

IMPROVING THE ELECTION DAY SURVEY

R. Michael Alvarez

Professor, Political Science

California Institute of Technology

Thad E. Hall

Assistant Professor, Political Science

University of Utah



CALTECH/MIT

VOTING TECHNOLOGY PROJECT

March 31, 2006

Executive Summary

The collection of detailed election administration data from local and state jurisdictions across the United States has proven difficult. The problem is partly due to the decentralized nature of election administration in the United States but is exacerbated by the lack of a centralized reporting of election information. Following calls made for better reporting of election administration data in recent years from many groups, including the Caltech/MIT Voting Technology Project, the U.S. Election Assistance Commission mounted an ambitious effort in 2004-2005 to collect detailed and consistent election administration data in the Election Day Survey. This effort was comprehensive and ambitious, but the less than perfect result illustrates many problems associated with collecting meaningful data regarding election administration.

In this report, we briefly discuss some of the general problems seen when we have attempted to use the Election Day Survey data in our research. We then provide twelve recommendations for improvement as a conclusion to this report:

- 1. Define the terms and concepts to promote uniformity across jurisdictions.**
- 2. Refine and clarify the survey instrument.**
- 3. Seek additional important information.**
- 4. Reduce the cost and complexity of the survey instrument.**
- 5. Improve the technology of reporting.**
- 6. Insure interoperability across datasets.**
- 7. Automate data collection where possible.**
- 8. Improve compliance.**
- 9. Improve data distribution.**
- 10. Provide data at precinct level.**
- 11. Methodological development of tools for fixing and analyzing future datasets.**
- 12. Build collaborations.**

Below we discuss some of the problems in the 2004 Election Day Survey and these recommendations in more detail.¹

¹ We thank Kim Brace for his feedback on an earlier version of this report. We also thank Melissa Slemin for her assistance, and Delia Bailey, Sarah Hill, and Betsy Sinclair for their help with the 2004 Election Day Survey database.

Introduction

For democratic elections to function as expected, they must be completely open to public scrutiny. The procedures of democratic elections should be transparent and it should be possible for members of the public to completely and accurately audit the outcome of any democratic election. Openness and transparency are vital so that winners of elections know that they were placed into office due to the support of participants in the election and so that losers of elections know that their opponents were not given office due to mistakes or malfeasance. Openness and transparency are thus critical for the development of trust and confidence in a democratic electoral process.

Unfortunately, since the 2000 presidential election, it has proven difficult for the public, the research community, and even the federal Election Assistance Commission (EAC), to obtain precise, detailed, and complete accounting of election administration practices and outcomes. Examples of the problems associated with obtaining precise, detailed and complete information about recent elections abound, some of them documented in recent studies.²

Having detailed and complete information about election administration practices and outcomes is important for other reasons, in addition to public trust and confidence. As election officials throughout the nation move to improve their administrative practices and processes, they need information so that they can evaluate their efforts and study the efforts of improvements in other jurisdictions as well. In both the public and private sectors, the collection of performance measurement data is becoming widespread and is closely associated with quality improvement efforts. Thus a robust data collection effort, continuing through future federal elections, will provide critical information that will allow for detailed evaluation of the effects of changes in administrative practice and procedures.

Prior to the 2004 presidential election, the Caltech/MIT Voting Technology Project (VTP) issued a report, “Insuring the Integrity of the Electoral Process: Recommendations for Consistent and Complete Reporting of Election Data.” In this report, the VTP team called for the collection, distribution, and retention of election administration data from before, during, and after every election, at the lowest possible level of administration. In an important effort, the EAC attempted the first national effort to collect and distribute many of the types of data called for in the VTP report, and to make these data available to the interested public and research community in the form of their “2004 Election Day Survey.”

