CALTECH/ MIT VOTING TECHNOLOGY PROJECT
A multi-disciplinary, collaborative project of the California Institute of Technology - Pasadena, California 91125 and the Massachusetts Institute of Technology - Cambridge, Massachusetts 02139

TITLE Measuring Election Performance

Name R. Michael Alvarez
University Caltech

Key words:

VTP WORKING PAPER #94
December 11, 2009
Measuring Election Performance

R. Michael Alvarez*
Professor of Political Science
Co-Director, Caltech/MIT Voting Technology Project

December 11, 2009

*This paper is based on a presentation provided as part of “Democracia Electoral: Hacia Una Nueva Agenda”, Instituto Electoral Veracruzano: Universidata Cristobal Colon, Veracruz, MX, October 15, 2009. I thank Alfonso Ayala and Nathanial Persily for their comments about the presentation, and thank workshop participants for their questions and comments. I also thank the Carnegie Corporation of New York, John S. and James L. Knight Foundation, and the Pew Charitable Trusts for their support of the research discussed in this paper. Thanks as well to my colleagues from the Caltech/MIT Voting Technology Project, especially Charles Stewart III, Stephen Ansolabehere, and Thad Hall; and to Lonna Atkeson. Finally, thanks to Gloria Bain.
Introduction

The 2000 presidential election was one of the closest elections in American history. A margin of about 550,000 votes separated Al Gore from George Bush, only about 0.52% of votes cast. And despite the fact that Gore received more of the popular vote than Bush, after a contentious situation in Florida and a U.S. Supreme Court decision, Bush was the recipient of more Electoral College votes than Gore (271 to 266) and Bush became president.\(^1\)

The controversy surrounding the 2000 presidential election, especially in the state of Florida where Bush ended up defeating Gore by 537 of the 5,963,110 votes cast, focused a great deal of attention on the administration of elections in that state, and on the technologies that voters used to cast their ballots. In that context, the Caltech/MIT Voting Technology Project (VTP) was formed. This academic research effort that was given the goal of determining how election technologies performed in the 2000 presidential election, where they did not perform well, and how to improve those technologies.

The VTP was then, and continues to be, a unique research project. It involves faculty, students and staff at two universities that are normally quite competitive (MIT and Caltech), which are located on opposite sides of the United States (MIT on the East Coast, and Caltech on the West Coast). Furthermore, the VTP has always involved the participation of a transdisciplinary research group: computer scientists, mechanical engineers, political scientists, human factors and usability experts, economics, and business and management scholars. The collaboration of scholars from so many academic disciplines has allowed VTP research to take a broad perspective on the many complex issues that arise in the study of election administration and technology. The VTP has also worked with scholars are other universities in throughout the world, and has developed strong relations with election officials on several continents.

However, when the VTP research team came together right after the November 2000 general elections in the United States to focus on the problems seen with election technology in that election (especially in Florida), we immediately grappled with a significant dilemma. There simply was not a great deal of previous research on election technologies; specifically, there was not an established scientific approach to studying the performance of election technologies.\(^2\)

---

\(^1\) Much has been written about the Florida situation in 2000, and the legal battle thereafter. See Posner (2001) for an early study of the legal issues involved in the 2000 presidential election.

\(^2\) There was some prior research on voting machines, political participation and voting outcomes (Fraser 1985; Mather 1964, 1986; Saltzman 1988; Thomas 1968; and White 1960). There was also some research on “ballot rolloff”, and other how voting machines might be related to various measures of “rolloff” (Bullock and Dunn 1996; Darcy and Schneider 1989; Nichols and Strizek 1995; Shocket et al. 1992; Walker 1966).
This paper discusses the recent development of methods to measure election performance. I begin with the question that motivated some of the early work of the VTP, namely the development of the residual vote measure to study the performance of election technologies. The discussion then moves to the broader question of studying the overall performance of elections, through the use of voter surveys. Then I turn to the “election ecosystem” approach to studying election performance, and conclude with a sort set of ideas for how the study of election performance can be improved.

