Graphical visualization of polling results^{*}

Susanna Makela, Yajuan Si, Andrew Gelman

 $21 \ {\rm Feb} \ 2015$

Abstract

It's wasteful to do a big expensive poll and then just report a few percentages. Graphs allow us to convey more information more directly, both to general audiences and to specialists. In addition, graphical methods can be useful to survey researchers to understand weighting and other aspects of survey construction and analysis. We demonstrate with several examples.

Keywords: Graphics, visualization, polls, statistical modeling

Introduction

Graphics are an integral part of modern statistics and political science. In Gelman and Unwin (2013) we propose several goals for statistical graphics divided into "discovery" goals and "communication" goals. Discovery goals for graphics include giving an overview of the content of a dataset, a sense of its scale and complexity, and exploration for any unexpected aspects. Communication goals are useful for both a general audience and for specialists. Compared to tables, graphs allow many more comparisons to be visible at once, and thus can make even complex statistical reasoning more accessible to a general audience. In addition, graphs can help statisticians better evaluate their assumptions and interpret their inferences, and they help social scientists to better extract and evaluate the substantive claims and conclusions of models.

Polling is expensive, and falling response rates necessitate the most effective use of available data. Modeling allows us to obtain better estimates, especially for small cells defined by demographic groups of interest, by borrowing strength across available data. New polling methods using nonprobability samples also require statistical modeling for generalizability; see, for example, Wang et al. (2014).

Graphs can and should be used in each step of the modeling process, from exploring raw data to presenting and explaining final model results; in this chapter, we describe their use in each of these steps and illustrate with examples that arise from several previously published works, which we

^{*}Chapter entitled "Visualization and presentation," for the Oxford Handbook on Polling and Polling Methods, edited by Lonna Atkeson and Michael Alvarez. We thank the National Science Foundation for partial support of this work.

now briefly summarize. We encourage the reader to refer to these publications for greater detail on the data and models behind the graphics shown here.

Gelman et al. (2007) use multilevel modeling to explain the apparent paradox of poor voters favoring Democrats and rich voters favoring Republicans, while poor states overall tend to support Republican candidates and rich states support Democratic ones. Rothschild et al. (2014) seek to understand large swings in election polls, arguing that reported swings are often likely due to sampling bias rather than true changes in vote intention. Ghitza and Gelman (2013) use multilevel regression and poststratification to estimate election turnout and voting patterns among subsets of the population defined by multiple demographic and geographic characteristics. In Ghitza and Gelman (2014), the same authors develop a generational model of presidential voting, finding that political events in voters' teenage and young adult lives are important in shaping their long-term partisan preferences. With response rates to traditional polls rapidly declining, Wang et al. (2014) demonstrate the potential of a highly non-representative dataset of presidential vote intention, collected via the Xbox gaming platform, in obtaining accurate election forecasts via multilevel modeling and poststratification. Finally, Makela et al. (2014) demonstrate how statistical graphics can be used to better understand the survey weights that come with many surveys with complex sampling designs.

Exploring raw data

Large polls and complex public opinion surveys have a great deal of structure and patterns that can be difficult to summarize concisely. Tables of numbers and percentages quickly become unwieldy and unreadable, and comparisons between groups and quantities of interest are much more difficult to make with tables than with graphs. When exploring a raw dataset, graphics help give a clearer understanding of its characteristics by illuminating the qualitative content, allowing us to check assumptions (for example, whether outcomes between particular subgroups conform to subject matter knowledge), confirm expected results, and find distinct patterns (Gelman and Unwin, 2013).

For example, the left panel of Figure 1 from Ghitza and Gelman (2014) plots the relationship between age and Republican vote share in 2008 among non-Hispanic whites, which is complex and nonmonotonic. This plot uses only the raw data (with lowess curves for clarity), not model estimates. While subject matter knowledge may lead us to assume that Republican vote share is lower among younger people than older people, this graph complicates that assumption and forces us to consider alternative explanations.

