

Oregon Election Performance Auditing

Pilot Project: Analysis and Recommendations

April 13, 2020

Paul Gronke
Professor of Political Science, Reed College
Director, Early Voting Information Center

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

Paul Manson
Research Director
Early Voting Information Center

Silvia Kim
Ph.D. Candidate
California Institute of Technology



April 13, 2020

To: Steve Trout, Director
Elections Division
Oregon Secretary of State

From: Paul Gronke Director, Early Voting Information Center at Reed College
R. Michael Alvarez, Professor of Political Science, Caltech, and Co-director
of the Caltech/MIT Voting Technology Project

Dear Director Trout:

We were recently retained by the Oregon Secretary of State's Election Division (S-16500-08) to conduct a pilot project regarding the feasibility of a performance audit of the Oregon statewide elections system, and to collect polling data to establish statewide levels of confidence in the system and identify areas that need improvement. Our final report and report appendices are attached.

While we did not find any immediate vulnerabilities in the system, and found high quality and accurate data in the system, the most important and pressing conclusion of our report is that the Oregon Centralized Voter System, like many voter registration systems nationwide, is difficult to audit as it is based on outdated technology that needs to be modernized.

We also found that Oregonians express a high level of confidence in the integrity of the state's election system, but there are significant gaps in knowledge about processes and procedures currently in place that should enhance confidence and trust.

We provide a pathway forward to conduct audits for the January 2020, May 2020, and November 2020 elections, and use the auditing system as a method to help the Secretary of State make fully informed decisions about how to modernize the OCVR and associated databases. We also recommend a strategic communications plan and ongoing monitoring of public attitudes to assure that the public understands how the state assures the security and integrity of the system.

At monitoringtheelection.us, we provide the rationale for a trusted third-party elections performance audit system, how these systems can help to improve election administration help election administrators assure the integrity of their registration and elections database systems.

This project involved the first collaboration between the Early Voting Information Center at Reed College and the Caltech/MIT Voting Technology Project at the California Institute of Technology. While our institutions have very different educational missions, we share an interest in helping to leverage our scientific research capacity to improve American elections and to mentor new generations of scholars of election science and administration.

We want to highlight the contributions of Seo-Young Silvia Kim, who is finishing her PhD in Social Science at Caltech, and Paul Manson, who is finishing his PhD in Public Affairs and Policy at Portland State University and serves as Research Director at EVIC. We are fortunate to be able to work two smart and creative junior scholars.

Questions or comments about the report can be directed to Dr. Paul Gronke, gronkep@reed.edu, or Dr. Michael Alvarez, alvarez.research@gmail.com.

Contents

1	Executive Summary	2
2	Pilot Project Scope	4
2.1	Specific Aims of the Project	5
2.1.1	Accessing Key Elections Data Flows	6
2.1.2	Examples of Database Changes And How Performance Auditing Can Operate	7
2.1.3	Public Opinion About Oregon’s Elections System	8
2.1.4	Pilot Project Outcomes	9
3	Voter Files and Election Workspace Auditing Pilot Project Results	11
3.1	Best Case Scenario: The Orange County Example	13
3.2	Results: Data Access and Retrieval	13
3.3	Results: Data Schema and Database Design	15
3.3.1	Oregon Centralized Voter Registration System (OCVR) and Elections Workspace (EW)	15
3.4	Results: Data Quality (Static)	17
3.5	Results: Data Quality (Dynamic)	18
3.6	Results: Election Workspace (EW)	19
4	June 2019 Oregon Election Audit and Voter Confidence Survey	22
4.1	Public Opinion and Voter-Centric Election Administration	22
4.2	Survey Research Methodology	22
4.3	Questionnaire, Top Line Results, and PowerPoint Summary	24
4.3.1	Results: Voter Confidence in Oregon is High	24
4.3.2	Results: Oregonians Are Somewhat Concerned About Fraud and Worry About Hacking	26
4.3.3	Results: Election Audits May Increase Confidence, Only If Voters Know Audits Occur	28
4.4	Summary and Recommendations from the Public Opinion Survey	29
5	A Path Forward For A Comprehensive Performance Audit of the Oregon Elections System	31
5.1	Auditing Oregon’s Voter Registration Database: Immediate Next Steps	31
5.2	Vote-By-Mail Transactions Data Auditing: A Pilot Project for January 2020	32
5.3	Public Awareness and Engagement on Election Auditing: Monitoring the Public Pulse Through November 2020	33

1 Executive Summary

In this pilot project, we established a proof of concept for a data-driven, comprehensive performance audit of the Oregon election system. The project consisted of two parts: (1) accessing and evaluating key data flows in the Oregon Centralized Voter Registration system and Elections Workspace; and (2) completing a public opinion survey to establish baseline estimates of voter knowledge of the state's election administration system and their confidence in Oregon elections.