The Residual Vote

When the VTP researchers set out our first analysis of election technologies, in particular examining their reliability and accuracy (as those were the primary concerns coming out of the 2000 presidential election in Florida), we had to work quickly to gather useful evaluation data. We also had to use that data to develop measurement approaches that would let us determine which election technologies were more reliable and accurate, and which were less reliable and accurate.

The primary responsibility for administering elections in the United States falls to state governments. In practice in most states, counties and sometimes municipalities have responsibility for conducting elections. The multiplicity of election administration authorities makes collecting any type of election performance data complicated. In addition, the regulations and procedures for the systematic collection and reporting for data about any particular election are inconsistent across states and are sometimes inconsistent across local election jurisdictions within a particular state (Alvarez, Ansolabehere and Stewart 2005). This made it very difficult to collect a critical piece of data needed for the evaluation of election technology --- what type of voting system was used by the local election officials in a particular jurisdiction. As luck would have it, a private election data company (Election Data Services, Inc.) had been collecting this information in the United States before the 2000 presidential election, and we were able to reach an arrangement to get access to the data they had collected.3

Knowing each election jurisdiction’s voting technology (typically at the county level), we then needed to devise a measure of the voting system’s performance as observed at that same geographic level. Although data reporting was far from perfect, many state and local election jurisdictions in 2000 either reported, or were willing to provide to VTP researchers, the number of ballots cast in the 2000 election and the number of ballots counted in races on the ballot. Armed with this data, VTP researchers were able to develop a performance measure

3The relative difficulty the VTP faced collecting this important data in 2000 led us to recommend the creation of a federal entity that would collect and disseminate data like this. Such an entity was created by the 2002 Help America Vote Act (HAVA) --- the U.S. Election Assistance Commission (EAC). Since 2004, the EAC has worked to collect and distribute data about election administration (see http://www.eac.gov/program-areas/research-resources-and-reports/completed-research-and-reports/election-day-survey-results).
termed the "residual vote" --- computed simply as the percentage of ballots cast that
did not record a vote in a particular race (the number of votes in a given race
divided by the total number of votes cast). In our initial analyses, the VTP studied
the two races at the top of the ballot in each location, the presidential race, and then
either the gubernatorial or U.S. Senate race (VTP 2001a, 2001b).

Table 1 reproduces the estimates of the residual vote, by voting technology,
reported by the VTP in 2001 (VTP 2001b, page 21). This simple analysis led to a
number of important conclusions. Generally, we found that paper ballot systems
had lower residual vote rates, especially paper and optical scan voting systems. The
exception to this result was for punch card voting systems, which had the highest
residual vote in presidential elections of any voting technology. As we estimated
that more than 30 million voters used punch card voting systems in the 2000
presidential election, these results imply that had these voters used an optical scan
voting system 300,000 more votes would have been counted in the presidential
race, and 420,000 more votes would have been counted in the Senate and
gubernatorial races. The high residual vote rate that we found for punch card and
lever voting machines led us to recommend that they be phased out, a
recommendation that was incorporated into the 2002 Help America Vote Act
(HAVA).

<table>
<thead>
<tr>
<th>Machine Type</th>
<th>President</th>
<th>Governor &amp; Senator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Ballot</td>
<td>1.8%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Punch Card</td>
<td>2.5</td>
<td>4.7</td>
</tr>
<tr>
<td>Optical Scan</td>
<td>1.5</td>
<td>3.5</td>
</tr>
<tr>
<td>Lever Machine</td>
<td>1.5</td>
<td>7.6</td>
</tr>
<tr>
<td>Electronic (DRE)</td>
<td>2.3</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Subsequently, the concept of a residual vote has been widely used in many
studies. For example, researchers have used the residual vote to study how factors
other than voting technology are related to residual vote rates (e.g. Ansolabehere
to assess the effects of voting technology innovations; between 2000 and 2004
many counties got rid of old voting systems, and Stewart (2006) took advantage of
those innovations to assess the reduction in residual votes when election
jurisdictions acquired new voting systems. Stewart found that the largest change in
the residual vote rate came from the transition from punchcards to electronic voting
systems (-1.61%), the second greatest reduction was from the change in optical scan
to electronic voting systems (-1.23%), while the third greatest reduction was the
transition from punchcards to optical scanning (-1.09%).