While an important effort, the EAC’s “2004 Election Day Survey” demonstrates that there is still much progress that needs to be made to collect, distribute, and retain precise, detailed, and

² See R. Michael Alvarez, Stephen Ansolabehere and Charles Stewart III, “Studying Elections: Data Quality and Pitfalls in Measuring the Effects of Voting Technologies.” *Policy Studies Journal*, 2005, 33(1), 15-24. See also the discussion in Charles Stewart III, “Measuring the Improvement (or Lack of Improvement) in Voting Since 2000 in the U.S.”, Caltech/MIT Voting Technology Project Working Paper 36, http://vote.caltech.edu/media/documents/wps/vtp_wp36.pdf. These same issues were also discussed in the 2001 report of the Caltech/MIT Voting Technology Project, “Voting: What Is, What Could Be”, <http://vote.caltech.edu/reports/2001report.htm>.

complete information about the conduct of federal elections in the United States. The EAC, and the election jurisdictions that participated in the “2004 Election Day Survey” data collection effort, should be applauded for what they managed to establish in 2004 with this first effort. However, as we are now in the next federal election cycle (and soon will be in the 2008 presidential election cycle), it is time to provide some initial evaluation of the EAC’s 2004 data collection and distribution effort, and to suggest ways that this effort can be improved for 2006 and beyond.³

Data Quality Issues in the Election Day Survey

The EAC’s “2004 Election Day Survey” was an ambitious effort, but like many ambitious projects, there is room for considerable improvement in future elections. The EAC and all of those involved in the collection and distribution of the Election Day Survey (EDS) information deserve praise for the fact that they worked hard to collect these data and to provide these data to the public (including election officials and the EAC’s contractor for this project, Election Data Services, Inc.). The problems associated with the survey relate primarily to data quality and management at the state or local levels. Specific problems with the EAC’s 2004 effort include incomplete coverage of election jurisdictions, missing data, disparate definitions of how to interpret election terms (e.g., what is early voting?), heterogeneous survey responses and data, thinness of administrative data, poor distribution of the data, and a lack of data at the lowest level of election administration (the precinct). Each of these problems is discussed below, and many of these echo problems with the data discussed in the Election Day Survey report.⁴ This report concludes with a discussion of proposals for improving reporting and distribution of election administration data in future elections.

Incomplete Coverage of Election Jurisdictions

The 2004 EAC data had two coverage problems. The first coverage problem is that two territories (Guam and American Samoa) did not respond to requests for data from the EAC; while troubling that some territories did not respond at all to the EDS, the fact that these two territories did not respond at all to the request for information should not render many analyses of these data problematic (unless, of course, the analysis is focused on the territories).

The second coverage problem is more severe, and it concerns varying coverage rates from jurisdictions across the questions in the survey. According to summary data provided in the EAC’s report for selected survey questions, coverage rates vary from a high of 98.8% for the number of ballots counted in election jurisdictions to a low in this table of 8.2% for the number of polling places where visually impaired voters can cast private ballots. Particularly problematic were the coverage rates for information on voting machine malfunctions, were only

³ In 2004 and 2005 the EAC also conducted data collection efforts in two other survey-style projects, a survey of jurisdictions regarding their implementation of “Motor Voter” and of implementation of UOCAVA voting procedures. In this report we are not addressing those other survey efforts, other than to note in our recommendations that the EAC may consider consolidation of these survey efforts, to the extent that they cover identical or similar territory.

⁴ The complete report of the 2004 Election Day Survey effort is available at http://www.eac.gov/election_survey_2004/intro.htm.

485 of 6,567 jurisdictions provided information, so few that the survey report does not provide the voting machine malfunction data by jurisdiction (as was provided for the other topics covered by the Election Day Survey).

Practically speaking, this coverage problem means that considerable care must be placed on the interpretation of EAC survey data, especially those questions with relatively low rates of election jurisdiction coverage. Continuing with the example of the number of polling places where visually impaired voters can cast private ballots, the poor coverage across election jurisdictions in many states makes the provided data difficult if not impossible to use: only three of Arizona's 15 election jurisdictions provide data for this question, one of California's 58 election jurisdictions provided data, but all 169 of the election jurisdictions in Connecticut provided data. The fact that the data are sometimes reported without giving any indication of which election jurisdictions are providing information makes it very difficult to determine if the election jurisdictions covered in a particular state are representative or not. This problem exists for each question in the EAC survey, and leads us to caution users of this data to be quite careful in their interpretation of any survey question with low overall jurisdiction coverage and highly variable coverage across states.

Missing Data

Missing data is a different problem than survey coverage. Missing data refers to information that is not available for certain questions in the Election Day Survey data; while a jurisdiction might provide information in some areas, many did not provide information for all questions. The missing data problem in the Election Day Survey data has four consequences for researchers, each of which needs to be considered by any user of these data.

