Interactive Voice Response Polling in Election Campaigns: Differences with Live
Interview Surveys

Alexander C. Brunk

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State
University in partial fulfillment of the requirements for the degree of

Master of Arts
In
Political Science

KAREN M. HULT, chair
CRAIG L. BRIANS, co-chair
WAYNE D. MOORE
XAVIER MEDINA VIDAL

December 5, 2014
Blacksburg, VA

Keywords: polling, public opinion, interactive voice response, U.S. elections,
political campaigns, survey research, IVR

Interactive Voice Response Polling in Election Campaigns: Differences with Live Interview
Surveys

Alexander Crowley Brunk

ABSTRACT

Since the early 2000s, Interactive Voice Response (IVR) has become a widely popular
method of conducting public opinion surveys in the United States.   IVR surveys use an
automated computer voice to ask survey questions and elicit responses in place of a live
interviewer. Previous studies have shown that IVR polls conducted immediately before elections
are generally accurate, but have raised questions as to their validity in other contexts.

This study examines whether IVR polls generate measurably different levels of candidate
support when compared to live interviewer polls, as a result of non-response bias owing to lower
response rates in IVR surveys. It did so by comparing polling in 2010 U.S. gubernatorial and
U.S. Senate elections that was conducted using both live interviewers and IVR. The findings
suggest that in general elections, IVR polls find fewer undecided voters compared to surveys
conducted using live interviewers. In primary elections, IVR polls can show larger support than
live interview polls for a more ideologically extreme candidate who has high levels of support
among more opinionated and engaged voters.

Implications are that journalists and other consumers of polling data should take into
account whether a poll was conducted using IVR or live interviewers when interpreting results.
IVR polls may tend to over-sample more engaged and opinionated voters, often resulting in
smaller percentages of undecided respondents, and higher levels of support for specific
candidates in certain contexts.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

In recent years, public opinion polls have become a critical part of public discourse in American politics. In the twenty-four hour news environment poll results are a constant source of discussion and debate. Cable news shows, newspapers, blogs, and other media sources constantly report on the latest poll numbers. Pundits on talk shows argue for or against the validity of individual polls, and political professionals and casual observers of politics study poll results in an attempt to glean valuable information. Journalists have found that reporting on the latest poll data is a quick and easy way to write a story and meet a deadline, particularly in the face of declining staff, revenue, and circulation at many news organizations (Goidel 2011, 14). At the same time, cuts to newsrooms mean that these same organizations often lack the resources to thoroughly analyze the reported poll numbers (Dunaway 2011, 75). The result is that we read and hear a lot about poll results but not enough about what they mean.

Technology has lowered the barriers of entry to the polling industry, leading to an explosion of polls. Decades ago, only a few major firms and university research centers had the resources and equipment to conduct regular public opinion polls for wide release. Advances in computer technology have made nearly every aspect of conducting a poll easier. Automated machines dial telephone numbers, eliminating the need for interviewers to dial by hand to reach respondents. Sophisticated statistical software packages are available at fairly low cost, or even for free. Email allows poll results to be transmitted within minutes of calls finishing. In 2014, it is easier to commission a poll than it has ever been.

In hotly contested, high profile political campaigns, polling is done for many months leading up to the election. Some pollsters release polls on a regular schedule, such as once a month. Others may conduct polls in response to specific events, such as a scandal or the entry of

a new candidate into the race. In many cases dozens of polls are released over the course of a campaign. These polls feed the demand for news stories and can help shape the narrative of campaigns.

The validity of polls conducted well in advance of an election is difficult to gauge because it is unclear what such polls are actually measuring. These polls purport to give the public a glimpse of how the campaign is taking shape at particular times, but without an election immediately following there is no way to tell whether they are correct. A poll of likely voters conducted two days before a major election can be fairly evaluated based on the election results. A poll conducted weeks or months earlier is not necessarily predicting an outcome, since the levels of support for the candidates will often change between the date the poll was conducted and the date of the election. The resulting poll results are very difficult to grade. As pollster Mark Mellman puts it, "there is no objective standard for being 'right' prior to Election Day" (Mellman 2010). When these issues are discussed in this thesis, I will use the term "accuracy" to indicate the predictive accuracy of a poll, or how well the poll result matches the ultimate election result. On the other hand, I will sue the term "validity" to describe how well the poll does in measuring the current state of voting preferences at the time it was taken.

A relatively new method of polling that has become widely used in recent years is Interactive Voice Response, or IVR.   Interactive Voice Response polls are conducted solely through the use of a computer system that reads an automated message to respondents on the phone, eliminating the need for live interviewers. Some of the most commonly cited polling organizations in the news, including Rasmussen Reports (Rasmussen Reports 2012) and Survey USA (Survey USA 2012), use IVR to conduct their polls. Without the need to pay interviewers, IVR polls are much less expensive to conduct, lowering the barriers to entry for conducting

them. (Goidel 2011, 22) Between 2002 and 2010 the percentage of publicly reported election

polls conducted using IVR more than doubled (Blumenthal 2011a).

Many in the survey research industry initially viewed Interactive Voice Response polls

with a high degree of skepticism. Scott Keeter, Director of Survey Research for the Pew

Research Center, summarizes these concerns:

> [Interactive Voice Response] polls violate many elements of the survey canon's best
> practices – they have only a rudimentary ability to select a respondent within a
> household, no live interviewer to establish a rapport with respondents, little assurance
> that respondents are actually adults and eligible to participate, no ability to provide
> answers to questions about the purpose of the study or who is sponsoring it (Keeter 2011,
> 43)

However, past research has found little evidence that these and other concerns about IVR

methodology affect the accuracy of polls conducted using this approach. Studies conducted by

the American Association for Public Opinion Research and the National Council on Public Polls

have found that IVR surveys conducted immediately prior to elections are no less accurate than

those conducted using live interviewers (AAPOR Ad Hoc Committee on the 2008 Presidential

Primary Polling 2009) (National Council on Public Polls 2011). The AAPOR report in particular

provided a strong endorsement of IVR from the public opinion research industry's leading

professional association.

Analyses of the accuracy of IVR polls have focused on polling conducted immediately

prior to elections and have not considered polls conducted weeks or months in advance. Indeed,

nearly all analyses of aggregate polling data focus on pre-election polls because the similarity of

a poll to the actual election result serves as the best indicator of a poll's validity. Determining

the validity of polls conducted earlier in a campaign is a difficult task: as previously noted, there

is no benchmark for what it means for a poll to be "right" three months before an election

(Mellman 2010). This limits our understanding of IVR methodology, a fact not lost on those in the industry. Pollster Mark Blumenthal argues that though IVR polls conducted immediately prior to elections have been shown to be accurate, that does not mean that this finding can be generalized for all IVR polling (Blumenthal 2009). It is not well understood if IVR polls are valid in other situations.

This gap in understanding presents a problem in a world where polls drive campaign narratives and serve as an impetus for key decisions. Campaigns send out press releases touting favorable poll numbers and send them to journalists in the hope of generating favorable news coverage. Journalists write articles about poll results and use their knowledge of the latest polls to frame other news stories about a campaign. Professional fundraisers show poll results to donors to convince them that a donation could make a difference in a tight race. Websites such as Pollster.com, RealClearPolitics, or FiveThirtyEight include individual polls as part of aggregates or models to predict election outcomes. This information is in turn read by journalists, political professionals, and casual observers of politics. If IVR and live interviewer polls do sometimes produce different results, then the decision making of all of these actors should reflect this.

There is very little scholarship in political science or on survey methods more generally that discusses IVR and live interviewer public opinion polling. As more and more public opinion polling is being conducted using IVR, this is an important area of focus for researchers. The availability of years of compiled poll results on websites like Pollster.com and RealClearPolitics.com have allowed researchers to do research on these aggregate data that would have been impossible only a decade ago. Without a solid understanding of the differences

between IVR and live interviewer polls, researchers may include both types of polls in reports and models without accounting for potential differences in the data owing to the technique used.

This thesis examines whether Interactive Voice Response polling produces different estimates than traditional live interview polling in situations other than those immediately prior to an election. It considers existing research on IVR survey methodology in general as well as applications to public opinion polling specifically. I want to determine whether any systematic differences between IVR and live interview polling can be identified and to provide some possible explanations for these differences. This study also reviews some of the primary methodological challenges for public opinion polling in 2014, and how these challenges may be of particular concern to pollsters using Interactive Voice Response.

The outline for this thesis is as follows. Chapter 2 discusses survey research methods as they have been traditionally applied in public opinion polling in the United States, with a specific discussion of non-response bias. Chapter 3 examines Interactive Voice Response as a survey method in general and as a tool for public opinion polling specifically. Chapter 4 lays out the research methodology and statistical techniques that were used for the analysis of aggregate polling data. Chapter 5 contains an analysis of undecided rates in general election polling data in the 2010 U.S. Senate and gubernatorial elections, and finds support for the expectation that IVR polls tend to find fewer undecided voters. Chapter 6 includes an analysis of both undecided rates and candidate preference in four 2010 primary elections for which sufficient data are available, finding that in some cases differences in candidate preference can be explained by whether or not IVR was used to conduct a poll. Chapter 7 concludes by summarizing the findings from the previous chapters and how they should impact the interpretation of IVR polling results going forward.

CHAPTER 2: Survey Research Methods in Public Opinion Polling

This chapter discusses the methodology underlying survey research, how it has been used to conduct political polls, and some of the challenges that pollsters have faced over time as they have sought to measure public opinion. It contains a discussion of poll accuracy and concludes with a specific review of the problem of survey non-response and the literature examining its impact on political polling.

Sampling

In order to examine the attitudes and opinions of any large population, a researcher will draw a sample from that population and then survey them. A sample, by definition, is simply a set of cases drawn from a larger population that is smaller in number than the total population (Underwood, et al. 1954, 106). A random sample drawn from a target population for the purposes of measuring the total population is known as a probability sample. Basic statistical sampling theory tells us that a random sample of size n drawn from a population p will accurately represent the population within a specified level of confidence. The Central Limit Theorem states that as n grows larger, the means of samples of the same size will approach a normal distribution. Thus a larger sample will allow for a higher level of confidence and a smaller margin of error (Knoke and Borhnstedt 2002, 80-82). A margin of error is simply the standard error of a survey proportion that is multiplied by a chosen number of standard deviations to ensure that at a selected level of confidence, the survey proportion is identical to the "true proportion." The formula for the standard error of a proportion is:

$$\sigma_{prop} = \sqrt{\frac{pq}{N}}$$

[here p = any hypothetical true proportion; q = 1-p; N = sample size (Underwood, et al. 1954, 118)].

These results assume simple random sampling, which often is difficult to achieve in practice. This assumption is that every member of the target population has an equal chance of being included in the sample being surveyed (Knoke and Borhnstedt 2002, 70). Another way of stating this is that every member of the population is equally likely to be contacted by the interviewer. In virtually every real world scenario, however, there are problems with drawing a random sample. In the real world, the odds of each person in the target population being included in the sample that is surveyed is not the same. Furthermore, the sample frame that the sample is drawn from will never completely represent the entire population. For example, voter files, residential addresses, or lists of telephone numbers might be treated as being the total population, but such lists are never 100% complete or accurate. Thus, there is the potential for a biased sample if the people who end up taking the survey are not representative of the target population.

In the context of elections and political campaigns, a survey of a sample conducted in order to measure voting preferences in an upcoming election is usually referred to as a public opinion poll, or just a poll. The researchers who conduct these surveys are referred to as pollsters. If a poll is conducted only a few days prior to an election, it is referred to as a pre-election poll. Polls can measure many things, including attitudes toward particular issues and responses to information (positive or negative) about specific candidates. Most importantly, they often include questions about who each respondent intends to vote for in an upcoming election, which is known colloquially as the "horse race" or ballot question. In the results of these polls, responses to horse race questions are represented as proportions indicating the level of support

for each candidate (e.g. 49% for Barack Obama and 48% for Mitt Romney). Here, when I refer

to a poll, I am referring to a survey of public opinion in the United States that is designed to

measure voting intent in an upcoming election, and poll results refer to the specific result of the

horse race ballot question in the survey.

Pollsters have struggled with sampling since the early days of the industry, and they have

had to adjust their methods to meet these challenges. In 1936 *The Literary Digest* surveyed more

than two million of its subscribers about the upcoming presidential election without attempting

to create any kind of random sample. As a result, the publication predicted that Republican Alf

Landon would defeat President Franklin Roosevelt for the presidency. The results proved to be

quite inaccurate, as the subscriber list skewed heavily Republican and was not representative of

the American voting population. At the same time, pollster George Gallup developed a fairly

accurate estimate of the election results by sampling only about two thousand voters nationwide.

Twelve years later in 1948, the quota sampling methodology Gallup developed failed to predict

the reelection of President Harry Truman over challenger Thomas Dewey. Pollsters adjusted

their methods again, with a renewed effort to use random sampling to choose which voters to

interview (Traugott and Lavrakas 2008, 14).

In addition to difficulties with obtaining random samples, pollsters also face the issue of

non-response bias. Non-response bias refers to any situation where the responses given by those

who agree to take a survey vary from those who do not. This complicates the work of a survey

researcher, because it calls into question the validity of the statistical theory cited above for

determining the level of confidence in the sample. The math behind survey research essentially

assumes that non-response does not exist (Groves 2006, 647). A researcher can draw a random

sample from a target population that is representative of that population, but unless every single person in the sample takes the survey, then non-response is a concern.

In the context of public opinion polls, non-response bias can occur for a variety of reasons. Not all voters answer the phone or can be easily reached. Some are not home during the times that most polls are conducted. Older voters generally are easier to reach than younger voters. Having a higher percentage of respondents who answer a survey does not necessarily eliminate bias. A survey can have a 90% response rate and still have problems of non-response bias if the ten percent who do not agree to take the survey would have answered questions differently. "The key determinant of bias is whether the propensity to respond to a survey is correlated with important variables of interest" (Keeter 2011, 31). These "variables of interest" can be many things, including demographics, party affiliation, and level of interest in politics.

Polling Procedures

Pollsters have continuously adjusted their sampling methods over the years based on both the availability of new technologies and the evolution of human behavior. Prior to the early 1970s, most polls were conducted by sending interviewers door-to-door and speaking with respondents' in-person. However, the near-universal adoption of the landline telephone provided a much easier – and safer and cheaper – means of conducting public opinion polls. People were also much more likely to be home for a door-to-door interviewer to speak with decades ago than in 2014. Since the 1970s, the vast majority of public opinion polling has been conducted using telephone interviews. Only recently has this begun to change, with the introduction of surveys conducted over the Internet. Firms such as Zogby, Harris Interactive, and Knowledge Networks now regularly release the results of public opinion polls conducted online. However, the

difficulty of drawing a random sample of voters using the Internet has meant that thus far, it has largely failed to displace telephone interviewing as the preferred method of conducting public opinion polls (Traugott and Lavrakas 2008, 140).

Traditionally, telephone polls are conducted by dialing random phone numbers in a sample drawn from a target population, and then attempting to interview the people who answer and record their responses to survey questions. A pollster begins by acquiring a large sample, either a list of phone numbers or a phone-matched voter file and then attempts to contact a pre-determined number of voters and survey them as to their voting intentions. An auto dialer machine dials phone numbers and connects live interviewers to a call only when the system detects that a person has answered the phone. If there is no answer or the system reaches a voice mail message, it simply hangs up and moves on to the next number. Once connected, the interviewer asks the person on the line if he or she would be willing to take a survey, and if the person agrees the interviewer commences with the survey questions. During the interview, the interviewer reads the questions from a computer screen and types the interviewee's responses into the computer. This system is termed Computer Assisted Telephone Interviewing (CATI). Depending on the pollster and the data he or she wishes to collect, the poll can be just a few questions, or it can last for ten or fifteen minutes.

Public opinion polls vary both in length and content, depending on who is running the poll and what its purpose is. Pollsters who conduct polls for media organizations tend to run relatively short surveys that chiefly measure the horse race, which simply estimates where the candidates are on a ballot test at the moment the poll is taken, and the public images of officeholders and candidates. On the other hand, internal campaign pollsters conducting research for use in planning campaign strategy tend to run longer surveys with a wider variety of

questions to assist the campaign with message testing and voter segmentation. Interested third parties such as political action committees, interest groups, and business associations may ask questions related to their areas of concern, as well as ballot tests to see if their preferred candidate is likely to win.

The survey and sampling methods used for public opinion polling are not fixed or universally agreed upon. Every pollster, whether a commercial firm or an academic institute, uses a distinctive approach. A common method of conducting public opinion polls is through Random Digit Dialing (RDD), in which random phone numbers from telephone exchanges within the political district being surveyed are sampled and dialed. RDD is widely used by academics and media pollsters. Most internal campaign pollsters, on the other hand, rely on Registration-Based Sampling (RBS) that involves sampling from phone-matched voter files. Some firms use a mix of RDD and RBS. There is some evidence that RBS polling is more accurate in lower turnout elections, but there is little consensus on which method is superior (Green and Gerber 2006, 216).

Different pollsters also take different approaches to the issue of non-response bias. Polls done by academic institutes are often conducted over several nights or even weeks in order to attempt as many contacts as possible before giving up on not reaching specific voters. Commercial polling firms tend to conduct surveys in one to three nights, meaning that there are limited opportunities for re-contact attempts for voters who do not respond (Traugott and Lavrakas 2008, 119). Partly due to the limitations involved in conducting surveys over shorter periods of time, media pollsters often do substantial post-weighting to make their samples more representative (Traugott and Lavrakas 2008, 79). Depending on the size of the weights, post-weighting can in many cases substantially alter the results of the poll.

A further challenge for accurately polling voters is the abandonment of landline telephones by large segments of the voting population. With the widespread adoption of cellular phones over the last several years, many households have dropped landline service entirely and exclusively use cell phones. This is a recent phenomenon: in early 2003 only 3% of American households were "cell phone only;" by the end of 2006 it had risen to nearly 13% (Keeter, et al. 2007, 773). By fall 2010 the Pew Research Center estimated that 25% of U.S. households had moved to using only cell phones (Keeter, Christian and Dimock 2010). A June 2012 estimate by the Centers for Disease Control put the number at over 35% (Blumberg and Luke 2012, 1). This complicates things for pollsters, since it is against federal law to use an automated dialing machine to contact cellular phones (U.S. Code 2012).[1] As a result, in order to survey respondents who are "cell phone only," call centers must dial cell phones by hand rather than using an auto dialer. This makes surveying cell phone only voters more difficult and more expensive. Not all pollsters include cell phones in their samples, though the number that do is growing. Two Pew studies from after the 2008 and 2010 elections showed a partisan gap between a "landline only" poll of likely voters and a poll that included both landlines and cell-phone only households, with polls that do not include cell phones showing a bias toward Republican candidates that grew from 2008 to 2010 (Keeter, Christian and Dimock 2010).

---

[1] "It shall be unlawful for any person within the United States, or any person outside the United States if the recipient is within the United States to make any call (other than a call made for emergency purposes or made with the prior express consent of the called party) using any automatic telephone dialing system or an artificial or prerecorded voice to any telephone number assigned to a paging service, cellular telephone service, specialized mobile radio service, or other radio common carrier service, or any service for which the called party is charged for the call "

Accuracy of Public Opinion Polls

Despite numerous potential sources of error, public opinion polls in recent years generally have predicted national and statewide election results with great accuracy. A study of national presidential pre-election polls in the 2004 general election showed that on average, 19 final pre-election polls were within about 1.7% of predicting the correct percentages of the two candidates, with similar values for all presidential elections in the last 20 years (M. W. Traugott 2005, 648-649). A 2008 study using the same methodology found that polls conducted in 2008 were even more accurate, with twenty final pre-election polls averaging within about 1.5% of accurately predicting both the percentages of the two presidential candidates and the margin between them. (Panagopoulos 2009, 898).

Studies have consistently shown that polls predicting U.S. Senate and gubernatorial outcomes also are quite accurate. The 2008 study cited above found that though polls in 2008 U.S. Senate elections were not quite as accurate as statewide presidential polls, they still were on average accurate within a few points (Panagopoulos 2009, 902). Statistician Nate Silver reported an analysis on his blog in 2010 of average poll results leading up to Senate and gubernatorial elections from 1998-2009. In it, he simply averaged all polls conducted between 30 and 60 days prior to Election Day in races where at least two different firms published poll results. He found that even 30 days out, averaging poll results was very effective at predicting the election outcome. Silver found that in general elections for the U.S. Senate, since 1998 no candidate who led by an average of more than 5.5 points 30 days before the election was defeated, a record of 68-0 (Silver 2010). A study by the National Council on Public Polls of polls conducted in the 2010 U.S. Senate and gubernatorial elections found that 77% of polls conducted after October

20th of that year had results within the margin of error for the actual election outcome, with an average candidate error of 2.4 percentage points (National Council on Public Polls 2011).

Despite their generally good track records, pollsters have had some famous misses in recent years. In the 2008 New Hampshire Democratic primary, every single poll conducted in the days immediately prior to the election showed Barack Obama leading. Ultimately, Hillary Clinton prevailed by about three points. In the 2010 U.S. Senate election in Nevada, polls prior to the election consistently showed Republican nominee Sharron Angle leading the Democrat, Senator Majority Leader Harry Reid. Yet on Election Day, not only did Reid win, but he won by more than five points. Such mishaps are evidence that although polls are remarkably effective at predicting election results, they are not perfect, and the presence of unseen sources of error occasionally cause the survey responses to be very far from the ultimate election results.

When pollsters do fail to predict major election outcomes, amateurs and professionals alike attempt to discern the reasons. The American Association for Public Opinion Research went as far as to commission a report on 2008 primary polling in response to the New Hampshire results. The results of the report were inconclusive: while the committee conducting the analysis had several hypotheses regarding the reasons for the large amount of error, they were unable to discern a single clear reason for the inaccurate results (AAPOR Ad Hoc Committee on the 2008 Presidential Primary Polling 2009, 73). Following the 2010 elections, an article on Pollster.com reported that the only polls that accurately predicted Harry Reid's victory in Nevada were the internal tracking polls conducted for the Reid and Angle campaigns. Yet since the exact methodology campaign pollsters used is a closely guarded secret, it is impossible to ascertain for certain what the reasons were that their polls were right while others were wrong (Blumenthal 2010).

Response Rates and Poll Accuracy

It is unclear how large of a problem non-response is in public opinion polling or indeed in survey research as a whole. The accuracy of polls in predicting election results has led many to believe that non-response may improve survey samples by excluding persons unlikely to vote (Blumenthal 2011a), since respondents willing to participate in surveys are also more likely to be voters (Voogt and Saris 2003, 175). Voting seems to be correlated with a propensity to take surveys. The broader literature on survey methodology has found links between non-response rates and non-response bias on certain statistics (Traugott and Lavrakas 2008, 78), but no general relationship between the two (Groves 2006, 663). Although the original theories behind survey research assume that response rates are 100%, it is unclear how much it matters if they are not since different response rates do not necessarily mean different survey results (Groves 2006, 647). While declining response rates are a concern, how much of a concern is a matter of debate.

Non-response only matters if it affects the distribution of a variable of interest to the researcher. For example, if a pollster surveys likely voters in a district with only a five percent response rate, but the percentages of Republicans and Democrats in the survey sample mirror the voter registration numbers in that district, then non-response bias on the party registration variable is clearly not an issue despite the low response rate. However, the same survey might have significant differences on other variables. Response rates to surveys have been in decline for decades, regardless of mode and topic (Blumenthal 2011a). In politics at least, the continuing ability of polls to predict election results with a high degree of accuracy means that low response rates are not *necessarily* a problem.

A notable study conducted in 2000 found that response rate had very little impact on measures of political attitudes and party affiliations. The study compared surveys of samples

drawn from the same population that varied the level of effort used to contact possible respondents. A "standard" level attempted to contact respondents over five days in which the interviewer simply spoke to the person who answered the telephone, while a "rigorous" level attempted to contact respondents over an eight-week period and randomly selected the respondent among household members. The resulting response rates were 36% for the "standard" study and 61% for the "rigorous" study. However, despite the very different levels of non-response, the two groups of respondents were nearly identical on party identification, ideology, the candidates they voted for in recent elections, and "a whole host of other political attitudes" (Keeter, et al. 2000, 139).

A nearly identical survey six years later found that although measures of attitude and ideology were similar, engagement and interest in politics differed. The second study, conducted in 2006, resulted in response rates of 25% for the "standard" sample and 50% for the "rigorous" sample (Keeter, et al. 2006, 763) – a measurable decline in response rate over the previous study. The research found that on 77 out of 84 items, the two surveys were statistically indistinguishable. However, among the measures that were different were level of engagement and interest in politics. For example, while 56% in the standard sample claimed to have voted in 2002, only 48% in the rigorous sample did (Keeter, et al. 2006, 772). The study found that "don't know" answers in the survey were higher among those included in the rigorous sample, and especially high among respondents who were the most difficult to contact (Keeter, et al. 2006, 773).

Significant evidence suggests that those most willing to participate in surveys are more interested in and engaged in politics than the general voting population. A comprehensive study of survey non-response conducted by John Brehm in 1993 found that respondents who were

more interested in taking surveys also were better informed about political topics. Brehm looked

at data from two National Election Studies conducted in the late 1980s and compared

respondents who were "amenable" (i.e. those willing to take the survey upon being asked by the

interviewer) with those who were "reluctant" (those who initially refused to participate but

agreed on a subsequent contact). Brehm found that 28% of amenable respondents could recall

the name of the incumbent in their congressional district, while only 7% of reluctant respondents

could name the incumbent (Brehm 1993, 62-63). Interviewers also asked respondents to identify

the jobs of a list of major political figures. While the amenable respondents were able to identify

40% the jobs of those named, the reluctant respondents were only able to identify 13%.   Brehm

concluded that "respondents who are amenable to be interviewed in the first place manifest

greater interest in politics and more information about politics than those who resist being

interviewed" (Brehm 1993, 62-63).

In his study of non-response, Brehm identified four main factors that influence whether

someone will choose to participate in a survey: the respondent's relationships to strangers, to the

interviewer, to the survey, and to self-image (Brehm 1993, 51-52). According to this theory, a

person will decide whether to participate in a survey based on the balance of those factors. For

example, a person who is suspicious of strangers might be less likely to agree to an interview,

while a person who enjoys speaking to strangers might be more likely to. A person more

interested in a particular survey topic might be more interested, and a person with more self-

confidence might be more willing to share his or her opinions. Although Brehm was writing

specifically about in-person studies that were conducted door-to-door, the framework is likely

similar for live interviewer telephone surveys, though the balance of influences may be different.

The experience of speaking with an interviewer on the phone rather than in person is

qualitatively different, so for example a person who is suspicious of strangers and is unlikely to open a door for an interviewer might be more likely to speak to one on the phone.

In summary, as survey response rates decline, while most survey statistics remain constant, members of actual samples tend to be more politically engaged and informed. Voters who are less amenable to taking surveys in the first place tend to be less informed about politics. When public opinion polls fail to include these less-amenable respondents, the result is a sample that over-represents better informed and more highly engaged voters.

In the following chapter, I will consider how these issues – response rates, non-response bias, and poll accuracy - are impacted by the use of Interactive Voice Response polling in place of live interviewers to conduct public opinion polls.

CHAPTER 3: Interactive Voice Response

This chapter introduces Interactive Voice Response (IVR) polling. It begins by discussing the recent emergence of IVR and its widespread use in politics, followed by a review of the survey methods literature comparing IVR and live interviewers in non-political contexts. This chapter then examines recent attempts to compare the accuracy of IVR and live interviewer political polls, followed by an exploration of potential reasons for differences in poll results

Overview of IVR

An Interactive Voice Response survey works similarly to a traditional live interviewer telephone survey, in that it begins by using a computer-controlled system to call a list of phone numbers. However, while traditional telephone surveys connect people who answer the phone to live interviewers, in IVR surveys a computer both dials the number and conducts the interview. When respondents answer the phone, they hear a pre-recorded message that asks them questions and provides instructions on how to respond. Respondents record their answers to questions by either using their touch-tone phone (e.g., press 1 to vote for Barack Obama) or by speaking their response, with voice recognition software interpreting their answers. (Traugott and Lavrakas 2008, 51). Sometimes the voice used is a recording of a celebrity, such as a local television anchor (Keeter 2011, 43). As the respondent speaks or punches keys, his or her responses are automatically entered into a computer system that collects survey data. The system then moves onto the next question, until the respondent hangs up or completes the survey.

The widespread use of Interactive Voice Response for conducting public opinion polls is a fairly recent phenomenon. Pollster SurveyUSA was the first to widely use IVR for polls beginning in the mid-1990s (Survey USA 2012). An analysis of publicly released pre-election

polls conducted in 2002 found that 17% were conducted using Interactive Voice Response (Blumenthal 2011b). By 2010, a similar analysis conducted by the National Council on Public Polls found that 41% of the polls conducted in the twenty days leading up to Election Day were done exclusively using IVR (National Council on Public Polls 2011).

