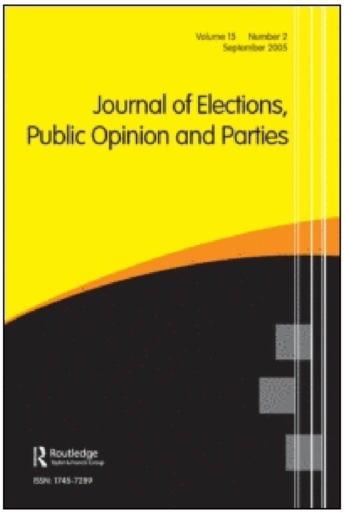

Journal of Elections, Public Opinion & Parties
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713727959

The 2006 Cooperative Congressional Election Study

Lynn Vavreck [a]; Douglas Rivers [b]
[a] University of California, Los Angeles and YouGov/Polimetrix [b] Stanford University and YouGov/Polimetrix,

Online Publication Date: 01 November 2008

To cite this Article Vavreck, Lynn and Rivers, Douglas(2008)'The 2006 Cooperative Congressional Election Study',Journal of Elections, Public Opinion & Parties,18:4,355 — 366

To link to this Article: DOI: 10.1080/17457280802305177

URL: http://dx.doi.org/10.1080/17457280802305177

Journal of Elections, Public Opinion and Parties
Vol. 18, No. 4, 355–366, November 2008

The 2006 Cooperative Congressional Election Study

LYNN VAVRECK* & DOUGLAS RIVERS**
*University of California, Los Angeles and YouGov/Polimetrix; **Stanford University and YouGov/Polimetrix

ABSTRACT *In 2006 Polimetrix, Inc. of Palo Alto, CA. fielded the Cooperative Congressional Election Study, the largest study of Congressional elections ever fielded in the US. The project was a joint venture of 38 universities and over 100 political scientists. In this paper, we detail the design and execution of the project, with special attention to the method by which the sample was generated. We show that the estimates from the Common Content of CCES outperform conventional estimates based on RDD phone surveys. We also argue that opt-in panels, internet surveys, and cooperative ventures like CCES provide cost-effective alternatives for social scientists under certain conditions. These types of surveys can provide reductions in RMSE over conventional methods when sample matching is used to ameliorate the biases that come with sampling from an opt-in panel.*

The 2006 midterm elections are viewed by scholars of Congressional politics as a pivotal set of contests (Jacobson, 2007). The Democrats took control of Congress promising the beginning of large programmatic efforts to hasten recent trends in American politics. Also in 2006, teams from 38 universities around the country joined together to execute the largest study of Congressional elections ever fielded, the Cooperative Congressional Election Study (CCES). The study began with a pool of over 138,000 Americans and delivered a final dataset consisting of 36,501 respondents, which made up a nationally representative dataset consisting of two interviews per person. Each respondent in the study was a part of the larger common project (the Common Content – the first ten minutes of the survey), but also participated in one of the team's studies (the Team Content – the second ten minutes of the survey). Every time a new team joined CCES another 1,000 cases were added to the overall sample. Through this design, CCES delivered a common dataset with the items common to all respondents, profile data on all the respondents, and team datasets of 1,000 representative cases to each team. Over 100 political scientists advanced their research through participation in CCES. Given the importance of the 2006 elections, we believe the data from CCES will reveal valuable patterns of behavior and serve as an important baseline for similar studies in years to come.[1]

Correspondence Address: Lynn Vavreck, Department of Political Science, UCLA, Los Angeles CA 90095, USA. Email: lvavreck@ucla.edu

ISSN 1745-7289 Print/1745-7297 Online/08/040355-12
DOI: 10.1080/17457280802305177

The Idea of CCES

Traditionally, political scientists interested in studying US midterm elections relied on the National Election Study (NES) for data on voters in these elections. The traditional NES design consisted of generating a nationwide representative sample using area probability sampling with clustering. Respondents were mainly interviewed in person. Over time, however, this design produced interviews with a relatively small number of actual midterm voters – most of whom, by design, lived in non-competitive congressional districts. An inability to stray from this basic design increasingly produced datasets that were of minimal value to scholars interested in contemporary questions of elections and representation. The NES could not increase the size of the sample due to cost constraints and a reluctance to move to other less expensive modes of interview only exacerbated this dilemma. The 2002 midterm election study was the last one the National Science Foundation funded. Thus, without CCES and Ted Carmine's CSR study, scholars would have missed the opportunity to study an important alteration of the political landscape.