Of course, the residual vote measure is far from perfect, as has been pointed
out by researchers who employ it (Alvarez, Ansolabehere and Stewart 2005). It
represents undervotes and overvotes, and it cannot separate intentional nonvoting
on a ballot choice from an error or voting machine malfunction. Despite these flaws, it does represent a measure that can be computed using comparable data, and data that is increasingly available from election officials.

**Other Measures of Election Performance**

But while much of the focus in the wake of the 2000 presidential election was on the apparent technological failures of voter machines, there also were other concerns that the larger electoral process was not performing as well as many had expected. Throughout the nation, but in particular in states that were closely contested in the presidential election, there were reports of long lines in polling places, of problems with voter registration, issues with vote by mail ballots, and other issues with the conduct of the election. These reports led VTP researchers to look at ways to quantify these problems, so that we could better direct our future research efforts and provide guidance to state and federal reform efforts.

When it came to some of these other election performance concerns, it was the case that some research had been done on voter registration problems in the past, and that there was one systematic data source that could be used to study some of these problems --- the U.S. Census Bureau's Current Population Survey Voting Supplement (CPS). The CPS is a survey conducted every month, but in Novembers of even-numbered years, the CPS has for the past few decades included a handful of questions about voting and the voting experience. In addition to having these questions, the CPS is also an important data resource as it has a very large sample (approximately 50,000 households, allowing for both state-level and other detailed demographic analysis). The CPS Voting Supplement has been used for important research that links state and local procedures to voter turnout, for example the seminal research of Wolfinger and Rosenstone (1980).

Importantly, the CPS Voting Supplement asks eligible citizens if they voted, and if they say they did not vote, they ask if they were registered. If registered, the non-voters are asked “What was the main reason you did not vote”, and they are then given the opportunity to provide a single answer from a long list of reasons for nonvoting. Table 2 gives the results from this question, for the 2000 presidential election.4

<table>
<thead>
<tr>
<th>Table 2: Reasons for Not Voting, 2000 Presidential Election</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too busy</td>
</tr>
<tr>
<td>Illness or emergency</td>
</tr>
<tr>
<td>Not interested</td>
</tr>
<tr>
<td>Out of town</td>
</tr>
<tr>
<td>Other reasons</td>
</tr>
<tr>
<td>Didn’t like the candidates</td>
</tr>
</tbody>
</table>

---

Many of the reasons for not voting, while interesting, are not ones that election reformers can easily resolve through changes in policy or technology. For example, 12.2% were not interested in the election, and 7.7% did not like the candidates. Some of these responses are very important, for example the 6.9% of registered non-voters who said they did not vote because they had a registration problem. As there were almost 150 million active registered voters in the 2000 election, and about 107 million who turned out to vote, that implies that there were approximately 43 million registered nonvoters. If 6.9% of them could not vote due to registration problems, that is nearly 3 million “lost votes” due to voter registration problems alone. The exact nature of these voter registration problems is not clear from this data, as the CPS questionnaire does not follow-up with these respondents, but it is likely that some combination of problems associated with errors in voter registration lists, incomplete or inaccurate voter registration information, or other voter registration mixups produce these problems.

Data like the CPS Voting Supplement, in combination with residual vote estimates, allowed the VTP in the wake of the 2000 presidential election to produce some overall estimates of election performance, in the form of these “lost votes” measures. In particular, the VTP estimated that 1.5 to 2 million votes were lost due to bad ballots and faulty voting machines; between 1.5 and 3 million votes were lost due to voter registration problems, that as many as 1 million votes were lost due to problems in polling places, and that unknown number was lost due to problems with absentee ballots. These estimates were important guides for policy makers who worked at both the state and federal levels on election reforms after the 2000 presidential election.