Survey methodologists have historically viewed Interactive Voice Response polls as inaccurate and unscientific. In 2000, former AAPOR President Michael Traugott stated with regards to IVR: "[u]ntil there is more information about their methods and a longer track record to evaluate their results … we shouldn't confuse the work they do with scientific surveys and it shouldn't be called polling." IVR polls "violate many elements of the survey canon's best practices" (Keeter, Public Opinion Polling and Its Problems 2011, 43) [Keeter, 2011, 43 is sufficient…] and face tremendous skepticism. Many media outlets have refused to report their results. Indeed, the *Washington Post* banned the reporting of any IVR poll results until late 2010, and still limits their use in articles (Cohen 2012).

A further problem with IVR polls is their inability to include cell phones in their samples. While live interview polls can include cell-phone only households with some difficulty and at greater cost, it is illegal to use Interactive Voice Response to contact cell phones (U.S. Code 2012). Thus, as the "cell phone only" population has continued to grow over the last several years, IVR polls should be at a disadvantage in their ability to accurately survey the likely voting population. An IVR pollster thus must rely heavily on stratification or post-weighting in order to accurately construct a sample of likely voters while drawing from a pool that does not include a large segment of the target population.

Despite concerns about methodology, Interactive Voice Response polls have gained popularity largely because they are inexpensive and easy to conduct. Without the need to pay

interviewers, the only expenses associated with running an IVR poll are for the equipment,

sample, and carrier fees. An IVR system can conduct a poll of hundreds of likely voters in just a

few hours (Traugott and Lavrakas 2008, 114). Their lower cost is especially appealing to clients

(such as media organizations) who want to see poll results without paying tens of thousands of

dollars for live interviewers. The economic advantages of a faster and cheaper polling

methodology have been a powerful driving force for the adoption of IVR (Goidel 2011, 22).

Effects of Interactive Voice Response vs. Live Interviewers on Survey Results

It is well established in the survey methodology literature that contextual effects can

impact survey results. For example, respondents tend to answer some types of questions

differently when they are asked in a telephone interview versus when they are displayed on a

computer screen (Dennis, et al. 2005). Giving respondents a choice of more than one way of

taking a survey may result in higher response rates (Converse, et al. 2008). Even within the

same survey mode, subtle changes in questionnaire design can shift survey responses. Variations

in how a question is displayed on a screen during an online survey can produce significantly

different answers (Dillman and Christian, Survey Mode as a Source of Instability 2005).

Changing the order in which the same survey questions are asked can have a significant impact

on the responses received (McFarland 1981). Thus, it is unsurprising that the use of a recorded

voice in place of a live interviewer might impact both a person's willingness to take a survey and

the responses that he or she might give.

Outside the realm of electoral politics, a small body of literature compares Interactive

Voice Response and live interview survey methods. This research generally has found that IVR

surveys are less susceptible to social desirability bias but are subject to increased non-response.

Studies have repeatedly found that respondents give more honest responses to sensitive questions on IVR surveys. The same studies also have shown that IVR surveys tend to have substantially lower response rates than surveys conducted using live interviewers.

Social desirability bias refers to the reluctance of respondents to answer honestly in surveys when they believe that their response may be viewed as embarrassing or otherwise undesirable. Social desirability bias arises when a respondent perceives that there is a preferred or expected response to a survey question, referred to as a "socially desirable" response (Traugott and Lavrakas 2008, 100). If a significant number of respondents are affected by the desire to give a socially desirable answer, then it can affect the results of the survey. For example, people who are surveyed as to whether they voted in a recent election are more likely to report voting even if they did not actually vote, which can produce an overestimate of voter turnout (Traugott and Lavrakas 2008, 100). Research also has documented that respondents may be reluctant to indicate support for candidates of a certain race or sex if the interviewer conducting the survey is of a different race or sex than the respondent or of the same race or sex of the candidate being asked about (e.g. a white respondent might be less willing to tell a black interviewer that he or she plans to vote against a black candidate) (Traugott and Lavrakas 2008, 23).

Because IVR eliminates the need for a live person on the line conducting an interview, it has been proposed as a valuable tool for reducing social desirability bias. Intuitively it makes sense that absent an actual live person on the line, respondents might give more honest answers to an IVR system. With a live interviewer, a number of factors can influence the respondent, such the race and gender of the interviewer or whether the respondent feels unwilling to disclose a sensitive opinion or piece of personal information. With an IVR system, where respondents

are speaking with a computer rather than a live person, interviewer effects should be less likely to be a problem. This is true of self-administered surveys in general and also applies to online surveys (Greenberg 2011, 123).

Research finds that social desirability bias can be reduced through use of IVR systems in place of live interviews. A split-sample survey of patients with lower back pain in 1999 was conducted with half the sample being surveyed by live interviewers and the other half using an IVR system. The study found that those surveyed with IVR were much more likely to admit to mental health issues compared to those speaking with a live interviewer (Millard and Carver 1999, 157). The report concluded that because IVR eliminates the need for a respondent to speak with a live interviewer, it might be a preferable means of doing research on sensitive topics (Millard and Carver 1999, 153). A split-sample study of customers at a bank by the Gallup Organization found that respondents surveyed via IVR were more likely to say they were not satisfied with their experience at the bank than were those surveyed by live interviewers, leading the study's authors to conclude that IVR surveys result in more honest answers (Tourangeau, Steiger and Wilson 2002, 272). A University of Michigan study with a similar split-sample framework found that IVR respondents were much more willing to disclose sensitive information, such as the number of sex partners (Couper, Singer and Tourangeau 2004, 564). The study's authors concluded that "such differences suggest that the IVR experience (even with the recorded voice of a human interviewer) is a qualitatively different experience for respondents than interacting with a real interviewer" (Couper, Singer and Tourangeau 2004, 567).

The other major finding in scholarship comparing Interactive Voice Response surveys to those conducted with live interviewers is that IVR studies are susceptible to increased non-

response. The previously cited Gallup study of bank customers had an overall response rate of 49 percent for the live interview sample but only 24 percent for the IVR sample (Tourangeau, Steiger and Wilson 2002, 267-269). A similar analysis of Maryland residents with a split sample had a 38 percent response rate for the live interviewer group but only a 29 percent response rate for those surveyed with IVR (Tourangeau, Steiger and Wilson 2002, 270). A 2008 study that compared IVR and live interview surveys along with mail and Internet surveys found that the response rate for IVR was 28% while the response rate for live interviewers was 44% (Dillman, Phelps, et al. 2009, 7). A meta-review of published studies on Interactive Voice Response from 1989 to 2000 found that IVR response rates were consistently low (Corkrey and Parkinson 2002, 347). Although response rates vary from survey to survey, the literature indicates that IVR surveys have consistently higher non-response than live interview surveys in the same circumstances.

The finding that IVR surveys result in consistently lower response rates means that the decision to participate in an IVR survey works differently than the decision to participate in a live interview survey. Recall that John Brehm's four criteria for the decision of a respondent to participate in a survey were the respondent's relationships to strangers, to the interviewer, to the survey, and to her or his self-image (Brehm 1993, 51-52). When comparing IVR to surveys involving human contact, there are fewer factors to consider in explaining the motivation of an individual respondent to participate or not to participate. Brehm's first two factors both depend on the interaction of the respondent with another human being conducting an interview. Without a live person on the line, there is no relationship with strangers in general or with the interviewer specifically to consider. Thus, it is possible that a person who would agree to participate in a live survey because of those factors would simply hang up on an IVR survey. For example, a person

who would agree to a live interview simply because he or she simply enjoys speaking to people on the phone might refuse an IVR interview in which he or she is only responding to an automated system. The decision to participate thus will rely entirely on the respondent's relationship to the survey and relationship to self-image. A person who is more interested in the survey topic might be more amenable to taking an IVR survey. A person who is more confident in themselves and their opinions might be more willing to participate in an IVR survey. With consistently lower response rates for IVR surveys, it is very possible that IVR surveys encompass a narrower subset of respondents who are interested in participating for these reasons alone.

Accuracy of IVR Polls vs. Live Interviewer Polls

In the context of public opinion polling, multiple studies have shown that Interactive Voice Response polls are just as accurate at predicting election outcomes as polls conducted using live interviewers. The consistently accurate results of IVR polls as predictors of election outcomes as measured by multiple comprehensive studies of poll results have led many to believe that IVR polls are just as accurate as live interviewer polls. Comparisons have been done of both primary and general election poll results in several different election cycles, all indicating that IVR polls are very accurate at predicting election results.

A 2009 study that the AAPOR conducted on polling accuracy in the 2008 presidential primary elections found no measurable difference in the accuracy of polls conducted using IVR compared to those conducted using live interviewers (AAPOR Ad Hoc Committee on the 2008 Presidential Primary Polling 2009). The report did note that the comparisons involved small sample sizes and could be impacted by other factors. Despite the caveat, the finding was widely

reported in various publications and political blogs as an endorsement of IVR by the most prestigious opinion research association in the industry (Blumenthal 2009).

Other analyses also have found that pre-election IVR polls are just as likely to accurately predict election results as pre-election live interview polls. One study found that in the state-level presidential election results for 2008, there was only a one tenth of one percent difference in the candidate error rate between polls conducted using IVR and those conducted with live interviewers (Keeter 2011, 43). Another study found that the accuracy of IVR and live interviewer polls in 2008 were not significantly different (Panagopoulos 2009, 905). After the 2009 elections, Stuart Rothenberg of the Rothenberg Political Report gave IVR pollster SurveyUSA his "Pollster of the Cycle Award" for accuracy in each of the major off-year elections that year (Rothenberg 2009). As early as 2003, Joel Bloom praised SurveyUSA's accuracy in a presentation at that year's AAPOR convention, noting that the firm performed at above average levels compared to other non-partisan polling organizations (Bloom 2003). In 2008 statistician Nate Silver did an analysis on his website FiveThirtyEight that included past poll results from 171 elections for president, governor, and U.S. Senate over the previous ten years. The study found that among thirty-two different pollsters, the second and third most accurate were IVR pollsters SurveyUSA and Rasmussen Reports (Silver 2010). Another analysis Silver conducted showed that after comparing eight daily tracking polls on the presidential election from fall 2008, Rasmussen was the most accurate (Silver 2008b). Despite fears about methodology and the lack of cell phones in IVR samples, they appear to be quite effective as predictors of election outcomes.

Although many studies conclude that IVR polls are just as accurate as live interview polls, in a few notable instances IVR and live interview polls have produced widely varying

results. In two recent elections, the 2010 U.S. Senate special election in Massachusetts and the

2010 U.S. Senate Republican primary election in Delaware, IVR and live interviews produced

diverging estimates that seem unlikely to have been the result of influences other than mode of

interview.

In the run-up to the 2010 U.S. Senate special election in Massachusetts, pollsters

produced widely varying estimates about the race. In polls conducted by Interactive Voice

Response (IVR) pollsters, the race was tight for weeks before the election. However, polls

conducted using traditional live interviews showed a large lead for the Democratic nominee,

state Attorney General Martha Coakley, which disappeared only in the last few weeks as

Republican Scott Brown's campaign gained momentum. The results seemed to converge as the

election approached, with both IVR and live interview polls conducted immediately before the

election showing similar results, accurately predicting a narrow win for Brown. Although the

interviewer mode was certainly not the only difference between the polls, differences in

sampling or likely voter models would explain consistently divergent results, but not results that

are identical at one point during the campaign but widely divergent at another point.

In the aftermath of the race, Democratic pollster Mark Mellman observed that the pattern

present in the Massachusetts special election also could also be found in other elections, such as

the 2006 U.S. Senate elections in Washington and Connecticut. In the Washington race, an IVR

poll conducted several months before the election showed a narrow four point lead for the

incumbent, while live interview polling conducted at about the same time showed the incumbent

leading by more than twenty points. In Connecticut, early IVR polling showed a neck-and-neck

race between Senator Joe Lieberman and challenger Ned Lamont, while polls conducted using

live interviewers showed a large lead for Lieberman (Mellman 2010). Mellman hypothesized

that Interactive Voice Response polls are only valid when conducted immediately before an election.

An experimental poll conducted in the Republican primary election in 2010 in Delaware provides even stronger evidence for Mellman's hypothesis. In that election, Christine O'Donnell unexpectedly defeated longtime Republican Congressman and former Governor Mike Castle for the Republican nomination for the U.S. Senate. Jan Van Lohuizen, who served as Mike Castle's pollster, conducted two identical polls of Republican primary voters approximately one week prior to the primary election, one using live interviewers and the other using Interactive Voice Response. He reported that the IVR poll showed a fourteen point lead for O'Donnell over Castle (53 percent to 39 percent) while the live interviewer poll showed a lead for Castle of twenty-three points (56 percent to 33 percent) (van Lohuizen and Samohyl 2011). O'Donnell ultimately prevailed in the election, winning by six points (53 percent to 47 percent).

At face value, the IVR poll that showed a lead for Christine O'Donnell would appear to be the more accurate. However, this ignores the fact that the dynamic of the campaign changed dramatically in the final week. O'Donnell had been given very little chance against Castle, a fixture in Delaware politics for nearly three decades. However, in the final week, O'Donnell received a series of major endorsements, including Republican Senator Jim DeMint and former Alaska Governor Sarah Palin (van Lohuizen and Samohyl 2011). It was not until the final week of the campaign that she received significant media coverage. Although certainly not impossible, it is highly improbable that O'Donnell held a larger lead on the ballot than her ultimate margin of victory prior to her campaign gaining momentum in the media. Thus the IVR survey that had greater predictive accuracy may in fact have been a less valid representation of the mood of the electorate when it was taken.

The Delaware study provided further evidence of a difference between IVR and live samples on a ballot test for the Republican primary for the U.S. House of Representatives. The U.S. House primary included the exact same pool of voters as the U.S. Senate race since Delaware has only one seat in the House. In that campaign, the IVR poll showed the more conservative candidate Steve Urquhart leading by twenty-two points (54 percent to 32 percent). The live interview poll, on the other hand, showed that he led the more moderate candidate Michelle Rollins by only two points (38 percent to 36 percent). Urquhart ultimately won the election by one point (48 percent to 47 percent) (van Lohuizen and Samohyl 2011). As in the Senate race, the live interview poll results for the primary for Delaware's House seat appears to make sense in the context of the final results, while the IVR poll result tells a story that is inconsistent with the outcome.

Causes of IVR / Live Interviewer Election Poll Differences

Based on the literature on Interactive Voice Response, four possible explanations can be considered for discrepancies between IVR and live interview survey methods when used to conduct public opinion polls. First, differences between IVR and live interview poll results may be produced by other methodological issues involved in taking the surveys and not the type of interviewer. Second, IVR poll results are simply not valid measures of public opinion. Third, IVR polls produce different results as a result of social desirability bias. And finally, the differences between IVR and live interview poll results may be caused by increased non-response in IVR samples.

The first explanation is that IVR poll results do not differ systematically from live poll results and only appear to do so because of other technical concerns, such as the absence of cell

phones, likely voter models, the use of registration-based or random digit dial sampling, post-weighting, or sample stratification. The polling divergence in the Massachusetts special election, for example, could have been caused by other factors; I do not have access to all the details of the methods of the polls that were conducted. The Delaware study does provide some evidence contradicting this hypothesis, since in a controlled experiment with survey language, timing, and sample held constant, there were substantial differences between the two sets of data (van Lohuizen and Samohyl 2011). It is clear, however, that methodological differences other than the use of IVR and live interviewers cannot be ruled out as potential causes of differences between poll results.

The second explanation is that IVR poll results are simply not valid measures of public opinion, and that they only appear to be in some cases by coincidence. As noted, many leading figures in academic survey research were once highly skeptical that IVR polls could accurately measure public opinion. The substantial number of studies showing that IVR polls are just as accurate in predicting election results as live interview polls would seem to disprove this hypothesis (AAPOR Ad Hoc Committee on the 2008 Presidential Primary Polling 2009) (National Council on Public Polls 2011). IVR polls conducted immediately before elections are clearly accurate measures of public opinion when compared against election outcomes.

The third hypothesis is that IVR polls do produce different results, reflecting social desirability bias. The literature on IVR versus live interviews identified decreased social desirability bias as a benefit of conducting surveys using IVR. Again, the study of the 2010 Delaware primary provides some evidence contradicting this. If voters were less likely to tell a live interviewer that they would be voting for the more conservative candidate, then the IVR poll should have more accurately predicted the final election results. However, the horserace

numbers for the U.S. House primary were far more accurate in the live interview poll than in the IVR poll, and the greater "accuracy" of the Senate primary numbers is suspect because of how much the race changed in the final week. The presence of significant social desirability bias would also lead IVR polls to be more accurate predictors of election outcomes compared to polls conducted using live interviewers. However, as has been noted, all studies comparing pre-election poll results have found that neither IVR nor live interviewers are consistently more accurate in predicting outcomes.

The final explanation – and the one that this thesis focuses on - is that the differences between live and IVR samples in public opinion polls are caused by increased non-response in the IVR samples. If non-response bias is producing the differences, then there will be significant differences between samples for the two survey methods and a lower response rate for the IVR polls. Van Lohuizen and Samohyl found two notable differences between the IVR and live interview samples that they collected for their polls. First, the IVR poll results had a higher percentage of self-identified "very conservative" voters (53% compared to 40%) and a smaller percentage of moderate or liberal voters (19% compared to 32%) when contrasted with the live interview poll results. More importantly, and consistent with the literature on Interactive Voice Response in other contexts, they found that the response rate for the IVR poll was less than half the response rate for the live interview poll (9% compared to 23%). They conclude that "it may very well be that the type of person who is more likely to respond to a robo-poll has a stronger desire to be heard, is less likely to be undecided, and is more likely to express an opinion" (van Lohuizen and Samohyl 2011). This finding is supported by a study conducted in 2007 indicating that in 2006 statewide election polls, IVR surveys tended to have fewer undecided voters than

those of live interviewers in the same month (Blumenthal and Franklin, Methods and Horse Races: Phone, IVR and Internet Polls in the 2006 Elections 2007).

This thesis will investigate the proposition that IVR polls consistently over-sample more engaged and opinionated voters due to non-response bias. Thus, results of IVR and live interview polls will be similar when highly engaged and informed voters are representative of the larger voting population and different when they are not. Just before an election, when nearly all voters are engaged and informed, the differences between live interview and IVR polls tend to disappear. But polls conducted earlier in campaigns show demonstrable differences when comparing the two types of telephone interview methods.

This study does not seek to make claims about whether IVR or live interview polls are more accurate. As previously noted, there is no standard for a poll being "right" prior to Election Day. It is possible, as appears to have been the case in the 2010 Republican Senate primary in Delaware, that a poll that was probably less representative of the mood of the electorate on the day it was taken ended up being closer to the ultimate outcome. I believe that Interactive Voice Response polls and live interview polls do not generally tap the same groups of voters. What impact this has on the validity of individual polls will vary from election to election.

More specifically, this research tests two hypotheses. The literature indicates that IVR surveys tend to have higher rates of non-response (Tourangeau, Steiger and Wilson 2002) (Dillman, Phelps, et al. 2009), and voters who are more difficult to contact for surveys tend to have lower levels of engagement and interest in politics (Brehm 1993) (Keeter, Kennedy and Dimock, et al. 2006). Interactive Voice Response polls have been shown in at least one instance to overstate support for a more conservative candidate in an ideologically-charged Republican primary election (van Lohuizen and Samohyl 2011). Thus, my hypotheses are:

h1. Interactive Voice Response poll samples contain more engaged and opinionated voters, resulting in fewer undecided voters when compared with live interviewer polls conducted approximately the same length of time prior to an election.

h2. In situations where a specific candidate tends to receive more support among highly engaged and opinionated voters (such as an ideologically-charged primary election), Interactive Voice Response polls will show significantly greater support for that candidate when compared to live interviewer polls.

The following chapter will discuss my approach to testing these hypotheses, including data sources and statistical techniques. This analysis will attempt to isolate the impact of survey type (IVR vs. live interviewer) in public opinion polls by constructing models that test the impact of this variable on support for specific candidates and the proportion of voters that remain undecided, controlling for the distance between the date(s) that a poll was conducted and the date of the election.

CHAPTER 4: Research Design and Methods

To test whether Interactive Voice Response polls over-sample highly engaged and opinionated voters, I focused on looking for differences in estimates between polls conducted using the two techniques during the course of an election campaign. This is challenging because sometimes highly engaged voters nearly identically resemble the larger voting public in their opinions. This is particularly true just before an election, the point at which polls are generally evaluated for their accuracy. Evaluating pre-election polls is relatively simple due to the presence of the actual election results as a clear barometer for accuracy. Studying polls conducted at other points in the campaign is more of a challenge, because of the need to measure proximity to the election date as an independent variable along with any other variables being considered. As a result, the focus of this analysis is not whether polls are "accurate" predictors of election outcomes. Rather, it examines whether there are statistically significant differences between live interviewer and IVR surveys.

The aggregate poll results I used for this analysis are fully accessible and publicly available for download via the Huffington Post's Pollster.com website. The editors of Pollster.com have collected the poll results over the course of the last several years. The use of aggregate polling data for research is well established, as data collected via Pollster.com have been used in numerous studies of public opinion polls, including those presented in academic journals (Traugott and Wlezien 2009, 878) as well as less formal analyses by non-academics (Silver 2010). The scope of the data used for this study was all publicly released polls of likely voters for each of the following U.S. Senate and gubernatorial elections in 2010.

The 2010 elections for U.S. Senate and governor were chosen because they were the most recently held national elections (at the time this study began) and because they fit the profile of

elections in which I expected to find differences between IVR and live interview polls. The

elections chosen for this analysis needed to strike a good balance between having sufficient data

available for a meaningful analysis and yet not being so high profile as to render the data useless

for testing my hypothesis. For example, hundreds of polls were conducted and released in the

2008 and 2012 U.S. presidential campaigns, because those elections were so high profile I would

not expect highly engaged and informed voters to vary as much with less engaged voters as they

would in elections for other public offices in the United States. However, while I believe this

phenomenon would likely be present in even lower profile elections for offices such as the U.S.

House and state and local positions, simply not enough publicly available polling data are

available for analysis of poll results in those cases.

For the 2010 general elections, each election is included for which there are data,

regardless of the competitiveness of the election, in order to have the maximum number of both

types of polls to compare. For primaries, the analysis focuses on elections for which there are

sufficient polls to run a regression model.[2]

Comparing poll results that track electoral campaigns is a very imperfect science due to

the limited information available on the methodology used to conduct each poll. The lack of

available data on each pollster's methodology limits how well these differences can be accounted

for in comparing poll results. It is unclear what individual pollsters' likely voter models are,

whether they surveyed using Registration Based Sampling (RBS) or Random Digit Dialing

---

[2] All surveys of likely voters conducted using IVR or live interviewers were included in the
analysis, with the exception of surveys conducted by the firms Strategic Vision and Research
2000. Both these pollsters have been accused of fraud and have been shown to be highly
inaccurate and have very questionable results. As a result, poll results from those two polling
firms are excluded from this dataset to reduce the likelihood that fraudulent data could be
skewing the results.

(RDD), what the response rate was, whether they included cell phones in their samples, or whether there was post-weighting or stratification involved in creating the final top line result that is reported to the public. Although some pollsters publish this information, most do not. This complicates doing complete analyses of the differences between polls, and increases the likelihood that the results may be affected by a variable that cannot be accounted for. Statistical techniques will be used to examine how much variance in poll results can be accounted for simply by polling with live interviewers versus Interactive Voice Response.

Even though the information available online is limited, a number of key data points are available for each poll. Pollster.com collects and reports several important characteristics of each poll whose results they publish. These include:

1) The type of survey – Telephone, Internet, or Mail

2) The type of telephone survey – Interactive Voice Response or Live Interviewers

3) The type of survey sample – Likely Voters, Registered Voters, or Adults

4) The day or dates that the survey was in field

5) The number of interviews conducted

Since the analysis here focuses on telephone polls of likely voters conducted using Interactive Voice Response and Live Interviewers, it excludes all polls conducted using the Internet or mail, any polls that used multiple methods, and surveys that were of Registered Voters or Adults. All polls in this analysis are telephone surveys of likely voters. For each poll, in addition to the poll result, there are three available data points: the number of interviews conducted, the dates that the poll was in field – which provides a data point for the proximity of when the poll was taken to the actual election being held, and whether it was conducted using Interactive Voice Response or Live Interviewer.

Because the polls included in this analysis have varying sample sizes, I used Weighted

Least Squares (WLS) regression to determine the impact of telephone survey mode on the results

of each poll. WLS differs from Ordinary Least Squares (OLS) regression in that it adjusts the

model based on weights given to each individual data point. In this case, the data points are the

individual poll results included in the analysis. . WLS regression is designed for situations in

which heteroscedasticity arises – that is, when the variance of each data point in a model is not

constant (Shalizi 2013, 155). A data set containing many surveys of varying sample sizes is such

an instance, because the estimates will contain less variance when sample sizes are larger than

when sample sizes are smaller. WLS "tries harder to match observations where the weights are

big and less hard to match them where the weights are small" (Shalizi 2013, 158).

One potential issue with WLS is that it can potentially over-estimate the R^2 value for a

regression model. Willet and Singer found that if WLS weights result in a better fit for the linear

model than would be the case with an OLS regression, the R^2 value can be substantially higher

(Willett and Singer 1988). They counsel researchers not to rely exclusively on the R^2 value to

determine the predictive validity of the model. This analysis will focus chiefly on the impact of

one independent variable – the use of IVR or live interviewers – rather than the predictive

validity of the entire model, especially due to the large number of unknowns about the

methodology employed by each polling firm that cannot be accounted for in the models. When

R^2 values are discussed, I will calculate the R^2 using OLS for comparison to ensure that the

results do not differ drastically.

The weights for the WLS regression were developed using the data that are available for

each survey. Shalizi recommends that when dealing with surveys of varying sample sizes, the

sample sizes themselves can be used as the weights (Shalizi 2013, 162). However, such weights

are not the best measure of differences in the precision of each survey. Doubling the sample size does not make a survey twice as precise, and the percentages themselves impact the precision of each data point. Recall that the formula for the standard error of a proportion is:

Equation 4.1

$$\sigma_{prop} = \sqrt{\frac{pq}{N}}$$

The differences in precision for each data point can be estimated by taking the inverse of the standard error for each survey. Then the weight *w* for each data point will be:

Equation 4.2

$$w = \frac{1}{\sqrt{\frac{pq}{N}}}$$

This formula takes advantage of knowing not only the sample size N for each data point, but also the proportion p that contributes as well to the precision of the data point. The standard error for a proportion is smallest as p approaches 1 (or 100%) and is largest when p is .5 (or 50%).

In the analysis of general election polls, I slightly altered the formula for the weights since some of the proportions being compared are zero. Some undecided percentages in general election polls are zero. This made it impossible to incorporate the proportion into the formula for determining the weight for each data point, since the formula for the weight would involve dividing by zero. As a result, for general election polls, the weights were calculated by determining the maximum standard error for the particular sample size (which occurs when the proportion is .5 or 50%) rather than using the exact statistic.

Equation 4.3

$$w = \frac{1}{\sqrt{\frac{.25}{N}}}$$

The WLS regression sought to fit a line to the plot of all poll results during the course of

a campaign. The dependent variable, Y, is the proportion of voters from the survey who are

undecided or support a particular candidate; the model includes two independent variables: a

dummy variable X_1 (where IVR polls are coded 1 and live interview polls 0) and X_2, the number

of days prior to the election that the poll was conducted.