Confronted with this new pattern, the authors construct corresponding curves for the 2000-2012 elections (Figure 1, center panel). Nonmonotonic patterns are apparent in each election year, but there is no clear trend across elections, and the peaks and valleys in different election years do not line up by age. Graphing period trends in the left and center panels of Figure 1 have revealed an unexpected pattern but did not help us understand it.

Perhaps graphing generational or cohort trends—that is, changing the x-axis from age to birth year—may further illustrate the situation. These trends are graphed in the right panel of Figure 1, and indeed, the peaks and valleys are nearly perfectly aligned, providing strong evidence for generational trends in presidential voting. As Ghitza and Gelman (2014) note, "this relationship remains clear and strong over the course of 12 years, measured across multiple surveys conducted

by different organizations, and unaltered by any complicated statistical model. This appears to be no statistical artifact." These three simple plots clearly illustrate a striking pattern that is the foundation of the entire paper.



Figure 1: Raw data and LOESS curves, indicating the relationship between age and presidential voting preferences among non-Hispanic white voters for the 2000-2012 elections. (L) The relationship is clearly non-monotonic and quite peculiar in 2008; instead of a linear or even quadratic relationship, the curve changes directions multiple times. (C) Non-monotonicity is a feature of the other elections as well, though no clear pattern is apparent from this graph alone.
(R) The true relationship emerges when the curves are lined up by birth year instead of age. The peaks and valleys occur in almost identical locations, strongly suggesting a generational trend.
(For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this chapter.) (Ghitza and Gelman, 2014)

Graphics can also help us understand the design and construction of polls and surveys, particularly with the rise of nontraditional polling methods. Wang et al. (2014) generate election forecasts using data collected through the Xbox gaming system in the 45 days before the 2012 U.S. presidential election. Their panel dataset consists of over 750,000 interviews with more than 345,000 unique respondents. However, the sample is clearly nonrepresentative and biased most severely with respect to age and sex; this bias is shown in Figure 2, which plots the demographic composition of the Xbox sample to the 2012 electorate as estimated from national exit polls. Similarly, Figure 3 plots daily estimates of two-party Obama support across the 45 days before the 2012 election for the Xbox data compared to averages from traditional polls, clearly displaying how time trends in the Xbox data compare to time trends in a representative sample.

Many polls and public opinion surveys have complex sampling schemes and come with weights that correct for known differences between the sample and population. Here again graphics are useful in understanding survey weights and their relationship to the data, as demonstrated by Makela et al. (2014). Figure 4 plots binned survey weights against the design variables used to calculate the weights. Such figures can be helpful when deciding how to incorporate sampling weights in a model—whether they should be included directly or indirectly through the design variables.



Figure 2: A comparison of the demographic, partisan, and 2008 vote distributions in the Xbox dataset and the 2012 electorate (as measured by adjusted exit polls). As one might expect, the sex and age distributions exhibit considerable differences. (Wang et al., 2014)

Furthermore, it is useful to know how survey weights are related to outcomes of interest, as shown in Figure 5. Here we see that the proportion of children who are overweight or have asthma varies weakly with the survey weights, while household income varies much more strongly, indicating that not accounting for survey weights in a model of household income could result in biased estimates. Finally, since large weights can lead to highly variable estimators, understanding the relationship between weights and sample size is important. Figure 6 shows binned weights plotted against sample size to illustrate that, although the vast majority of observations have weights with small magnitude, there are a small number of observations with large weights that can lead to noisy estimates.

Model building

When working with large datasets, graphs are instrumental in iteratively building models of increasing complexity. Figure 7 from Ghitza and Gelman (2013) illustrates one way of comparing raw data to estimates from a simple model and an incrementally more complex model.

The left panel plots raw 2008 McCain vote share by state and income for non-Hispanic whites. We can immediately see that there is much variation in McCain vote share across states, as we would expect. However, these raw estimates are quite noisy, and a clear structure is difficult to discern, even with a sample size exceeding 15,000 (Ghitza and Gelman, 2013).



