's (2000) advice has been widely adopted. About four out of five articles included at least one post-estimation figure, and nearly three out of five featured more than one. This appears to continue the growth seen in our time series analysis, in which post-estimation visualizations grew from thirty percent before 1996 to nearly forty percent afterwards. The second most common purpose is descriptive visualization, which accounts for about twenty percent. Relationship-oriented visualizations represent just under ten percent. No other category exceeds five percent. These findings demonstrate a substantial growth in the use of figures to interpret estimation results over the two decades since Kastellec and Leoni (2007) and at a higher rate than reported by Traunmüller (2020) for the period 2003–2018.

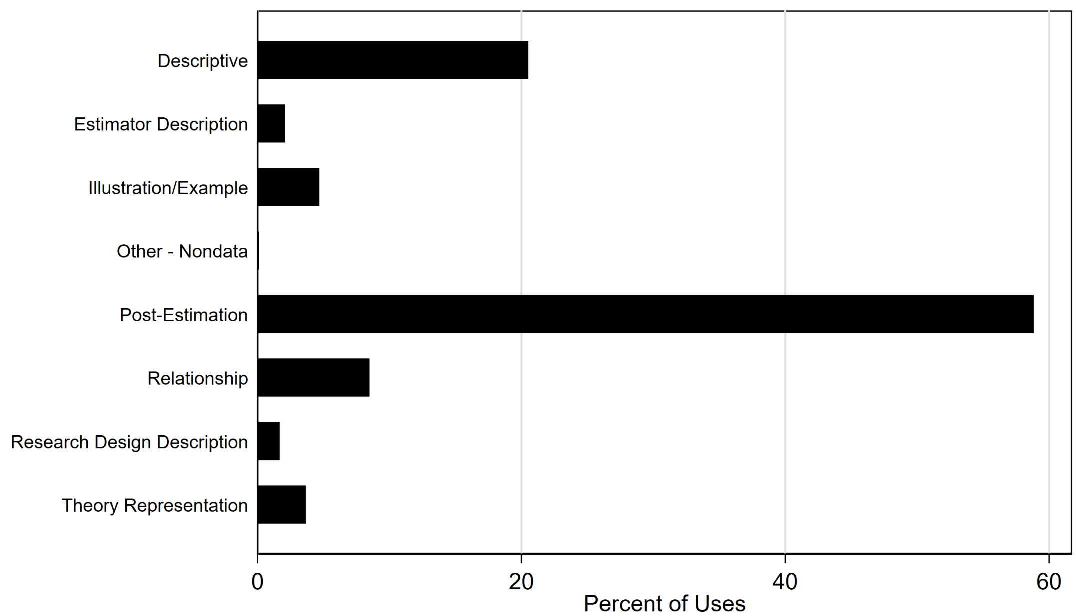

Figure 9: Usage Rates of Visualization Purposes

Note: Purposes are weighted by the total number of purposes per figure.

We also examine the visual representation forms used to achieve these purposes. Most figures use at least two forms: over 50% of the figures combine multiple representation types. This is likely driven by the frequent usage of visualizations for post-estimation, where multiple forms are used to present the results. Over 40% of the figures use only a single representation to deliver the researcher's purpose.

A few particular types stand out. The most frequently chosen plot types are line graphs, spike plots, and dot plots. As shown in figure 10, these three are the only figure types with more than 50 instances of usages. This finding makes sense because these three forms are typical elements of displaying post-estimation results, such as predicted probabilities or marginal effects. For instance, spikes are frequently used to display confidence intervals alongside dots that represent the most likely values, while lines may be utilized to connect values across the x-axis.

The next most frequent forms are histograms, bar charts, scatterplots, and area chart, each with nearly around 50 usages as shown in figure 10. They are often utilized as descriptive tools, either to showcase the distribution or frequency of data or tools to highlight relationships between variables. After accounting for the most frequent used purposes, our analysis shows that political scientists have also used quite diverse visualization forms, including choropleth maps, waffle plots, and tree diagrams. However, some visualization types are too rare to be commonly used. In fact, almost half of our representation categories did not appear even once in the sampled articles. We did not find many exploding cartograms, Gantt charts, Nightingale, or Merimekko plots, but the large number of unused forms suggests that political scientists could expand their visualization repertoire by considering alternative representation forms to depict the same data and information.

Figure 11 connects the types of representations to the motivations for which political scientists create visualizations, showing the percentage of each representation type for each purpose. To keep the number of representations manageable, we report percentages only for those that appeared among the top five representations for any single purpose, which captures 19 total, and then grouped all others into a residual category call "Not in Any Top 5".

For theory representation, the most frequently used forms are flowcharts (27%) and game trees/matrices (22%). Research design descriptions often use screenshots (35%), e.g., to showcase prompts or visuals that respondents may see in an online survey, but also make use of flowcharts (18%), tree diagrams (12%), or timelines (12%). To visualize relationships, articles largely rely on line charts and scatterplots. Together, nearly half of the representations for this purpose are based