When the NES announced it would no longer run a midterm election study, scholars interested in work on Congress, representation, and state politics were left wondering how they could facilitate their work in the future. While the limited power and coverage of the traditional NES midterm studies constrained the kinds of analyses that could be done, the data – and the time series in particular – were valuable to the community. There were, however, new questions to answer and fresh puzzles begging for a different kind of project. Quickly, people began to talk about what could be done to improve upon what the NES had started and expand it so that more complicated questions and more precise answers could be examined.

In spring of 2003, we began thinking about better ways to conduct political surveys leveraging developments in modern technology. We liked the idea of using Secretary of State Voter Registration Files to help with sampling challenges and we knew that somehow the internet or cell phones would be involved. Quickly, the idea of matching members of a large volunteer panel to random samples drawn off voter registration files emerged as a way to overcome the sample quality problems that most internet survey firms experienced. Rivers worked through the theoretical underpinnings of the methodology and determined that if there were a large number of people in the panel, the matching method would produce representative samples with low levels of total error (Rivers, 2007).

With the help of a small team of people and under the banner of a new research company called Polimetrix, we began to recruit people to join the *PollingPoint* panel in winter 2004.[2] By winter 2005 there were over 1 million people in the panel and we began to think about ways we might conduct an election study in 2006. We decided we would try to leverage the size and coverage of the panel to interview people from every state, or if possible every Congressional District. Through this design, we could ask respondents to "vote" on some of the same bills their representatives in Congress voted on in the last session, and then we could scale voters the same way political scientists had scaled politicians (Poole & Rosenthal, 1997; Clinton *et al*., 2004). Moreover, questions about the effectiveness of campaign advertisements, ground campaigns, and various other forms of electioneering could be sorted out if the design included congressional districts that differed in their levels of campaign intensity and effort. The study could have a very large number of respondents allowing scholars to answer questions that had previously been impossible to address using NES data.

Over roughly the same period of time, Steve Ansolabehere (MIT) began talking with American politics scholars about fielding a midterm election study that could fill the void left by the NES. He started a program at MIT called the Public Opinion Research Training Lab (PORTL), the purpose of which was to give students a chance to work with a survey from start to finish over the course of a semester. For this project to work, the survey instrument had to be written, fielded, and the data returned extremely quickly. The internet seemed like the perfect mechanism to deliver data under these circumstances. The confluence of Polimetrix and PORTL resulted in the design of CCES with Steve Ansolabehere filling the role of Principal Investigator (PI) on the project.

The Structure of CCES

The organizational structure of CCES consisted of the PI and a Design Committee responsible for the content of the common portion of the study. Members of the Design Committee included the PI (Steve Ansolabehere), Study Director (Lynn Vavreck), CEO of Polimetrix (Doug Rivers), and also Bob Erikson (Columbia), Don Kinder (Michigan), Liz Gerber (Michigan), John Sides (George Washington), Jeremy Pope (BYU), and Wendy Rahn (Minnesota). The role of the Study Director was to recruit participants to the project, facilitate the design of team questionnaires, and interface with the staff at Polimetrix, which would field the project. The PI and design team were responsible for the content of the study that would be common across all respondents. They also consulted with Polimetrix about profile and background data. The project was funded by the participating teams, each of which paid $15,000 for background data, profile data, voter registration data, 15 minutes of unique team-driven survey content on 1,000 people interviewed twice, 15 minutes of survey content on a 36,443 person representative sample of the general population. The general invitation to participate in the project was extended in spring 2006 and word of the opportunity spread quickly. By August, there were over 30 participating teams and the design committee had finalized the content of the common portion of the survey. As teams joined the project they raised many requests and ideas that were unanticipated. Mainly, we underestimated people's appetite for survey experiments and learned that many of the people who joined CCES were doing so because they could design and field their own experiments.