Figure 3: Daily (unadjusted) Xbox estimates of the two-party Obama support during the 45 days leading up to the 2012 presidential election, which suggest a landslide victory for Mitt Romney. The dotted blue line indicates a consensus average of traditional polls (the daily aggregated polling results from Pollster.com), the horizontal dashed line at 52% indicates the actual two-party vote share obtained by Barack Obama, and the vertical dotted lines give the dates of the three presidential debates. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this chapter.) (Wang et al., 2014)

The middle panel depicts estimated McCain vote share plotted against income from a model in which the effect of income is restricted to be the same across states. As in the raw data, there is wide variation in the estimates of McCain vote share across states. The left panel plots estimates from a model in which the effect of income is allowed to vary by state. The inferences drawn from the model in the middle panel now seem simplistic when compared to estimates from the right panel.

Increasing the complexity of the model by allowing the effect of income to vary by state gives a more complete picture of voter behavior and adds an important new dimension to the story told by the middle panel, namely that the effect of individual income on McCain vote share depends on state-level income. Importantly, simply comparing predicted probabilities or tables of model coefficients would have made this conclusion difficult to come by, while the appropriate graphs make it nearly impossible to miss.

A similar story is told by the set of graphs in Figures 8-10, originally published in Gelman et al. (2007). Figures 8 and 9 are analogous to the middle and right panels of Figure 7, respectively: estimates of Bush support in Figure 8 are from a model in which the effect of individual income is the same across states, while those in Figure 9 are from a model allowing the effect to vary by state. The size of the hollow circles represents the proportion of households in each income category relative to the national average, while the solid circles represent the average state income.

The full story is shown in Figure 10, which plots the probability of voting Republican against



Figure 4: For four discrete raking variables in the Fragile Families study, we plot the proportion of respondents at each level of the given variable vs. binned baseline survey weights (log scale). The binned averages are smoothed by lowess curves. Sample size is high so we use a large number of bins (as indicated by the tick marks on the x-axes). A few of the tick marks are labeled to indicate the log weights in some of the bins; the total range of the weights is large, varying by a factor of approximately $\exp(8.5)$ or 5000.

HS = high school. (Makela et al., 2014)



Figure 5: Sample proportions of (A) children who are overweight, (B) children with asthma, and (C) families receiving welfare benefits, and (D) annual household income, all plotted vs. binned survey weights. Data for (A) are from the nine-year wave, and data for (B)–(D) are from the one-year wave. (Makela et al., 2014)

individual income for the six presidential elections between 1984 and 2004. This graph allows us to examine how the effect of individual income changes not only across states, but across elections as well.

Graphs dividing model estimates into small multiples are also instructive in understanding the structure captured by a model. One good example of this is Figure 11 from Ghitza and Gelman (2013), which plots the 2008 two-party McCain vote share against income for all voters and non-Hispanic whites by state as estimated from pooled Pew surveys and a multilevel model. For



Figure 6: Sample sizes by weight bin for baseline weights in the Fragile Families study for (A) all weight bins, (B) weight bins with sample size less than 100. (Makela et al., 2014)



Figure 7: The first panel shows the raw data; the middle panel is a hierarchical model where state coefficients vary, but the (linear) income coefficient is held constant across states; the right panel allows the income coefficient to vary by state. Adding complexity to the model reveals weaknesses in inferences drawn from simpler versions of the model. Three states – Mississippi (the poorest state), Ohio (a middle-income state), and Connecticut (the richest state) – are highlighted to show important trends. (Ghitza and Gelman, 2013)

most states, the relationship between income and McCain vote share is similar for all voters and non-Hispanic whites, but there are several states—Louisiana, South Carolina, Mississippi, and Maryland among them—in which the pattern for non-Hispanic whites deviates notably from all



Figure 8: The paradox is no paradox. From the multilevel logistic regression model fit to Annenberg poll data from 2000 to 2004: probability of supporting Bush as a function of income category, for a rich state (Connecticut), a middle-income state (Ohio), and a poor state (Mississippi). The open circles show the relative proportion (as compared to national averages) of households in each income category in each of the three states, and the solid circles show the average income level and estimated average support for Bush for each state. Within each state, richer people are more likely to vote Republican, but the states with higher income give more support to the Democrats. (Gelman et al., 2007)

voters, particularly for lower income quintiles. These plots emphasize the importance of accounting for interactions between income, state, *and* ethnicity, not just income and state, when modeling McCain vote share.