The dependent variable is the percent of voters who express a preference for a particular

candidate in an election or who do not express a preference for any candidate. When analyzing

general elections, there will only be one dependent variable for each campaign – the percent of

voters who are undecided. However when analyzing primary elections, there will be multiple

dependent variables: the proportion of voters who are undecided as well as the proportion of

voters supporting each candidate. Because some polls include multiple values indicating no

preference (such as "undecided," "no opinion," "not sure," or "don't know"), I calculated a

"modified undecided" value for each poll, in which Y_u, the percent of voters without a preference

for any candidate, is a function of all respondents taking the survey minus those expressing a

preference for all of the candidates combined. If C is the total number of candidates for whom

voters expressed a preference and Y_{1-C} are the proportions indicating the support for each

candidate in the survey, the equation for percentage of undecided voters is:

Equation 4.4

$$Y_u = 1 - Y_1 - Y_2 \dots - Y_C$$

The variable X_2 taps the number of days prior to the election that the poll was conducted. Because some polls are conducted in a single day and others are conducted for multiple days, the date of the poll will be estimated by using the midpoint of the time period during which the poll was conducted. Let D_1 be the first day that a poll was in field, D_L be the last day it was in the field, and E be the date of the election. X_2, the number of days before the election the poll was conducted, was calculated by:

Equation 4.5

$$X_2 = E - D_L + (D_L - D_1)/2$$

With Y, X_1 and X_2 thus defined, the equation for the regression is:

Equation 4.6

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_1$$

I hypothesized that both X_1 and X_2 should be statistically significant predictors of Y, the proportion of undecided voters in a poll. Since voters make up their minds as an election approaches, the closer a poll is conducted before an election the fewer undecided voters there will be. Thus, a larger value of X_2 will mean a lower value for Y. If hypothesis h1 is correct and IVR polls survey more engaged and opinionated voters, then the use of Interactive Voice Response ($X_1 = 1$) will be associated with a smaller proportion of undecided voters, and the hypothesis predicts a negative relationship. In the regressions in which Y taps the proportion of support for a specific candidate, the impact of IVR will vary depending on the candidate and campaign. According to the second hypothesis, the use of IVR polling, X_1, should be positively

associated with higher values of Y for more ideologically extreme candidates due to greater support for those candidates among highly engaged and opinionated voters.[3]

As previously discussed, this research is limited by the lack of available information on the methodology employed by each polling firm for which I have data. There are many methodological factors that cannot be accounted for when determining how a polling firm got to a particular estimate. For this reason, as well as due to the natural amount of random error in any survey data, I do not expect R^2 values for the models to account for most or all of the variation, but will chiefly focus on whether there was a statistically significant impact of IVR or live interviewers on the poll results. The following two chapters lay out how this analysis was conducted and my principal findings.

---

[3] Note: all regression analysis and scatterplots were done using the R statistical package. The regressions were running using the standard linear model function available in base R, while the scatterplots were created using the lattice package.

CHAPTER 5: General Election Polling in 2010

This analysis uses general election polling in 2010 U.S. Senate and gubernatorial

elections to test the first hypothesis:

> h1. Interactive Voice Response poll samples contain more engaged and opinionated
> voters, resulting in fewer undecided voters when compared with live interviewer polls
> conducted approximately the same length of time prior to an election.

I found that Interactive Voice Response polls report statistically significant smaller

percentages of undecided voters compared to those conducted using live interviewers. This

supports my expectation that IVR polls tend to survey a narrower, more opinionated segment of

the electorate.

This chapter elaborates on these findings. My main objective will be to discuss the

statistically significant differences in undecided rates between IVR and live interview polls at the

same points in a general election cycle. If IVR polls do systematically over-sample highly

engaged and better informed voters, undecided rates will be consistently lower in IVR samples

when compared with live interview samples. The analysis will include both a global look at

undecided rates across all 2010 U.S. Senate races and across all 2010 gubernatorial races, as well

as individual races for which there are sufficient polls for analysis

In November 2010, there were 37 elections for the U. S. Senate and 37 elections for state

governorships. These included regularly scheduled elections for that election cycle, as well as

special elections to fill unexpired terms for U.S. Senate seats in Delaware, New York, and West

Virginia, as well as the governorship in Utah. Some of these elections were hotly contested, with

the candidates separated by only a few percentage points, while others were landslides. (See

Tables 5.1 and 5.2.)

The 2010 midterm elections resulted in across-the-board gains for the Republican Party. Republicans gained a net of six U.S. Senate seats, defeating incumbents in Wisconsin and Arkansas and winning open Democratic seats in Illinois, Indiana, North Dakota, and Pennsylvania. Republicans also gained a net of six governorships, defeating Democratic incumbents in Iowa and Ohio and winning open Democratic seats in Florida, Kansas, Maine, Michigan, New Mexico, Oklahoma, Pennsylvania, Tennessee, Wisconsin, and Wyoming. Democrats failed to defeat any Republican incumbents, but picked up open Republican-held governorships in California, Connecticut, Hawaii, Minnesota, and Vermont. The Republican-held governorship of Rhode Island was won by Independent candidate Lincoln Chafee.

Table 5.1: Contested United States Senate Elections in 2010

State	Republican	%	Democrat	%	Other Major Candidate	%
AL	***Richard Shelby***	**65**	William G. Barnes	35		
AK	Joe Miller	35	Scott McAdams	23	***Lisa Murkowski***	***39***
AZ	***John McCain***	**59**	Rodney Glassman	35		
AR	**John Boozman**	**58**	*Blanche Lincoln*	37		
CA	Carly Fiorina	43	***Barbara Boxer***	**52**		
CO	Ken Buck	47	**Michael Bennet**	**48**		
CT	Linda McMahon	43	**Richard Blumenthal**	**55**		
DE	Christine O'Donnell	40	**Chris Coons**	**57**		
FL	**Marco Rubio**	**49**	Kendrick Meek	20	Charlie Crist	30
GA	***Johnny Isakson***	**58**	Michael Thurmond	39		
HI	Campbell Cavasso	22	***Daniel Inouye***	**75**		
ID	***Mike Crapo***	**71**	Tom Sullivan	25		
IL	**Mark Kirk**	**48**	Alexi Giannoulias	46		
IN	**Dan Coats**	**56**	Brad Ellsworth	38		
IA	***Chuck Grassley***	**65**	Roxanne Conlin	33		
KS	**Jerry Moran**	**70**	Lisa Johnston	26		
KY	**Rand Paul**	**56**	Jack Conway	44		
LA	***David Vitter***	**57**	Charles Melancon	38		
MD	Eric Wargotz	36	***Barbara Mikulski***	**62**		
MO	**Roy Blunt**	**54**	Robin Carnahan	41		
NV	Sharron Angle	45	***Harry Reid***	**50**		
NH	**Kelly Ayotte**	**60**	Paul Hodes	37		
NY (1)	Joseph DioGuardi	36	***Kirsten Gillibrand***	**62**		

NY (2)	Jay Townsend	33	*Chuck Schumer*	**65**		
NC	*Richard Burr*	**55**	Elaine Marshall	43		
ND	**John Hoeven**	**76**	Tracy Potter	22		
OH	**Rob Portman**	**57**	Lee Fisher	39		
OK	*Tom Coburn*	**71**	Jim Rogers	26		
OR	Jim Huffman	39	*Ron Wyden*	**57**		
PA	**Pat Toomey**	**51**	Joe Sestak	49		
SC	*Jim DeMint*	**62**	Alvin Greene	28		
UT	**Mike Lee**	**62**	Sam Granato	33		
VT	Len Britton	31	*Patrick Leahy*	**64**		
WA	Dino Rossi	48	*Patty Murray*	**52**		
WI	**Ron Johnson**	**52**	*Russ Feingold*	47		
WV	John Raese	43	**Joe Manchin**	**54**		

Table 5.2: Contested U.S. Gubernatorial Elections in 2010

State	Republican	%	Democrat	%	Other Major Candidate	%
AL	**Robert Bentley**	**58**	Ron Sparks	42		
AK	*Sean Parnell*	**59**	Ethan Berkowitz	38		
AZ	*Jan Brewer*	**55**	Terry Goddard	42		
AR	Jim Keet	34	*Mike Beebe*	**65**		
CA	Meg Whitman	42	**Jerry Brown**	**53**		
CO	Dan Maes	11	**Dan Hickenlooper**	**51**	Tom Tancredo	37
CT	Tom Foley	49	**Dan Malloy**	**50**		
FL	**Rick Scott**	**59**	Alex Sink	48		
GA	**Nathan Deal**	**53**	Roy Barnes	43		
HI	J.R. Aiona Jr.	41	**Neil Abercrombie**	**58**		
ID	*Butch Otter*	**59**	Keith Allred	33		
IL	Bill Brady	46	*Pat Quinn*	**47**		
IA	**Terry Branstad**	**53**	*Chet Culver*	43		
KS	**Sam Brownback**	**63**	Tom Holland	32		
ME	**Paul LePage**	**38**	Libby Mitchell	19	Eliot Cutler	37
MD	Robert Ehrlich	42	*Martin O'Malley*	**56**		
MA	Charlie Baker	42	*Deval Patrick*	**48**		
MI	**Rick Snyder**	**58**	Virg Bernero	40		
MN	Tom Emmer	43	**Mark Dayton**	**44**		
NE	*Dave Heineman*	**74**	Mike Meister	26		
NV	**Brian Sandoval**	**53**	Rory Reid	42		
NH	John Stephen	45	*John Lynch*	**53**		
NM	**Susana Martinez**	**54**	Diane Denish	46		
NY	Carl Paladino	34	**Andrew Cuomo**	**61**		
OH	**John Kasich**	**49**	*Ted Strickland*	47		

OK	**Mary Fallin**	**60**	Jari Askins	40		
OR	Chris Dudley	48	**John Kitzhaber**	**49**		
PA	**Tom Corbett**	**54**	Dan Onorato	46		
RI	John Robitaille	34	Frank Caprio	23	**Lincoln Chafee**	**36**
SC	**Nikki Haley**	**51**	Vincent Sheheen	47		
SD	**Dennis Daugaard**	**62**	Scott Heidepriem	38		
TN	**Bill Haslam**	**65**	Mike McWherter	33		
TX	***Rick Perry***	**55**	Bill White	42		
UT	***Gary Herbert***	**64**	Peter Corroon	32		
VT	Brian Dubie	48	**Peter Shumlin**	**50**		
WI	**Scott Walker**	**52**	Tom Barrett	47		
WV	**Matt Mead**	**72**	Leslie Petersen	25		

* Candidate in **bold** denotes winner. Incumbents in *italics*.

The 2010 U.S. Senate and gubernatorial elections provide a wealth of available polling data. In many of these elections, substantial polling was done over the course of the campaigns, going back to early 2009.   These included polls conducted for independent organizations, newspapers, TV stations, political action committees, official Republican and Democratic committees, and the campaigns themselves. Pollster.com collected all publicly released horse race ballot numbers for each race and made them available through its website. In all, 755 polls were conducted for U.S. Senate elections, of which 263 used live interviewers and 492 used Interactive Voice Response. A total of 686 polls were conducted in gubernatorial elections; 243 used live interviewers and 443 IVR. The distribution of polls in these categories was not uniform in either the time of year or the states in which they were conducted. This was especially the case for polls conducted using live interviewers: most polls several months before the elections used IVR. Table 5.3 shows the total number of polls conducted each month in both gubernatorial and U.S. Senate contests, and indicates how many of them used live interviewers versus IVR. Tables 5.4 and 5.5 show the total number of polls conducted in each state by survey mode.

Table 5.3: U.S. Senate and Gubernatorial Election Polls by Month in 2010

	IVR Senate	Live Senate	IVR Governor	Live Governor	Total IVR	Total Live
January	16	5	7	6	23	11
February	19	5	18	5	37	10
March	29	4	28	7	57	11
April	24	8	24	5	48	13
May	32	8	28	5	60	13
June	45	7	44	7	89	14
July	48	15	37	9	85	24
August	59	18	52	17	111	35
September	91	62	76	63	167	125
October	139	133	129	119	268	252

An early test of hypothesis h1 can be done by averaging all poll results of each type and breaking the results down by month. This was the method used by Blumenthal and Franklin in their 2007 paper, "Methods and Horse Races." If Interactive Voice Response polls sample more engaged and opinionated voters, then such polls on average should contain a lower percentage of undecided voters than live interview polls conducted during the same month in the year leading up to the November 2010 election. This is a crude estimate that treats a poll taken on the first day of a month as the same as one taken on the last day, but it should provide some evidence about the hypothesis. Figures 5.1, 5.2, and 5.3 report poll results for undecided voters in U.S. Senate and governors' races by month.

Figure 5.1: 2010 U.S. Senate Elections - Percent of Undecided Voters by Month

Month	IVR	Live
January	11.8	21.6
February	11.0	17.6
March	11.1	15.5
April	11.4	22.8
May	9.1	11.8
June	9.8	15.0
July	8.3	16.2
August	8.2	15.2
September	6.2	9.5
October	4.5	7.7
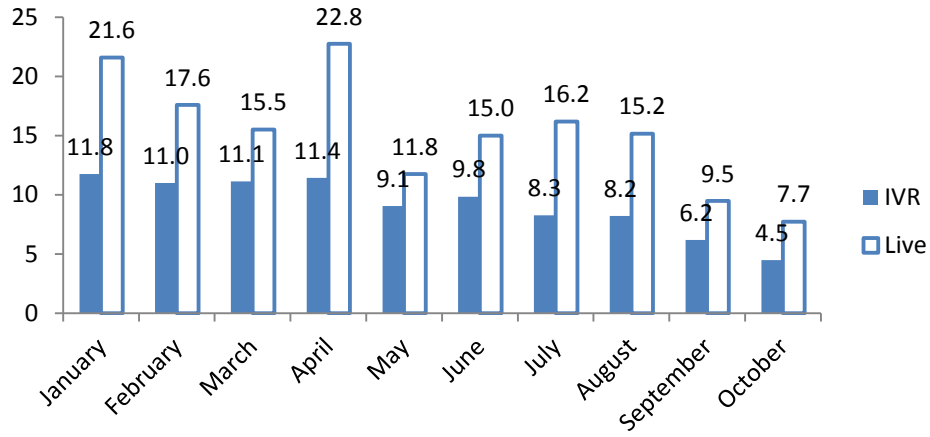
Figure 5.2: 2010 Gubernatorial Elections - Percent of Undecided Voters by Month

Month	IVR	Live
January	15.4	16.3
February	12.3	15.4
March	11.9	25.6
April	11.0	19.8
May	11.3	17.0
June	10.7	16.4
July	8.3	17.2
August	8.6	15.3
September	6.7	11.5
October	5.2	9.0

Figure 5.3: 2010 U.S. Senate & Gubernatorial Elections - Percent of Undecided Voters by Month

Month	IVR	Live
January	12.9	18.7
February	11.6	16.5
March	11.5	21.9
April	11.2	21.6
May	10.1	13.8
June	10.2	15.7
July	8.3	16.6
August	8.4	15.2
September	6.4	10.5
October	4.8	8.3
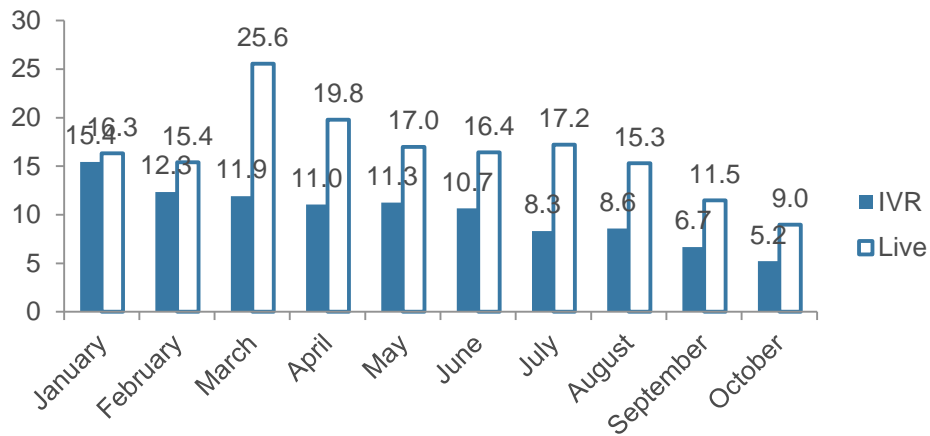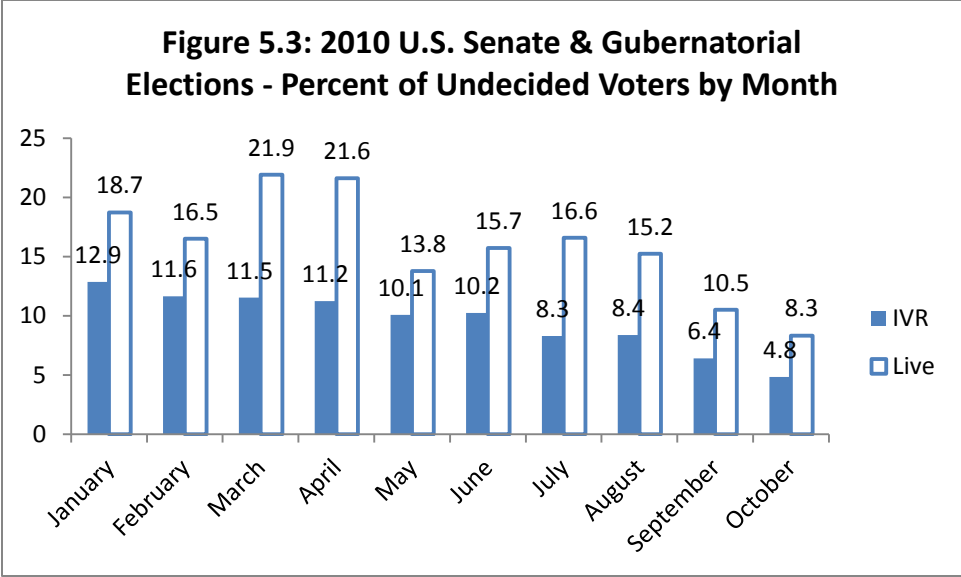
Note: Averages in Figures 5.1-5.3 are un-weighted means of all Live Interview surveys conducted during each month in 2010, calculated using the aggregate function in R.

These figures show that on average, in both U.S. Senate and gubernatorial elections, polls conducted using Interactive Voice Response have substantially lower proportions of undecided voters than those conducted using live interviewers. The differences between percentages of undecided voters in live and IVR polls vary substantially from month to month, especially in the early months of the year. This is likely due in part to the relatively small number of total polls each month conducted using live interviewers from January through July. That the averages are built on only a few surveys likely skews the results due to the limited number of races that are being polled.

In the last two months before the election, where each group includes over 100 polls, the data strongly support a difference in undecided rates between the two survey modes. In September, Interactive Voice Response polls averaged 6.2 percent undecided respondents, while live interview polls averaged 9.6 percent. In October, IVR polls averaged 4.7 percent undecided and live interviewer polls 8.0 percent. These averages are problematic and not fully comparable, since they do not account for differences in variance between each poll or the differing points during the month that the polls were conducted. Nevertheless, they confirm that the findings of Blumenthal and Franklin in 2006 held true in 2010, and provide evidence that Interactive Voice Response polls consistently found fewer undecided voters compared to polls done with live interviewers.

In order to more precisely compare the effects of live interviewers with Interactive Voice Response, I used weighted least squares regression to test a model that includes the potential influences of differences in the races in each state, the approximate dates on which the poll was conducted, and differences in variance among individual polls.

Equation 4.6, permits examination of whether the use of IVR survey methods is a statistically significant determinant of a smaller proportion of undecided voters in a poll compared to polls with live interviewers. The key variables are Y_u, the proportion of undecided voters in an election poll; X_2, the number of days prior to an election that a poll was conducted; and X_1, a dummy variable for the use of IVR or live interviewers.

$$Y_u = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_1$$

The number of polls conducted in each Senate and gubernatorial race varied substantially, depending on the level of interest and the competitiveness of the race. Tables 5.4 and 5.5 show the total number of publicly available telephone poll results of likely voters for each poll conducted for the 2010 U.S. Senate and gubernatorial elections, as well as the number of polls that each of the two telephone polling techniques, live interviewers and Interactive Voice Response. Three U.S. Senate races had at least 20 polls of both types conducted, and six gubernatorial races also had at least 20 polls of each type. On the other hand, several races had no live interviewer polls at all, making a comparison of the two types of polling in those races impossible. Most states fell somewhere in between, with enough polls to make comparisons between the two types of telephone polls but still rather few for observing statistically significance differences.

Table 5.4: General Election 2010 U.S. Senate Polls by State

	IVR	Live	Total
AK	6	13	19
AL	6	0	6
AR	14	6	20
AZ	9	0	9
CA	34	19	53
CO	30	9	39
CT	22	10	32
DE	12	6	18
FL	25	22	47
GA	11	0	11
HI	2	0	2
IA	8	3	11
ID	3	1	4
IL	33	17	50
IN	10	2	12
KS	7	0	7
KY	27	12	39
LA	13	8	21
MD	5	3	8
MO	21	8	29
NC	32	5	37
ND	9	0	9
NH	14	9	23
NV	27	21	48
NY (1)	10	11	21
NY (2)	7	9	16
OH	29	17	46
OK	4	1	5
OR	12	3	15
PA	32	27	59
SC	5	0	5
SD	0	0	0
UT	2	5	7
VT	2	1	3
WA	24	9	33
WI	15	10	25
WV	17	5	22

Table 5.5: General Election 2010 U.S. Gubernatorial Polls by State

	IVR	Live	Total
AK	7	4	11
AL	5	1	6
AR	7	4	11
AZ	16	0	16
CA	32	21	53
CO	20	7	27
CT	19	7	26
FL	22	20	42
GA	18	1	19
HI	5	2	7
IA	9	7	16
ID	4	2	6
IL	26	9	35
KS	8	0	8
MA	13	10	23
MD	10	9	19
ME	9	7	16
MI	8	9	17
MN	13	6	19
NE	3	0	3
NH	10	8	18
NM	12	8	20
NV	20	21	41
NY	12	12	24
OH	32	22	54
OK	5	4	9
OR	14	5	19
PA	25	21	46
RI	9	4	13
SC	7	2	9
SD	9	0	9
TN	4	1	5
TX	15	5	20
UT	3	5	8
VT	4	1	5
WI	20	9	29
WY	4	0	4

The results of testing the model indicate that in 10 U.S. Senate races and 12 gubernatorial races, the use of Interactive Voice Response was associated with a significantly (p<.05) lower percentage of undecided voters (See Tables 5.6 and 5.7). In those ten U.S. Senate races, the use of Interactive Voice Response was associated with polls showing between two and sixteen percent fewer undecided voters compared to those conducted with live interviewers. Similarly, in the twelve gubernatorial races, the use of IVR was associated with between three and thirteen percent fewer undecided voters than in polls done using live interviewers.

Although the use of Interactive Voice Response was not associated with significantly fewer undecided voters in many races, this may be simply the result of an insufficiently low sample size of polls conducted for those elections. Type of telephone poll had a statistically significant impact on the proportion of undecided respondents in five of the six gubernatorial races with 20 polls or more of each type and in all three of the U.S. Senate races with 20 or more polls of each type. This suggests that the lack of statistically significant differences in undecided rates between IVR and live interview polls in many of the other races is due to the limited number of polls available for analysis.

The number of days before an election that a poll was conducted also was found to have a statistically significant impact on poll results in many of the elections examined. In 16 U.S. Senate races and 27 gubernatorial races, polls conducted closer to Election Day showed a lower percent of undecided voters compared to those conducted earlier in the cycle. The magnitude of the effect is reflected in the percent of fewer undecided voters associated with a poll being taken one day closer to the Election Day. Among the elections in which the impact was significant, this effect was as low as .012 percent per day and as high as .108 percent per day.

It is noteworthy that while a poll's proximity to Election Day was a significant predictor of undecided rates in many races, it was not in some of them. It is not entirely clear why this was the case, but it could be explained by some elections matching up two candidates who were already well-known in their respective states, meaning many voters already had strong opinions of them and could have made up their minds early in the campaign.

Among gubernatorial elections with more than 15 polls in the sample, only Texas, Maryland, and Iowa did not show significantly fewer undecided voters as the election approached. In both Maryland and Iowa, incumbent Democratic governors were being challenged by former Republican Governors who were already well-known among voters in those states. In Texas, Republican Governor Rick Perry was being challenged by Bill White, the prominent Democratic Mayor of Houston. All six candidates were already well established figures in their states at the outset of the campaigns. However, there were two prominent candidates in Ohio with former Congressman John Kasich matching up against incumbent Governor Ted Strickland, and there undecided rates declined significantly as Election Day approached.

Campaigns for U.S. Senate in which there was no significant change in undecided rates as Election Day approached are a little bit more difficult to explain. The Missouri U.S. Senate race matched up prominent GOP Congressman Roy Blunt against Secretary of State Robin Carnahan. Similarly in Washington, two-time gubernatorial nominee Dino Rossi was the Republican candidate against incumbent U.S. Senator Patty Murray. However, the same phenomenon was present in the New Hampshire U.S. Senate race, which featured appointed State Attorney General Kelly Ayotte making her first run for elected office. It is also noteworthy that more U.S. Senate campaigns (18) featured no statistically significant impact of proximity to

the election on undecided rates than gubernatorial campaigns (11), indicating that it is possible that voters are more likely to make up their minds earlier in elections for the U.S. Senate compared to elections for governor.

In both the gubernatorial and U.S. Senate polls, Ohio stands out as being the only state for which both a large number of IVR and live interviewer polls were available and no statistically significant differences appeared in undecided rates between the two sets of polls. It is unclear why this was the case. One possible explanation is that all four candidates for the two offices – then-Governor Ted Strickland, then-Lieutenant Governor Lee Fisher, then-former Congressman John Kasich, and then-former Congressman Rob Portman -- were well known in state politics prior to the beginning of the campaigns; this may have meant that highly engaged and opinionated voters who were over-represented in the IVR polls had less of an advantage over other voters with regards to their familiarity with the candidates than was the case in other states.

The results of the analysis support hypothesis h1, for U.S. Senate and gubernatorial races overall, and in the most hotly contested specific races in 2010. In total, the use of Interactive Voice Response was associated with statistically significant lower percentages of undecided voters in U.S. Senate races in Pennsylvania, California, Nevada, Florida, Kentucky, North Carolina, Washington, Arkansas, Oregon, and Utah, as well as gubernatorial races in California, Pennsylvania, Florida, Nevada, Illinois, Massachusetts, Texas, Minnesota, Oregon, New Hampshire, Iowa, and Maine. In many of those elections, Interactive Voice Response polls produced survey results with on average five or six percent fewer undecided voters than those conducted with live interviewers.

In summary, both of the two methods I used to compare Interactive Voice Response and live interviewer public opinion polls found that IVR polls contain fewer undecided voters on

average. Looking at an un-weighted average of all U.S. Senate and gubernatorial polls taken in 2010 by month shows that IVR polls conducted during the same months tend to have fewer undecided voters than those conducted with live interviewers. A more sophisticated regression analysis of the polling data of each of the individual elections found that in most elections with sufficient polling data, IVR polls tended to find feature undecided rates several points lower than those conducted with live interviewers. This finding indicates that IVR polls likely survey a more highly engaged and opinionated segment of the electorate than do polls conducted using live interviewers. The next chapter turns to the second hypothesis.

Table 5.6: 2010 Gubernatorial General Election Polls

* p <.05, ** p < .01, *** p < .001

State	Number of Surveys			Overall Model		Survey Type (IVR=1)		Days Before Election	
	IVR	Live	Total	Multiple R^2	Significance	Estimate	Significance	Estimate	Significance
OH	32	22	54	0.112	*	-0.087		0.012	*
CA	32	21	53	0.591	***	-5.841	***	0.035	***
PA	25	21	46	0.466	***	-6.909	**	0.072	***
FL	22	20	42	0.572	***	-6.559	***	0.108	***
NV	20	21	41	0.800	***	-6.909	***	0.022	***
IL	26	9	35	0.472	***	-10.014	***	0.031	***
WI	20	9	29	0.457	***	-2.010		0.038	***
CO	20	7	27	0.345	**	-1.344		0.062	**
CT	19	7	26	0.700	***	-0.059		0.065	***
NY	12	12	24	0.328	*	-2.368		0.037	**
MA	13	10	23	0.634	***	-4.761	*	0.039	***
NM	12	8	20	0.624	***	-3.318		0.055	***
TX	15	5	20	0.376	*	-5.355	**	0.006	
GA	18	1	19	0.316	*	-1.081		0.027	*
MD	10	9	19	0.272		-2.350		0.016	
MN	13	6	19	0.823	***	-9.870	***	0.077	***
OR	14	5	19	0.541	**	-5.091	**	0.040	**
NH	10	8	18	0.674	***	-3.690	*	0.045	***
MI	8	9	17	0.563	**	-1.282		0.078	**
AZ	16	0	16	0.743	***	NA		0.030	***
IA	9	7	16	0.522	**	-5.724	**	0.002	
ME	9	7	16	0.699	***	-12.934	***	0.072	*
RI	9	4	13	0.401		-4.521		0.032	*
AK	7	4	11	0.291		1.188		0.023	
AR	7	4	11	0.724	**	-3.462		0.070	**
OK	5	4	9	0.571		-2.520		0.038	*
SC	7	2	9	0.177		2.514		0.030	
SD	9	0	9	0.359		NA		0.032	
KS	8	0	8	0.113		NA		-0.004	
UT	3	5	8	0.723	*	-1.108		0.020	*
HI	5	2	7	0.234		-0.015		0.015	
AL	5	1	6	0.670		-10.383		0.011	
ID	4	2	6	0.564		-10.977		0.040	
TN	4	1	5	0.943		-7.560		0.076	*
VT	4	1	5	0.653		-5.018		0.025	
WY	4	0	4	1.000	***	NA		0.073	***
NE	3	0	3	0.995	*	NA		0.025	*

Dependent variable is percent of voters in each poll who do not express support for a candidate.