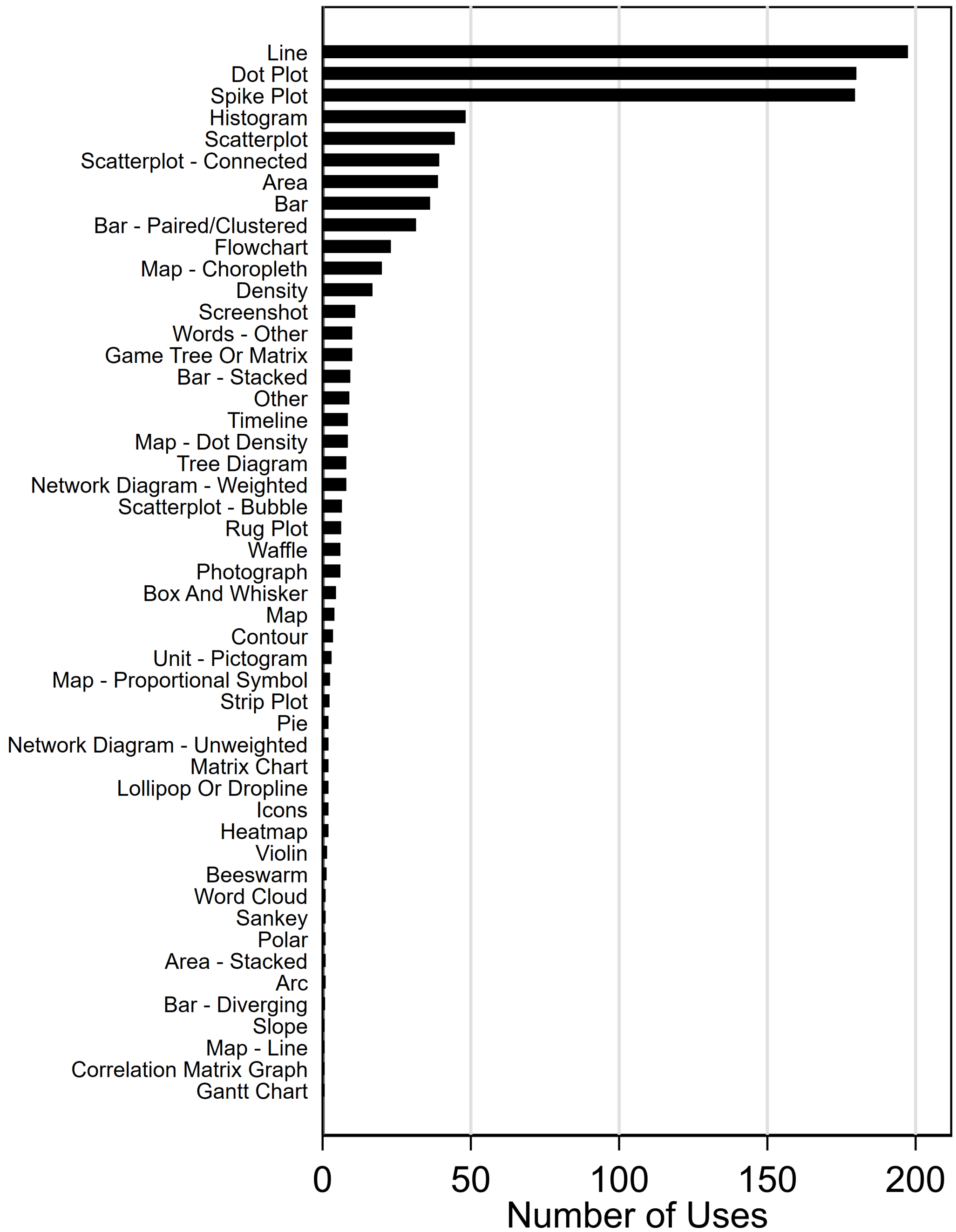

Figure 10: Frequency of Visualization Representations

Note: Representations are weighted by the total number of representations per figure.

on these two forms. Post-estimation representations are largely comprised of line charts (20%), spike plots (28%), and dot plots (27%), which are often used together in a single visualization as discussed earlier.