The first problem that missing data produces is that it leads to an attrition of the sample size for particular questions. In a situation where there is a lot of missing data for some particular survey question, we can be in a situation where any inference drawn from that data relies upon relatively few observations, meaning that differences we may observe across jurisdictions may not be statistically reliable. This lack of statistical precision or power is the inevitable result of missing data, especially in situations where there is a great deal of missing data, which is true of some of the Election Day Survey questions.

Secondly, missing data means that for some of the survey questions we lack the ability to necessarily know that when we examine information from a question that the inferences we draw from that analysis are representative of the entire population of jurisdictions. If the missing data arises from certain types of jurisdictions, then as we lack information from those jurisdictions any analysis of the data will not be able to shed light on procedures or outcomes from those jurisdictions. This could lead analysts to make incorrect generalizations about analyses using the Election Day Survey data.

Third, worse yet, the missing data may arise because of deliberate non-reporting by election jurisdictions. This does not mean that election officials are acting maliciously, not reporting information to cover up problems. Instead, perhaps certain types of jurisdictions (say small jurisdictions with few resources) are unable to retain some of the more detailed information

requested by the Election Day Survey; this means that some jurisdictions are not responding to Election Day Survey questions intentionally, which can produce a problem worse than the lack of representativeness discussed above. Instead, this could produce what social scientists call selection bias, whereby some units of analysis are systematically not responding to the Election Day Survey based on certain attributes of those jurisdictions. If this is the case, then results of analyses of these data may be incorrect, unless steps are taken to deal with the selection bias.

Fourth, state law often forces jurisdictions to conduct elections in a manner resulting in the loss of important data. This is most common in the area of ballot counting. Many states cannot report data on ballots cast early versus absentee versus in-person because state law requires that all early and absentee ballots be counted in precincts “as if the voter cast a ballot in person.”

Of course, social scientists and statisticians have encountered problems of missing data before in their research, and there are well-established techniques for dealing with problems of missing data.⁵ One widely used approach, for situations where the missing data problems are not the product of a selection process, is to use some form of data “imputation” to plug the holes in the dataset to provide more precision for statistical analysis. These techniques are well known and widely used, and could easily be applied to the Election Day Survey data, but only if the missing data was not the result of a selection process.

Things get more complicated if the data is missing due to some form of selection process. While “selection models” have been studied in the social science and statistics literature, they are complicated and can be difficult to utilize. But if the missing data arises from a selection process, the ideal approach is to statistically deal with the selection process and to alleviate the associated bias using standard techniques.⁶

Another strategy that might be utilized is to provide supplemental data to alleviate missing data problems in the Election Day Survey. There are other organizations and private firms that often collect election data; perhaps the EAC could ascertain whether or not those other datasets have information in them that might be integrated into Election Day Survey data, where necessary to alleviate missing data problems. The only caveat to such data supplementation is that there should be clear documentation in user datasets as to the supplementation procedure, where the supplementation data came from, and how users of the data can separate original survey responses from states or counties from the supplemented data.

Heterogeneous Survey Responses and Data

This is a particularly serious problem in the Election Day Survey, a problem that may trip up casual users of this data. The issue is how survey respondents, in this case state and county

⁵ A reference from the political methodology literature for how political scientists can deal with missing data problems is Gary King, James Honaker, Anne Joseph and Kenneth Scheve, “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation”, *American Political Science Review*, 95(1), March 2001, 49-69. A standard reference from the statistics literature is Roderick J.A. Little and Donald B. Rubin, *Statistical Analysis With Missing Data*, second edition, Wiley, 2002.

⁶ Approachable treatments from the political methodology literature with direct application to data like that found in the Election Day Survey are: Christopher H. Achen, *The Statistical Analysis of Quasi-Experiments*, The University of California Press, 1986; John Brehm, *The Phantom Respondents*, University of Michigan Press, 1993.

election officials, interpreted the meaning and intent of the questions asked in the Election Day Survey instrument. Given varying administrative practices across states—and even within states, but across counties—it is clear that election officials interpreted some of the questions in the survey instrument very differently, and that this has produced a situation where the information provided is not consistent across jurisdictions.

For example, one of the most important data sections in the survey instrument concerned voting equipment, as this is an area of active research and public interest. Here, the survey instruments asked officials to “provide a listing of the types of voting equipment in use in each county of the State including the type of voting system, manufacturer, number of units used in each county/local election jurisdictions, the software version (if applicable), and an indication as to whether the voting system has or has not previously been used in a Federal election in that jurisdictions.” How election officials interpreted “the number of units used” is the source of substantial heterogeneity in the Election Day Survey, making the data provided in the survey database difficult if not impossible for statistical analysis.