The key findings of the report are:

- ✓ The quality of data in Oregon's election system, based on a static analysis and using established data quality standards, is generally high;
- ✓ Development of diagnostic tools and forensic indicators that would allow election administrators to pro-actively identify data anomalies and vulnerabilities is feasible, but requires immediate financial and personnel investment directed at system modernization;
- ✓ Oregonians express a high level of confidence in the election system, but are not well aware of existing audit procedures and other system protections.

The goal of the project was to generate forensic tools that could help Oregon elections officials identify anomalous patterns in data flowing into and out of key databases in the election administration system. Unfortunately, an outdated and archaic database system hampered data accessibility to such a significant degree that while we could demonstrate how such tools could be developed, it was impossible for us to obtain sufficient access to produce statistically valid measures of key items like the frequency (and range) of new registrations, the number of address changes, the frequency of registration removals, the frequency of cross-county moves, the rates of undeliverable mail ballots, and the rates of ballot returns (among some of the most important quantities for registration auditing). This is not meant to indicate that these quantities cannot be calculated or are not available, but rather that we were unable to estimate *statistical parameters* (averages, standard deviations) of these quantities, and by use of these statistical quantities, develop forensic tools which would identify anomalies.

Our inability to quickly and reliably access necessary auditing information from the Oregon Centralized Voter Registration system is in our opinion a consequence of a system that was implemented in 2004, and which has not been substantially modernized since implementation. The system was designed to solve a specific set of federal policy objectives, which unfortunately did not include data quality analysis and auditability.

We recommend that Oregon's database management system, which is now more than fifteen years old, undergo an extensive overhaul. A modernized system will (1) allow for flexible and stable access, and more responsiveness to external clients; (2) make more efficient use of staff and resources in the Elections Division; and (3) create a more secure and efficient data pipeline within the Division and between the Division and Counties. With better data access, we can perform a

more detailed audit of the election databases than we have showed in this pilot report.

2 Pilot Project Scope

The Oregon Election Performance Auditing Project provides a proof of concept for election performance auditing in Oregon. The project design has its roots in the “election ecosystem” approach to understanding election systems (Alvarez, Atkeson and Hall, 2012; Huefner et al., 2007), and more recently, a performance audit pilot project implemented during the 2018 election cycle in Orange County, CA, the results of which are reported at <https://monitoringtheelection.us/2018-orange-county-election-performance-project>.

The ecosystem approach recognizes that “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” (Huefner et al., 2007). Following the ecosystem approach, comprehensive auditing needs to be broader than checking for anomalies in the registration list or implementing a post-election ballot audit. Both of these are important steps to assure election integrity, but the ecosystem approach encourages a broader perspective.

Using data-driven and quantitative methods, *many* components of a county’s or state’s system of registration, election administration, and election technologies can be audited. As *ecological or comprehensive performance audit* can include the accuracy and quality of voter registration databases, vote-by-mail transactions data, election results during the immediate post-election tabulation and canvass process, and voter experiences and confidence, just to name a few. The particular components and design of the performance audit depends on specific features of the jurisdiction’s election system.

The need for ecological and holistic election performance auditing is especially pressing in Oregon, due to the interconnectedness of the elections system, as reflected in the component parts of the Oregon Centralized Voter Registration system (OCVR) and Elections Workspace (EW). Key data flows connect the state’s automatic registration system, traditional registration activity, precinct assignments, ballot design and delivery, mail ballot flows, and post-election results. As we explore below, Oregon’s OCVR includes some non-elections related activities such as service on juries, which may hinder the efficient operation of the system.

We have not explored, as yet, how this ecosystem operates within and among Oregon’s diverse thirty-six counties, ranging from coastline to range land, and from urban to rural. As we will return to below, a full scale performance audit should examine county-level election data flows, and how variations in the elections technology and staff expertise in various counties may impact the state databases. This was beyond the scope of this pilot project, but is something we recommend for the future.

This project also set the groundwork for establishing linkages between the performance of Oregon’s election system with public trust and voter confidence. Establishing this groundwork is importation because of research showing that “(p)rocesses and procedures than can increase the

transparency of election administration ... can also help instill in the minds of voters and other stakeholders a sense of confidence in the integrity of the election administration process in their county, state, and nation.” (Alvarez, Atkeson and Hall, 2012). A first-ever election administration focused public opinion survey provides a baseline estimate of the state of public knowledge about and confidence in Oregon’s election system, with an eye to monitoring the public barometer going forward. This survey is notable in that it was carried out after significant changes have been made to the state’s election system, following a highly competitive presidential election, and eighteen months after Oregon was the first in the nation to implement automatic voter registration.