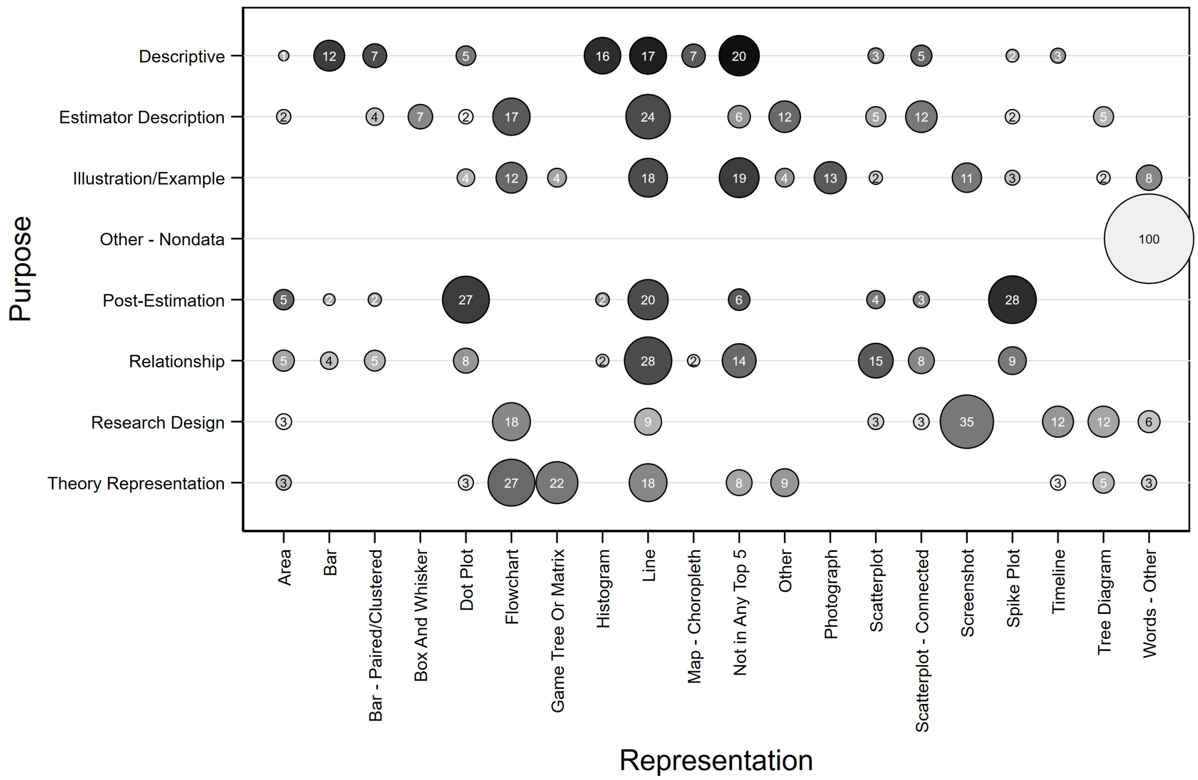

Figure 11: Different Purposes Use Different Representations

Note: Plot shows the conditional percentage of representation types for each purpose. Markers are sized by the conditional percentage, which is also reported by the text in the markers. Markers are shaded according the unconditional percentage for the purpose-representation combination. Any representation in the top five for any purpose is included, all other representations outside the top five for all purposes are combined into the "Not in Any Top 5" Category. Representations are weighted by the total number of representations per figure.

Visualizations for illustration or examples make use of a wide variety of representation types. Words (8%), screenshots (11%), photographs (13%), line graphs (18%), or flowcharts (12%) are frequently used, while a considerable portion of figures (18%) lie outside of the top five representations for any purpose. Besides "other - nondata" category, estimator description is the purpose with the highest proportion of figures classified as "other" with 12% of its representations relying on miscellaneous forms not captured by our main categories. Otherwise, estimator description is most often showcased through line graphs (24%) or flow charts (17%).

Lastly, representations for descriptive purposes make the most use of visualization types outside of the top five for any purpose, with 20% of descriptive figures using representations not listed in figure 11. This may be driven by the use of forms such as choropleth maps, dot density maps, or density plots. They can be useful for descriptive purposes and are not infrequently used (as shown in figure 10), but they still do not appear among the top five representations for any purpose. Apart from these, descriptive visualizations in our sample frequently use line charts (17%), histograms (16%), and some form of bar chart (12%).

6 Conclusion

Our survey of data visualization in both temporal and cross-sectional contexts shows that political scientists and academic journals increasingly recognize and use the benefits of visualization. Over time, our time-series survey of *APSR* articles shows that both the number of visualizations and the diversity of their forms have been increased. Recently, visualizations are often utilized for the same purposes across fields. Outside of Political Theory, post-estimation visualizations are extremely common, followed by descriptive and relationship purposes. Their representations are also concentrated, with line graphs, spike plots, and dot plots dominating — likely driven by the frequent use of post-estimation techniques.

While the use of visualization has been increased over time and representations vary by purpose, the dominance of post-estimation suggests opportunities for future visualization work: using visualizations to show their theoretical argument, describe their models, or present data dynamics. Visualizations can communicate more effectively than text alone, whether by clarifying statistical (in)significance, demonstrating substantive findings, or making a clear delivery of a theoretical argument and the construction of samples from a population.

References

Arriola, Leonardo R, Donghyun Danny Choi and Matthew K Gichohi. 2022. “Increasing intergroup trust: Endorsements and voting in divided societies.” *The Journal of Politics* 84 (4):2107–2122.

Cairo, Alberto. 2016. *The truthful art: Data, charts, and maps for communication.* New Riders.

Cleveland, William S. 1985. *The elements of graphing data.* Murray Hill, NJ: Bell Telephone Laboratories, Inc.

Few, Stephen. 2009. *Now You See It: Simple Visualization Techniques for Quantitative Analysis.* Oakland, CA, USA: Analytics Press.

Franconeri, Steven L., et al. 2021. “The science of visual data communication: What works.” *Psychological Science in the public interest* 22 (3):110–161.

Friendly, Michael. 2008. “A Brief History of Data Visualization.” In. *Handbook of Data Visualization*, 15–56. Berlin, Heidelberg: Springer.

Friendly, Michael, and Howard Wainer. 2021. *A history of data visualization and graphic communication.* Cambridge, Massachusetts: Harvard University Press.

Gordon, Sanford C., and Sidak Yntiso. 2022. “Incentive Effects of Recall Elections: Evidence from Criminal Sentencing in California Courts.” *The Journal of Politics* 84 (4):1947–1962.

Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. “Using Graphs Instead of Tables in Political Science.” *Perspectives on Politics* 5 (4):755–771.

King, Gary, Michael Tomz and Jason Wittenberg. 2000. “Making the most of statistical analyses: Improving interpretation and presentation.” *American journal of political science*:347–361.

Sauer, Carl Ortwin. 1918. “Geography and the Gerrymander.” *American Political Science Review* 12 (3):403–426.

Schwabish, Jonathan. 2016. *Better Presentations: A Guide for Scholars, Researchers, and Wonks.* Columbia University Press.