Consider the responses to this question from Southern California. Orange County (California), which used an electronic voting device for precinct voting, reported the use of 9,000 units for their 1273 reported polling places. Los Angeles County, which used the InkaVote system for precinct voting, reported 245 units for 3581 polling places. Clearly these two jurisdictions have interpreted the question about the number of units used quite differently. Orange County appears to be reporting the actual number of voting devices used in precinct voting, while Los Angeles County appears to be reporting the number of tabulation units used in the election. Thus, any researcher interested in using these data to study the distribution of voting units across precincts needs to exercise great caution using these data. This problem likely permeates throughout the Election Day Survey, an issue which will likely only be uncovered as researchers look very carefully at these data.

Defining Election Terms

One large problem in the EDS is that states and localities clearly define the same term or concept differently based on state law, regulation, or custom, a problem that is closely related to heterogeneous responses. One example is the concept of voting in-person at a special polling place prior to an election. In some states, this is called “early voting” but other states refer to this as “in-person absentee voting.” When the EDS asks about early voting or absentee voting, different states are reporting different results based on their interpretation of the term. Similarly, the reporting of voter registration data likewise reflects how states treat or define inactive voters. Some states have no inactive voters and other states report sizable populations. The lack of common data definition is a problem that makes interpretation of many components of the EDS difficult, and potentially plagues the analysis of these data across states. As we discuss in more detail below, there are a number of strategies that should be utilized in future EDS efforts to develop better and more consistent data definitions to avoid this problem.

Thinness of Administrative Data

This is an area where there is considerable room for improvement in future data collection efforts. There is a great deal of administrative data that is needed to better audit the election process in the United States. Some of this is performance data, some of this is data that is needed to understand the security of the process and to audit election outcomes, and some is needed to provide a better understanding of the basic costs of conducting elections in the United States. Here a useful reference is the Caltech/MIT VTP recommendations issued in October 2004, where a wide range of administrative data was discussed and recommended for collection and distribution. To facilitate improved administrative data collection we attach as an appendix to this report the complete list of data that the VTP report called for.

Some of this data (for example, detailed information on the training of polling place workers or procedures associated with overseas and uniformed services voters) may be collected as part of independent EAC research contracts. However, it is unclear whether those data will be provided in formats that can easily be integrated with the Election Day Survey data, or that they will even be made available to the interested public at the jurisdiction level. An example here is the UOCAVA survey data, which were recently released only on a state-by-state basis. Either these data should be collected as part of a unified future Election Day Survey, or efforts should be made to provide these additional data in a way that makes them easily integrated into future Election Day Survey efforts.

Better Data Distribution

A common complaint from the research community, and the interested public, is that data on elections is often difficult to obtain, and when it is available, is often provided in formats that are not user-friendly for a secondary user of the data.⁷ For example, when data is provided in many cases, it is sometimes available only as a PDF file, requiring that secondary users of the data undertake substantial work to transform that data into a useable database format (for example, Microsoft Excel, or other database formats that other statistical software packages can use). This is despite the fact that these PDF files containing useful data were themselves produced by some software package that likely could also have produced the data in more useable formats, instead of solely PDF.

The Election Day Survey data was originally available only in the form of html tables on the EAC's website, something that was relatively remedied when the data were also made available in the form of downloadable html or Excel state-by-state files. However, the Election Day Survey data were never made available in one large, easy-to-use and well-documented database, of the sort that is common in the research community. The relative difficulty associated with the decentralized manner of data provision may account for the lack of use of these data.

⁷ See, for two examples, Alvarez, Ansolabehere and Steward (2005) and Stewart (2005), both referenced earlier.

As part of our study of the EAC’s “2004 Election Day Survey” members of our research team took the data as distributed from the EAC’s website, and manipulated those datasets into a variety of forms that might be more easily used by researchers and the public. These datasets are available the VTP website at <http://www.votingtechnologyproject.org/data.htm>.⁸

Lack of Precinct Data

No doubt, election officials, researchers, and the public now have access to more comprehensive data about election administration as the result of the 2004 Election Day Survey. But for the type of detailed election auditing or forensic analysis that many argue is necessary for the American electoral process, we need to have access to more finely grained data: information about the practices and outcomes at the lowest level of election administration, the voting precinct.