Scatterplots for all 2010 Governors elections included in Appendix

Table 5.7: 2010 U.S. Senate General Election Polls
* p <.05, ** p < .01, *** p < .001

State	Number of Surveys			Overall Model		Survey Type (IVR=1)		Days Before Election	
	IVR	Live	Total	Multiple R^2	Significance	Estimate	Significance	Estimate	Significance
PA	32	27	59	0.307	***	-4.818	*	0.036	***
CA	34	19	53	0.255	***	-3.996	**	0.013	**
IL	33	17	50	0.266	***	-2.837		0.027	***
NV	27	21	48	0.506	***	-4.341	***	0.026	***
FL	25	22	47	0.651	***	-3.399	***	0.052	***
OH	29	17	46	0.430	***	-0.575		0.036	***
CO	30	9	39	0.792	***	-2.030		0.039	***
KY	27	12	39	0.449	***	-6.593	***	0.039	***
NC	32	5	37	0.663	***	-9.631	***	0.043	***
WA	24	9	33	0.233	**	-2.133	**	0.009	
CT	22	10	32	0.292	**	-0.465		0.015	**
MO	21	8	29	0.064		-0.985		0.003	
WI	15	10	25	0.315	*	-4.470		0.081	**
NH	14	9	23	0.189		-5.066		0.009	
WV	17	5	22	0.514	**	-1.975		0.108	***
LA	13	8	21	0.333	*	-3.329		0.019	*
NY(1)	10	11	21	0.402	**	-2.866		0.057	**
AR	14	6	20	0.495	**	-4.451	**	0.022	**
AK	6	13	19	0.171		-6.029		0.031	
DE	12	6	18	0.286		0.702		0.083	
NY(2)	7	9	16	0.062		-1.767		0.023	
OR	12	3	15	0.868	***	-13.894	***	0.015	
IN	10	2	12	0.179		0.021		0.838	
GA	11	0	11	0.011		NA		0.003	
IA	8	3	11	0.264		-2.110		0.013	
AZ	9	0	9	0.581	*	NA		0.039	*
ND	9	0	9	0.211		NA		0.009	
MD	5	3	8	0.490		-5.360		0.064	
KS	7	0	7	0.137		NA		0.011	
UT	2	5	7	0.942	**	-15.825	**	0.081	**
AL	6	0	6	0.251		NA		-0.012	
OK	4	1	5	0.928		-9.811		0.010	
SC	5	0	5	0.798	*	NA		0.055	*
ID	3	1	4	0.970		-13.272		0.029	
VT	2	1	3	1.000		-3.000		NA	
HI	2	0	2	1.000		NA		0.191	

Dependent variable is percent of voters in each poll who do not express support for a candidate.

Scatterplots for all 2010 U.S. Senate elections included in Appendix

CHAPTER 6: Polling Differences in Primary Election Polls

After comparing live interview and interactive voice response polls in the 2010 general election, the analysis focused on primary elections that year for the United States Senate and governorships, examining the second hypothesis:

> h2. In situations where a specific candidate tends to receive more support among highly engaged and opinionated voters (such as an ideologically-charged primary election), Interactive Voice Response polls will show significantly greater support for that candidate when compared to live interviewer polls.

Consistent with the findings of Van Lohuizen and Samohyl (2011), in several primary elections statistically significant differences appeared between the estimates of support for individual candidates when comparing Interactive Voice Response and live interview polls. Some candidates appeared to be doing better in horse race metrics in IVR polls than in those conducted by live interview pollsters.

Primary elections are notoriously difficult to poll. Turnout can vary substantially from one primary election to the next, and this makes it hard to judge who actually is a "likely voter." States have different rules for primary participation – some states have no party registration and allow any voter to cast a ballot in either party's primary, while in other states one must be registered with a particular party in order to vote in that party's primary elections. Other states have something in between, with registered partisans restricted to voting in their respective primaries but independent voters not registered with either party given the choice of which primary to cast a vote in. Primary elections often have a far larger pool of persuadable voters that can be swayed to support different candidates, compared to a general election in which candidates are competing for the votes of a small proportion of the voting public known colloquially as "swing" voters. These factors make polling primary elections a far more difficult task than polling general elections.

A wealth of data exists on general election polling in 2010, but much less primary polling data are available. In 2010, most states held primary elections for the offices of U.S. Senate and governor. However, most of these elections have few publicly released polls. Many races were not competitive, and in others there was simply not enough interest from national polling organizations to conduct more than a few polls. The result is that in contrast with the general election analysis in Chapter 5, the small number of primary polls available significantly limits the analysis that can be done. Only a handful of races have enough polls to even attempt an analysis. Tables 6.1 and 6.2 summarize the polls taken for primary elections in 2010.

Table 6.1: Primary Election 2010 U.S. Senate Polls by State, Party, and Survey Type

State	Party	IVR	Live	Total
Pennsylvania	D	9	13	22
California	R	6	15	21
Florida	R	9	10	19
Kansas	R	11	3	14
Kentucky	D	7	4	11
Kentucky	R	8	3	11
Arizona	R	10	3	13
Florida	D	3	10	13
Arkansas	D	1	3	4
Nevada	R	0	11	11
North Carolina	D	7	3	10
Colorado	D	3	5	8
Colorado	R	5	2	7
Illinois	D	2	5	7
Illinois	R	3	3	6
Connecticut	R	0	5	5
Ohio	D	0	5	5
Arkansas	R	1	0	1

Table 6.2: Primary Election 2010 Gubernatorial Polls by State, Party, and Survey Type

State	Party	IVR	Live	Total
California	R	8	14	22
Florida	R	3	15	18

Georgia	R	14	3	17
Texas	R	9	4	13
Georgia	D	12	1	13
Pennsylvania	D	1	10	11
Nevada	R	0	10	10
Illinois	R	2	4	6
Illinois	D	2	4	6
California	D	0	4	4
Colorado	R	1	1	2
Connecticut	D	0	3	3
Connecticut	R	0	2	2
Minnesota	R	2	0	2
Minnesota	D	2	0	2
Nevada	D	1	1	2
Pennsylvania	R	1	1	2

Although a number of primary races for the U.S. Senate and governorships had many polls conducted, the distribution of live interviewer and Interactive Voice Response polls across each election is uneven. The majority of elections had polls that were almost entirely Interactive Voice Response or live interviewer. For example, 13 polls were conducted in the Democratic primary for the governorship of Georgia, but only one used live interviewers. On the other hand, in Pennsylvania there were 11 polls done for the Democratic primary for governor, but 10 used live interviewers and only one Interactive Voice Response. As Chapter 5 demonstrated, attempting an analysis with a very small number of polls of one type or the other is unlikely to generate statistically significant results.

The analysis in this chapter will focus exclusively on primary elections for which at least 10 polls were conducted, including at least five polls of each type – live interviewer and IVR. Four primary elections met these criteria:

1. Republican primary for governor of California

2. Republican primary for U.S. Senate in California

3. Republican primary for U.S. Senate in Florida

4. Democratic primary for U.S. Senate in Pennsylvania

   I analyzed the poll results in each of these four races using WLS regression. The model below examines whether the use of IVR is a statistically significant determinant of the proportions of voters who either express a preference for a candidate in a poll or are undecided. The key variables are differing operationalizations of Y, the proportions of voters expressing a preference for a candidate or indicating no preference in a poll; X_2, the number of days prior to an election that the poll was conducted; and X_1, a dummy variable for the use of IVR or live interviewers.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_1$$

Table 6.3 shows the results for each of the four primary elections, which the following discussion elaborates on.

Table 6.3: Influences on Candidate Preferences: Selected 2010 Primary Elections

*** p < .001

	Overall Model		Survey Type (IVR=1)		Days Before Primary	
	Multiple R^2	Significance	Estimate	Significance	Estimate	Significance
CA GOP Gubernatorial Primary						
Undecided %	0.882	***	-9.819	***	0.139	***
Whitman %	0.306	0.031	2.330	0.606	-0.080	0.026
Poizner %	0.635	***	-0.017	0.948	-0.098	***
FL GOP Senate Primary						
Undecided %	0.793	***	-10.299	***	0.018	0.098
Rubio %	0.954	***	6.353	0.011	-0.113	***
Crist %	0.792	***	0.172	0.958	0.089	***
CA GOP Senate Primary						
Undecided %	0.875	***	-11.397	***	0.187	***
Fiorina %	0.622	0.001	6.428	0.134	-0.168	0.002
Campbell %	0.237	0.150	1.052	0.694	0.062	0.059
DeVore %	0.844	***	-0.330	0.682	-0.070	***
PA DEM Senate Primary						
Undecided %	0.184	0.146	-0.345	0.055	0.002	0.827
Specter %	0.512	0.001	-0.841	0.653	0.032	***
Sestak %	0.391	0.009	1.176	0.617	-0.032	0.003

California – Gubernatorial Republican Primary

The 2010 GOP primary for Governor of California was not a highly ideological race. Both candidates were moderates with non-political backgrounds: Steve Poizner was the State Insurance Commissioner who had previously been a Silicon Valley executive. Meg Whitman was the former CEO of Ebay. The similarities between the two candidates thus extended beyond ideology: both were multimillionaires from northern California with backgrounds in the high-tech industry. Unlike other primaries in 2010, this one was not characterized by the type of ideological battle in which different polling methods had previously shown to product divergent results (van Lohuizen and Samohyl 2011). Whitman won the primary election with 64% compared to 27% for Poizner after spending tens of millions of dollars of her personal wealth on the campaign. She ultimately was defeated in the general election by California Attorney General and former Governor Jerry Brown.

In the GOP primary for California governor, the use of Interactive Voice Response had a strong and statistically significant impact on the proportion of undecided voters but not on preference for either candidate. Interactive Voice Response polls in this primary had on average about 10 percent fewer undecided voters than those conducted using live interviewers. This finding is consistent with hypothesis h1 – that IVR polls over-sample highly engaged and opinionated voters. However, the results do not indicate that IVR had any statistically significant impact on support for candidates Meg Whitman or Steve Poizner.

Figure 6.1: 2010 CA GOP Gov Primary Polls - Undecided
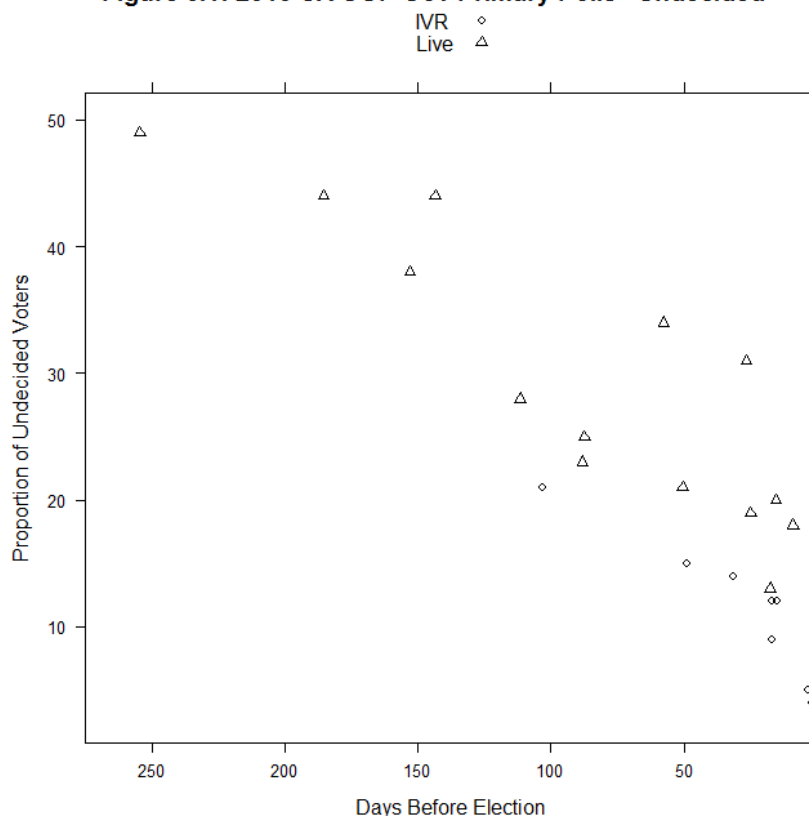
Poll results for the 2010 California Republican primary for governor do not appear to have been affected by the phenomenon found in the Delaware study. The levels of support for the two candidates do not vary significantly depending on the type of survey conducted. In a primary election that was not charged with ideological differences, it is unsurprising that the use of IVR did not have a statistically significant impact on the support levels for either candidate. Given a choice between two candidates with similar ideologies and professional histories, a pool of more opinionated and engaged voters is not likely to be substantially different than the broader set of primary voters.

Figure 6.2: 2010 CA GOP Gov Primary Polls - Poizner
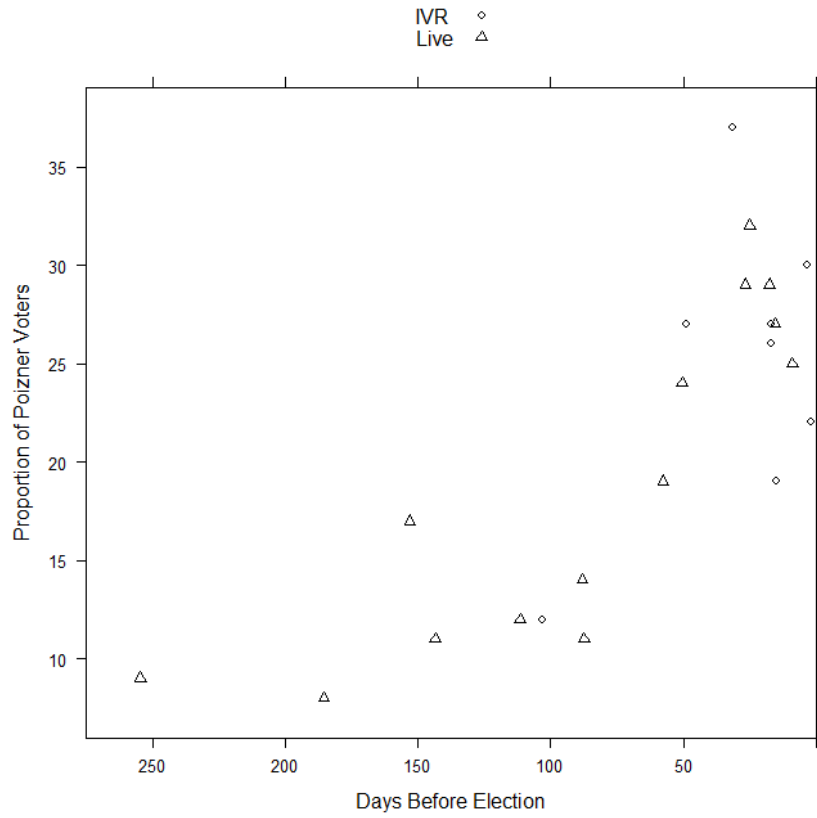
Figure 6.3: 2010 CA GOP Gov Primary Polls - Whitman

65
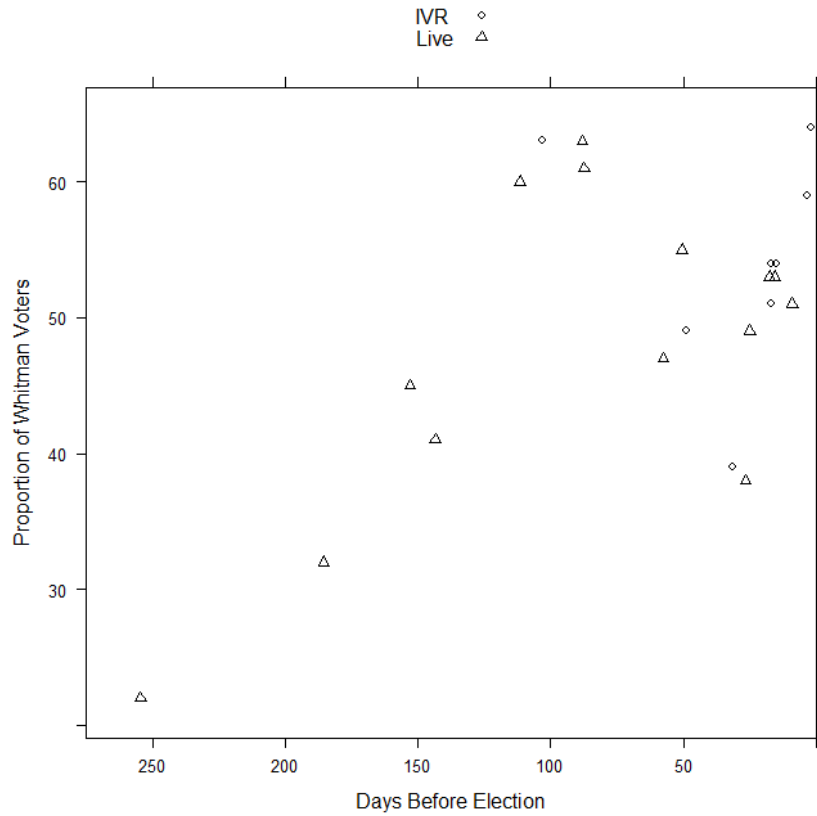
Overall, while Interactive Voice Response did not have an impact on candidate preference, the California primary data strongly supports hypothesis h1. Proximity to the election had a strong and statistically significant impact on undecided rates as well as support for each candidate. For every day closer to the primary election, the number of undecided voters declined by .14%, and support increased for both candidates, though at a higher rate for Meg Whitman. The overall model for undecided rates had an R^2 value of .88, indicating that most of the variation in undecided rates in this primary election can be explained by the date the poll was conducted and whether it was done using live interviewers or IVR. Next I will consider another election that took place at the same time, the California GOP U.S. Senate primary.

California U.S. Senate Republican Primary

The primary for U.S. Senate in California was a more ideologically charged race than the gubernatorial contest. It included three candidates: Tom Campbell, a former Congressman; Carly Fiorina, the former CEO of Hewlett Packard; and Chuck DeVore, a California State Assemblyman. DeVore was considered the most conservative of the three candidates, with Campbell the most liberal and Carly Fiorina somewhere in the middle. Although Campbell had an early lead, Fiorina eventually won the primary in a landslide with 56% compared to 22% for Campbell. DeVore's candidacy never really caught on, and he finished with 19% of the vote.

In the Republican primary for U.S. Senate in California, the use of Interactive Voice Response was associated with significantly fewer undecided voters and may have had an impact on the proportion of support for one of the candidates, Carly Fiorina. Similar to the gubernatorial primary, Interactive Voice Response polls in the U.S. Senate primary showed on average about 11 percent fewer undecided voters compared with those conducted using live interviewers. The scatterplot in Figure 6.4 tells a pretty clear story that IVR polls consistently

showed fewer undecided voters than did live interviewer surveys over the course of the campaign.

Figure 6.4: 2010 CA GOP Sen Primary Polls - Undecided
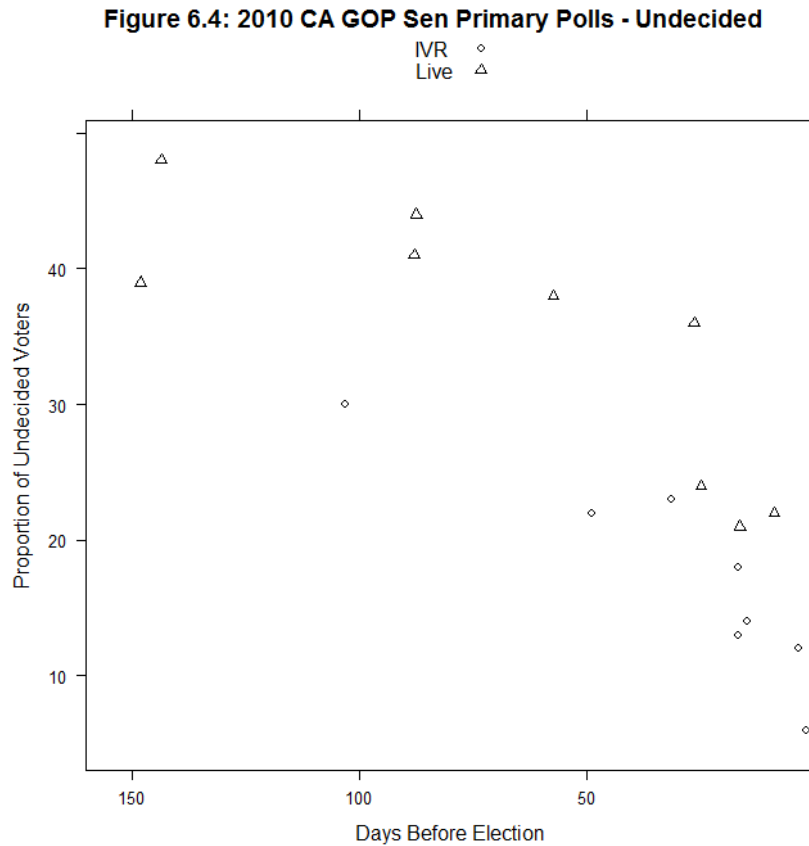
Among the candidates, the use of Interactive Voice Response was not associated with different levels of support for Campbell or DeVore, but may have had an impact on support for Fiorina. Polls conducted with IVR tended to find about six percent more support for Fiorina than those done using live interviewers. This would seem to support hypothesis h2 that as the more conservative candidate, Fiorina would have higher levels of support in IVR polls. However, the significance level for this finding is relatively low (p=.13) suggesting that this could be statistical noise in the data. The scatterplot for Fiorina's support in the primary, Figure 6.5, shows that in the last few weeks of the campaign, IVR polls indicated that she had consistently higher support than live interviewer polls did. The progression in support for her candidacy appears to be

exponential rather than linear, which could also affect the ability of the linear regression model

to accurately predict the outcome of interviewer mode on the results.

Figure 6.5: 2010 CA GOP Sen Primary Polls - Fiorina
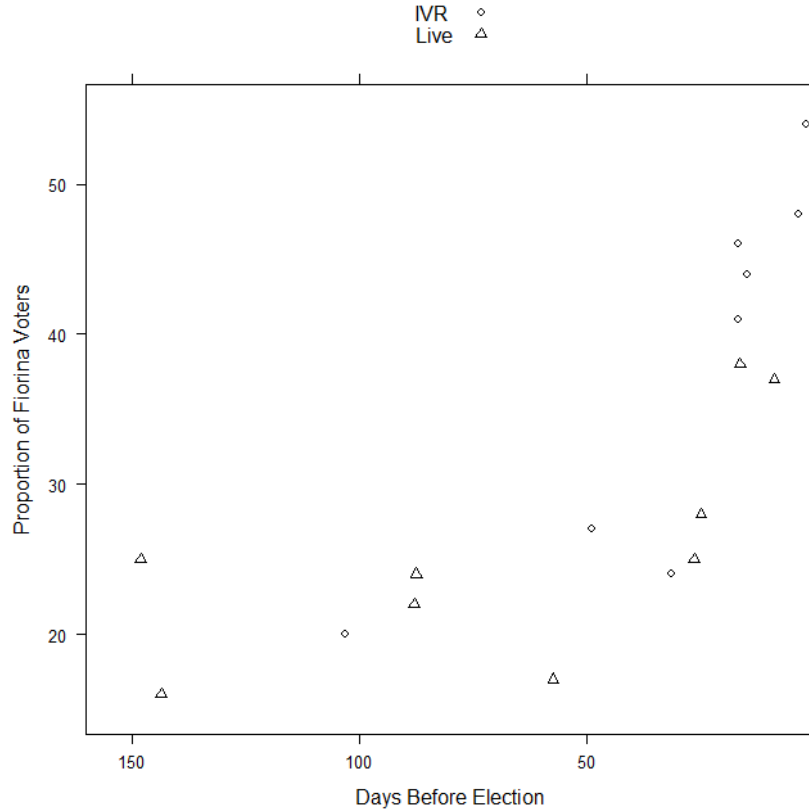
Consistent with the findings from the California gubernatorial primary, as Election Day

grew closer polls found fewer undecided voters and more voters supporting candidates. For each

day closer to the election, about .19% fewer voters indicated that they were undecided.

Proximity to Election Day was also associated with significantly higher levels of support for

Fiorina and DeVore. However, proximity to the election was actually associated with lower

levels of support for Campbell, consistent with him beginning the campaign with a lead in polls

before losing support.

Figure 6.6: 2010 CA GOP Sen Primary Polls - Campbell
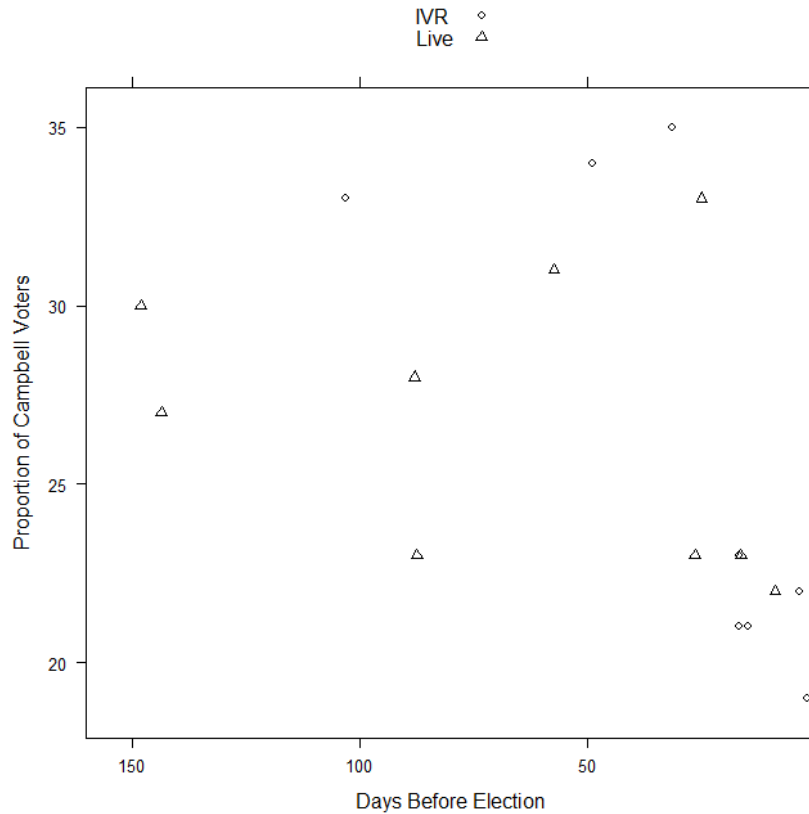
The biggest surprise in the California U.S. Senate primary is that the use of IVR was not associated with higher levels of support for the most conservative candidate in the field, Chuck DeVore. DeVore was the type of candidate who should have received strong support from the type of highly engaged and opinionated conservative primary voters who propelled Christine O'Donnell to victory in Delaware. However, this may be more reflective of the fact that DeVore's campaign failed to catch on with voters. The conservative base of support he was counting on for victory for the most part supported Carly Fiorina.

Figure 6.7: 2010 CA GOP Sen Primary Polls - DeVore

IVR ◇
Live △
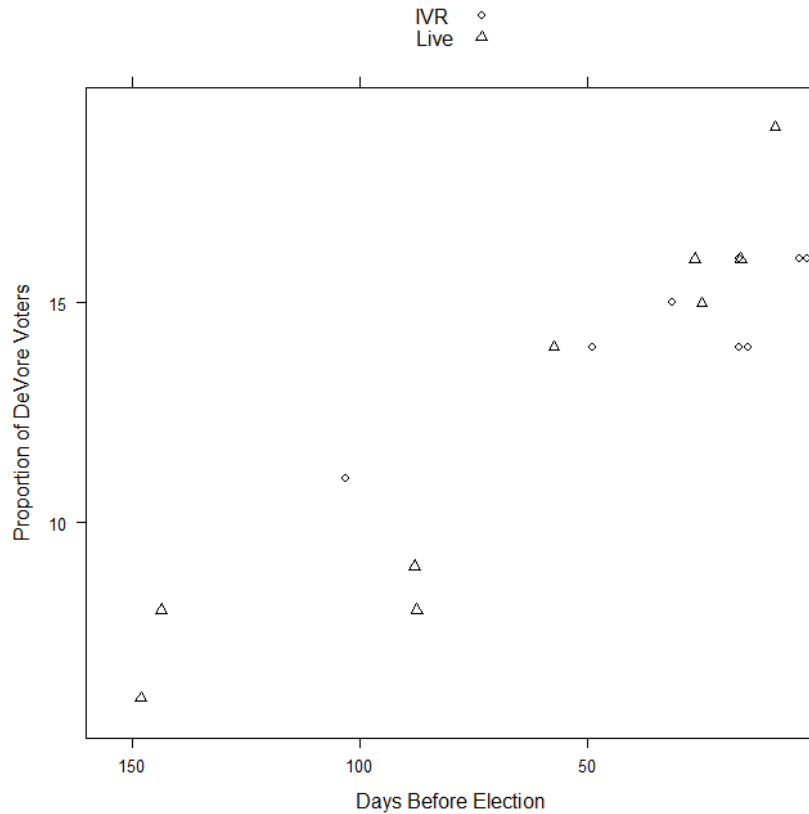
Overall, the California GOP U.S. Senate primary results lend some support for the hypothesis that Interactive Voice Response finds higher levels of support for a more ideologically extreme candidate in a partisan primary election. The higher levels of support for Carly Fiorina among IVR polls, especially at the end of the campaign, indicate that IVR polls may have been surveying a more highly engaged and opinionated slice of the Republican primary electorate. However, the impact on Fiorina's support levels was only significant with 87% confidence and support for the most conservative candidate Chuck DeVore was not higher in IVR polls, so the results are inconclusive.