Schwabish, Jonathan. 2022. “Categorizing and Ranking Graphs in Economics Research: The American Economic Review over the Last Century.” *Journal of Benefit-Cost Analysis* 13 (1):20–33.

Story, Russell M’Culloch. 1909. “The Problem of the Chinese in the Philippines.” *American Political Science Review* 3 (1):30–48.

Traunmüller, Richard. 2020. “Visualizing Data in Political Science.” In. *The SAGE Handbook of Research Methods in Political Science and International Relations*, 436–458. SAGE Publications Ltd.

Tufte, Edward R. 1983. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

Tukey, John W. 1977. *Exploratory Data Analysis.* Reading, MA: Addison-Wesley.

Supplemental Information

A Details on Fields, Purposes, and Post-estimation Details

Table 1: List of values for coding article features (Cross Sectional Data)

Article Type	**Article Field**
Regular	American
Note	Comparative
Forum	Formal Theory
Replication	International Relations
Correction	Methodology
Letter	Methods
	Other
	Pedagogy
	Political Theory

Table 2: List of values for coding article features (Time Series Data)

Article Type	APSR Article Type	Article Field
Comment	American Government and Politics	American
Controversies	Foreign Governments and Politics	Comparative
Correction	Instruction and Research	Formal Theory
Forum	International Affairs	International Relations
Letter	Legislative Notes and Reviews	Methodology
Note	Notes and Memoranda	Other
Notes from the Editor	Notes on Administration	Pedagogy
Presidential Address	Notes on Current Legislation	Political Theory
Regular	Notes on Current Legislation	
Replication	Notes on Municipal Affairs	
Report	Notes on Rural Local Government	
Response	Other	
	Political Science and Political Power	
	Politics and Ethics–A Symposium	
	Public Administration	
	Report	
	Report of the Third National Conference on the Science of Politics	
	Reports of Round Table Conference	
	Research Article	
	Rural Local Government	
	The European Scene	

Table 3: List of values for coding figure purposes (Cross Sectional and Time Series Data)

Purpose	Post-estimation Details
Descriptive	Coefficient plot
Estimator description	Distribution
Illustration/example	First differences
Other - data	Marginal effects
Other - nondata	Model Fit/Evaluation
Post-estimation	Other
Relationship	Predicted values/probabilities
Research design description	
Theory representation	

B List of Visualization Representations Coded

Table 4: List of Representation Types Coded with Coding Details/Instructions

Representation	Details
Arc	displaying inter-relationships between data arrayed in a single horizontal axis; observations are located in a single line and connected by arcs
Area	line graphs with the filled/shaded areas between lines and a baseline, could be multiple variables if not stacked; if the figure is clearly another form (e.g., density with shading), it should be considered as the form (e.g., density not area).
Area - Stacked	several area series stacked on top of each other - must be stacking, not overlapping areas charts.
Bar	bars with different heights or lengths to represent values for some set of categories; could be vertical or horizontal
Bar - Circular/Spiral	bars with a spiral format – starts at the center of a spiral and progresses around the circle; typically for time-based data
Bar - Diverging	each bar can go in opposite directions from a central line; not used for bar charts with negative values (i.e., each bar can only go in positive or negative directions but not both)
Bar - Paired/Clustered	bars grouped together to show two or more values for each category
Bar - Radial	bars plotted on a polar coordinate system, i.e., bars radiate outwards from center circle
Bar - Stacked	bars stacked with different component values for each category; must be stacked rather than overlapping.
Beeswarm	shows individual data points (typically as dots) without overlaps
Box and whisker	shows the summary of distribution with a box corresponding to the first quartile and third quartile, a horizontal line inside the box corresponding to the median, whiskers from the box to show the higher and lower extremes, and outliers
Bubble - nested	smaller bubbles into larger bubbles to show two or more levels of hierarchy or relative sizes of a variable
Bubble - nonnested	Bubbles used to show relative size of a variable, but without x/y axes - relative location of bubbles does not matter.
Bullet	bars (or lines) with three layers – observed value, target value, and background range
Bump	lines which present the rank order and changes in ranks over time
Candlestick	bars with lines extending upward and downward from the bars – each representing a specific time interval (e.g., an hour or a year)
Cartogram - Contiguous	adjusting the size of geographic unit according to the values – with approximate geographic locations and distorted shapes but still connected

Continued on next page

Table 4 – *Continued from previous page*

Representation	Details
Cartogram - Graphical	adjusting the size of geographic unit according to the values – with other shapes; includes, for Dorling (circles) and DeMers (squares)
Cartogram - Gridded	using cells (typically hexagons or squares) for each geographical unit to approximate major geographical shapes – with colors or shapes to indicate magnitues of each unit
Cartogram - Nonarea	distorting physical geography to present relative time and distance, e.g., as in a subway map that does not respect relative distances or locations
Cartogram - Noncontiguous	adjusting the size of geographic unit according to the values – with units separated and likely distorting actual geographic locations
Chord Diagram	displaying inter-relationships between data arrayed in a circle; observations are located around the circle and connected by arcs within the circle
Contour	shows a third variable as contours; each contour line (or level curve) represents the common third-values
Correlation matrix graph	a graph of correlation matrix that shows the strength of correlation with shapes, colors, or shades
Cycle chart	shows a series of repeating actions or processes with a circular format
Dendrogram	shows a hierarchical clustering where individual items/objects are connected at different levels depending on their similarity; more similarity leads to make two items meet sooner in the deodrogram; looks similar to a tree diagram but it shows the connection from bottom (individual) to top (upper group/cluster).
Density	uses a line (typically smooth curve) to show the distribution of data
Dot Plot	uses symbols (typically dots) to show each data point, typically over a nominal variable
Fan	lines with colors and/or color saturations of the shaded area to show changes in uncertainty over time, see Gradient for similar plot with bars.
Flowchart	shows the paths of a process or sequence; doesn't necessarily need to have clear "roots"; could go back and forth in more than one direction; could skip levels (typically with arrows)
Game Tree or Matrix	A game tree (extensive form) shows every possible state of the game as roots/nodes with branches/edges connected from one stage to the next stage(s); A game matrix (normal form) shows actions as rows/columns. These can also include subgames of each type or when the same structure is used to show part of the game, e.g., payoffs, probabilities, etc.

