Resolving the Controversy

Chapter 3: Should Election Polls Be Banned?

Arguments *against* public preelection polls charge that they influence voter behavior. Voters may decide to stay home if the polls predict a landslide–why bother to vote if the result is a foregone conclusion? Exit polls are particularly worrisome because, in effect, they report actual election results before the election is complete. The U.S. television networks agree not to release the results of their exit surveys in any state until the polls close in that state. If a presidential election is not too close to call, the networks may know (or think they know) the winner by midafternoon, but they forecast the vote one state at a time as the polls close across the country. Even so, a presidential election result may be known (or thought to be known) before voting ends in the western states. Some countries have laws restricting election forecasts. In France, no poll results can be published in the week before a presidential election. Canada forbids poll results in the 72 hours before federal elections. In all, some 30 countries restrict publication of election surveys.

The argument *for* preelection polls is simple: democracies should not forbid publication of information. Voters can decide for themselves how to use the information. After all, supporters of a candidate who is far behind know that fact even without polls. Restricting publication of polls just invites abuses. In France, candidates continue to take private polls (less reliable than the public polls) in the week before the election. They then leak the results to reporters in the hope of influencing press reports.

One argument *for* exit polls is that they provide a means for checking election outcomes. Discrepancies between exit polls and reported election outcomes invite investigation into the reasons for the differences. Such was the case in the 2004 U.S. presidential election. Were the exit polls flawed, or were the reported election results in error?

Chapter 4: The Harris Online Poll

The Harris Online Poll uses probability sampling and statistical methods to weight responses and uses recruitment to attempt to create a panel (sampling frame) that is as representative of the population of interest as possible. But the panel also consists of volunteers and hence suffers to some extent from voluntary response. In addition, panel members are Internet users, and it is not clear how such a panel can be representative of a larger population that includes those who do not use the Internet.

Thus, the verdict is out on whether the Harris Poll Online provides accurate information about well-defined populations such as all American adults.

Chapter 6: Is It or Isn't It a Placebo?

Should the FDA require natural remedies to meet the same standards as prescription drugs? That's hard to do in practice, because natural substances can't be patented. Drug companies spend millions of dollars on clinical trials because they can patent the drugs that prove effective. Nobody can patent an herb, so nobody has a financial incentive to pay for a clinical trial. Don't look for big changes in the regulations.

Meanwhile, it's easy to find claims that ginkgo biloba is good for (as one website says) "hearing and vision problems as well as impotence, edema, varicose veins, leg ulcers, and strokes." Common sense says you should be suspicious of claims that a substance is good for lots of possibly unrelated conditions. Statistical sense says you should be suspicious of claims not backed by comparative experiments. Many untested remedies are no doubt just placebos. Yet they may have real effects in many people—the placebo effect is strong. Just remember that the safety of these concoctions is also untested.

Chapter 7: Hope for Sale?

One issue to consider is whether bone marrow transplant (BMT) really keeps patients alive longer than standard treatments. We don't know, but the answer appears to be "probably not." The patients naturally want to try anything that might keep them alive, and some doctors are willing to offer hope not backed by good evidence. One problem was that patients would not join controlled trials that might assign them to standard treatments rather than to BMT. Results from such trials were delayed for years by the difficulty in recruiting subjects. Of the first five trials reported, four found no significant difference between BMT and standard treatments. The fifth favored BMT—but the researcher soon admitted "a serious breach of scientific honesty and integrity." The *New York Times* put it more bluntly: "he falsified data."

Another issue is "smart" compassion. Compassion seems to support making untested treatments available to dying patients. Reason responds that this opens the door to sellers of hope and delays development of treatments that really work. Compare children's cancer, where doctors agree not to offer experimental treatments outside controlled trials. Result: 60% of all children with cancer are in clinical trials, and progress in saving lives has been much faster for children than for adults. BMT for a rare cancer in children was tested immediately and found to be effective. In contrast, one of the pioneers in using BMT for breast cancer, in the light of better evidence, now says, "We deceived ourselves and we deceived our patients."