Pennsylvania U.S. Senate Democratic Primary

The 2010 Democratic primary for U.S. Senate in Pennsylvania was a highly ideological contest between a longtime incumbent and a challenger who was also a sitting member of Congress. In early 2009, former Republican Congressman Pat Toomey announced that he would challenge five-term Republican U.S. Senator Arlen Specter for the GOP nomination for U.S. Senate. Specter had previously defeated Toomey for the GOP U.S. Senate nomination by a narrow margin in 2004. Less than two months after this announcement, Arlen Specter switched parties and announced he would run for reelection as a Democrat. Shortly after, Congressman Joe Sestak announced that he would challenge Arlen Specter for the Democratic nomination. Despite trailing in early polls, Sestak ultimately defeated Specter in the May 2010 primary by a margin of 54% to 46%.

Based on the nature of the campaign, if IVR polls survey a more engaged and opinionated portion of the electorate, they should show greater levels of support for the lesser-known and more liberal Joe Sestak. A broader sample of likely Democratic primary voters surveyed by live interviewers should show higher levels of support for the more moderate Arlen Specter. However, no statistically significant differences in support levels for either of the two candidates appeared between live interview and Interactive Voice Response polls. There was a statistically significant difference between undecided rate between the two kinds of polls (p=.055) but that was the only result for which the type of telephone poll made a difference.

A possible reason that the model did not find a significant difference is that most early polls done using live interviewers in this race were of registered voters rather than likely voters. As a result, they were excluded from the model. A number of early live interview surveys of registered voters showed remarkably lower levels of support for Joe Sestak than did IVR surveys

of likely voters conducted during the same time period. For example, a live interviewer

Susquehanna poll of registered voters conducted in October of 2009 found 44% of voters

supporting Arlen Specter but only 16% supporting Joe Sestak. An IVR Rasmussen Reports poll

of likely voters taken the same week found 46% for Arlen Specter and 44% for Joe Sestak

(Huffington Post: Pollster.com 2011). While it is likely that this difference in level of support is

partially due to survey method, because the Susquehanna poll did not try to identify likely voters

it is not an apples to apples comparison. The result of the exclusion of this and other surveys is

illustrated by Figures 6.8-6.10, the scatterplots of the Pennsylvania Senate primary. They show

no live interview surveys of likely voters taken from 350 days before the primary until about 50

days before the primary.

Figure 6.8: 2010 PA DEM Sen Primary Polls - Undecided
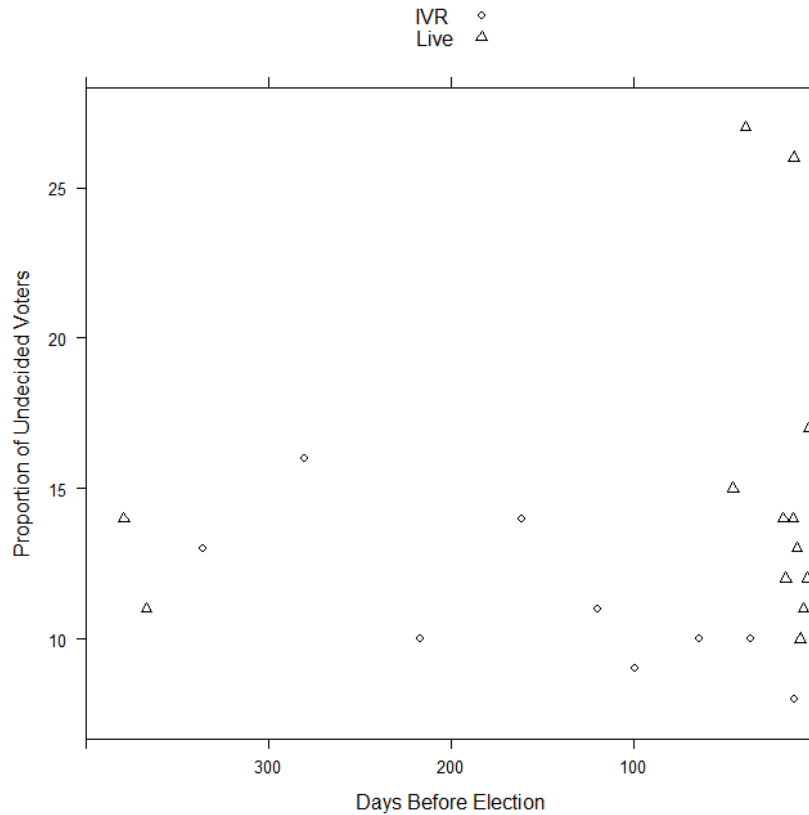
Figure 6.9: 2010 PA DEM Sen Primary Polls - Specter

IVR ◇
Live △

Proportion of Specter Voters

Days Before Election
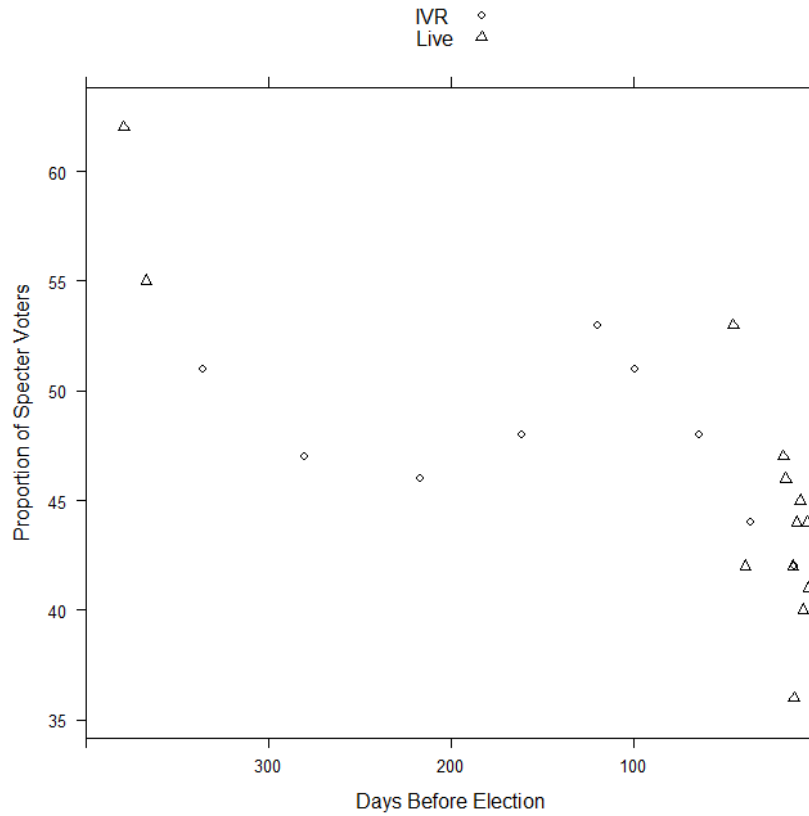
Figure 6.10: 2010 PA DEM Sen Primary Polls - Sestak

IVR ◇
Live △

Proportion of Sestak Voters

Days Before Election

73
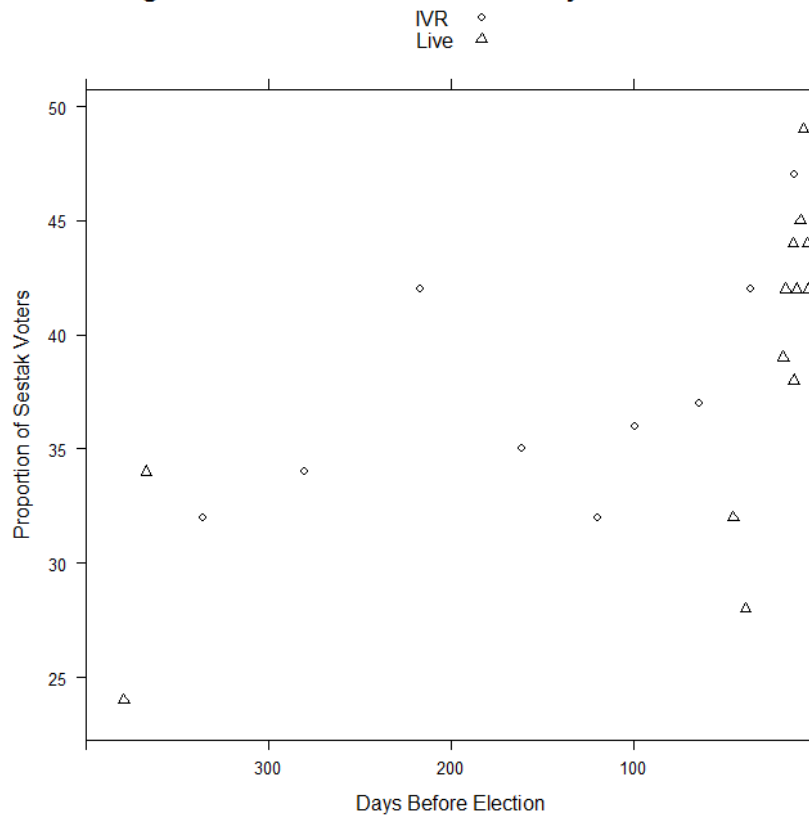
Unlike the California primaries, proximity to Election Day was not associated with any reduction in the proportion of undecided voters. However, the closer polls were to Election Day, the more support they showed for Joe Sestak and the less for Arlen Specter. This is consistent with the result of the election, in which Sestak closed a deficit to ultimately defeat the incumbent.

In light of the concerns regarding the absence of any live interviewer surveys during most of the campaign, I re-ran both the regression models and the scatterplots for this election focusing exclusively on polls that were conducted over the final few months of the campaign, specifically any within 125 days of the election. Table 6.4 reflects the revised data.

Table 6.4: Influences on Candidate Preferences: Pennsylvania 2010 Democratic Primary

*** p < .001

	Overall Model		Survey Type (IVR=1)		Days Before Primary	
	Multiple R^2	Significance	Estimate	Significance	Estimate	Significance
Undecided %	0.353	0.059	-6.862	0.028	0.038	0.347
Specter %	0.598	0.003	-2.775	0.243	0.125	0.002
Sestak %	0.639	0.001	7.373	0.018	-.181	***

Overall, the Pennsylvania poll results did not support hypothesis h2. This may largely be due to limitations of the dataset, in particular the absence of any live interviewer surveys of likely voters until very late in the campaign. Expanding the analysis to include polls of registered voters would undoubtedly find higher levels of support for Joe Sestak among IVR surveys, however it impossible to say whether that higher level is due to the difference in interviewer mode or the difference in sample composition.

Florida Republican U.S. Senate Primary

In early 2009, former Speaker of the Florida House of Representatives Marco Rubio announced his candidacy for the U.S. Senate seat in Florida that opened up following the retirement of Senator Mel Martinez. Shortly after, Florida Governor Charlie Crist announced his own candidacy after being recruited to run by national Republicans. Crist, a moderate, was perceived as a lock to win the general election. Rubio was perceived as the more conservative candidate, and began the race much less well known than the sitting governor. Despite trailing early, Rubio gained support over the course of the campaign, and by the spring of 2010 it was clear from polls that he would win the primary. Governor Crist then withdrew from the Republican primary and ran for the U.S. Senate as an Independent. Marco Rubio won the general election in a three way race against Crist and the Democratic nominee, Congressman Kendrick Meek. Although Rubio ultimately won the Republican nomination essentially unopposed, extensive polling was done during the course of the primary campaign between Rubio and Charlie Crist.

First, looking at undecided rates, Interactive Voice Response polls showed significantly lower levels of undecided voters compared to those conducted using live interviewers. On average, IVR polls in the Florida primary had about 10% fewer undecided voters than live interviewer polls did. While undecided rates did appear to decline as the election approached, the association was not as strong as in California. On average each day closer to the election meant only .018% fewer undecided voters (p=.098).

Figure 6.11: 2010 FL GOP Sen Primary Polls - Undecided
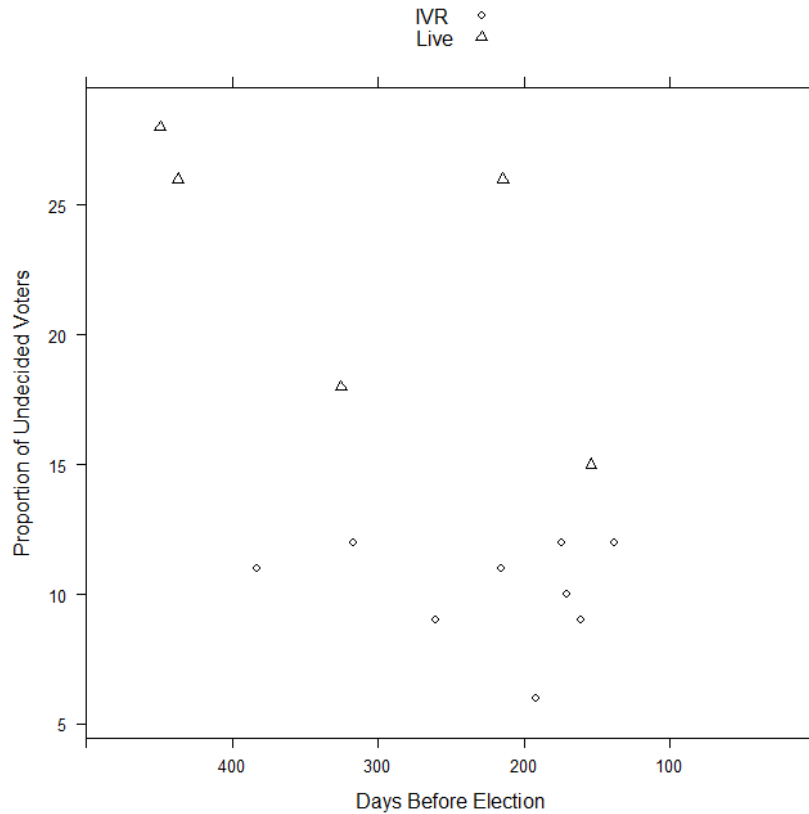
The regression model shows that IVR polls found significantly higher levels of support

for Rubio's candidacy than did live interviewer polls. On average, IVR surveys found six

percentage points more support for Rubio than did live interviewer surveys (p=.011). As support

for Rubio climbed during the course of the campaign, there was also a strong association

between proximity to Election Day and support for Rubio. For each day closer to the election,

polls found .113% more support for Rubio's candidacy.

Figure 6.12: 2010 FL GOP Sen Primary Polls - Rubio

IVR ◇
Live △
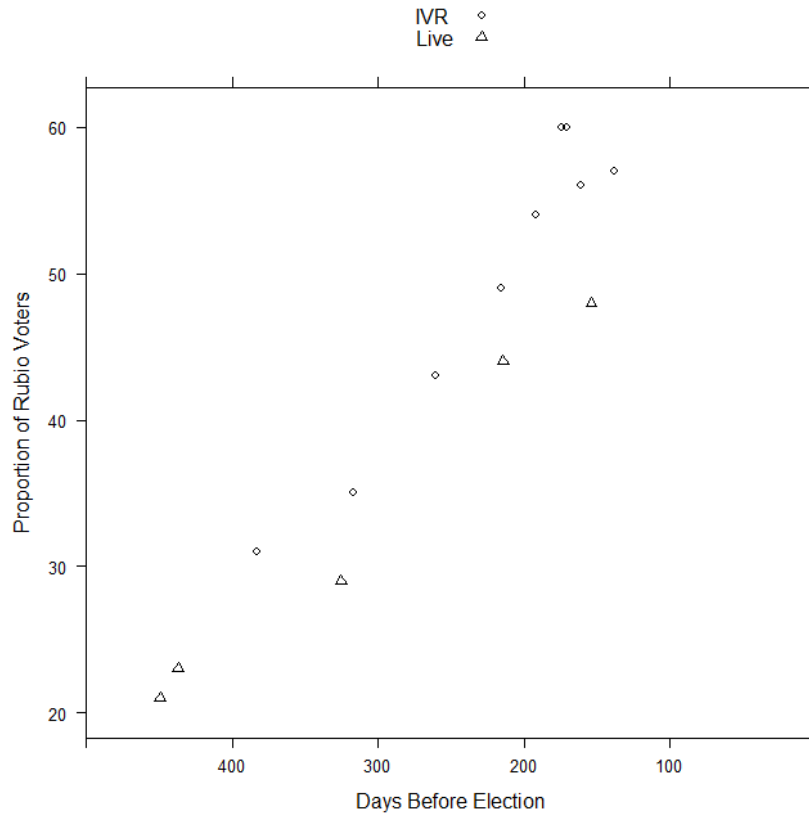
Perhaps most remarkably, the overall R^2 for Rubio's candidacy was .954, indicating that

over 95% of the variation in support for his candidacy over the course of the primary campaign

could be accounted for just with the date the poll was conducted and whether the poll was done

using IVR or live interviewers. While Figure 6.12 shows that Rubio's support increased in a

linear manner over the course of the campaign, this is still an amazing finding considering that

the polls were conducted by eight different pollsters all of whom have their own methodologies

and research techniques. Despite the differences in how those firms did their research, they only

accounted for five percent of the variation in poll support for Marco Rubio's candidacy.

I noted in Chapter 5 that some researchers have expressed concerns with the use of R^2

when using WLS regressions, as I have done here (Willett and Singer 1988). In order to confirm

this finding then, I ran out a regular OLS regression on support for Marco Rubio in the Florida

primary election. The OLS regression without assigning weights to account for the differences in sample size finds an R^2 value of .953, nearly identical to the results of the WLS regression that included the weights. Thus, the high amount of variance accounted for by the model is not simply a result of the regression weights skewing the data.

The proportion of poll support for Charlie Crist was unaffected by the use of IVR or live interviewers, though it was affected by the proximity to Election Day. Polls conducted closer to the election found significantly lower levels of support for Crist. It is interesting that while IVR was associated with 10 percent fewer undecided voters and 6 percent more Rubio voters, support for Crist was not affected. Because no primary polls were done after Charlie Crist switched to running as an Independent, I cannot say for sure whether this would have changed had the final months of the primary election been allowed to play out.

Figure 6.13: 2010 FL GOP Sen Primary Polls - Crist
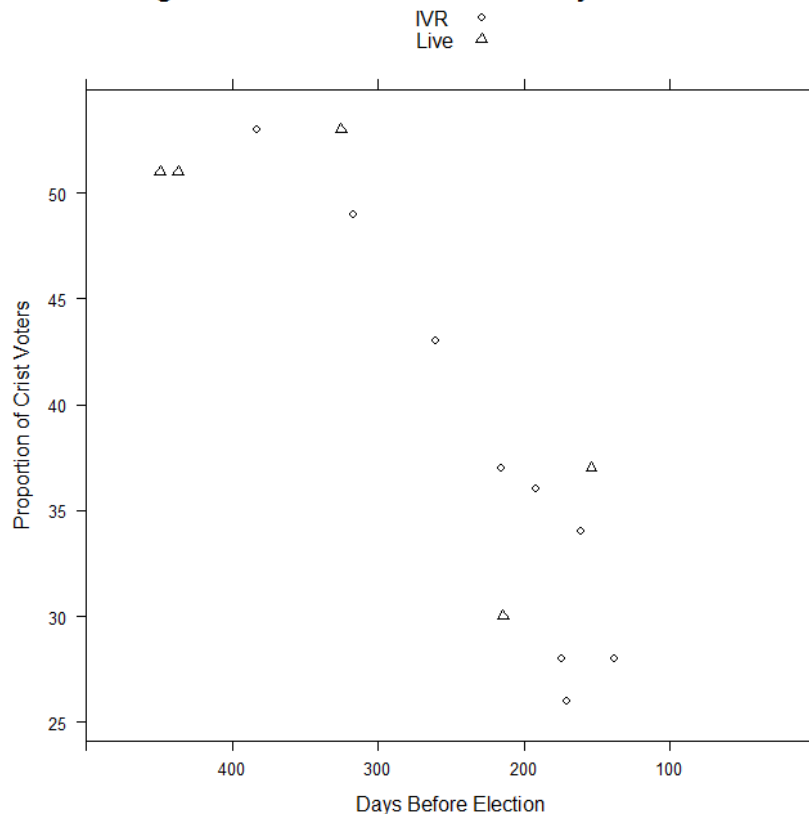
Overall, the results for the Florida Republican U.S. Senate primary do support hypothesis h2. Similar to the Van Lohuizen and Samohyl study in Delaware, an analysis of the Florida GOP primary results shows that in a primary between a lesser known, ideologically conservative Republican and a better known, ideologically moderate Republican, IVR polls tend to indicate greater levels of support for the more conservative candidate than do polls conducted using live interviewers. On average, IVR polls found 10 percent fewer undecided voters and six percent more support for Marco Rubio compared to live interview polls. That indicates that there may be a similar phenomenon at play in the Florida primary as was present in the Delaware primary.

Summary of Primary Polling Findings

Among the four primary elections I looked at, there was some support for hypothesis h2 but it was not as clear or as strong as I anticipated. Based on the nature of the four primaries, I expected that there would be little impact of survey mode in the California gubernatorial primary but it would be visible in the other three. Instead, only the Florida GOP primary showed a clear and consistent impact of survey mode on support for one of the candidates.

The primary results also provided some additional support for h1 in the form of much lower undecided rates among IVR poll results. Across the three Republican primaries in California and Florida, the use of Interactive Voice Response was associated with about ten percent fewer undecided voters when compared to surveys conducted using live interviewers. This figure was both consistent across these three primaries as well as substantially higher than the differences found in undecided rates in general elections in Chapter 5.

However, the lack of congruity between the Pennsylvania data and the other three primaries presents a challenge from an analytical standpoint. There was no statistically significant effect of interviewer mode on either support for the candidates or undecided rate in the Pennsylvania data. This may be a result of the data challenges I discussed, but I cannot say this with certainty. It is also the only Democratic primary among the four that I analyzed, so I have no evidence to support my thesis that IVR surveys sample a more engaged and opinionated voter group in Democratic primaries as well as Republican ones.

The final chapter will review the findings from this research and what it means for the broader literature on survey methods and the future of political polling.

CHAPTER 7: Conclusions

In advance of nearly all major contested elections in the United States, public organizations conduct polls of "likely voters" that purport to let the public know the status of the election horse race at that moment in time. These polls, from a variety of sources, cannot be fairly judged to be accurate or inaccurate because there is nothing to compare them against other than an election weeks or months away when the state of the race might be very different. There is no imminent election outcome and no way to say with certainty whether the pool of voters being surveyed reflects what the eventual electorate will look like. A great deal can be learned, however, by comparing these poll results to each other. This study focused on whether there are consistent differences between poll results from surveys conducted using Interactive Voice Response (IVR) and those done using live interviewers.

Although it is difficult to say whether one approach or the other is more desirable, Interactive Voice Response surveys consistently find fewer undecided voters in general election surveys than do surveys conducted using live interviewers. At least in U.S. Senate and gubernatorial races in 2010, Interactive Voice Response surveys show significantly lower percentages of undecided voters independent of the date of the survey and the state in which the survey was conducted. Even though this pattern was not present in every election, it appeared in the vast majority for which I had sufficient numbers of polls of each type for comparison. In these elections, Interactive Voice Response polls showed anywhere between three and fifteen percent fewer undecided voters when compared to live interview surveys. The difference leads to the conclusion that Interactive Voice Response surveys poll a narrower, more opinionated segment of the electorate compared to surveys conducted using live interviewers.

I also hypothesized that in contested primary elections, which normally are lower information environments; there would be differing results in some primaries due to the narrower subset of voters surveyed by IVR exhibiting difference levels of preferences for candidates compared to the larger pool of primary voters. The results of analyzing U.S. Senate and gubernatorial primary elections were less clear. Due to the very small number of surveys conducted in advance of most primaries, only a handful of races had sufficient surveys for analysis. Of the four races I looked at, only the Republican Senate primary in Florida showed consistent differences along the expected lines: support was higher for the more ideologically extreme candidate (Marco Rubio) in the IVR surveys. However, the Florida data are not without problems: because Governor Crist dropped out of the race, there was not the usual large number of surveys taken in the final weeks before the primary.

That no similar pattern appeared in the Democratic primary in Pennsylvania may mean that the phenomenon of IVR surveys showing more support for more ideologically extreme candidates may apply only in Republican primaries. It is also possible that this may have been a feature of the 2010 election cycle and the mood of Republican primary voters during that specific time period. Van Lohuizen and Samohyl suggested when looking at the 2010 Delaware GOP primary that the same phenomenon likely would be found in Democratic primaries as well. It is also possible that the phenomenon exists in Democratic primaries as well, but it was not observed in Pennsylvania due to issues specific to the polling in that case, as well as the peculiarities of that race in which a longtime incumbent switched parties.

This study was limited by the use of a single midterm election to provide the data. At the time this study began, the 2010 midterm election was the most recent national election and

provided a broad pool of polling data for many competitive elections across many different states.

The frequent release of public polling on statewide elections during the course of campaigns is a relatively new phenomenon, as is the collection of each of these poll results in easy to navigate online aggregators such as Pollster.com and RealClearPolitics.com. In elections between 1990 and 2002, there were on average fewer than 200 likely voter polls conducted for all U.S. Senate elections (McGhee 2014). However, this number increased rapidly starting in the following years, with around 500 polls being conducted in 2006 and 2008. The 2010 elections had nearly 700 published U.S. Senate polls, substantially more than any of the preceding elections, and more than in 2012 (McGhee 2014).

With only the last few elections providing large numbers of polls to analyze, I chose to focus entirely on the 2010 midterms. Polling from the 2008 elections might have been included in this research, but I chose not to do so. One reason was that I was concerned that the distinctive political environment in 2008 might have been problematic for analyzing poll results. The country was largely focused on the presidential election, and there were few elections for governorships that year. While there were several competitive U.S. Senate races, there were relatively few early polls conducted in those races, possibly due to the focus that spring on the lengthy Democratic presidential primary. For example, an analysis by FiveThirtyEight indicated that only 40 percent of 2008 U.S. Senate races were polled between January and March of that year compared to 59 percent of 2010 U.S. Senate races that were polled during the same period in 2010 (Enten 2014). As such, I believed that the 2010 elections provided the best opportunity for a wealth of easily available and recent polling data to study.

Overall, although the results of this study support the claim that IVR surveys sample a more opinionated subset of the voting population, more research is needed to confirm this finding. A broader study in which a survey is administered using a split sample of live interviewers and Interactive Voice Response, across multiple elections at different points during the campaigns, holding other factors constant, would be more effective than using existing data as this analysis did. Above all, this study is limited by what is unknown about the methods that pollsters employed in every individual survey studied. Unless pollsters begin releasing much more detailed information about their methodologies, aggregate studies of poll results will be limited. Based on this research, a comprehensive study would likely find that substantial differences exist between the samples in IVR polls compared to those in live interview polls despite other differences in polling methodology.

Consumers of public opinion polling should remember that there is no universally accepted definition of a poll being "right" weeks or months before Election Day. A poll released two months prior to an election might show proportions of support for candidates in line with the eventual outcome, but that does not mean that it reflected levels of support for those candidates at that moment in time. Thus the predictive accuracy of a poll might not be the best measure of its validity. A poll might show a close race between two candidates months before an election because those two candidates are well-known, established figures with long track records in public life. Or, it might show a close race because the voters measured by the poll reflects only the most engaged and opinionated voters who are paying attention to the campaign so far in advance of an election.

As long as polling firms release poll results to the public without fully disclosing their methods, experience and intuition will continue to be of utmost importance for the interpretation

of results. When two or more polls are released that are markedly different, there are any number of questions that should be asked. Did the polls survey likely voters or registered voters? Did they include cell phones? Were they conducted by a partisan polling firm, a media pollster, or an academic institute? Was it paid for by an organization with a vested interest in portraying a certain narrative about the campaign?