Continued on next page

Table 4 – *Continued from previous page*

Representation	**Details**
Gantt Chart	horizontal bars or lines to show the duration of different values or actions or events
Gauge	typically has a something between half-circle to circle form with a pointer to indicate the data point within a particular range
Geometric Diagram	a visual representation of geometric concepts and shapes. These diagrams can range from simple shapes like circles and triangles to more complex figures.
Gradient	bars where each bar plots a primary number of importance and add a color gradient on one or both sides to show distributions or differences in uncertainty. Don't use in place of Contour plot. See fan for similar plot with lines.
Heatmap	uses colors and/or color saturations to show magnitudes of individual values within a dataset; typically looks like a table with colored cells
Histogram	uses a series of bars to show the distribution of data
Horizon Chart	a compact area chart equally sliced hotizontally and collapsed into single band; typically looks like a combination of area and heatmap
Icons	a symbol or graphic representation to deliver simplified but communicative information; Use Unit - Pictogram or Unit - Isotype if conveying data values.
Instance Chart	individual moments or instances of categorical "activities", typically over time
Line	values are connected by line(s) to show values over a continuous variable
Line - Circular	a line chart wrapped around a circle
Lollipop or dropline	lines with dots (or other shapes) at the end
Map	a map that presents a geospatial representation of reality or hypothetical situation
Map - Choropleth	a map with colors or patterns on geographic units to present quantities or magnitudes of the values of that unit
Map - Dot Density	a map with dots other symbols to indicate the presence of data values
Map - Flow	a type of thematic map that uses linear symbols to represent movement between locations.
Map - Grid	a map with values overlaid graphically in a regular grid, e.g., as squared 10km by 10km colored to show values.
Map - Isarithmic	a map with connected lines and areas (usually with different colors) with the same values to depict smooth and continous phenomena; often do not follow physical geographical units (e.g., states).
Map - Line	a map with lines to show geographical items (e.g., river or roads)

Continued on next page

Table 4 – *Continued from previous page*

Representation	Details
Map - Prism	the height of the geography is raised according to a variable
Map - Proportional Symbol	uses symbols on a map (e.g., dots, icons) that vary in size to represent a quantitative variable
Marimekko	bars with scaled widths corresponding to other variables in the graph, widths may add up to 100%, but if both widths and heights do, see also Marimekko.
Matrix Chart	a grid format which shows relationships between two or more variables
Mosaic	a special type of Marimekko where bars with both heights and widths add up to 100 percent; but doesn't necessarily need to show a hierarchical relationship like treemap
Network Diagram - Unweighted	no weights – no difference in node size and width of edges
Network Diagram - Weighted	weights can be indicated by size of nodes and/or width of edges
Nightingale	a pie chart with slices (typically the sequence follows a clockwise direction), where each slice extends in a different direction; shows both changes over time and part-to-whole relationship
Other	placeholder for miscellaneous representations
Parallel coordinates	shows high dimensional data by using multiple vertical axes and connecting by lines
Photograph	an image taken with a camera, capturing visible items in the real world.
Pie	shows data in a circular form, where each slice of the circle represents relative size of the data. Use this for donut chart, too.
Polar	Data points drawn on separate axes that radiate from the center. Use this if axes are interval-valued. See Radar chart if they are Nominal. Also known as spider or star chart.
Pyramid	a special type of diverging bar chart that is reserved for comparing distributions
Radar	Data points drawn on separate axes that radiate from the center. Use this if axes are interval-valued. See Poloar chart if the axes are interval-valued.
Raincloud	shows both the distribution of density of data and the actual data points; actual data points are placed directly below each density graph, making actual data points (density graphs) look like raindrops (clouds).
Ridgeline	a series of overlapped density plots or histograms for different groups presented in the same horizontal axis to show overall shapes of distribution among different groups
Rug plot	shows a frequency distribution with equal-height tick marks on the axis