Chapter 8: Six-Year Graduation Rates, High School GPA, and Standardized Tests

We see that high school grade point average (GPA) alone appears to be a better predictor of six-year graduation rates than does SAT or ACT scores–as high school GPA increases, graduation rates increase from 26% to 47%, but as SAT or ACT score increases, graduation rates increase from 26% to only 35%. Combining high school GPA with SAT or ACT score does a better job of predicting six-year graduation rate than either by itself. At the highest level of each variable together, the predicted graduation rate is 72%.

We need to remember that students with the same high school GPA and SAT or ACT scores often perform quite differently in college, which could impact time to graduation. Motivation and study habits matter a lot. Choice of major, choice of classes, and choice of college also impact time to graduation.

However, if a college decides to go test optional, we return to the fact that high school GPA alone is a better predictor of six-year graduation rate than is SAT or ACT score alone. While the graduation rates are lower than with high school GPA and SAT or ACT combined, it might be worth the expanded diversity, which will benefit students in more ways than can be measured by time to graduation.

Chapter 12: Income Inequality

These are complicated issues, with much room for conflicting data and hidden agendas. The political left wants to reduce inequality, and the political right says the rich earn their high incomes. We want to point to just one important statistical twist. Figures 12.4 and 12.5 report "cross-sectional" data that give a snapshot of households in each year. "Longitudinal" data that follow households over time might paint a different picture. Consider a young married couple, Jamal and Tonya. As students, they work part-time, then borrow to go to graduate school. They are down in the bottom fifth. When they get out of school, their income grows quickly. By age 40, they are happily in the top fifth. In other words, many poor households are only temporarily poor.

Longitudinal studies are expensive because they must follow the same households for years. They are prone to bias because some households drop out over time. One study of income tax returns found that only 14% of the bottom fifth were still in the bottom fifth 10 years later. But very poor people don't file tax returns. Another study looked at children under 5 years old. Starting in both 1971 and 1981, it found that 60% of children who lived in households in the bottom fifth still lived in bottom-fifth households 10 years later. Many people do move from poor to rich as they grow older, but there are also many households that stay poor for years. Unfortunately, many children live in these households.

Chapter 15: Gun Control

This study is observational in nature. States cannot be randomly assigned to have particular laws, and by focusing on data from a single year, the researchers were unable to assess whether firearm mortality changed in any way after certain laws went into effect. We therefore cannot conclude that specific state laws cause low firearm mortality, and we should be very cautious in terms of suggesting that changing laws will lead to lower firearm mortality.

When Kalesan and her colleagues published their work, critics acknowledged their efforts to control for different variables (like unemployment and gun ownership), but they were quick to point out several other variables that were not taken into account. For example, what about variables such as poverty level, alcohol consumption, and mental health? In a 2016 *Washington Post* article, columnist Carolyn Johnson questioned the findings of the study by noting that "People have very different inherent risks of death by firearm simply depending on whether they live in a rural or urban area. And while gun control laws may affect death from firearms, so will other factors, such as urban environments, poverty, and gang violence." Johnson went on to include comments from David Hemenway, a researcher from Harvard, and Daniel Webster, the director of the Johns Hopkins Center for Gun Policy Research. Although both Hemenway and Webster had concerns about other methodological issues in the study, each reached a very different conclusion about the importance of the findings. Hemenway lauded the study as a step in the right direction in terms of bringing scientific evidence to the debate on gun control, whereas Webster worried about how the flaws in the study might become politicized and ultimately lead the public to become skeptical of scientific research. Like many questions of causation, this one clearly remains open.

Chapter 16: Does the CPI Overstate Inflation?