Improving the Measurement of Election Performance and Voter Evaluations

After the VTP published our major report in 2001, and after HAVA was passed into law in 2002, some of our research effort was devoted towards new methodologies for studying the performance of election administration and voting systems. However, as the VTP team discussed in our 2001 report, before we could advance the science of studying elections, we needed to also work with election officials to enhance the reporting and accuracy of elections data. Because many counties and states did not report the basic information needed to compute residual

---

<table>
<thead>
<tr>
<th>Reason</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refused, don't know</td>
<td>7.5</td>
</tr>
<tr>
<td>Registration problems</td>
<td>6.9</td>
</tr>
<tr>
<td>Forgot</td>
<td>4.0</td>
</tr>
<tr>
<td>Inconvenient</td>
<td>2.6</td>
</tr>
<tr>
<td>Transportation problems</td>
<td>2.4</td>
</tr>
<tr>
<td>Bad weather</td>
<td>0.6</td>
</tr>
</tbody>
</table>

5 The Federal Election Commission found that the states reported 149,476,705 active registered voters in 2000 (FEC N.D., page 1). The turnout estimate (107,390,107) is from Michael McDonald’s “General Election Turnout Rate” database, http://elections.gmu.edu/Turnout_2000G.html.
votes was concerning, the VTP worked with others in the research community to assist local, state and federal election officials in efforts to improve the reporting and accuracy of election administration data.\(^6\)

At the same time, VTP researchers began to experiment with different forms of data collection and new methodologies aimed at improving our ability to determine whether or not an election system is performing well. One of these new approaches towards studying the performance of election systems came through efforts to use traditional methods of survey research, and to develop new questions that might allow researchers to understand the performance of election systems, as seen through the eyes of their primary customers --- voters. A number of studies have used the innovative questions from these surveys on voter confidence, voter registration problems, experiences with voter identification policies and voting technologies, perceptions of election administration, and opinions about electoral fraud (Alvarez, Hall and Llewellyn 2007, 2008a, 2008; Alvarez and Hall 2008a, 2008b; Atkeson et al., forthcoming; Atkeson and Saunders 2007; Hall et al., 2009). These research projects demonstrated the importance of gathering accurate and detailed data from voters about their experiences and perceptions of election administration and technologies.

These experiences were instructive for the development of the first major survey-based research effort to study election performance, the 2008 Survey of the Performance of American Elections. The research team that worked on this project --- Charles Stewart III, Stephen Ansolabehere, R. Michael Alvarez, Adam Berinksy, Thad Hall and Gabriel Linz --- first developed a questionnaire that was implemented on a pilot basis in the November 2007 gubernatorial elections in Mississippi, Kentucky and Louisiana. A revised and improved questionnaire was then implemented in the February 2008 “Super Tuesday” primary elections. The survey was then implemented nationally in the 2008 presidential election.\(^7\)

In addition to being the first systematic effort to study voter experiences using a large-scale survey, the 2008 Survey on the Performance of American Elections was innovative in its design. The survey used samples of 200 registered voters from every state; these voters were interviewed online. The survey was also implemented by phone, with 200 interviews, in ten states. Once appropriate

\(^6\) The problems the VTP researchers faced gathering data for our 2001 report are detailed on page 89 (Caltech/MIT VTP 2001); a long list of states did not provide sufficient information to compute residual vote rates. These problems were also discussed in Alvarez, Ansolabehere and Stewart (2005). The VTP issued recommendations in 2004 to the U.S. Election Assistance Commission (EAC) regarding resolving these data collection problems (Caltech/MIT VTP 2004), with additional detailed recommendations provided in 2006 (Alvarez and Hall 2006). See also the discussion in footnote 3, above.