Continued on next page

Table 4 – *Continued from previous page*

Representation	Details
Sankey	shows how categories (and their values) compare one another and flow into other categories over each transition/stage; arrows or lines to show the transition from one stage to another, and the width represent the magnitude of each transition
Scatterplot	uses dots to represent correlation (or lack of correlation) between two variables
Scatterplot - Bubble	a sactterplot with varying sizes of the circles according to a third variable
Scatterplot - Connected	a scatterplot with connected lines to show relationships of the points over stages (e.g., time or condition)
Screenshot	an image of the visible items displayed on the screen (e.g., computer, mobile, tablet).
Slope	plots each data point on a separate vertical axis and connecting the two with a line, showing changes in values between two stages (e.g., time or condition)
Sparklines	a small and simple line chart typically without coordinates; it shows general patterns and trends rather than specific values
Spike plot	a vertical line from the plot point to the spike base (often the x axis, but it can be other values such as the mean)
Stem and leaf	a table with stem and leaf columns, where the stem column contains the first or shared digit(s) and the leaf shows the last or unshared digit(s); typically with numbers but possible other values
Streamgraph	a variation of stacked area, where areas are depicted around a central horizontal axis (which doesn't have to be zero); shows flucuations in data over time
Strip Plot	each data point (typically with dots or circles but potentially with other symbols) is displayed along a single vertical or horizontal axis
Sunburst	two or more rings to present upper/major components and lower/sub components; inner rings indicate upper/major hierarchy
Surface plot	a graphical representation of a function of two variables, typically represented as a three-dimensional surface
Timeline	shows when certain events take place in chronological order
Tree Diagram	shows a level of hierarchy in a system or a group; has clear "roots"; could be vertical or horizontal but with single direction
Treemap	typically squares or rectangles are used to divide categories and sub-categories
Unit - Isotype	Use of images or icons, typically with sizing or shading used to represent counts or values. See Pictogram for simpler versions.

Continued on next page

Table 4 – *Continued from previous page*

Representation	**Details**
Unit - Pictogram	uses graphic symbols to show or compare data, but without variation in size or colors to represent variations in values. See Isotype Chart for more complex versions. See Icons if just a symbol/icon conveying no numeric values.
Venn Diagram	typically uses (overlapping and/or non-overlapping) circles to show the relationship between two or more different sets
Violin	shows the distribution of data with using density curves; the width of curves in each region represents frequency of data points
Voronoi	divides a space into some number of partitions
Waffle	Include waffle bar charts here
Waterfall	bars showing net changes in value between two or more points; by adding or subtracting values, it disaggregates the unique components that contributed to the net changes
Wheat	a combination of histogram and dot plots; actual data values are grouped along the horizontal axis and stacked vertically to show the total number of observations
Wilkinson Dot Plot	like a histogram with dots instead of bars (i.e., each data represents a data point and those with the same value or in the same bin are stacked on top of each other to show the relative frequency for that bin/value)
Word Cloud	visual representation of words by adjusting size of words according to their frequencies
Word Trees	visual branching of words which focuses on hierarchical structure or connection of words
Words - other	any other visual representation of words that do not fall into word cloud or word trees

C Additional Results from our Time-Series Sample

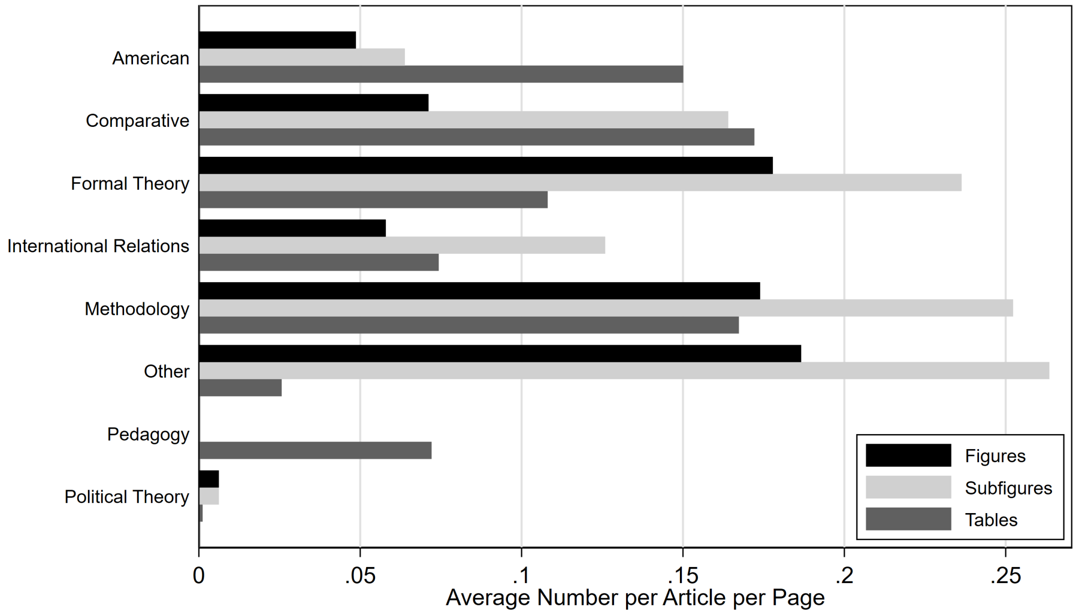

Figure 12: Average Number of Figures, Subfigures, and Tables per Article per Page, by Field