Often a more complex model leads to a new story that is more consistent with the data. Figure 12 from Rothschild et al. (2014) shows estimates of two-party Obama support over time for one model that adjusts only for demographics and another that adjusts for both demographics and partisanship. Under the first model, Obama support fluctuates sharply in the 45 days preceding Election Day, but adjusting for partisanship in addition to demographics greatly reduces this variation. Rothschild et al. (2014) interpret results from the latter model as "suggesting that most of the apparent changes in support during this period were artifacts of partisan nonresponse." In this case, graphing estimates from the two models in the same figure reveals a qualitatively different picture of Obama support prior to the 2012 election under the more complex model that adjusts for partisanship in addition to demographics compared to the simpler demographics-only model.

Another example of graphs illustrating the different stories two models can tell is Figure 13, also from Rothschild et al. (2014). Here, the authors plot changes in two-party Obama support before and after the first presidential debate across various subpopulations for the demographics-only and demographics plus partial product described above. The conclusions about the effects of the debate on Romney support on these subpopulations are different between the two models.



Figure 9: From the multilevel logistic regression model with varying intercepts and slopes fit to Annenberg poll data from 2000 to 2004: probability of supporting Bush as a function of income category, for a rich state (Connecticut), a middle-income state (Ohio), and a poor state (Mississippi). The open circles show the relative proportion (as compared to national averages) of households in each income category in each of the three states, and the solid circles show the average income level and estimated average support for Bush for each state. Income is a very strong predictor of vote preference in Mississippi, a weaker predictor in Ohio, and only weakly predicts vote choice at all in Connecticut. See Figure 5 for estimated slopes in all 50 states, and compare to Figure 3, in which the state slopes are constrained to be equal. (Gelman et al., 2007)

Understanding the results

Interpreting coefficients from even relatively simple models can be difficult. Adding interactions, nonlinear terms, and hierarchical structure to the model makes such interpretations even more challenging. Furthermore, in multilevel models, coefficients are modeled in batches, and we may be interested in the extent of partial pooling in the coefficient estimates, which is difficult to communicate via tables. Graphs can make regression results from even highly complex models easier to understand, summarize, and interpret.

One example of using graphs to understand model results comes from Ghitza and Gelman (2013). In describing models of election turnout and voting patterns, the authors note that "... we knew a priori that our estimates for Obama's vote share among African American groups needed to be high, over 90%, but we could not know what regression coefficient was plausible, as the coefficient could change drastically depending on functional form." In contrast, graphing the actual estimated Obama support for various demographic subgroups would immediately reveal whether the model captures this known aspect of the data and how the estimates behave as these subgroups are made finer and finer.

Figure 14 from their paper confirms that African Americans' predicted two-party McCain vote share is low. In addition, we see that adding more demographics reveals the heterogeneity within subgroups, but the overall estimates remain relatively stable (Ghitza and Gelman, 2013). Figure 14 also exemplifies how graphs can encode additional information in the color and size of plotting



Figure 10: Using exit poll data from 1984 to 2004, results for the varying-intercept, varying-slope multilevel logistic regression. The curves show the probability of supporting Bush as a function of income category, within states that are poor, middle-income, and rich. (Gelman et al., 2007)

symbols.