The CCES Survey Experience

The field period for CCES was the second week of October 2006 through Election Day for the pre-election wave. The post-election wave was fielded in the two weeks after Election Day, and began on Wednesday morning after the election. Every time a team joined CCES another 1,000 respondents were added to the total number of cases in the project. In this way, each respondent in the study belonged to one of the teams. A typical respondent's pre-election survey experience consisted of 10 minutes of survey content common to all respondents (Common Content) and then 10 minutes of content specific to a team (Team Content). Every respondent in the project received both parts of the survey. The total survey length was 20 minutes. The post-election survey had the same design but was half the total time (5 minutes each of Common and Team content).

The Common Content portion of the survey contained measures tracking respondent's levels of knowledge about the state and federal elections taking place in their state, knowledge of their Senator's voting record on specific pieces of legislation that the Senate voted on in the previous year, and the respondent's own "votes" on these same roll calls. Additionally, the Common Content contained measures of people's political predispositions and assessments of candidates and office holders, levels of attention and usage of media, and indicators for political knowledge and interest. Profile data, previously collected by Polimetrix, contained data on background information such as demographics, detailed media usage, and various issue positions.

The CCES teams were very creative in designing their studies. Topics of interest spanned subfields. Gary Jacobson investigated the war in Iraq and how it affected voting behavior in the 2006 elections. Walt Stone, Alan Abramowitz, Michael Herron, and Joe Bafumi tackled questions of ideological positioning. John Sides and Jack Citrin investigated misperceptions about immigrant populations; and Brian Gaines and Jim Kuklinski analyzed decisions about the Illinois state lottery. Several scholars leveraged the survey's large sample to study the role of religion in political decision-making (John Green, David Campbell, and Simon Jackman). Teams conducted experiments on television advertising (Christian Grose and Suzanne Globetti), radio advertising (John Geer and Lynn Vavreck), and background appeals (Brian Arbour). Others merged observational data on advertising buys with the survey data to estimate the decay of advertising effects (John Zaller, Lynn Vavreck, Seth Hill and James Lo); and local news received coverage from the Wisconsin team (Erika Fowler and Ken Goldstein). Two IR scholars separately joined the project in order to empirically estimate audience costs and the effects of reputation and political party on presidents' abilities to use diplomacy in the international arena (Robert Trager and Ken Schultz, respectively).

The internet provides many unique opportunities to leverage visual and audio content and several teams were motivated in designing their projects by the widgets and capabilities Polimetrix developed. Further, the internet allows interviews to be completed rapidly and invitations to complete surveys can go out simultaneously, unlike phone or face-to-face methods, which require manpower in order to complete the interviews. The speed with which the interviews can be completed means that samples can be mainly cleared within the period of a few days. This minimizes problems of differential effects due to time of interview.

Sampling Methodology

Prior to the 1970s, nearly all survey interviewing was conducted in person or by mail. High quality surveys were conducted in person using area probability samples, while much market research was performed on mail access panels using quota sampling. With the advent of random digit dialing (RDD), an intermediate possibility arose: it was possible to construct a respectable probability sample from phone numbers and save substantial amounts by conducting interviews over the phone. Nearly all media polling and most academic surveys, except a few large and generously funded projects such as the NES and the General Social Survey (GSS), quickly moved to the phone. Most households had telephones and initially response rates were quite high. However, over time response rates have deteriorated so that most media polls now have response rates around 20%. With enough time and effort, it is possible to achieve response rates of approximately 40% with RDD (largely by reducing non-contact), but there is little evidence that the additional time and expense is worth the effort (Holbrook *et al*., 2007).

The advent of the World Wide Web in 1991 created yet another possibility for interviewing. A number of large opt-in web access panels were created in the 1990s and now dominate survey data collection for market research. However, unlike phone numbers, there was no obvious way to sample email addresses, so that most web surveys were conducted using convenience samples (often involving quota sampling). Naturally, most surveys conducted on mail panels, which had long ago abandoned probability sampling, were among the first to migrate to the web and the traditional mail panel vendors now operate some of the largest web panels. Few in the academic or media world (at least in the US) were willing to sacrifice probability sampling (in the form of RDD) for the economies of web interviewing if it required a switch to quota sampling.