The template for the pilot project is the aforementioned performance auditing project, conducted in partnership between the Caltech/MIT Voting Technology Project and the Orange County, California Registrar of Voters. Orange County is the fifth largest county in the United States (Orange County Registrar of Voters, 2018), with over 1.5 million registered voters. This compares to 2.7 million registered voters in Oregon, as of July, 2019 (Oregon Secretary of State, 2019).

However, there were many differences between Orange County, CA and the State of Oregon that limited the scope of this project. Most obviously, Oregon is a state made up of 36 distinct and diverse counties, each with their own county commission and election director. On the other hand, Orange County, CA is a single county with a single Registrar of Voters. This meant that many of the data flows that were monitored in Orange County had a starting and ending point in that county, while Oregon’s statewide system, even though it is a “top down” system, must be successfully integrated with 36 distinct county databases, each with its unique maintenance scheme. We elaborate on some of the challenges that the pilot project encountered in Section 3 and make recommendations about how to address them in Section 5. A second key difference was that Oregon’s election data architecture is very different than that in Orange County, and various aspects of Oregon’s election data system introduced issues regarding efficient data extraction. We will also discuss these challenges below.

2.1 Specific Aims of the Project

The Oregon Election Performance Auditing Project had two specific aims, as described in the Scope of Work:

The first aim of the project was to build a secure process so that a research and auditing team could obtain and analyze important election data flows, in a timely manner, from the Elections Division of the Oregon Secretary of State.

The second aim of the project was to establish a baseline of public opinion about elections and election administration in Oregon, investigating both how well informed Oregon’s electorate is about the elections process and procedures, as well as how the public evaluates election administration in the state.

2.1.1 Accessing Key Elections Data Flows

Gaining secure access to key elections data flows was a critical first step for “confirming elections” (Hall, Atkeson and Alvarez, 2012), assessing election performance using an “ecosystem approach,” and more importantly, to design a *comprehensive election performance auditing system*. Furthermore, it is critical to have access to frequent and timely updates of the databases being audited, particular for large databases, where finding and detecting low-incidence errors, duplicate records, and other anomalies is a central feature of the audit. Timely updates are necessary so that the auditing process can operate in as near to real-time as possible, and so that errors and anomalies can be identified and resolved as quickly as possible.

At the outset of this project, it became clear that there were two related, yet separate, data sets that needed to be included, at a minimum, in a performance audit system developed for the state. (Additional datasets from counties may need to be added.) These are the *Oregon Central Voter Registration system (OCVR)* and *Oregon’s Elections Workspace (EW)*. From those important data sets, the performance audit system would examine:

1. From Oregon Centralized Voter Registration system:
 - New registrations.
 - Changes to old registrations.
 - Automatic voter registrations.
2. From the Elections Workspace:
 - Vote by mail ballot batch mailings and return processing.
 - Ballot counts and reporting systems.
 - Post-election data-driven reports, data displays, and documents.

Forensic tools that are developed based on these data would establish boundaries around “normal” rates of change in the OCVR and EW. In technical terms, our initial approach to determining the “normal” or non-anomalous rate of change would be defined as the mean plus or minus one-quartile, also known as the interquartile range (IQR) method. Anything outside of the IQR is identified as unusual or anomalous (Kim, Schneider and Alvarez, 2019). Identifying statistical outliers or anomalies helps to reduce “false positives”, and gives the analyst the ability to focus attention on outliers that require investigation. Statistical anomaly detection methods more sophisticated than the IQR method can be deployed once the auditing process has built a lengthy time-series, and these more sophisticated methods can help to further reduce the rate of “false positives.”¹

The approach to statistical anomaly detection used by Kim, Schneider and Alvarez (2019) requires the use of database snapshots provided at high frequency, so that the research team can construct

¹See (Chandola, Banerjee and Kumar, 2009) for discussion of different statistical anomaly detection methods.

a consistent time-series of the past behavior of the database. When this consistent and lengthy time-series of past database behavior has been built, statistical anomaly detection algorithms can then be trained to detect events that need investigation. These forensic tools provide information that allows an auditing team to find, diagnose, and resolve anomalous patterns of changes to these systems.

In non-technical terms, the forensic tools identify an average rate of change in registrations, along with upper and lower bounds around this rate. Depending on the detail of the data analyzed, daily, weekly, or monthly rates that fall outside of these boundaries generate an alert to the election administrator. The alert does not necessarily indicate any problems, but because they fall outside of predicted bounds, the potential cause for the alert should be investigated.