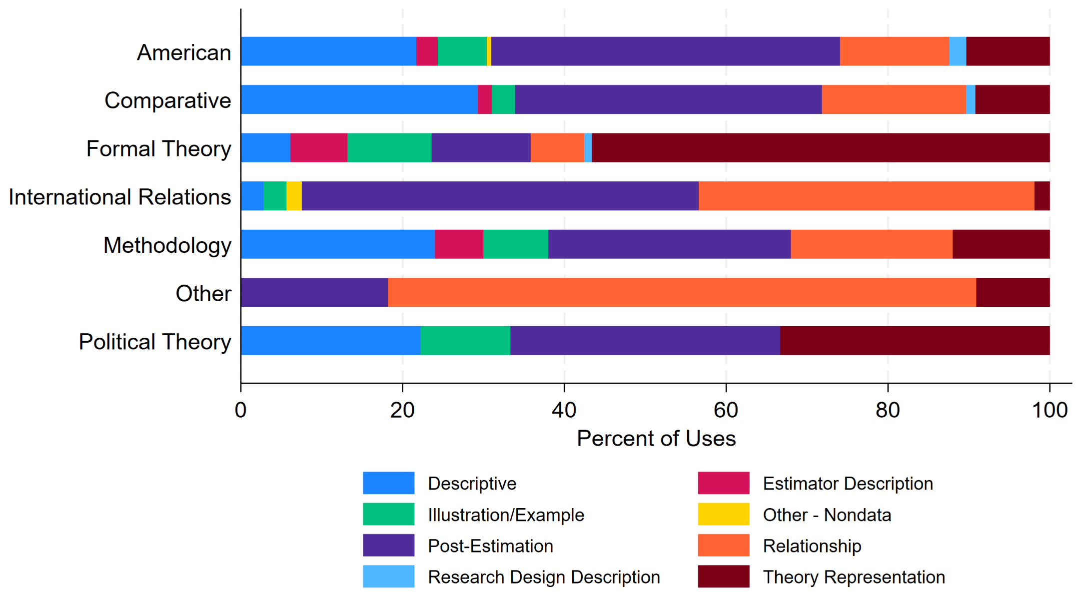

Figure 13: Purposes for Visualizations by Field

Note: Purposes are weighted by the total number of purposes per figure.

D Additional Results from our Cross-Sectional Sample

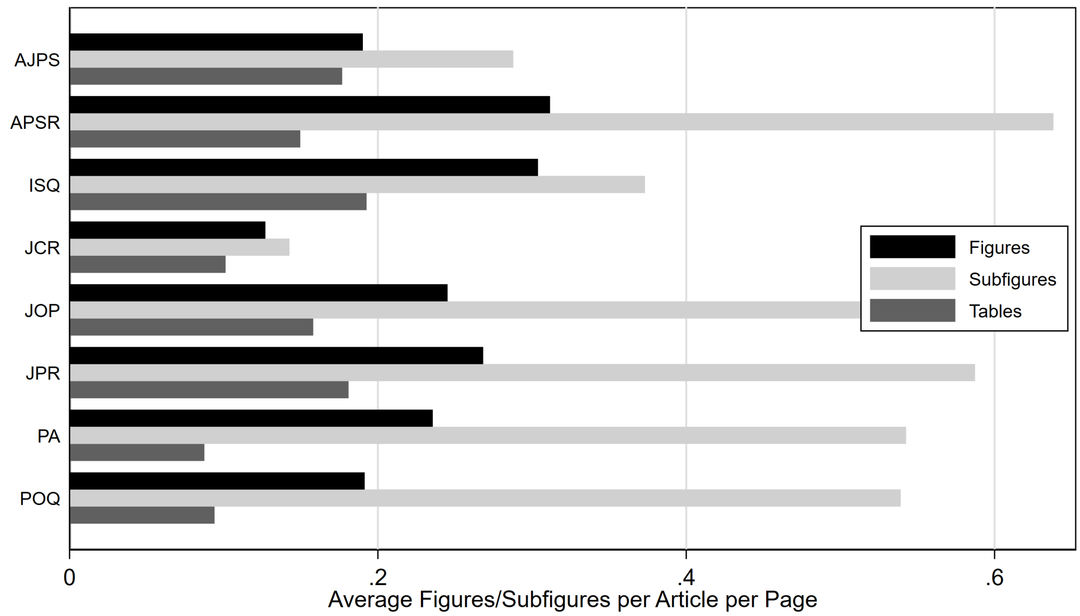

Figure 14: Average Number of Figures, Subfigures, and Tables per Article per Page, by Journal