Based on this research, I believe that the question of whether a poll was conducted using Interactive Voice Response or live interviewers should remain a central consideration when analyzing poll results. IVR polls have been established as effective last-minute predictors of election outcomes. However, the results of this research showing consistently lower levels of undecided voters in IVR surveys in both primary and general elections indicate likely differences in these poll methodologies that are not well understood. The analysis of the 2010 Florida GOP primary combined with previous research conducted on the 2010 Delaware GOP primary and anecdotal evidence from the 2010 Massachusetts special election indicate that IVR surveys may more accurately predict an outcome while misrepresenting the current snapshot of where things stand in the race.

This poses a particular challenge for journalists in an era when poll results can be self-fulfilling prophecies. The successful 2010 Scott Brown for U.S. Senate campaign gained significant media attention from the results of a single (IVR) Rasmussen poll showing him down by single digits, at a time when internal polls for his opponent's campaign showed him with much lower levels of support (Mellman 2010). If the Rasmussen survey had been conducted using live interviewers instead of IVR and the results showed Brown trailing by a larger margin, it might have significantly changed the media narrative surrounding the race. There are likely

many other elections in which early IVR polls have been important to media coverage of the campaigns.

While journalists and political elites are often the first to read and interpret poll results, polls can affect all citizens. Polls provide an important way for citizens to express their opinions and be heard by elites, so a different poll result due to a pollster's methodology can have an enormous impact. A politician may choose to vote for or against a piece of legislation based on poll results he or she receives. Furthermore, a citizen may choose to participate in an election based on a news story about a poll showing a tight race, or donate time or money to a campaign based on the perception of its competitiveness. The finding that polls can produce differing results depending on whether a pollster uses an IVR methodology can thus have a great deal of impact if it changes the way that poll results are interpreted. But whether a poll was conducted using live interviewers or IVR, it remains an open question as to which method produces the "better" result.

References

AAPOR Ad Hoc Committee on the 2008 Presidential Primary Polling. 2009. *An Evaluation of the Methodology of the 2008 Pre-Election Primary Polls.* American Association for Public Opinion Research.

Bloom, Joel David. 2003. "Reliable Compared to What? Empirical Tests of the Accuracy of Election Polls, 2002." *University of Oregon.* June 20. Accessed February 11, 2012. http://www.uoregon.edu/~jbloom.

Blumberg, Stephen J, and Julian V Luke. 2012. "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January - June 2012." *National Center for Health Statistics.* December. Accessed August 10, 2013. http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201212.PDF.

Blumenthal, Mark. 2011a. "Can I Trust This Poll?" Chap. 3 in *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*, edited by Kirby Goidel, 54-70. Baton Rouge: Louisiana State University Press.

—. 2009. "In Defense of Automated Surveys." *Pollster.com.* Edited by Mark Blumenthal. September 14. Accessed April 30, 2011. http://www.pollster.com/blogs/in_defense_of_automated_survey.php?nr=1.

—. 2010. "Not All Polls Were Wrong in Nevada." *The Huffington Post: Pollster.* November 5. Accessed January 15, 2012. http://www.huffingtonpost.com/2010/11/05/not-all-polls-were-wrong-_n_779705.html.

—. 2011b. "Poll Accuracy Held Steady in 2010." *The Huffington Post: Pollster.* February 14. Accessed January 15, 2012. http://www.huffingtonpost.com/2011/02/14/poll-accuracy-held-steady-in-2010_n_823181.html.

Blumenthal, Mark, and Charles Franklin. 2007. "Methods and Horse Races: Phone, IVR and Internet Polls in the 2006 Elections." *American Association for Public Opinion Research.* Anaheim. 4.

Brehm, John. 1993. *The Phantom Respondents.* Ann Arbor, MI: The University of Michigan Press.

Cohen, Jon. 2012. "Behind the Numbers." *The Washington Post.* February 16. Accessed August 10, 2013. http://www.washingtonpost.com/blogs/behind-the-numbers/post/covering-automated-surveys-in-2012/2012/02/01/gIQANvmJIR_blog.html.

Converse, Patrick D, Edward W Wolfe, Xiaoting Huang, and Frederick L Oswald. 2008. "Response Rates for Mixed-Mode Surveys Using Mail and E-Mail/Web." *American Journal of Evaluation* 29 (1): 99-107.

Corkrey, Ross, and Lynne Parkinson. 2002. "Interactive Voice Response: Review of Studies 1989-2000." *Behavior Research Methods, Instruments & Computers* 34 (3): 342-353.

Couper, Mick P, Eleanor Singer, and Roger Tourangeau. 2004. "Does Voice Matter? An Interactive Voice Response (IVR) Experiment." *Journal of Official Statistics* 20 (3): 551-570.

Dennis, J Michael, Cindy Chatt, Rick Li, Alicia Motta-Stanko, and Paul Pulliam. 2005. "Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone versus Internet." *Knowedge Networks.* January. Accessed November 10, 2014. http://www.knowledgenetworks.com/ganp/docs/research-0105.pdf.

Dillman, Don A, and Leah Melani Christian. 2005. "Survey Mode as a Source of Instability." *Field Methods* 17 (1): 30-52.

Dillman, Don A, Glenn Phelps, Robert Tortora, Karen Swift, Julie Kohrell, Jodi Berck, and Benjamin L Messer. 2009. "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR), and the Internet." *Social Science Research* 38: 1-18.

Dunaway, Johanna. 2011. "Poll-Centered News Coverage." Chap. 4 in *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*, edited by Kirby Goidel, 71 - 84. Baton Rouge, LA: Lousiana State University Press.

Enten, Harry. 2014. "Senate Races in 2014 Lack (Good) Polling." *FiveThirtyEight.* April 4. Accessed November 10, 2014. http://fivethirtyeight.com/datalab/senate-races-in-2014-lack-good-polling/.

Goidel, Kirby. 2011. "Publlic Opinion Polling in a Digital Age: Meaning and Measurement." Chap. 1 in *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*, edited by Kirby Goidel, 11-27. Baton Rouge: Louisiana State University Press.

Green, Donald P, and Alan S. Gerber. 2006. "Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts." *Public Opinion Quarterly* 70 (2): 197-223.

Greenberg, Anna. 2011. "Alternatives to Polling." Chap. 7 in *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*, edited by Kirby Goidel, 115-123. Baton Rouge: Louisiana State University Press.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70 (5): 646-675.

Huffington Post: Pollster.com. 2011. May 5. Accessed April 5, 2014. http://www.huffingtonpost.com/2009/06/09/10-pa-sen-dempr-svse_n_724727.html.

Keeter, Scott. 2011. "Public Opinion Polling and Its Problems." Chap. 2 in *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*, edited by Kirby Goidel, 28-53. Baton Rouge, LA: Louisiana State University Press.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64 (2): 125-148.

Keeter, Scott, Courtney Kennedy, April Clark, Trevor Tompson, and Mike Mokrzycki. 2007. "What's Missing from National Landline RDD Surveys: The Impact of the Growing Cell-Only Population." *Public Opinion Quarterly* 772-792.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. 2006. "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey." *Public Opinion Quarterly* 70 (5): 759-779.

Keeter, Scott, Leah Christian, and Michael Dimock. 2010. "The Growing Gap between Landline and Dual Frame Election Polls." *Pew Research Center.* November 22. Accessed January 22, 2012. http://pewresearch.org/pubs/1806/growing-gap-between-landline-and-dual-frame-election-polls.

Knoke, David, and George W., Mee, Alisa Potter Borhnstedt. 2002. *Statistics for Social Data Analysis.* 4. Belmont, CA: Thomas Wadsworth.

McFarland, Sam G. 1981. "Effects of Question Order on Survey Responses." *Public Opinion Quarterly* 45 (2): 208-215.

McGhee, Eric. 2014. "Why the polls are better than ever at predicting Senate races." *The Washington Post.* October 28. Accessed November 10, 2014. http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/10/28/why-the-polls-are-better-than-ever-at-predicting-senate-races/.

Mellman, Mark. 2010. "Robo-polls and human error." *The Hill.* February 2. Accessed May 9, 2011. http://thehill.com/opinion/columnists/mark-mellman/79351-robo-polls-and-human-error.

Millard, Richard W, and Joseph R Carver. 1999. "Cross-sectional Comparison of Live and Interactive Voice Recognition Administration of the SF-12 Health Status Survey." *The American Journal of Managed Care* 5 (2): 153-159.

National Council on Public Polls. 2011. "NCPP Analysis of Final Statewide Pre-Election Polls 2010." *National Council on Public Polls.* February 8. Accessed January 15, 2012. http://ncpp.org/?q=node/135.

Panagopoulos, Costas. 2009. "Preelection Poll Accuracy in the 2008 General Elections." *Presidential Studies Quarterly* 39 (4): 896 - 907.

Rasmussen Reports. 2012. *About Rasmussen Reports.* Accessed February 10, 2012. http://www.rasmussenreports.com/public_content/about_us.

Rothenberg, Stuart. 2009. "Learning the Right Lesson's from Tuesday's Results." *The Rothenberg Political Report.* November 9. Accessed February 11, 2012. http://www.realclearpolitics.com/articles/2009/11/09/learning_the_right_lessons_from_tu esdays_results_99077.html.

Shalizi, Cosma Rohilla. 2013. "Advanced Data Analysis from an Elementary Point of View." *Carnegie Mellon University Department of Statistics.* April 25. Accessed August 5, 2013. http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/ADAfaEPoV.pdf.

Silver, Nate. 2008a. *Pollster Ratings.* May 28. Accessed January 21, 2012. http://www.fivethirtyeight.com/2008/05/pollster-ratings-v311.html.

—. 2010. *The Uncanny Accuracy of Polling Averages, Part II: What the Numbers Say.* September 30. Accessed January 15, 2012. http://fivethirtyeight.blogs.nytimes.com/2010/09/30/the-uncanny-accuracy-of-polling-averages-part-2-what-the-numbers-say/.

—. 2008b. *Tracking Poll Primer.* October 21. Accessed January 21, 2012. http://www.fivethirtyeight.com/2008/10/tracking-poll-primer.html.

Survey USA. 2012. *Who is SurveyUSA?* Accessed February 10, 2012. http://www.surveyusa.com/index.php/who-is-surveyusa/.

Tourangeau, Roger, Darby Miller Steiger, and David Wilson. 2002. "Self-Admistered Questions by Telephone: Evaluating Interactive Voice Response." *Public Opinion Quarterly* 66: 265-278.

Traugott, Michael W, and Christopher Wlezien. 2009. "The Dynamics of Poll Performance During the 2008 Presidential Nomination Contest." *Public Opinion Quarterly* 73 (5): 866-894.

Traugott, Michael W. 2005. "The Accuracy of the National Preelection Polls in the 2004 Presidential Election." *Public Opinion Quarterly* 69 (5): 642 - 654.

Traugott, Michael W., and Paul J. Lavrakas. 2008. *The Voter's Guide to Election Polls.* Lanham, MD: Rowman & Littlefield.

U.S. Code. 2012. "Section 227 b(1)." *Legal Information Institute.* Accessed February 11, 2012. http://www.law.cornell.edu/uscode/search/display.html?terms=227%20b&url=/uscode/html/uscode47/usc_sec_47_00000227----000-.html.

Underwood, Benton J, Carl D Duncan, Janet A Taylor, and John W Cotton. 1954. *Elementary Statistics.* New York, NY: Appleton-Century-Crofts, Inc.

van Lohuizen, Jan, and Robert Wayne Samohyl. 2011. "Method Effects and Robo-polls." *Survey Practice.*

Voogt, Robert J J, and William E Saris. 2003. "To Participate or Not to Participate: The Link Between Survey Participation, Electoral Participation, and Political Interest." *Political Analysis* 11: 164-179.

Willett, John B, and Judith D Singer. 1988. "Another Cautionary Note About R^2: Its Use in Weighted Least-Squares Regression Analysis." *The American Statistician* 42 (3): 236-238.