\(^7\) Development of this research instrument continues; another version of this same survey methodology was just implemented in gubernatorial elections in early November 2009 in New Jersey and Virginia.
weights were used, the Internet sample compared well to the telephone sample; the sample also predicted the state-by-state presidential election vote quite well.

This innovative survey effort allowed for an unprecedented evaluation of the performance of the election process in the United States. The survey data gave the research team the opportunity to study in great depth the problems that voters encountered as they tried to participate in the 2008 presidential election; to estimate the number of “lost votes” in the election; and to evaluate the relative performance of the states on many different dimensions.

Beginning with the simple question of how many voters experienced problems when they tried to vote, the survey revealed that 11% of voters experienced at least one problem when they tried to vote. That percentage translates into more than 14 million voters who experienced a problem voting. For those who voted in person on Election Day or before the election, the survey found that the primary problem reported by voters was long lines.

Turning to lost votes (from other than faulty voting machines, as the survey did not attempt to study residual votes), from the problems reported by voters the survey indicated that:
- 2.2 million registered voters were prevented from voting because of the lack of proper identification;
- 1.9 million voters could not find their polling place
- 2.6 million did not vote because of long lines
- 2.2 million could not vote because of registration problems.

Importantly, these estimates for lost votes are not that different from what the VTP found for comparable categories in 2000, implying that despite the many reforms introduced after the 2000 election, the number of lost votes may not have changed dramatically. The importance of the survey approach is that it allows a much more nuanced perspective on the reported voter problems, where they are occurring (both geographically and by voting mode), and because this survey contains many more demographic and political questions that allow for more detailed study of who had problems in the 2008 election.

Finally, this survey allowed the research team to look at state-by-state issues in the 2008 elections, allowing a comparison across different election jurisdictions. One simple way to assess state-by-state election performance is to look at the percentage of voters reporting at a problem when they went to vote, by state. The range of reported problems is quite large, from 19.8% in Washington, 14.5% in Arizona, 13.9% in California, and 10.5% in Illinois, to 0.9% in Massachusetts and 0.4% in North Dakota. Such estimates might be particularly helpful to correlate with other measures of election administration and election reform; that research remains to be conducted.

**Studying the Election Ecosystem**
One of the clear conclusions from the development and analyses of these novel voter experience surveys is that the problems that voters face when they go to vote might be best characterized as important, but low-frequency events. For example, in the 2008 Survey on the Performance of American Elections, typically only a few percentages of voters experienced particular problems. Reports of problems were low in percentage terms. But in an electorate as large as what is common in an American presidential election, with over a hundred million voters, the number of voters affected is large. This indicates that these problems may call into question the outcomes of close and contested elections.

Furthermore, while surveys of voter experiences provide invaluable data about the problems that voters face, and their opinions and perceptions of the electoral process, these surveys are limited to only the perspective of voters. First, they do not consider the perspective of poll workers or election officials. Second, they do not allow for the study of the reliability and other aspects of the voting system used in a particular state or county. Third, they do not allow for a study of the context of the election (especially the polling place environment). And finally, they do not allow for the direct study of other aspects of election administration in a particular jurisdiction.

As pointed out in recent studies by Huefner, Tokaji and Foley (2007), Alvarez, Atkeson and Hall (2007), Hall (2008), and Atkeson, Alvarez and Hall (2009), a great deal can be gained by intensive study of the “election ecosystem” in particular counties, states, or regions. The “ecosystem” approach, currently a novel and developing methodology for studying elections in the United States, tries to examine the many different dimensions of election administration in a specific geographic region. As Huefner, Tokaji and Foley (2007) wrote, “... a state’s processes for administering its elections deserve to be understood as an ecosystem because the choices that a state makes about the procedures and requirements in one area inevitably affect the health and functioning of several other areas as well” (page 17). Studying the election ecosystem potentially allows for reinforcing conclusions; for example, if researchers find evidence from studying different aspects of the election ecosystem that some administrative or procedural problems are obvious, the reinforcing analyses might help clarify the problems and identify solutions.