Developing a comprehensive database of precinct-level data from any particular higher-level jurisdiction (municipalities, counties or states) will not be easy, and making such data usable will also be difficult. But this is a necessary next step in the evolution of our election data collection practices in the United States, as such data is necessary for more detailed analysis of how Americans vote, what affects the quality of their voting experience, how reliable precinct voting procedures and technologies are, for comprehensive audits, and for detection of anomalies and irregularities. Reporting jurisdiction-level data is wonderful, but it is insufficient, as much detail gets removed as the information from precincts is aggregated to the jurisdiction level.

Conclusions

Collecting data from a population as diverse as the local election community in the United States is a difficult proposition. The Election Day Survey is a critical effort that will allow the public—from election officials, to academics, and to the simply curious—to understand basic facts about elections in their community and in other communities nationally. However, it will take time, resources, and focused effort to make this survey *the* source of information on election administration. At this early stage is a good point to begin the process of making the important basic changes that will allow the survey to meet this goal. Thus in an effort to assist in reaching this goal, we offer the following recommendations for improving the Election Day Survey.

⁸ These databases were constructed with the assistance of Delia Bailey, Sarah Hill and Betsy Sinclair, and we thank them for their efforts. We have made no attempt to verify the data in these database files, other than minimal checks to insure that the merged data matches some fields in the original reports. We welcome any feedback that the research community may have about these databases.

Recommendations

- 1. Define the terms and concepts to promote uniformity across jurisdictions.** The EAC should bring together election officials and others to create a common set of data definitions for the concepts contained in the Election Day Survey. These definitions would ensure that everyone clearly understands what each question means and how to interpret each item. The EAC could distribute these definitions to provide standards for how election data are reported.
- 2. Refine and clarify the survey instrument.** The EAC, election officials, researchers, and other interested parties must insure that future EAC Election Day Survey instruments contain questions that will not be subject to different interpretations by those completing the survey. By providing a survey instrument that clarifies the intention behind each question, either by improved question wording or by providing additional clarification (perhaps by providing a “frequently asked questions” site, or by providing online or toll-free assistance), the EAC can insure that heterogeneous survey responses in the future are minimized.
- 3. Seek additional important information.** Without vastly increasing the complexity of future Election Day Survey instruments, the EAC should examine collecting new data in future surveys. There is little known about detailed precinct voting practices, election administration staffing or procedures, and about the costs of election administration. These and other important data were discussed in the 2004 Caltech/MIT Voting Technology Project report on improving election reporting, and we would here reiterate that report’s list of information that election jurisdictions should retain and report; that list is provided as an appendix to this report.
- 4. Reduce the cost and complexity of the survey instrument.** No doubt, one of the reasons that the 2004 Election Day Survey did not achieve full compliance --- and why there is much missing data --- is that many election officials may not have had the resources to retain and report the detailed information that the 2004 Election Day Survey requested. This may become an even more pressing issue if future Election Day Surveys request additional data, or more finely-grained data. So the EAC should consider taking steps to reduce the cost of completion for election officials, and to reduce the complexity of the survey instrument. For example, providing future survey forms online and in an interactive format might make it easier for election officials to complete a complex survey instrument (the online and interactive form could guide their responses, keep them from responding to irrelevant questions, and could perform simple checks on the information provided while the respondent is inputting data). Or the EAC could explore developing strategies for collecting some detailed supplemental information from a

sample of election jurisdictions, while continuing to collect complete information from all jurisdictions.⁹