Non-probability samples, however, are a reasonable approach for certain types of problems. There is little argument that convenience samples are adequate for experimental studies, even when the conclusions are intended to apply to some larger population. These are essentially model-based inferences that come from assuming that the experimental effects are homogeneous within the relevant population. Similarly, substantial levels of non-response require model-based adjustments. Any inferences from such samples depend as much upon the validity of unverifiable assumptions as on random selection. There is no logical difference between the type of modeling assumptions needed for non-response adjustments and those needed for self-selected samples.

In the case of web survey panels, all methods of recruitment (including those that start with some form of probability sampling) will inevitably involve some degree of self-selection. Without adjustment, survey estimates based upon such samples will be biased. Conventional methods of adjustment, such as quota sampling or post-stratification based upon a few demographic variables, are inadequate to address these biases (see, for example, Couper *et al.*, 2007).

At Polimetrix, and for CCES, we use a method called sample matching to construct a representative sample of the general population. The method of sample matching simultaneously reduces bias and improves efficiency. The availability of large amounts of auxiliary information from consumer and voter databases make it feasible to select a sample that is approximately balanced on a large set of variables. Sample matching is a cost-effective method for constructing samples with minimal bias. With sample matching, a population frame that includes large amounts of auxiliary information is used to select a target sample using known probabilities of selection. For each element of the target sample, the closest matching element from the panel is selected for interviewing. Because of imperfect matching, the resulting sample still needs to be weighted, but the weights are much smaller than would be needed for either a random sub-sample or a quota sample.

Constructing the Panel

There are fundamentally two problems with sampling for web surveys: lack of coverage for persons without internet access and non-random selection. The under-coverage problem, while not insignificant, is much less serious today, with roughly 70% internet penetration, than in 1998 when fewer than a quarter of US households had internet access. Eventually, the under-coverage problem is likely to disappear, much as it did for telephones in the 1950s and 1960s. However, non-random selection continues to be a problem as there is no ready analog of either area probability sampling or random digit dialing for internet users.

Further complicating things, before a company like Polimetrix can interview anyone, respondents have to give the company permission to send them email. For Polimetrix, everyone who does this is considered a member of the *PollingPoint* panel. Polimetrix attracts people to the panel by conducting short, entertaining surveys on the internet about a variety of popular topics that people choose to take because they are interested in sharing their opinions. In this way, the panel consists of people who are interested in NASCAR, cooking, knitting, movies, books, and a whole host of other topics. Polimetrix views this panel as a pool for *purposive* not random sampling. Typically, advertisements for these short surveys are placed on banners of popular web pages and people surfing the internet click on the banner because they want to share their thoughts on exercise or Harry Potter or gardening. Examples of recruitment advertisements can be viewed online at www.pollingpoint.com.

Sample matching is a purposive method for creating a sample when a large, but possibly unrepresentative, pool of respondents is available for interviewing that can be matched to units in the sampling frame according to some auxiliary variables. The fundamental idea is that one first selects a target sample from the sampling frame using some form of random sampling. However, instead of interviewing those people in the target sample, one finds the closest match in the pool of available respondents to each unit in the target sample. Collectively, the matched units are called the matched sample and they will resemble the target random sample in terms of the variables used for matching. The matching need not be exact – matching is usually performed using a distance function that measures the similarity between a pair of respondents, but if the pool of available respondents is sufficiently large and diverse then the matched sample is guaranteed to have approximately the same joint distribution of the matching variables as the target sample. Matched samples can be used *as if* they were random samples (Rivers, 2007). That is, the observations in the matched sample are nearly independent and have nearly the same distribution as a random sample from the target populations. However, the needed panel size grows rapidly as the number of characteristics used for matching increases.

For CCES, the target population was the general population and we wanted a 38,000-person sample. Following the process above, Polimetrix drew a random sample of this exact size from the 2004 American Community Study (ACS), conducted by the US Bureau of the Census, which is a probability sample of size 1,194,354 with a response rate of 93.1%. For each respondent in the Polimetrix-drawn ACS sample, the closest matching active Polimetrix panelist was selected using a weighted absolute distance measure on four Census variables – age, race, gender, and education, plus on imputed values of partisanship and ideology. The matching is based on joint distributions of the six variables. Following matching, the sample marginals are raked to the ACS marginals for age, race, gender, and education.[3]

Findings

Tables 1 and 2 use CCES data to present the percentage of likely voters in each state (with a sample of at least 300 likely voters) intending to vote Democratic for either Senator or Governor, along with the actual vote outcome (undecideds and minor party voters are deleted, except in Connecticut). Confidence intervals were computed assuming ignorable selection using the approximation given in Section 5 of Rivers (2007) and are shown in the accompanying figures.