Similarly, Figure 15 uses color and a grid of maps by age and income to display the heterogeneity in vote swing from 2004 to 2008 among non-Hispanic whites. While whites overall shifted toward Obama by 3.3%, poorer and older white voters in the South and Appalachia actually supported McCain in 2008 more than they did for Bush in 2004 (Ghitza and Gelman, 2013). This heterogeneity would be near impossible to determine from regression coefficients alone, and the use of color and repeated multiple graphs makes the variation by age, income, and geography immediately clear to the reader.

Regression coefficients from a complex model are summarized particularly clearly in Figures 16 and 17 from Ghitza and Gelman (2014). Full details of the model are given on Page 6-7 of that paper, but briefly, the model predicts the proportion of Republican presidential support by the birth year cohort, election year, and race/region group (non-Southern white, Southern white, and minority) to which a given survey respondent belongs. Specifically, Republican vote share is modeled as the sum of a generational effect—the importance of age in forming long-term presidential voting patterns and how this importance varies by race/region—and a period effect that captures election-to-election changes by race/region and the importance of these changes for different age groups.



Figure 11: All voters shown in black and non-Hispanic whites in gray. Dots are weighted averages from pooled June-November Pew surveys; error bars show +/-1 s.e. bounds. Curves are estimated using multilevel models and have a s.e. of about 3% at each point. States are ordered in decreasing order of McCain vote (Alaska, Hawaii, and Washington, DC, excluded). (Ghitza and Gelman, 2013)

Figure 16 summarizes the generational effects, which consist of an age-specific weight for ages 1 to 70 and an interaction term that allows the importance of these weights to vary by race/region group for each birth year and election year. While the actual numeric values of the age weights are difficult to understand, we can immediately see in the left panel that events occurring roughly between the ages of 14 and 24 have the largest impact on future vote preference. The interaction terms are



Figure 12: Among respondents who expressed support for either Barack Obama or Mitt Romney, estimated support for Obama (with 95% confidence bands) under two different poststratification models: the dark line plots results after adjusting for both demographics and partisanship, and the light line adjusts only for demographics (identical to Figure 2). The surveys adjusted for partisanship show less than half the variation of the surveys adjusted for demographics alone, suggesting that most of the apparent changes in support during this period were artifacts of partisan nonresponse. (Rothschild et al., 2014)

summarized in the right-hand panel of Figure 16, which displays their posterior distribution for each race/region group. Interaction terms are often difficult to interpret directly, and this graph allows us to ignore their exact numeric values and focus on understanding their substantive meaning, while also clearly displaying the uncertainty in their posterior estimates. The age weights are more important for whites, with the means of the estimates (denoted by the vertical lines) more than twice as high for whites than for non-whites. Ghitza and Gelman (2014) point out that these interaction terms were not restricted *a priori* to be positive by the model, but as we can see from the graph, each distribution is centered well away from zero; this is another feature of the estimates that would be difficult to discern without a graphical summary.

The period effects are displayed in Figure 17. These effects consist of an election-specific term that captures the effect of that election year for the three race/region groups, as well as an interaction term that allows this effect to be potentially stronger in some age groups than in others. Recall that Republican vote share is modeled as the sum of a generational and period effect, so a negative value for the election year effect indicates lower Republican vote share. Thus, the election year

effects plotted in the left-hand panel of Figure 17 show that non-whites have been consistently more likely to vote for Democratic candidates over the past 50 years, while Southern whites tend to vote more Republican, particularly in the four most recent elections. These results are consistent with subject matter knowledge, so we can be confident that the model is capturing expected patterns in the data.

The parameters governing the relative importance of the election year effects for different age groups are more difficult to summarize. One way to understand them is to calculate the ratio of the election year effect at ages 18 and 70, respectively the (approximate) peak and trough of the age-weight curve in Figure 16. The right-hand panel of Figure 17 plots the distribution of this ratio for the race/region groups. For Southern whites and minorities, there does not seem to be much of a differential age effect, and while there is a possible larger effect for young ages among non-Southern whites, the spread of the distribution is too wide to be conclusive. Again, these regression coefficients would have been difficult to interpret from a table, but we can easily understand them by graphing a clever transformation of the estimates that summarizes a relevant feature of the model. Furthermore, as in the right-hand panel of Figure 16, plotting the entire posterior distribution instead of a point estimate makes it easier to understand the extent of uncertainty in the estimated coefficients.