The Consumer Price Index (CPI) has an upward bias because it can't track shifts from beef to tofu and back as consumers try to get the same quality of life from whatever products are cheaper this month. This was the basis of the outside experts' criticisms of the CPI: the CPI does not track the "cost of living." Their first recommendation was that "the BLS should establish a cost of living index as its objective in measuring consumer prices." The Bureau of Labor Statistics (BLS) said it agreed in principle but that neither it nor anyone else knows how to do this in practice. It also said, "Measurement of changes in 'quality of life' may require too many subjective judgments to furnish an acceptable basis for adjusting the CPI." Nonetheless, a new kind of index that in principle comes closer to measuring changes in the cost of living was created in 2002. This new index is called the Chained CPI-U (C-CPI-U). It more closely approximates a cost-of-living index by reflecting substitution among item categories. This new index may be an improvement, but it is unlikely that the difficult problems of defining living standards and measuring changes in the cost of their attainment over time will ever be resolved completely.

Chapter 20: The State of Legalized Gambling

Opponents of gambling have good arguments against legalized gambling. Some people find betting addictive. A study by the National Opinion Research Center estimated that pathological gamblers account for as much as 15% of gambling revenue and that each such person costs the rest of us $10,000 over his lifetime for social and police work. Gambling does ruin some lives, and it does indirectly harm others.

State-run lotteries involve governments in trying to persuade their citizens to gamble. From the early days of the New York lottery, we recall billboards that said, "Support education—play the lottery." That didn't work, and the ads quickly changed to "Get rich—play the lottery." Lotteries typically pay out only about half the money bet, so they are a lousy way to get rich even when compared with the slots at the local casino. Professional gamblers and statisticians avoid them, not wanting to waste money on so bad a bargain. Poor people spend a larger proportion of their income on lotteries than do the rich and are the main players of daily numbers games. The lottery may be a voluntary tax, but it hits the poor hardest, and states spend hundreds of millions on advertising to persuade the poor to lose yet more money. Some modest suggestions from those who are concerned about state-run lotteries: states should cut out the advertising and pay out more of what is bet.

States license casinos because they pay taxes and attract tourists—and, of course, because many citizens want them. In fact, most casinos outside Las Vegas draw gamblers mainly from nearby areas. Crime is higher in counties with casinos—but lots of lurking variables may explain this association. Pathological gamblers do have high rates of arrest, but again the causal link is not clear.

The debate continues. Meanwhile, technology in the form of Internet gambling is bypassing governments and creating a new gambling economy that makes many of the old arguments outdated.

Chapter 23: Should Hypothesis Tests Be Banned?

It will probably not surprise you that the American Statistical Association (ASA) did not take kindly to the BASP band on hypothesis testing and confidence intervals. In 2016, the ASA published a formal response: https://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf

It is interesting to note that in 1999, the American Psychological Association (APA) appointed a Task Force on Statistical Inference. At that time, the task force did not want to ban hypothesis tests. The report that was produced by the task force was, in fact, a summary of good statistical practice:

- Define your population clearly.
- Describe your data production and prefer randomized methods whenever possible.
- Describe your variables and how they were measured.
- Give your sample size and explain how you decided on the sample size.
- If there were dropouts or other practical problems, mention them.
- "As soon as you have collected your data, before you compute *any* statistics, *look at your data.*"
- Ask whether the results of computations make sense to you.
- Recognize that "inferring causality from non-randomized designs is a risky enterprise."

The APA task force did say, "It is hard to imagine a situation in which a dichotomous accept-reject decision is better than reporting an actual *p* value or, better still, a confidence interval. ... Always provide some effect-size estimate when reporting a *p* value." But would the task force ban hypothesis tests altogether? "Although this might eliminate some abuses, the committee thought there were enough counterexamples to justify forbearance."

Sixteen years later, BASP banned hypothesis tests and confidence intervals. The controversy is not over. We encourage you to search the Web for the most up-to-date information about this issue. A good starting point would be to view the keynote address given by Ron Wasserstein (American Statistical Association) and Allen Schirm (Mathematica Policy Research; retired) at the 2019 USCOTS meeting: https://www.causeweb.org/cause/uscots/uscots19/keynote/2.