As can be seen from Tables 1 and 2 and Figures 1 and 2, the estimates appear to be approximately unbiased. However, the coverage of the 95% confidence intervals is somewhat below the nominal level.

In comparison, Mark Blumenthal and Charles Franklin at Pollster.com (2007) compared the CCES estimates with the results of conventional RDD telephone surveys (with live interviewers) and IVR (Interactive Voice Recording) interviews. The results are shown in Table 3. In this election, sample matching out-performed RDD samples (presumably using conventional weighting by either cells or raking),

Table 1. CCES senate election predictions

State	N	Predicted Vote (%)	Actual Vote (%)
Arizona	798	47.9	45.3
California	1015	67.8	63.1
Connecticut	401	47.8	44.4
Florida	1005	63.8	61.3
Massachusetts	799	71.3	69.5
Maryland	802	53.1	55.5
Michigan	800	57.9	58.0
Minnesota	501	59.4	60.5
Missouri	802	50.0	51.1
New Jersey	500	53.0	52.8
Nevada	402	44.2	42.5
New York	1011	72.9	68.0
Ohio	1003	59.2	55.9
Pennsylvania	1005	58.3	58.6
Tennessee	502	47.4	48.6
Texas	1004	30.9	36.9
Utah	402	34.0	33.0
Virginia	802	50.0	50.1
Washington	804	57.0	59.7
Wisconsin	502	74.2	69.5
West Virginia	301	67.0	65.7

whether a live interviewer was used or IVR. Another web survey (Zogby Interactive) using a different methodology was substantially worse than either the RDD samples or the matched Web sample. The sample sizes in the phone samples tended to be somewhat larger (typically between 600 and 1,000 interviews per state), so their standard errors before weighting would be smaller than the matched sample from Polimetrix. It is unclear whether the standard errors are larger or smaller after weighting. However, all of the other surveys have substantial amounts of bias compared to the matched sample. The N in Table 3 represents the number of predictions made using these data.

The Root Mean Squared Error (RMSE) of an estimate is a combination of both the sampling variability and the bias. What is perhaps most striking from Tables 1 and 2 is that the actual RMSEs for most of the samples are roughly three to four times the reported sampling error. This is because *all* of the methods are subject to some bias, which is not taken into account in the calculation of a margin of error. The reported standard errors appear to give an accurate measure of sampling vari-

Table 2. CCES gubernatorial election predictions

State	N	Predicted Vote (%)	Actual Vote (%)
Alabama	505	42.0	42.0
Arizona	798	58.2	64.0
California	1015	44.7	41.2
Colorado	500	60.0	58.0
Connecticut	401	33.3	35.9
Florida	1005	43.8	46.4
Georgia	804	41.6	39.9
Iowa	301	54.0	54.8
Illinois	800	61.0	55.6
Kansas	501	61.0	58.8
Massachusetts	799	65.6	61.2
Maryland	802	51.0	53.8
Michigan	800	56.1	57.1
Minnesota	501	52.1	49.5
Nevada	402	46.3	47.8
New York	1011	74.2	70.3
Ohio	1003	63.9	62.1
Oregon	502	54.0	54.2
Pennsylvania	1005	62.2	60.3
South Carolina	399	44.0	44.8
Tennessee	502	67.0	69.8
Texas	1004	39.4	43.3
Wisconsin	502	55.2	53.8

ability, *but ignoring bias means that reported confidence intervals are much too narrow*. Even for the matched estimator, which had the lowest level of bias, the nominal 95% confidence interval appears to have coverage closer to 90%. Even so, the matching estimator from CCES outperforms conventional estimates based on RDD phone surveys.