Model checking

Model checking is the process of understanding how well and to what extent the model fits the data and where could be improved. We first consider simple comparisons of model predictions to known outcomes or gold standard data as in Figures 18 to 20 from Wang et al. (2014), which used a non-representative dataset collected via the Xbox gaming platform to generate election forecasts for the 2012 presidential election and applied multilevel regression and poststratification to adjust the Xbox estimates. The 2012 exit polls are used as the benchmark or gold standard for evaluating the accuracy of the model-based forecasts.

Figures 18 and 19 show the discrepancies between two-party Obama vote share for various demographic subgroups obtained from the Xbox estimates and from exit polls. For simple onedimensional demographic groups such as sex and age, model estimates and benchmark values can be directly plotted on the same graph, as in Figure 18. However, as we further subdivide the population by considering two-dimensional demographic groups such as female moderates, white liberals, etc., directly plotting the two sets of estimates would render the plot difficult to read. Instead, plotting the *differences* and ordering them by magnitude allows us to easily see which subgroups' voting behavior is best captured by the model, as shown in the left panel of Figure 19. Here the authors have selected the 30 largest two-dimensional demographic subgroups for visual clarity. We can see the same comparison for all 149 two-dimensional demographic subgroups in the right panel of Figure 19. Encoding the relative size of the subgroup in the size of the dot allows an additional layer of information to be easily incorporated into the graph, making it clear to the reader that, as would be expected, the Xbox estimates are poorest for the smallest demographic subgroups and best for the largest ones.

Another way to check the fit of the model is to consider the posterior predictive distribution for a quantity of interest. In cases where benchmark data is unavailable, we can draw samples from this

distribution and calculate a test statistic to compare to the actual data. In the case of Wang et al. (2014) where benchmark data are available, Figure 20 plots the predicted distribution of electoral votes for Obama. The blue and green dashed lines represent, respectively, the actual number of electoral votes Obama captured (332) and the minimum number needed to tie (269). As most of the mass of this distribution is to the right of the minimum number needed to tie, we can see that the model estimates a high probability of an Obama victory (the estimated likelihood is in fact 88%). However, we also see that the distribution is quite variable, and the authors note that "extreme outcomes seem to have unrealistically high likelihoods of occurring." Graphs like Figure 20 are useful in revealing such possibly unexpected aspects of the model and prompting further investigation into which features of the data are not fully captured or misrepresented by the model, leading to another iteration in the cycle of data exploration and model building.

In addition to understanding the implications and meanings of regression coefficients, we also want to know how well the model fits the data overall. Figure 21 from Ghitza and Gelman (2014) plots R^2 , an elementary measure of the percent of variance in the outcome explained by a model, for their full model of vote choice, as well as a simpler model that includes only period/group effects. The importance of this graph is that it displays R^2 not only for the data as a whole, but also for the three race/region groups separately. Comparing the two models on the basis of the data as a whole may lead us to conclude that the simpler model is preferable, but the breakdown by race/region reveals that the advantage of the more complicated model is in its superior performance in explaining variance in vote choice among non-Southern whites.

Presenting results

Finally, graphs are essential in presenting and explaining the results of a poll or statistical model. A prime example is given in Figure 22 from Ghitza and Gelman (2014). The top panel shows the Gallup Presidential Approval series, the main covariate used to model presidential vote choice. The series is color-coded to highlight pro-Republican (red) and pro-Democratic (blue) years, with line thickness proportional to the age weights corresponding to white members of the cohort born in 1941. The bottom panel plots the cumulative generational effects—that is, the overall voting tendencies of the cohort at each age—excluding period effects so as to display general trends independent of the effects of any particular election.