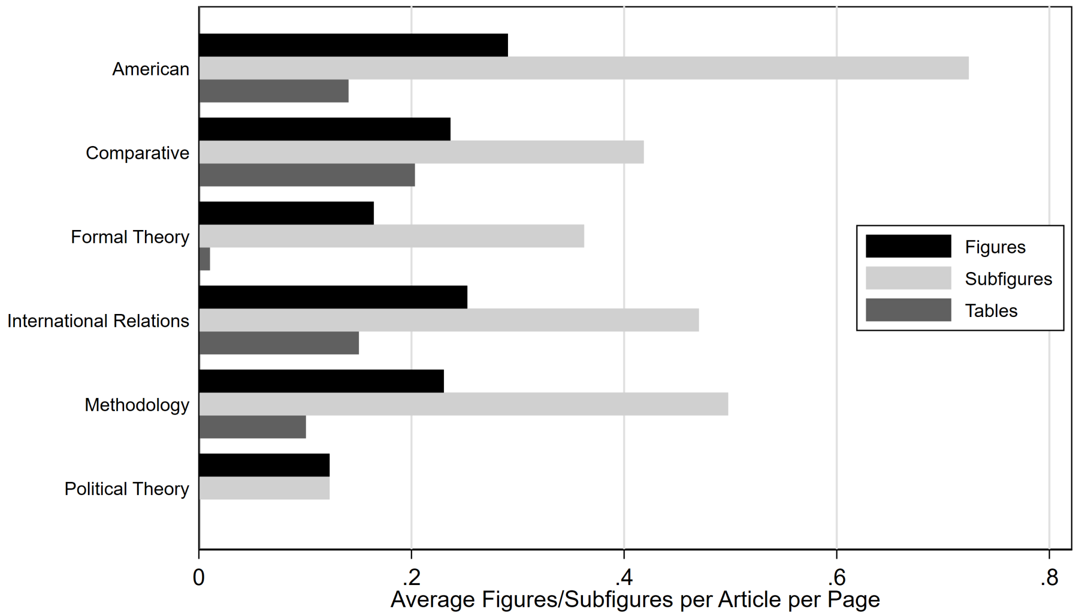

Figure 15: Average Number of Figures, Subfigures, and Tables per Article per Page, by Field

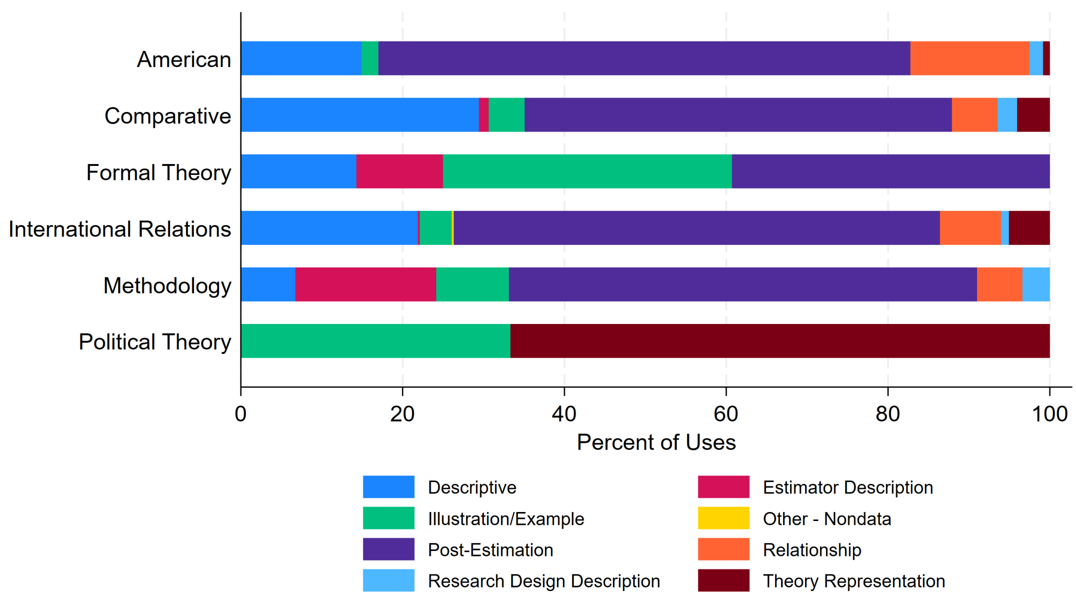

Figure 16: Purposes for Visualizations by Field

Note: Purposes are weighted by the total number of purposes per figure.

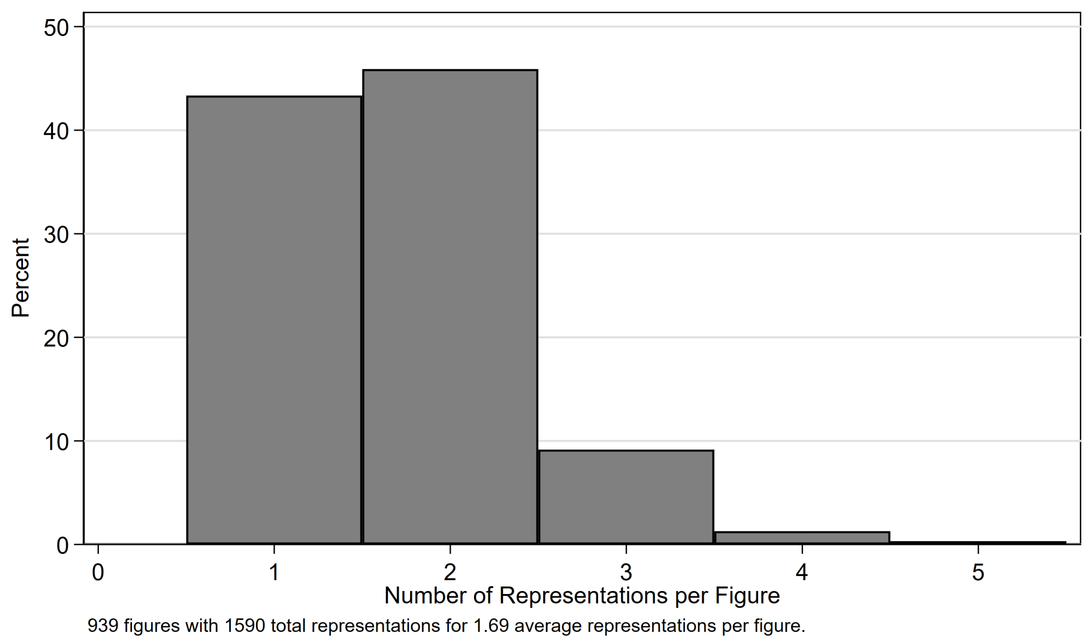

Figure 17: Number of Distinct Representation Types per Figure