- 5. Improve the technology of reporting.** As election officials increasingly move to electronic election management solutions, the EAC should work with election officials and vendors to make sure that the election management solutions store all of the information that will be requested from future Election Day Surveys, and that the election management solutions can produce complete reports of requested data. Even better, the EAC could work to develop and disseminate standard post-election survey reporting tools that could greatly alleviate the costs to election officials of collecting and distributing this data, and which would insure that the datasets produced would have identical formats. Such standards were discussed in a report that we co-authored.¹⁰
- 6. Insure interoperability across datasets.** Additionally, the EAC must work to insure that all of the various datasets that it provides in the future are completely interoperable—they should be easy to merge and should use common data formats, data definitions, and be fully documented. For example, the EAC could now work to make datasets from the 2004 Election Day Survey, the 2004 UOCAVA survey, and the 2004 NVRA report, and to make these three datasets fully interoperable. This could easily be accomplished if these survey databases were available in electronic form, with a common geographic identifier (for example, a “FIPS” code) that would allow users to merge the databases in useful ways, and if the datasets used common definitions.¹¹ This would allow for synergies that are impossible to predict, but which could be exploited by making all of these datasets easily interoperable.
- 7. Automate data collection where possible.** Automation could take a variety of forms. One way to automate data collection would be to work with vendors and election officials to insure that election management software routinely produce the requested information in a standard format, which is then forwarded to the state or federal levels, as a part of the routine post-election reporting of election results and administrative data. Another way to automate data collection would be to develop software tools that will routinely access data in electronic formats that are being made available by state and local election jurisdictions in future elections, and then to use that data either to pre-populate the survey form for those election jurisdictions or to use that data to check the information that the election official provides in response to the survey.

⁹ One model for this is how the U.S. Census Bureau implements the Decennial Census: all households are targeted with a “short form”, while a scientifically-selected sample of households receives a much longer and more detailed “long form.” The strategy is that the “short form” collects information that is needed from all households, while the “long form” collects supplementation information that can be used to provide estimates.

¹⁰ “The Next Big Election Challenge: Developing Electronic Data Transaction Standards for Election Administration.” IBM Center for the Business of Government, E-Government Series, July 2005.

¹¹ “FIPS” codes are the “Federal Information Processing Standards” codes (<http://www.census.gov/geo/www/fips/fips.html>). Where these commonly used and standard geographic identifiers used in the EAC survey databases, not only could researchers easily merge the various databases produced by the EAC they could also easily integrate other data with EAC survey data, or vice-versa. This would allow for synergistic use of the EAC survey data.

- 8. Improve compliance.** This should be a major goal for future Election Day Survey efforts. Better compliance, both in terms of overall survey compliance as well as efforts to increase the quality and extent of data provided by respondents, is a critical requirement for improvement in the Election Day Survey. Better compliance can be achieved by working with vendors and election officials to insure that they are aware of the rationale behind these data collection efforts, and that they understand the importance to their efforts of providing data that they can use in the future to evaluate their performance and to make more informed acquisition and procedural decisions. Compliance will also be improved by trying to reduce the costs to election officials of survey participation, perhaps through automation or other efforts, and perhaps by working more closely with election officials as they seek to provide the necessary information. It might also be necessary to provide some election jurisdictions some assistance in data retention and response efforts, to work collaboratively with them to insure that in the future they know what data is being sought and how they can most easily retain and provide that information.
- 9. Improve data distribution.** There are many potential users of Election Day Survey data: election officials, researchers, vendors, and the public, to name four. Future Election Day Survey datasets should be released in common database formats (Excel and comma-delimited), at the lowest level of aggregation (precincts or local jurisdictions). Future datasets should be distributed with common geographic identifiers, allowing each dataset to be merged with other EAC datasets.
- 10. Provide data at precinct level.** Where possible, data should be reported at the lowest level of aggregation, and for much of the basic data in the Election Day Survey, that level should be the precinct. Some data will likely be reported only at the jurisdiction level, but when possible all future data should be released at the lowest level of aggregation.
- 11. Methodological development of tools for fixing and analyzing future datasets.** In collaboration with the research community, the EAC should work to help develop and distribute methodological innovations for studying election administration data in the Election Day Survey. These tools should include methods for assessing data quality, imputation of missing data, tools for dealing with selection bias, and for analysis of these data for various substantive purposes (auditing or anomaly detection are two examples). The EAC could consider developing guidelines for the appropriate uses of election administration data, to avoid inappropriate or incorrect use or interpretations of data. Such collaboration could be facilitated by other state or federal government entities, like the National Science Foundation, or by collaboration with the research community. Additionally, the EAC should consider exploring how supplemental data might be provided to alleviate missing data in future EDS.
- 12. Build collaborations.** There are a variety of collaborations that the EAC can and should build to improve future Election Day Surveys. One key collaboration is with election officials, in the EAC should continue efforts to work with election officials to refine and clarify survey instruments. Also, by working with election officials, the EAC can better inform election officials of future data requests; election officials then may be in a better

position to track information that will be included in survey instruments. Another area that the EAC can improve collaborations is with the research community, where there is considerable expertise in both survey design and administration, as well as practical experience in the collection and analysis of election administration data. Many in the research community have resources available to them, either in terms of intellectual capital or physical resources (like research assistants), who may be able to assist the EAC in their Election Day Survey efforts. Finally, the EAC should establish collaborations with organizations and entities --- private and public --- that are now in the business of collecting and analyzing similar data. Examples of such organizations include the U.S. Census Bureau (in the public sector), academic organizations like the American National Election Studies or the General Social Surveys, and private entities that collect or use election administration data. These collaborations should eliminate overlapping efforts, and might provide for helpful synergies in future data collection efforts.