The top and bottom panels work in concert to tell the story of presidential voting for this cohort. Despite high approval ratings for President Roosevelt and during the first half of the Truman presidency, the members of this cohort were too young to be significantly affected by the popularity of these Democratic leaders. This lack of effect can be seen in the low age weights in those years (the thickness of the approval series) and the nearly zero values of the cumulative generation effect in the bottom panel. The most important years in terms of political socialization for this cohort occurred during the Eisenhower presidency, where the age weights for this cohort are at their largest. Eisenhower was a popular Republican president, reflected in the dark red of the approval series, and the 1941 birth cohort became steadily more pro-Republican over the course of his presidency.

The effects of subsequent presidents are described in more detail in Ghitza and Gelman (2014), and we pause here to summarize the many pieces of information incorporated in this graph. First, it displays the presidential approval series, with color to distinguish between pro-Republican and pro-Democratic years within a presidency; the measure of pro-Republican approval (equivalently, Democratic disapproval) is the main covariate used to model vote choice. Second, the graph incorporates the age weights, a substantively important aspect of the model, in the presidential approval series by making the width of the series proportional to these weights. Third, the bottom panel displays an easily interpretable summary of the model results in terms of generational effects. Finally, the juxtaposition of the two panels so that presidential administrations align with the age of the cohort neatly ties together the relationship between presidential approval and generational voting trends captured by the model. In short, this graph is useful because it shows the correspondence between the key covariates (presidential approval and age) and the outcome in a single figure and enhances the narrative that qualitatively ties the model together.

Figure 23 for Ghitza and Gelman (2014) plots cumulative generational trends for all white voters born between 1855 and 1994. The trends for each generation are shown in a solid line, with surrounding colored bands whose width is proportional to each generation's contribution to the total electorate in a given year. This plot allows us to easily visualize and understand the behavior of each generation over time and is an invaluable complement to the narrative given in the text of Ghitza and Gelman (2014).

Discussion

We have described the use of graphics in each step of the modeling process, from exploring raw data to presenting final results. The use of graphics helps us take advantage of all of the information available in a poll or dataset that was often conducted with considerable expense. These graphs allow us to communicate more information more directly, both to general audiences and to specialists.



Figure 13: Estimated swings in two-party Obama support between the day before and four days after the first presidential debate under two different poststratification models, separated by subpopulation. The vertical lines represent the overall average movement under each model. The horizontal lines correspond to 95% confidence intervals. (Rothschild et al., 2014)



Figure 14: Size = Subgroup population size 2007; Color by ethnicity: White = White, Black = Black, Red = Hispanic, Green = Other. Each bubble represents one demographic subgroup per state, with size and color indicating population size and ethnicity. As additional demographics are added, heterogeneity within subgroups is revealed by the dispersion of the bubbles, while estimates remain reasonable. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this chapter.) (Ghitza and Gelman, 2013)



Figure 15: State-by-state shift toward McCain (red) or Obama (blue) among white voters broken down by income and age. Red = McCain better than Bush; Blue = McCain worse than Bush. Only groups with ¿1% of state voters shown. Although almost every state moved toward Obama in aggregate, there are substantial demographic groups that moved toward McCain all over the map, specifically among older whites. (Ghitza and Gelman, 2013)



Figure 16: Estimates for the generational aspects of the model. (L) The rough age range of 14-24 is found to be of paramount importance in the formation of long-term presidential voting preferences. Political events at a very young age have very little impact, and after the age of 24, the age weights decrease, staying at a small steady magnitude from about the age of 45 onward. (R) These age weights, and the political socialization process implied by them, are substantially more important for non-Hispanic whites than for minorities as whole. (Ghitza and Gelman, 2014)



Figure 17: Estimates for the election-to-election period effects in the model. (L) Minorities are consistently more likely to vote for Democratic presidents, and Southern whites have steadily trended pro-Republican over the past 50 years. (R) Period effects are roughly similar between young and old voters among minorities and in the South; evidence is inconclusive for non-Southern whites. (Ghitza and Gelman, 2014)