Appendix: General Election Scatterplots

Polls Taken in 2010 Alaska Senate Election

Polls Taken in 2010 Alabama Senate Election

Polls Taken in 2010 Arkansas Senate Election

Polls Taken in 2010 Arizona Senate Election
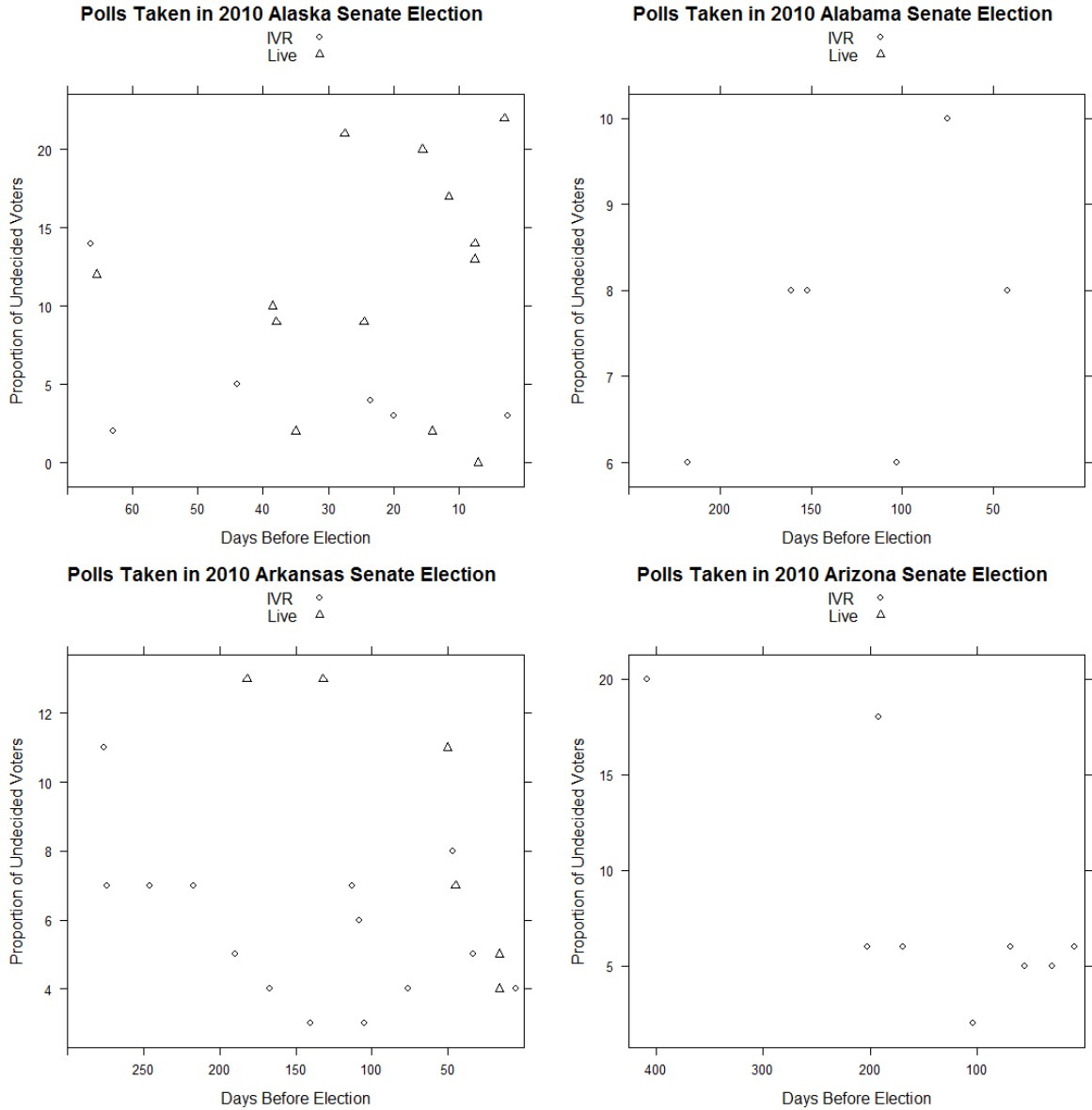
Polls Taken in 2010 California Senate Election
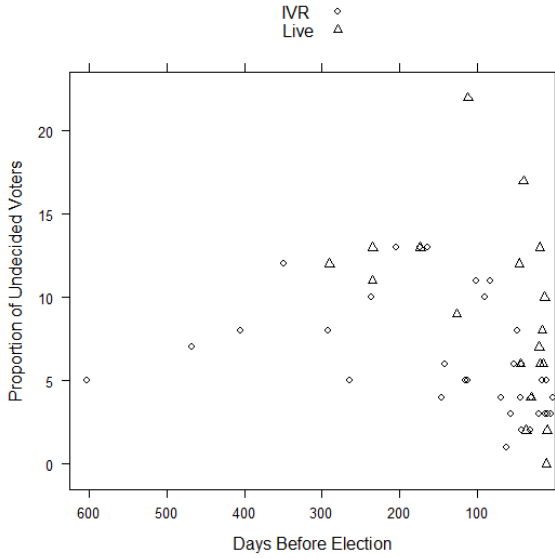
Polls Taken in 2010 Colorado Senate Election
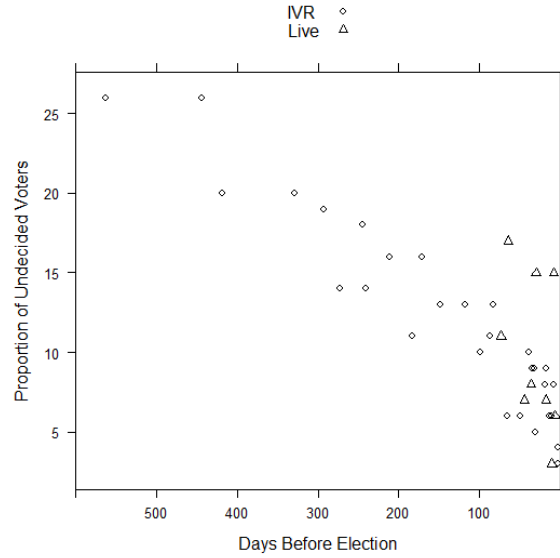
Polls Taken in 2010 Connecticut Senate Election
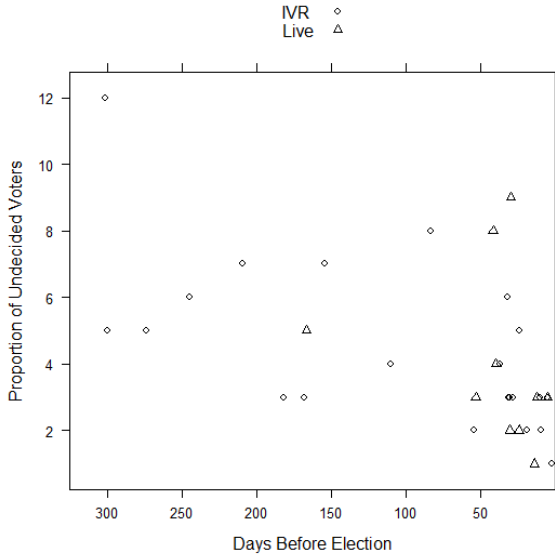
Polls Taken in 2010 Delaware Senate Election
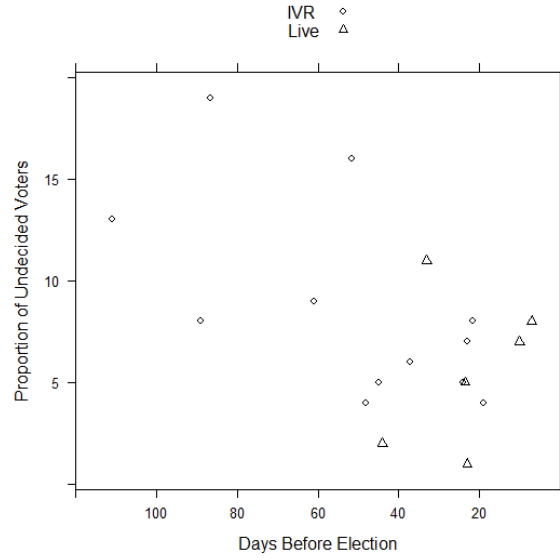
Polls Taken in 2010 Florida Senate Election
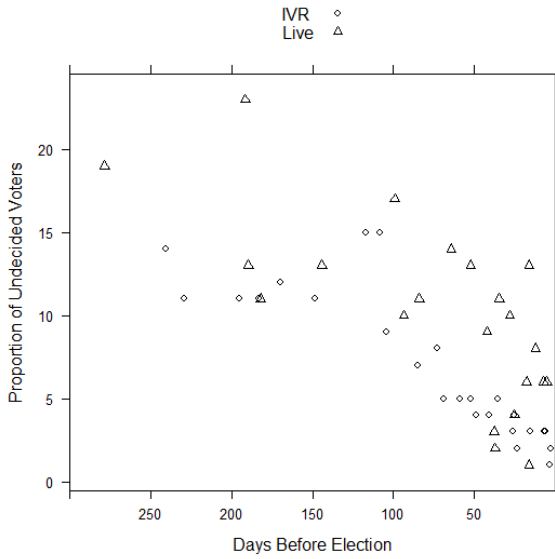
Polls Taken in 2010 Georgia Senate Election
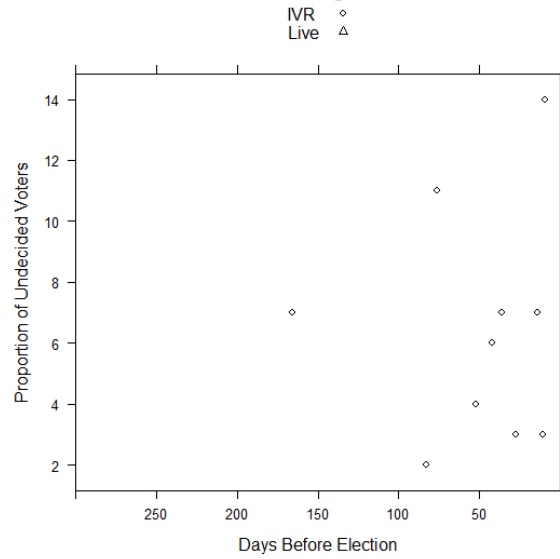
Polls Taken in 2010 Hawaii Senate Election
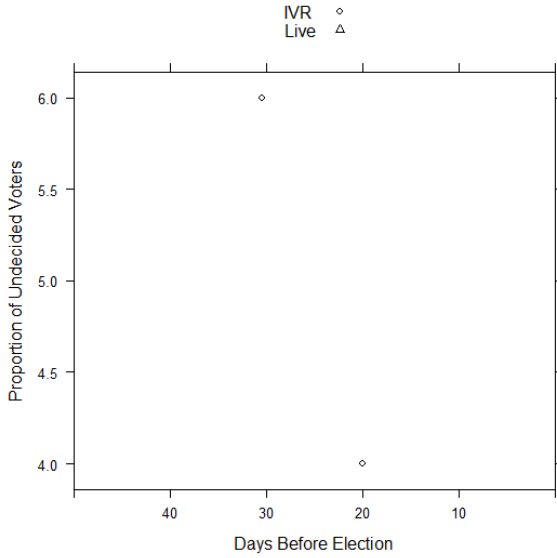
Polls Taken in 2010 Iowa Senate Election
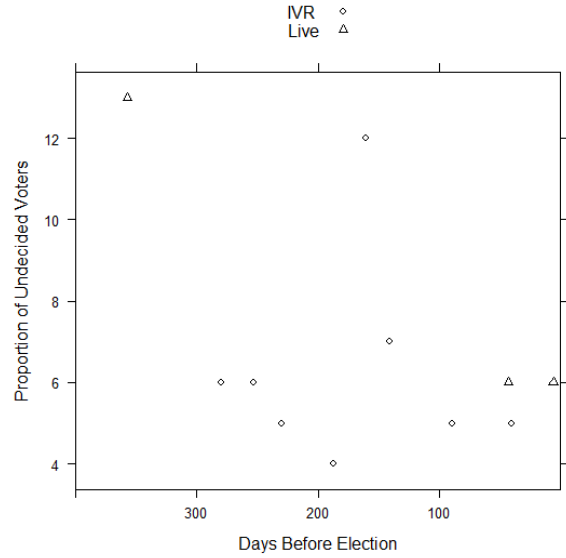
Polls Taken in 2010 Idaho Senate Election
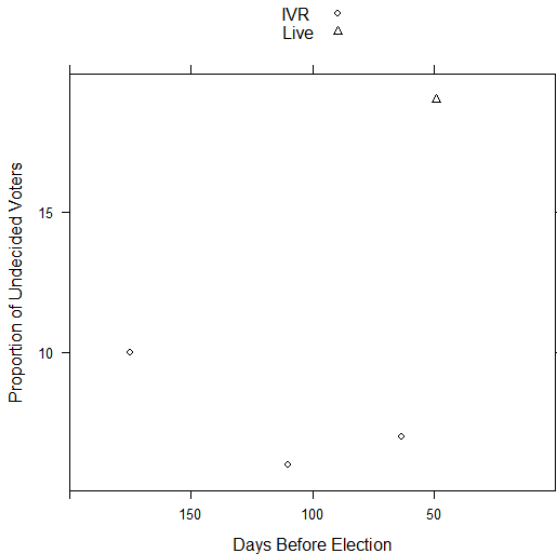
Polls Taken in 2010 Illinois Senate Election
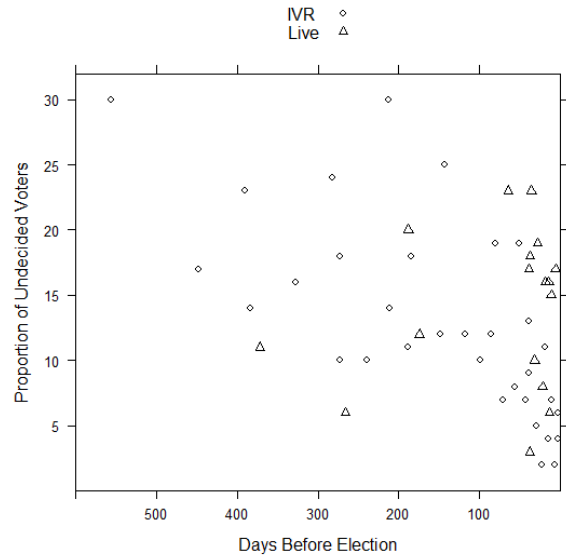
Polls Taken in 2010 Kansas Senate Election
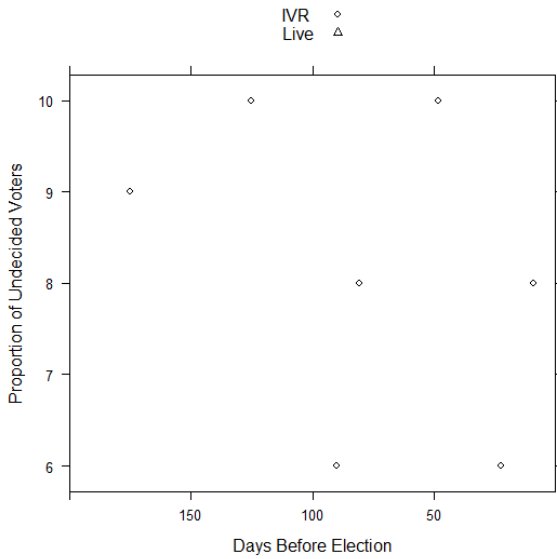
Polls Taken in 2010 Kentucky Senate Election
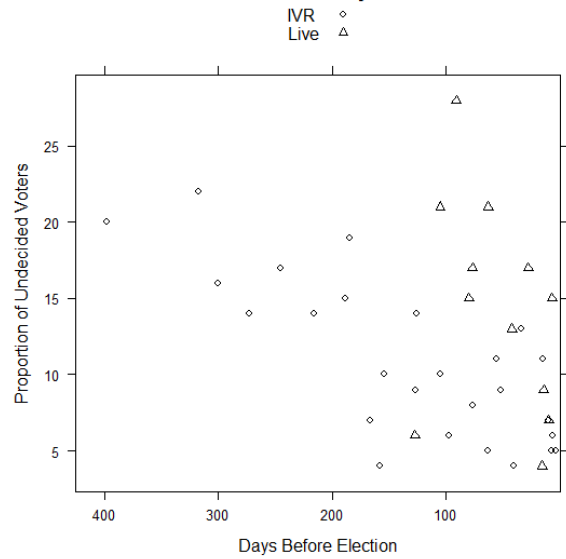
Polls Taken in 2010 Louisiana Senate Election
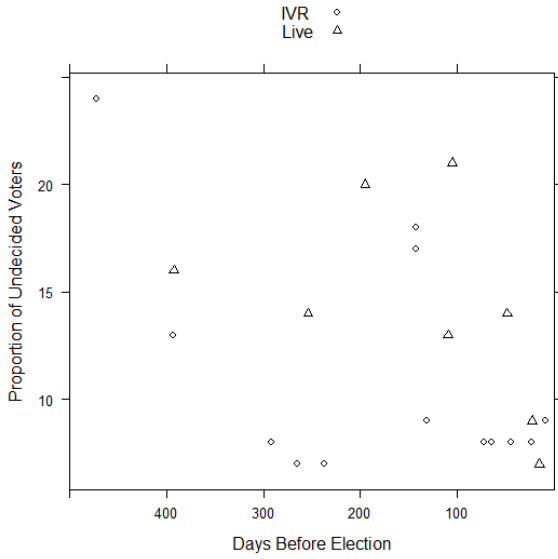
Polls Taken in 2010 Maryland Senate Election
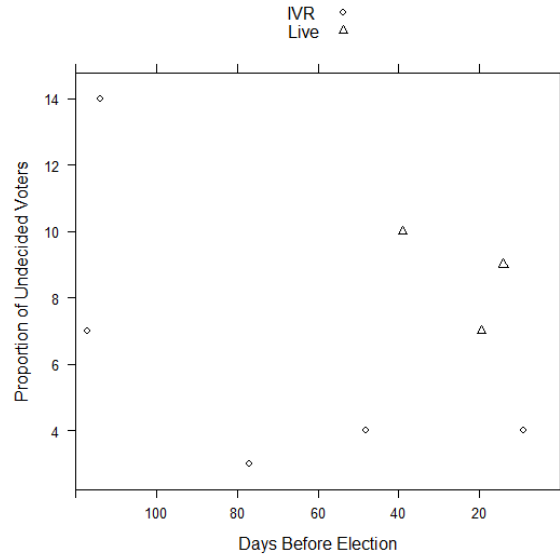
Polls Taken in 2010 Missouri Senate Election
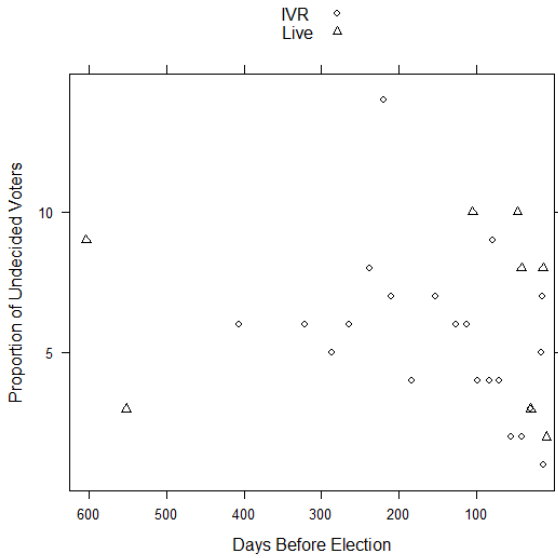
Polls Taken in 2010 North Carolina Senate Election
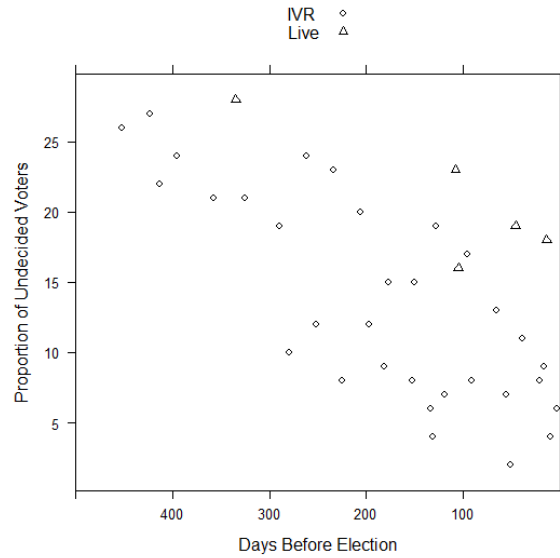
Polls Taken in 2010 North Dakota Senate Election
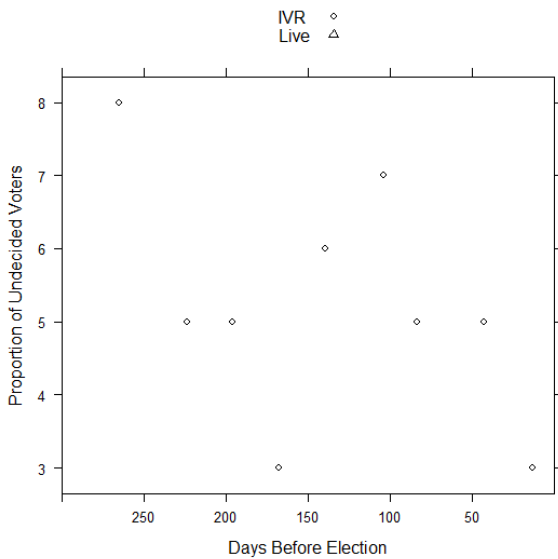
Polls Taken in 2010 New Hampshire Senate Election
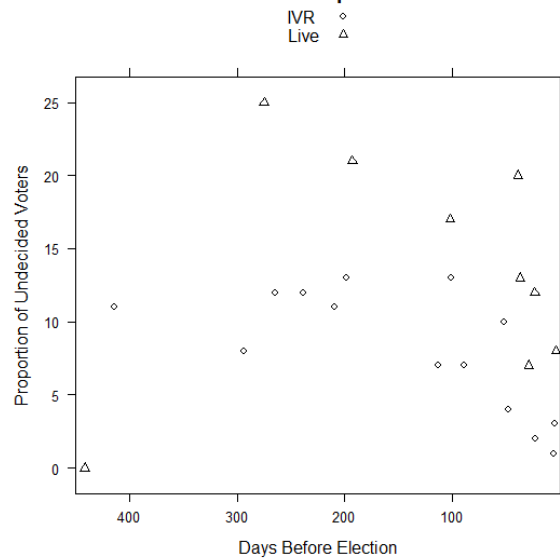
Polls Taken in 2010 Nevada Senate Election
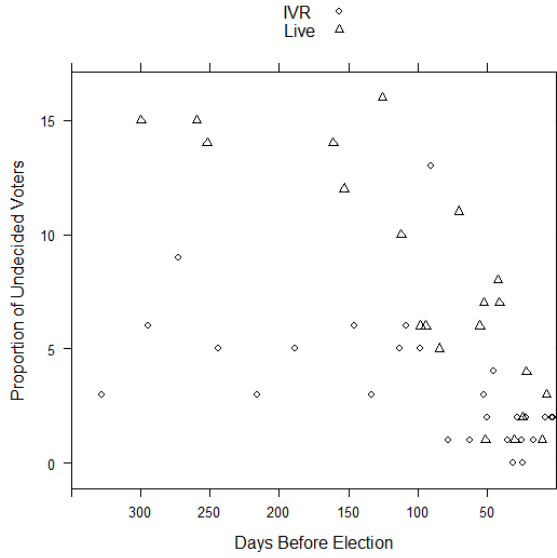
Polls Taken in 2010 New York Senate Election (Gillibrand)
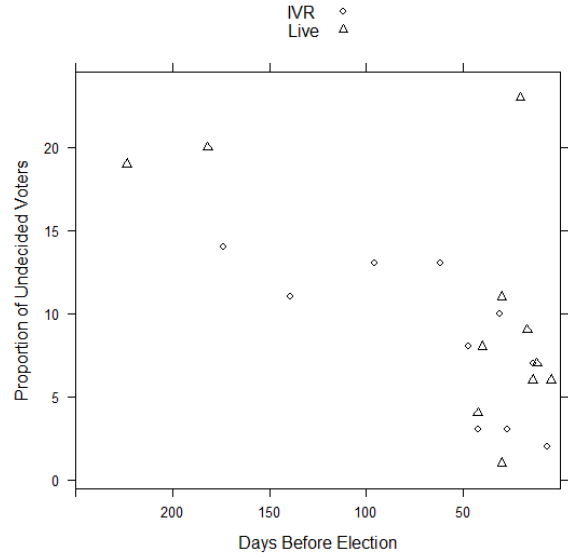
Polls Taken in 2010 New York Senate Election (Schumer)
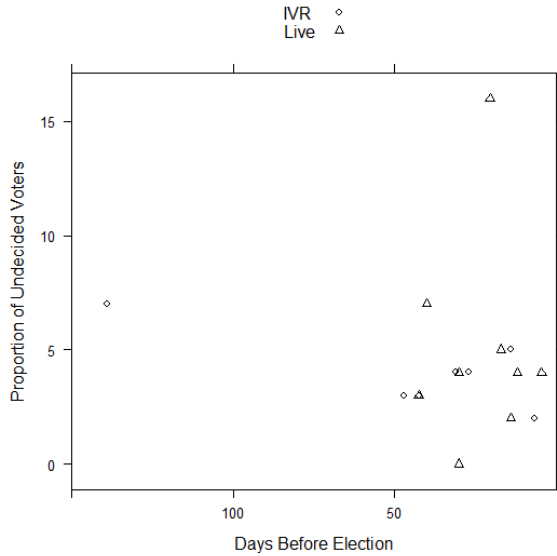
Polls Taken in 2010 Ohio Senate Election
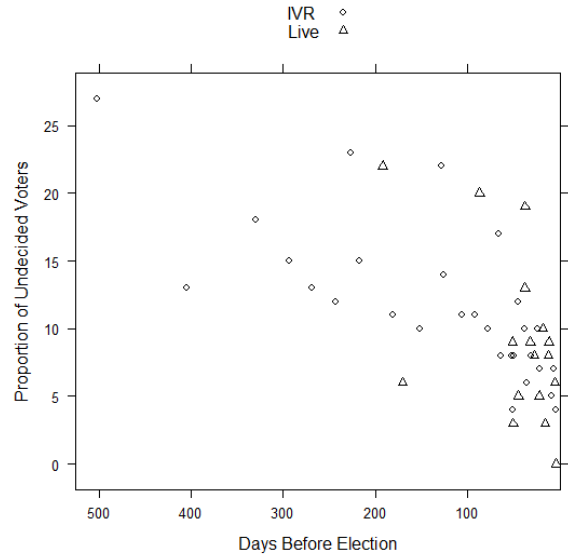
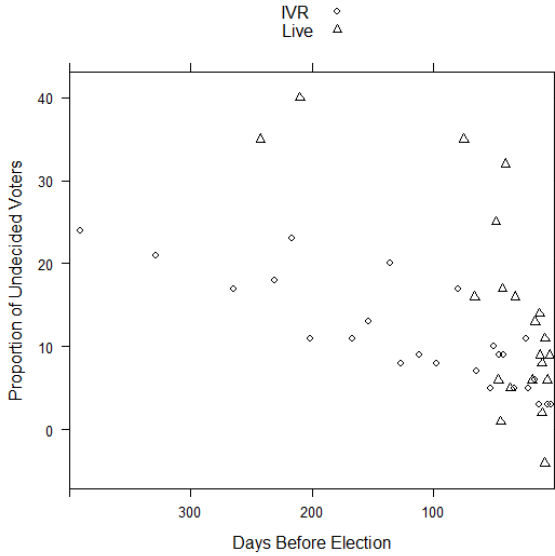Polls Taken in 2010 Oklahoma Senate Election
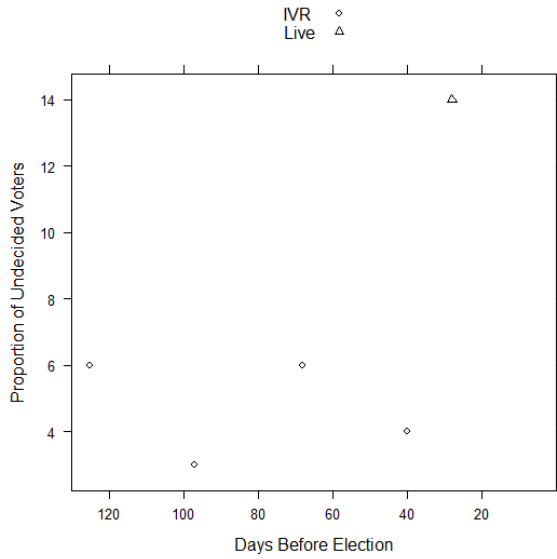
Polls Taken in 2010 Oregon Senate Election
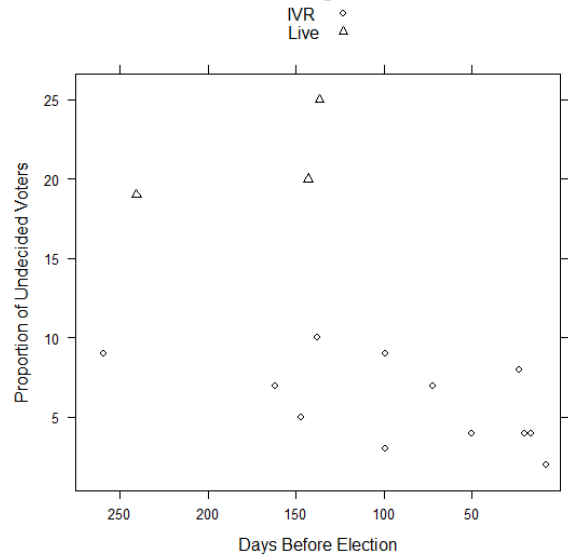
Polls Taken in 2010 Pennsylvania Senate Election
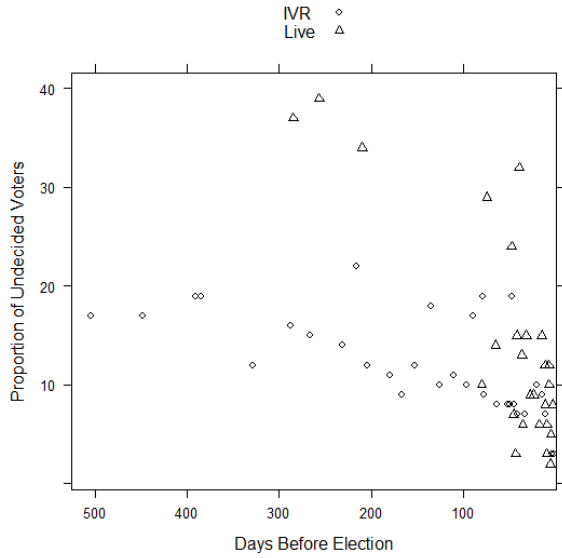
Polls Taken in 2010 South Carolina Senate Election
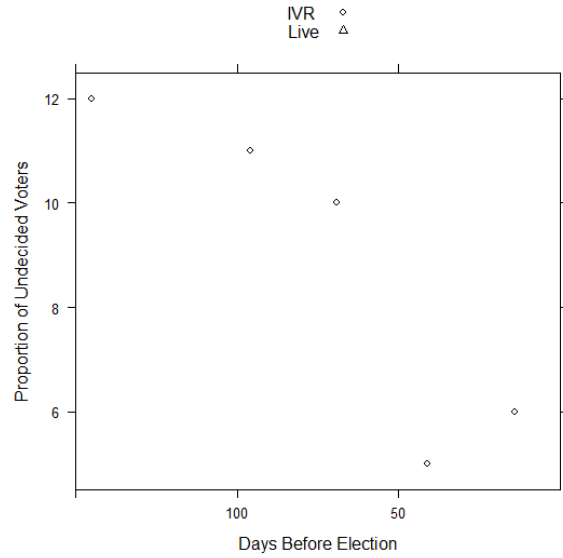
Polls Taken in 2010 Utah Senate Election
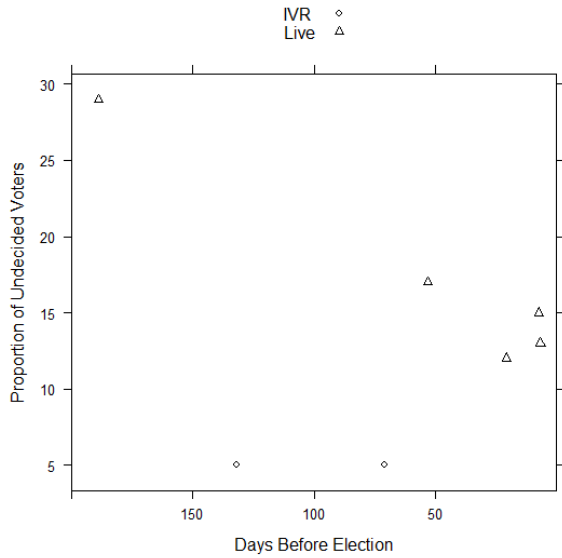
Polls Taken in 2010 Vermont Senate Election
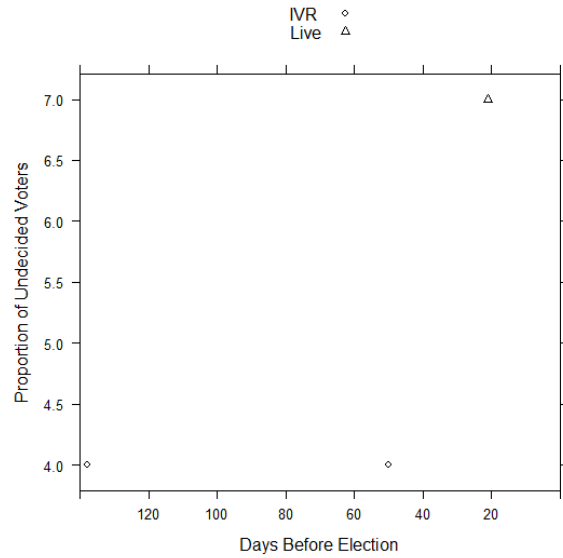
Polls Taken in 2010 Washington Senate Election
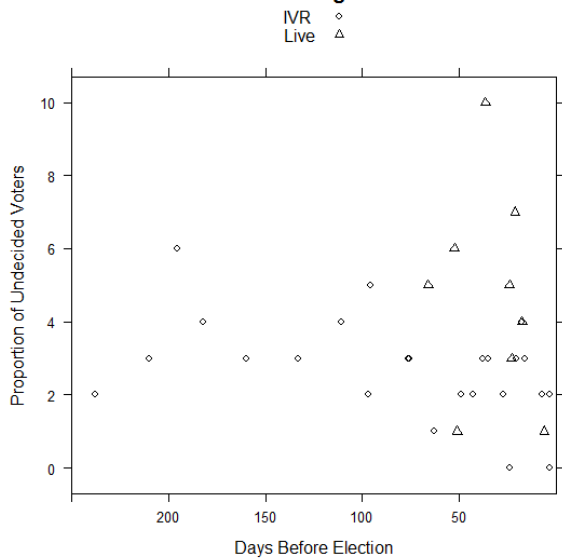
Polls Taken in 2010 Wisconsin Senate Election
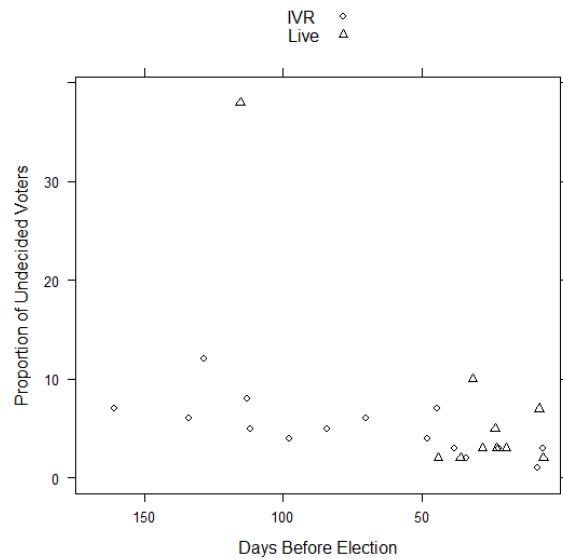
Polls Taken in 2010 West Virginia Senate Election
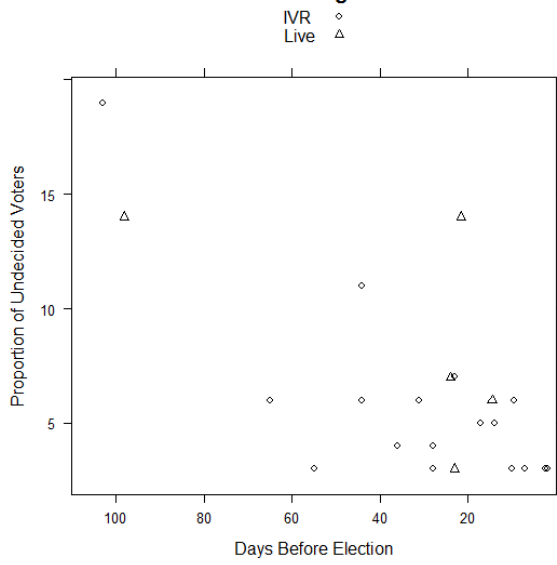