Appendix: Variable List from 2004 VTP Report, “Insuring the Integrity of the Electoral Process: Recommendations for Consistent and Complete Reporting of Election Data”

1. Data to collect before the election:

- a. Local voter registration numbers and lists. [P,S]
- b. Inventories of equipment and ballots upon acceptance (e.g., date of purchase, source, maintenance records, vendors, serial numbers, retain code versions in offsite escrow). [S]
- c. Seal numbers for ballots and machines and storage locations for voting equipment. [S]
- d. A record of personnel with access to equipment, including detail such as when and where. [S]
- e. Changes made to the equipment (e.g., oiling, charging, battery changes, memory upgrades, putting in a module, checking odometers, code drop). [S,P]
- f. A list of the times and modes by which voting equipment is transported (including license plate number and driver for chain of custody purposes). [S]
- g. Inventory of equipment and materials before and after transportation. [S]
- h. Inventory of equipment and materials before voting begins. [S]
- i. Pre-election equipment testing data, including the number of systems tested and problems observed during testing. [S,P]
- j. Number of training sessions held for poll workers, and a roster of poll workers attending each session. [P]
- k. Copies of sample ballots and voter information materials. [P]

This data helps assure that ballots, equipment and polling places are usable and also makes it possible to solve some problems and questions that may arise later.

2. Data to collect during the election:

- a. Number of poll workers at each poll, including the times at which poll workers arrive and leave. [S]
- b. Signatures (not check marks) of those present. [S]
- c. Signatures for inventory received election night, both in precincts and when inventory is returned to the central office. [S]
- d. Tally at precinct and time it was conducted. [S,P]
- e. The number of poll and early voting sites and any rents required to use these locations. The number of workers in each poll or early voting site, their rate of pay, and their required number of hours of work. [P]
- f. If “parallel testing” is conducted on Election Day, the number of voting machines tested, the way in which they were selected for testing, and the results of those tests. [S,P]
- g. Exact time when each poll site opened. [P]
- h. The number of poll sites that experienced significant problems, an explanation of the problems experienced, and a description of how these issues were resolved. [P,S]

These data will ensure that processes during the election are monitored. They also give the best possible means to later establish what voters' intentions were, and that they were allowed to vote correctly.

3. Data to collect after the election:

- a. Inventory of equipment and materials after polls close. [S]
- b. The total number of ballots cast (report absentee and poll site totals separately, if possible). [P,S]
- c. The number of votes cast for all candidates for federal office (reporting absentee and poll site totals separately, if possible). [P]
- d. The number of registered voters. [P,S]
- e. The number of people who voted as indicated on check in/check out lists. [P,S]
- f. The numbers of absentee ballots applied for, tabulated, and challenged. [P,S]
- g. The number of absentee ballots received, recorded by date received. [P]
- h. The number of absentee ballots returned from citizens residing outside the country, and the number of these that are challenged. [P,S]
- i. The number of tabulated provisional ballots provided to voters that were challenged. [P,S]
- j. The number of early voters. [P]
- k. Transportation records of equipment (consistent with above criteria). [S]
- l. Storage records of materials. [S]

These data establish the ability to know that votes were handled and reported correctly. Furthermore, they give people the ability to know how to improve processes for future elections.

4. Demographic and administrative data:

- a. The annual expenditures for election administration, including personnel and capital expenditures. [P]
- b. The number of physical voting sites and the number of precincts (if not the same because of consolidation) used in the election. [P]
- c. The number of days in which early voting is allowed, and the number of early voting sites operated. [P]
- d. Census demographics of voting precincts, if available. [P]
- e. Salary, by job category, of poll workers for the election, details of their job qualifications and hiring process, and years of experience [P,S]
- f. Type of election administration system (e.g., elected or appointed board, elected or appointed registrar). [P,S]