Figure 18: Comparison of the two-party Obama vote share for various demographic subgroups, as estimated from the 2012 national exit poll and from the Xbox data on the day before the election. (Wang et al., 2014)



Figure 19: Left panel: Differences between the Xbox MRP-adjusted estimates and the exit poll estimates for the 30 largest two-dimensional demographic subgroups, ordered by the differences. Positive values indicate that the Xbox estimate is larger than the corresponding exit poll estimate. Among these 30 subgroups, the median and mean absolute differences are 1.9 and 2.2 percentage points, respectively. Right panel: Two-party Obama support, as estimated from the 2012 national exit poll and from the Xbox data on the day before the election, for various two-way interaction demographic subgroups (e.g., 65+ year-old women). The sizes of the dots are proportional to the population sizes of the corresponding subgroups. (Wang et al., 2014)



Figure 20: The projected distribution of electoral votes for Obama one day before the election. The green vertical dotted line represents 269, the minimum number of electoral votes that Obama needed for a tie. The blue vertical dashed line indicates 332, the actual number of electoral votes captured by Obama. The estimated likelihood of Obama winning the electoral vote is 88%. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this chapter.) (Wang et al., 2014)



How Well Does the Model Explain Macro-Level Vote Choice?

Figure 21: The model accounts for 92% of the macro-level variance in voting trends over the past half century. That said, much simpler models, incorporating only period/group effects, would also eccount for much of the variance. The real substantive power of the model is in how it improves model fit within race/region groups, particularly among non-Southern whites. (Ghitza and Gelman, 2014)



Figure 22: The Presidential Approval time series, and the cumulative generational effects of that series, for Eisenhower Republicans, born in 1941. The series is drawn to emphasize this generation's peak years of socialization, according to the age weights found by the model. Dark blue indicates strongly pro-Democratic years, dark red for pro-Republican, and shades of grey in between. This generation missed most of the FDR years and were socialized through 10 straight pro-Republican years, spanning the end of the Truman presidency and eight years of the popular Republican President Eisenhower. Their partisan voting tendencies were somewhat stabilized back towards the neutral grey line by the pro-Democratic 1960s, and they reached a rough equilibrium by the end of the Nixon presidency. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this chapter.) (Ghitza and Gelman, 2014)



Figure 23: The cumulative preferences of each generation is shown, along with the weighted summation of the full white electorate. The generations are now more loosely defined, to allow the entire electorate to be plotted at once, with the width of each curve indicating the proportion of the white electorate that each generation reflects at any given time. The model – in this graph reflecting only the Approval time series and the age weights – can explain quite a bit about the voting tendencies of the white electorate over time. (Ghitza and Gelman, 2014)

References

- Gelman, A. E., Shor, B., Bafumi, J. and Park, D. K. (2007), 'Rich state, poor state, red state, blue state: What's the matter with Connecticut?', *Quarterly Journal of Political Science* 2(4), 345– 367.
- Gelman, A. and Unwin, A. (2013), 'Infovis and statistical graphics: Different goals, different looks', Journal of Computational and Statistical Graphics 22(1), 2–28.
- Ghitza, Y. and Gelman, A. (2013), 'Deep interactions with MRP: Election turnout and voting patterns among small electoral subgroups', *Americal Journal of Political Science* 57(3), 762–776.
- Ghitza, Y. and Gelman, A. (2014), 'The Great Society, Reagan's revolution, and generations of presidential voting', p. unpublished.
- Makela, S., Si, Y. and Gelman, A. (2014), 'Statistical graphics for survey weights', *Revista Colombiana de Estadstica*.
- Rothschild, D., Goel, S., Gelman, A. and Rivers, D. (2014), 'The mythical swing voter', p. unpublished.
- Wang, W., Rothschild, D., Goel, S. and Gelman, A. (2014), 'Forecasting elections with non-representative polls', *International Journal of Forecasting* p. to appear.