An excellent example of how powerful the ecosystem approach can be for studying the performance of an election system comes from the 2006 New Mexico project (Alvarez, Atkeson and Hall 2007). This project combined an exit poll survey that gauged the voter experience, a survey of poll workers, and an extensive Election Day observation effort. The combination of these three different analytic perspectives allowed the research team to identify a variety of important problems regarding New Mexico’s voter identification regulations:

New Mexico’s (voter identification) laws appeared to have been confusing to voters and poll workers alike. The law allows voters the
choice of several types of identification they could provide including a verbal statement of their name, address, birth year and the last 4 digits of their social security number. Although many poll workers asked for voter identification, many of them did not. The voter survey confirmed this finding indicating that almost 65% of voters showed some form of voter identification, while 35% did not (page 3).

More detailed analysis of the data indicated other problems with the application of New Mexico’s voter identification regulations in the 2006 election. Perhaps the most important of these additional analyses was one that found from the voter survey found that self-identified Hispanic and male voters were more likely to show some form of identification than non-Hispanics and women. When researchers examined from the perspective of the poll workers, it was determined that poll workers of all races --- including Hispanics --- asked Hispanic voters for identification more than they did of voters of other races. In short, the ecological approach is critical for understanding some of these issues.

Multi-method approaches for studying the performance of election systems have great potential. Their components (say a survey of voter evaluations, or poll worker evaluations) can produce important analyses and policy recommendations on their own; the integration of results across these different components can help reinforce conclusions drawn from election administration research as well as find areas where the interconnection between different administrative components need improvement.

Conclusion

The controversies associated with the 2000 US Presidential election exposed a number of weaknesses in the scholarly study of elections. Before 2000, there was not a great deal of academic research on election administration, on election technologies, or how many election procedures affected voting behavior and election outcomes. However, the research community has responded quickly to these weaknesses, by producing innovative new research approaches based on easily available data, and by developing new methodologies for studying election technologies and administration.

However, there is a great deal of additional research and development of new methodologies, that is needed in coming years. First, the types of election performance data that are collected today could be more quickly, accurately, and efficiently be gathered in the future. For example, real-time collection of voter and poll worker evaluations on Election Day could provide many insights into the conduct of a particular election. This is particularly true regarding the data that election officials gather and report; researchers need to work with election officials to insure that they gather and release the detailed data that is needed to evaluate the performance of election systems.
Second, the methodologies that have been developed to assess the performance of election systems have largely focused on certain aspects of those systems, especially the accuracy, reliability and accessibility of election systems. There have yet to be the implementation of methodologies to gather widespread data on the security, or usability, of election systems. Researchers have made some significant progress recently, using the residual vote concept to study ballot design problems (Kimball and Kropf 2005). Others have been using field trials of voting systems to also study usability (Herrnson et al. 2008). More development of these methodologies is needed in the near future.

Third, there is a great need for more study of these same issues internationally. By taking the study of election performance international, scholars will be able to collect vast quantities of additional data, and be in a much better position to understand how a wide variety of factors influence election performance. For example, many scholars are interested in how the governance of elections affects their relative performance; comparative analysis of this could shed a great deal of new light on whether partisan or non-partisan, unitary or commission-based, appointed or elected, governance systems lead to improved election performance.8

---

8 See, for example, Kimball et al. (2006), Alvarez et al. (2008a).
References


Atkeson, Lonna Rae, R. Michael Alvarez, and Thad E. Hall. Assessing Electoral Performance in New Mexico Using An Ecosystem Approach. Albuquerque, NM: The University of New Mexico, forthcoming manuscript.


Huefner, Steven F., Daniel P. Tokaji, and Edward B. Foley. 2007. From Registration to Recounts: The Election Ecosystem of Five Midwestern States. Columbus, OH: The Ohio State University Michael E. Moritz College of Law.