Concluding Thoughts

There are some who argue that non-probability samples are not usable for scientific inference. However, large portions of statistics are devoted to situations in which the data generating process is unknown and must be modeled. Every observational study is of this type. If we were to decline to make probability statements about anything but random samples, we could not make weather forecasts (for

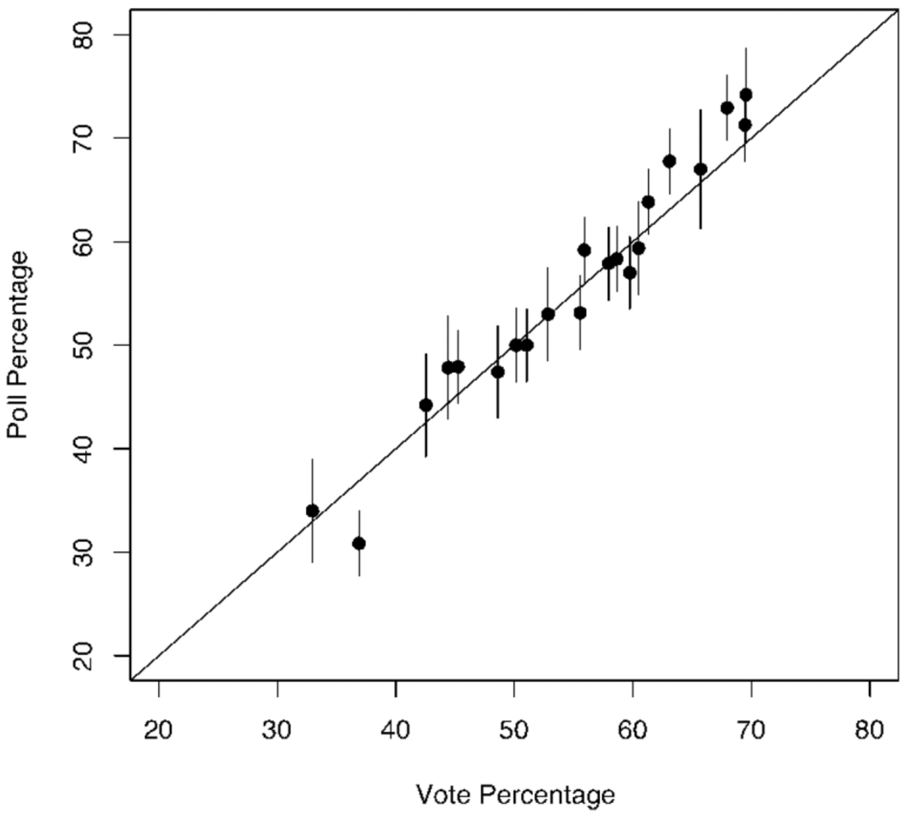

Figure 1. CCES senate estimates and 95% confidence intervals.

example, “the probability of rain tomorrow is 30%”). Most medical research, which involves randomization of treatment, but not random selection of participants, would be restricted to saying that the estimated treatment effect applies only to the small set of persons who participated in the experiment. In the case of web surveys like CCES, it is unlikely that even in-person recruitment will provide a sample without substantial amounts of self-selection. For many purposes, especially in the social sciences where incidence rates of the things we care about are likely to be small, opt-in panels represent a cost-effective alternative for increasing power when one has some confidence that the matching variables are sufficient to eliminate most of the potential bias.

As a discipline, we must continue to examine public attitudes and behavior even though the federal government may no longer fund this work. By coordinating our efforts as we did in CCES, we can generate greater numbers of observations and share the costs. The beauty of this design is that as power increases, the cost per participant goes down. By leveraging the extant literature in statistics and weighing