Polls Taken in 2010 Alaska Gubernatorial Election
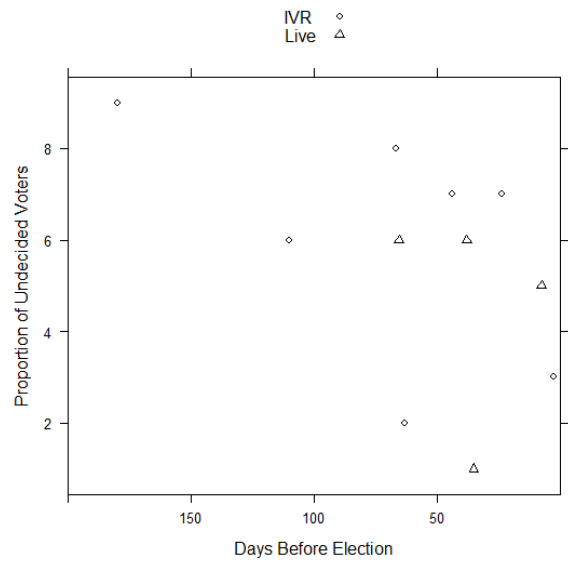
Polls Taken in 2010 Alabama Gubernatorial Election
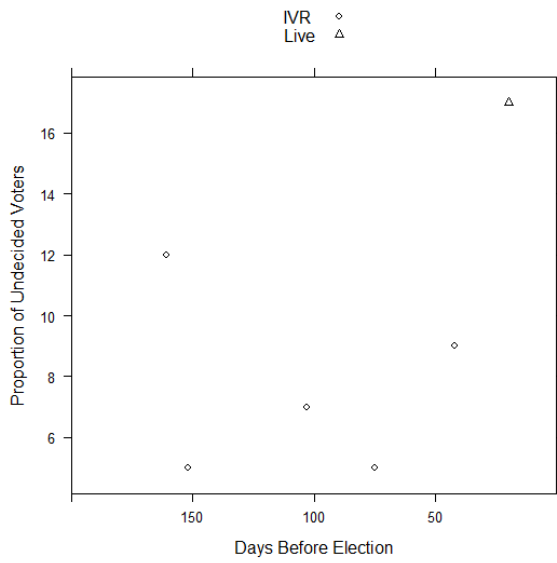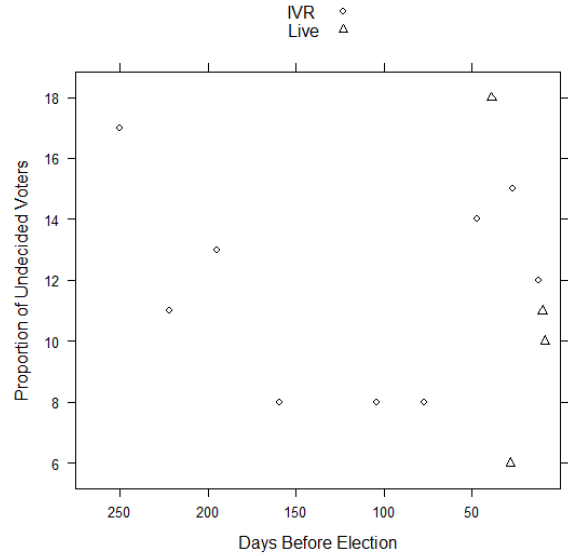
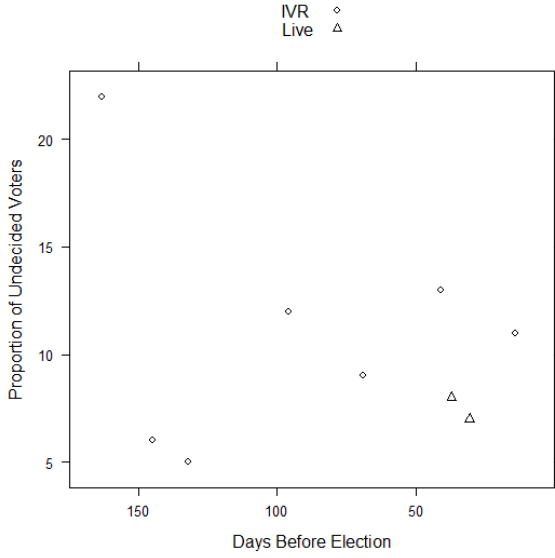Polls Taken in 2010 Arkansas Gubernatorial Election
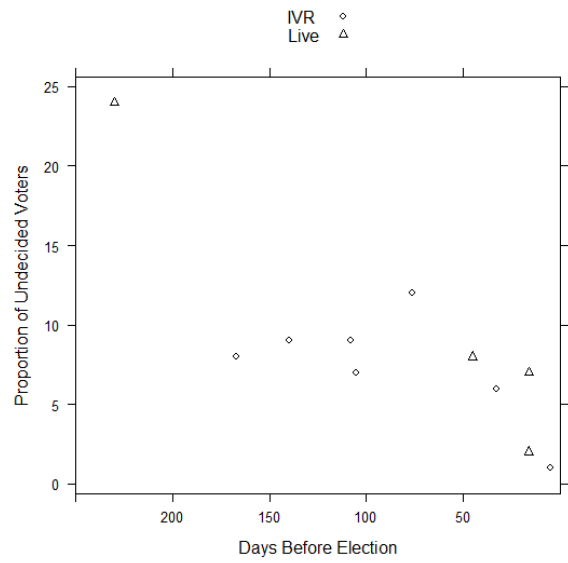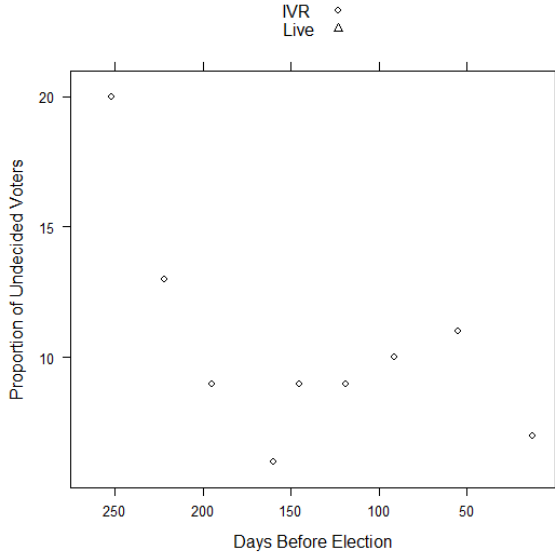
Polls Taken in 2010 Arizona Gubernatorial Election
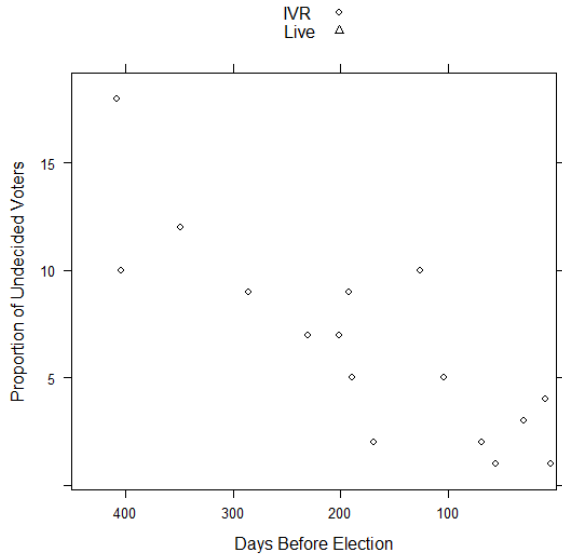
Polls Taken in 2010 California Gubernatorial Election
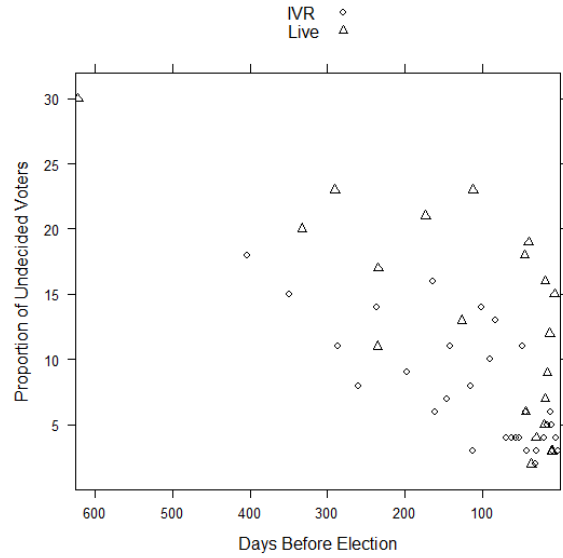
Polls Taken in 2010 Colorado Gubernatorial Election
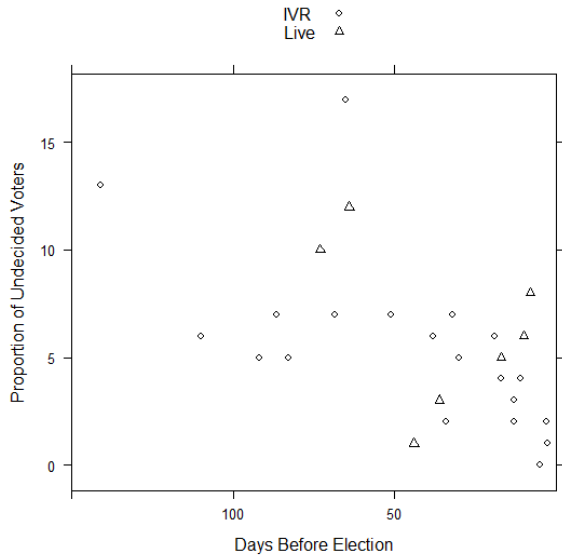
Polls Taken in 2010 Connecticut Gubernatorial Election
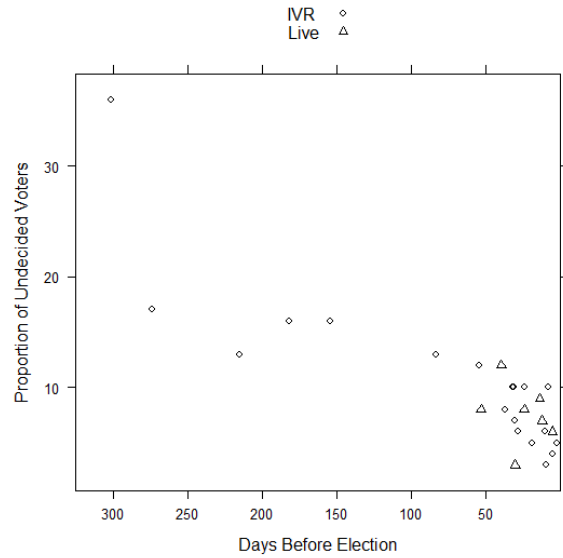
Polls Taken in 2010 Florida Gubernatorial Election
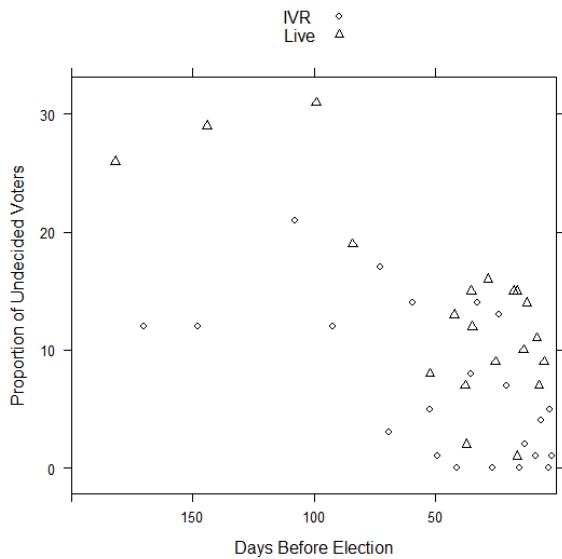
Polls Taken in 2010 Georgia Gubernatorial Election
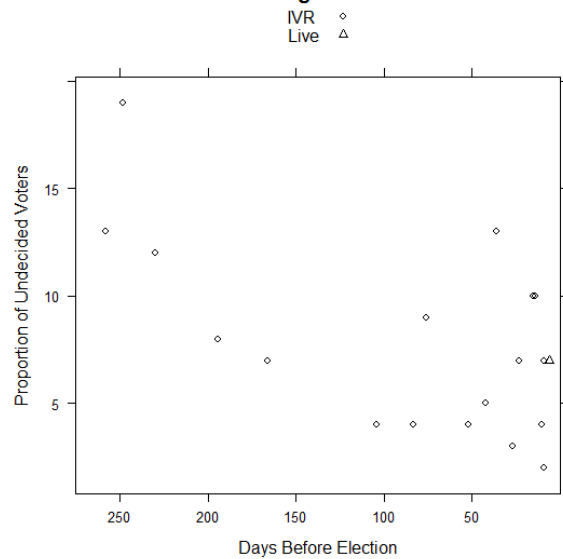
Polls Taken in 2010 Hawaii Gubernatorial Election
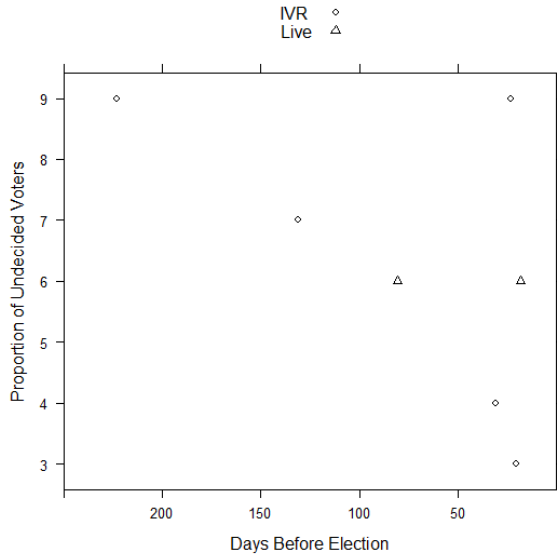
Polls Taken in 2010 Iowa Gubernatorial Election
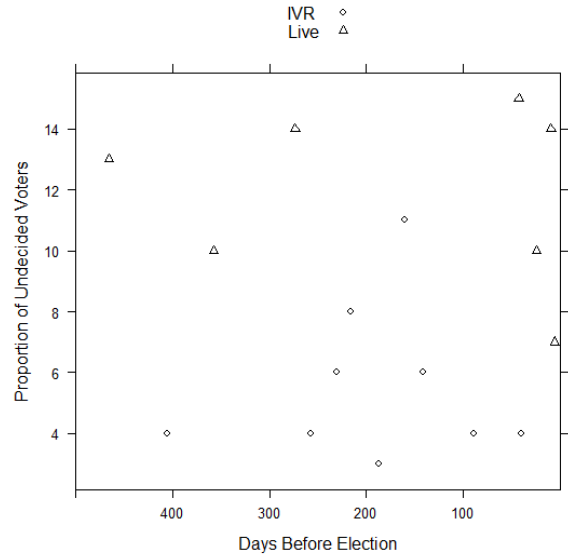
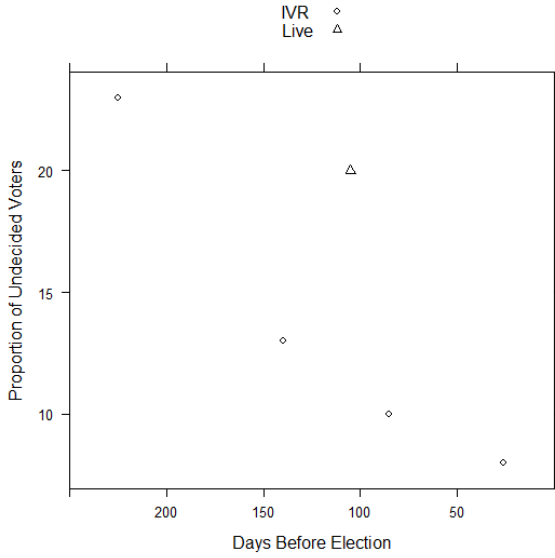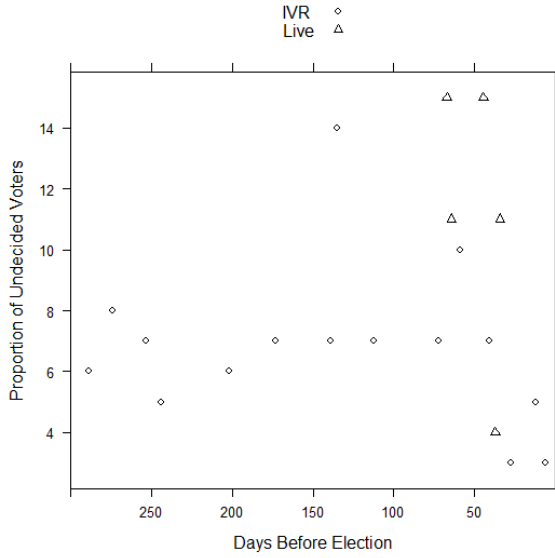Polls Taken in 2010 Idaho Gubernatorial Election
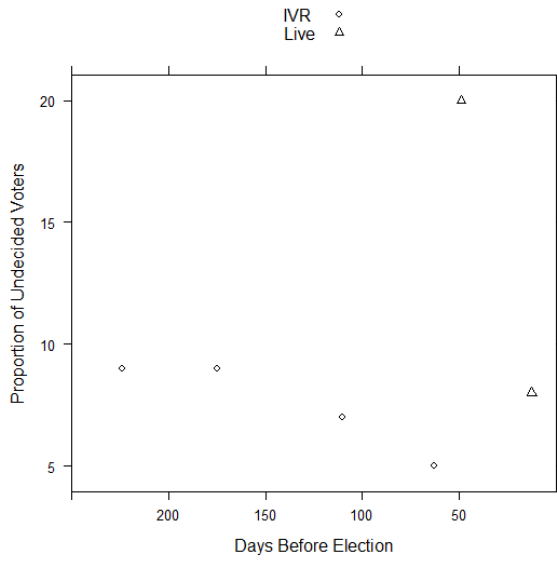
Polls Taken in 2010 Illinois Gubernatorial Election
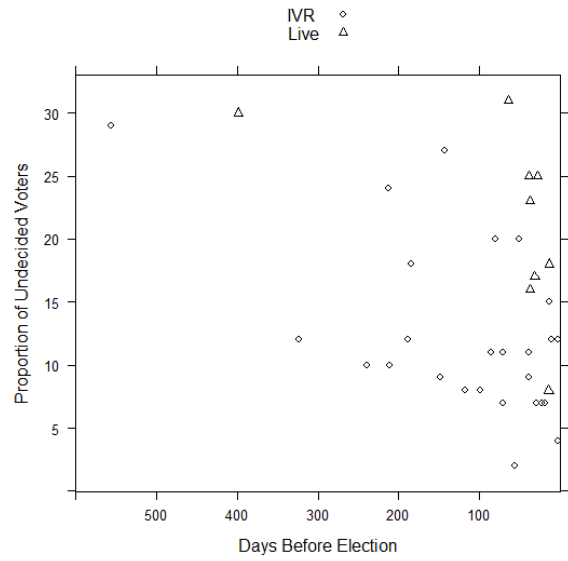
Polls Taken in 2010 Kansas Gubernatorial Election
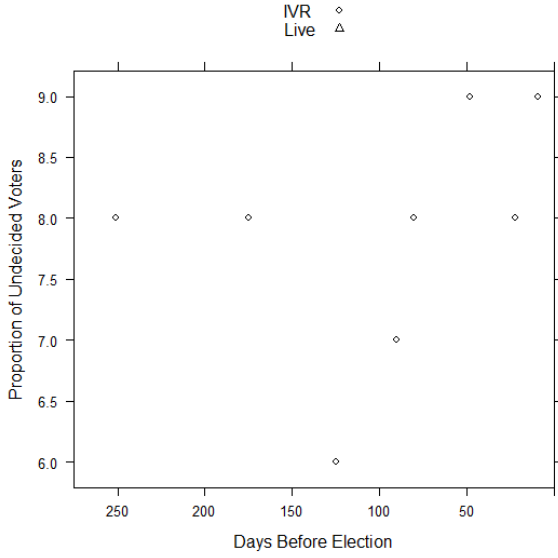
Polls Taken in 2010 Massachusetts Gubernatorial Election
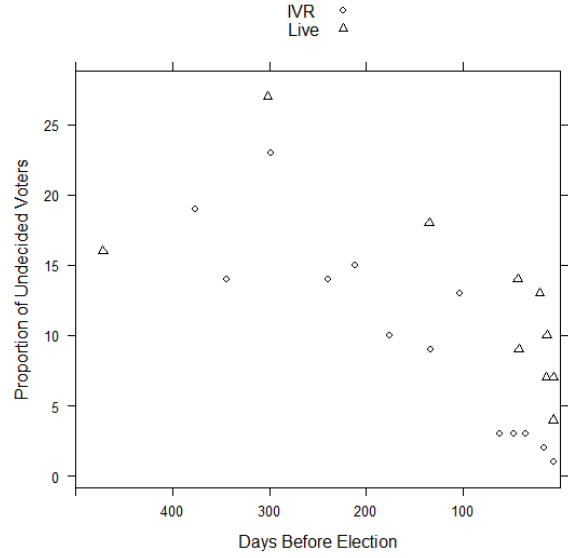
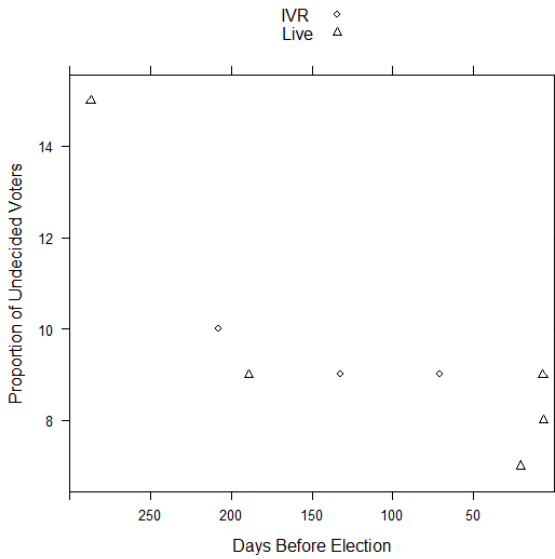Polls Taken in 2010 Maryland Gubernatorial Election
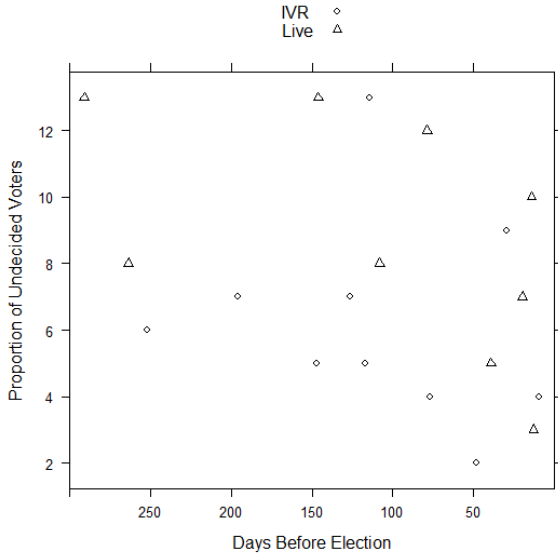
Polls Taken in 2010 Maine Gubernatorial Election
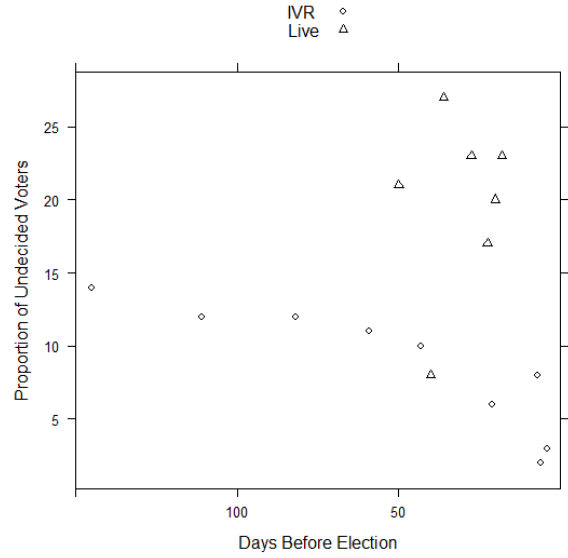
Polls Taken in 2010 Michigan Gubernatorial Election
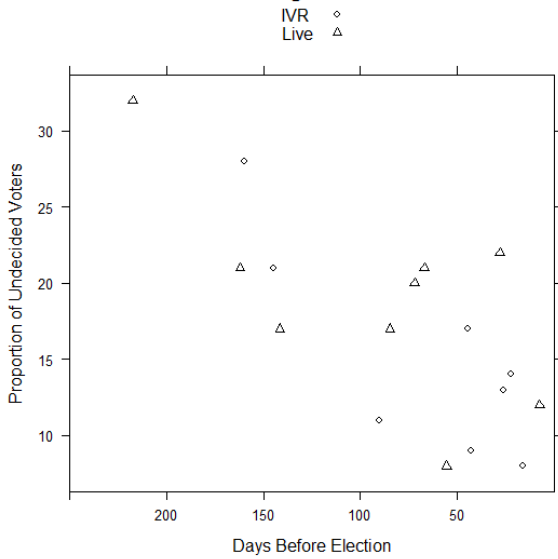
Polls Taken in 2010 Minnesota Gubernatorial Election
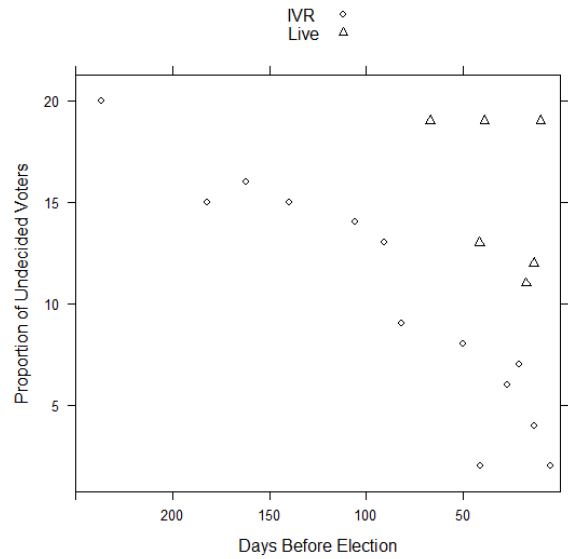
Polls Taken in 2010 Nebraska Gubernatorial Election
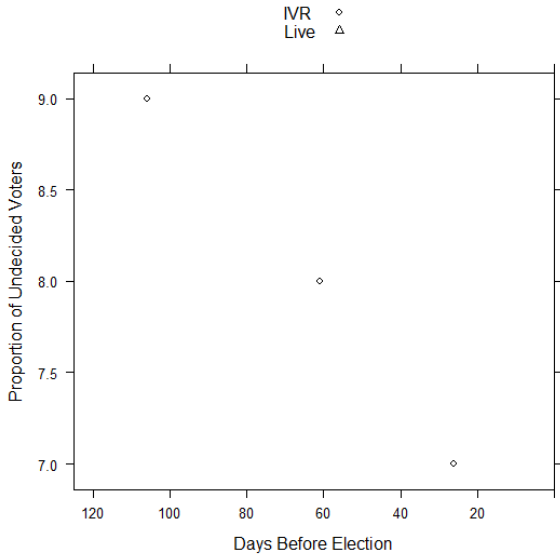
Polls Taken in 2010 New Hampshire Gubernatorial Election
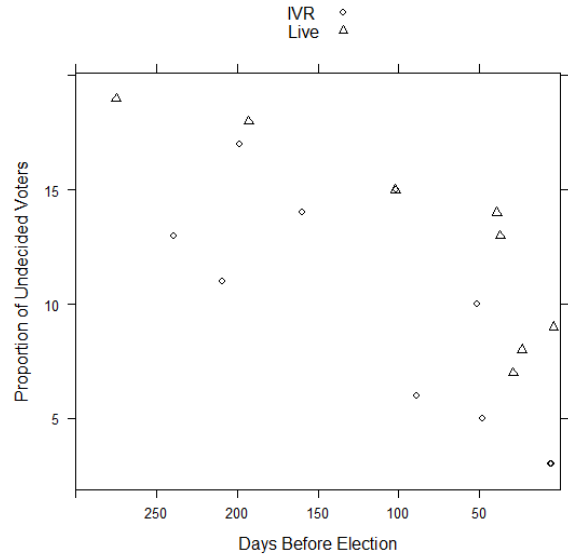
Polls Taken in 2010 New Mexico Gubernatorial Election
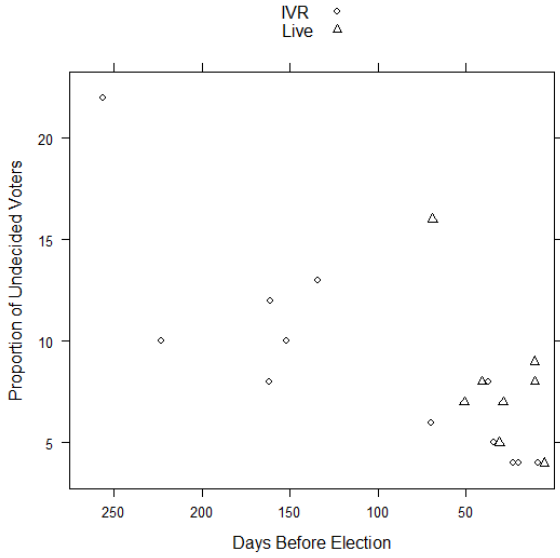
Polls Taken in 2010 Nevada Gubernatorial Election
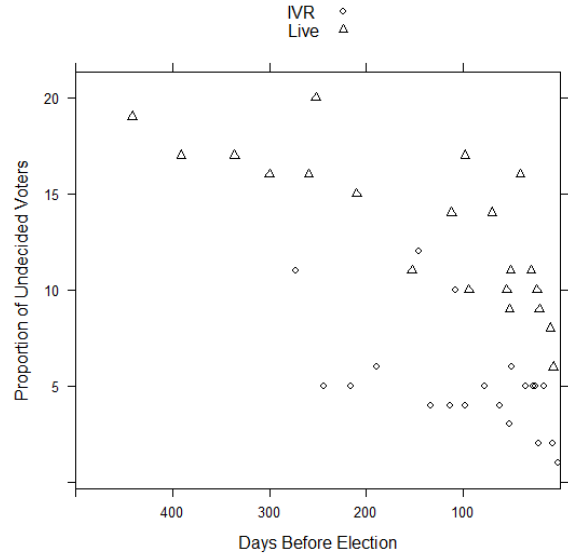
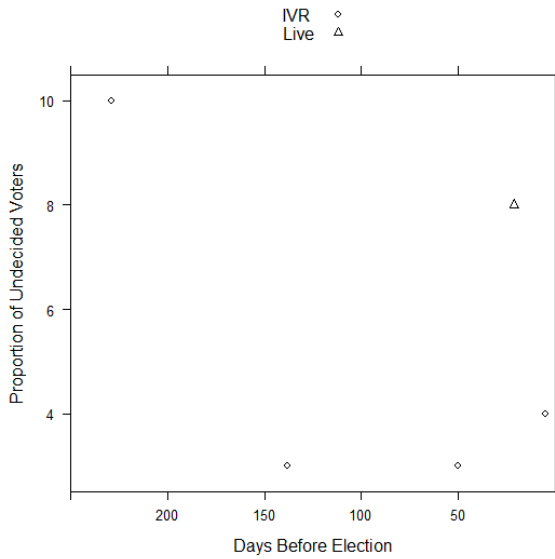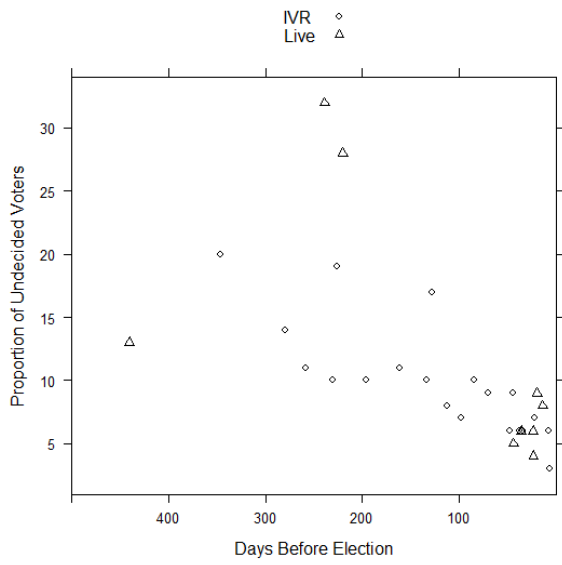Polls Taken in 2010 New York Gubernatorial Election
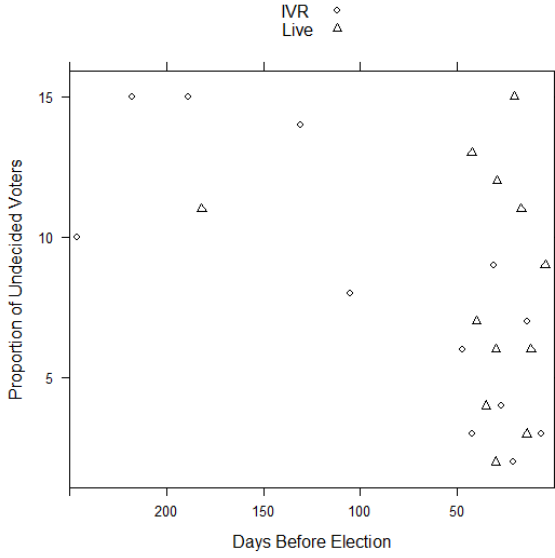
Polls Taken in 2010 Ohio Gubernatorial Election
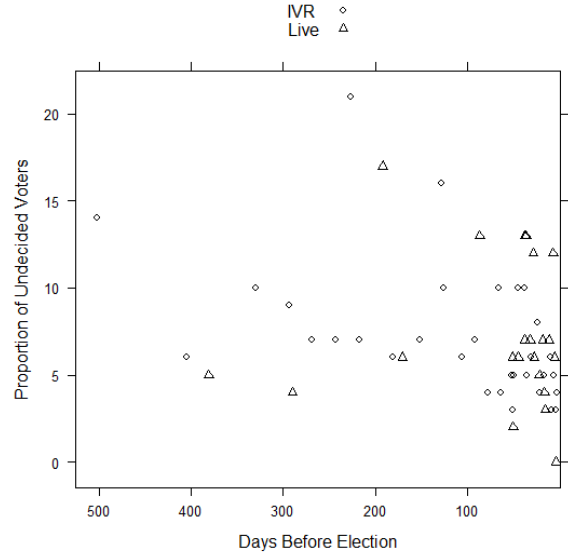
Polls Taken in 2010 Oklahoma Gubernatorial Election
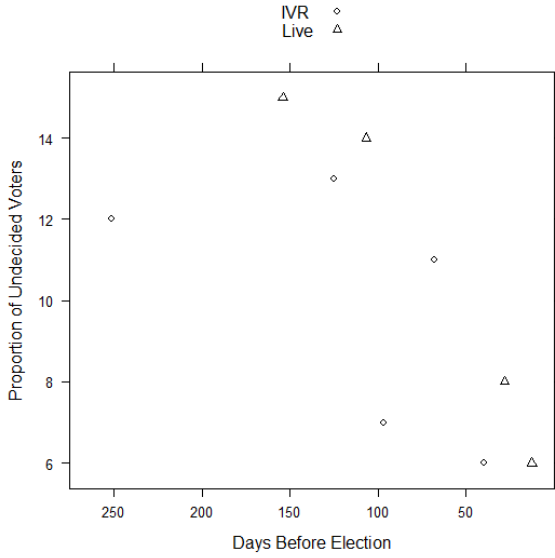
Polls Taken in 2010 Oregon Gubernatorial Election
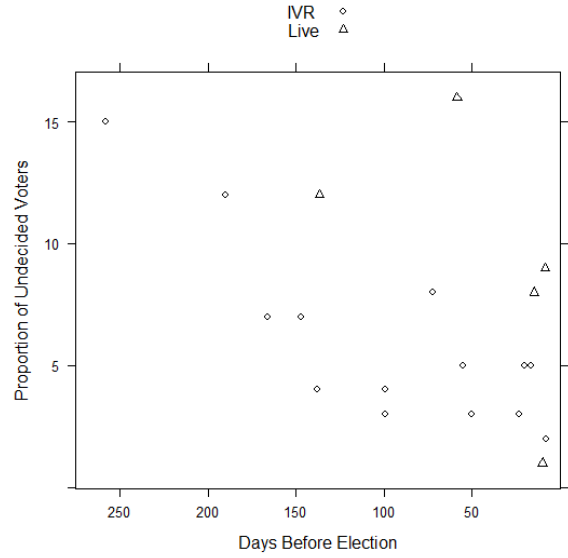
Polls Taken in 2010 Pennsylvania Gubernatorial Election

Polls Taken in 2010 Rhode Island Gubernatorial Election

Polls Taken in 2010 Wisconsin Gubernatorial Election

Polls Taken in 2010 Wyoming Gubernatorial Election