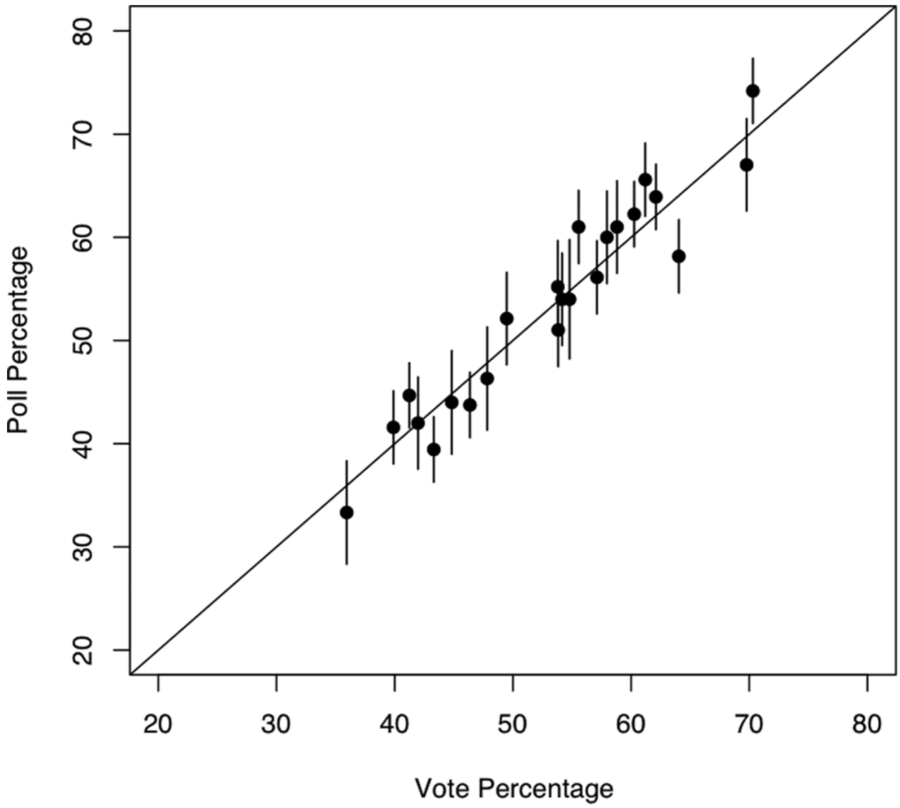

Figure 2. CCES governor estimates and 95% confidence intervals.

Table 3. Comparison of RDD and matched samples

Source	Polls (N)	Bias	RMSE
Phone	255	2.76	8.34
Rasmussen (IVR)	83	3.82	8.47
SurveyUSA (IVR)	63	3.4	7.25
Zogby (internet)	72	4.86	9.36
Polimetrix (internet)	40	–0.47	5.21

it in light of political relationships, as Polimetrix has done with the model assisted approach called sample matching, we can use new technologies to both increase the power of our tests and keep the costs of survey research affordable. These two innovations taken together provide promise for the future of observational research on political attitudes and behavior.

Notes

1. The other survey in the field during 2006 was a phone survey conducted by Edward Carmines by the Center for Survey Research at Indiana University.
2. These people included Brian Gaines (Illinois), Jeff Lewis (UCLA), Brad Palmquist, Drew Linzer (Emory), Jenny Blake, and Brian Stults. Everyone on the team was connected to political science in some way. Brad Palmquist was on the faculty at Harvard and Vanderbilt in Political Science, Brian Stults and Jenny Blake were undergraduate students of mine at Dartmouth College and UCLA, respectively; and Drew Linzer was a PhD student of Jeff Lewis's at UCLA.
3. Raking was performed using iterative proportional fitting. The final weights were trimmed to lie between .33 and 3.

References

Ansolabehere, Steve (2006) *Cooperative Congressional Election Study* (Palo Alto, CA).

Clinton, Joshua D., Jackman, Simon & Rivers, Douglas (2004) The statistical analysis of roll call voting: a unified approach, *American Political Science Review,* 98(2), pp. 355–370.

Couper, Mick P., Kapetyn, Arie, Schonlau, Matthias & Winter, Joachim (2007) Non-coverage and non-response in an internet survey, *Social Science Research*, 36, p. 131.

Holbrook, Allyson L., Krosnick, Jon A. & Pfent, Alison (2007) Response rates in surveys by the news media and government contractor survey research firms, in: James M. Lepkowski, Clyde Tucker, J. Michael Brick, Edith D. De Leeuw, Lilli Japec, Paul J. Lavrakas, Michael W. Link, Roberta L. Sangster, (eds) *Advances in Telephone Survey Methodology* (New York: Wiley).

Jacobson, Gary (2007) The president, the war, and the 2006-midterm elections. Paper prepared for presentation at the Annual Meeting of the Midwest Political Science Association, Chicago, Illinois.

Poole, Keith & Rosenthal, Howard (1997) *Congress: A Political-Economic History of Roll-Call Voting* (Oxford: Oxford University Press).

Rivers, Douglas (2007) Sampling for web surveys. Paper presented at Joint Statistical Meetings, Salt Lake City, Utah.