2.1.2 Examples of Database Changes And How Performance Auditing Can Operate

Two illustrations are provided to help understand how the forensic tools will operate.

Figure 1 is reproduced from McGhee et al. (2018), a research paper that examined the turnout effects of Oregon Motor Voter (OMV). This displays the month to month change in the percent of eligible voters who are registered in the state for the for the 2004, 2008, 2012, and 2016 presidential election years. The figure shows two things, first, that there was fairly dramatic month to month (and presumably, day to day) change in the rate of registrations prior to OMV, with notable extreme values in October 2004 and April 2008. Second, the figure shows that, since the implementation of OMV in 2016, the rate of change has reduced (as we would expect in a fully mature automatic voter registration system). There remains significant variation in the total number of registered voters, and by implication, variation in the *rates of change* each month, even after OMV.

A forensic toolbox may have flagged the spikes in registration in October 2004 and in April 2008, as both display values that are abnormally high when compared to the other years. There are likely to be administrative reasons for these spikes. Oregon was considered a tossup state in the 2004 presidential contest (Kerry eventually won by 4.2% over Bush). In 2008, the state had a highly competitive primary for an open Senate seat and the presidential primary contest between Obama and Clinton remained competitive through the state's May primary. In both elections, there were active third-party registration efforts underway. While there is no evidence of problems with voter registration during these two elections, one could easily imagine that a performance audit system would have generated an alert when these spikes occurred – in particular if the audit process was working with daily voter registration snapshots. These spikes would likely have been evident in the daily time-series, and could have been investigated at the time they were flagged.

A second example is drawn from Orange County, CA, where the research team at Caltech monitored the flow of vote by mail (no-excuse absentee) ballots, since these constitute such an impor-

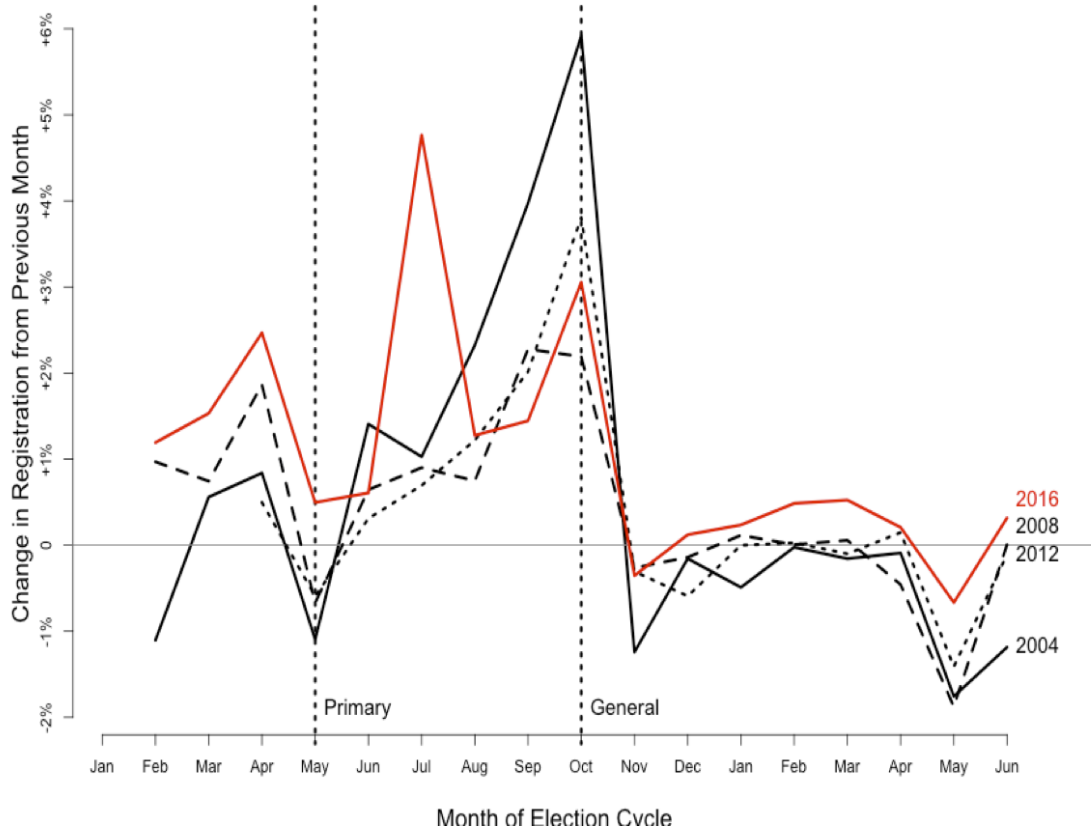


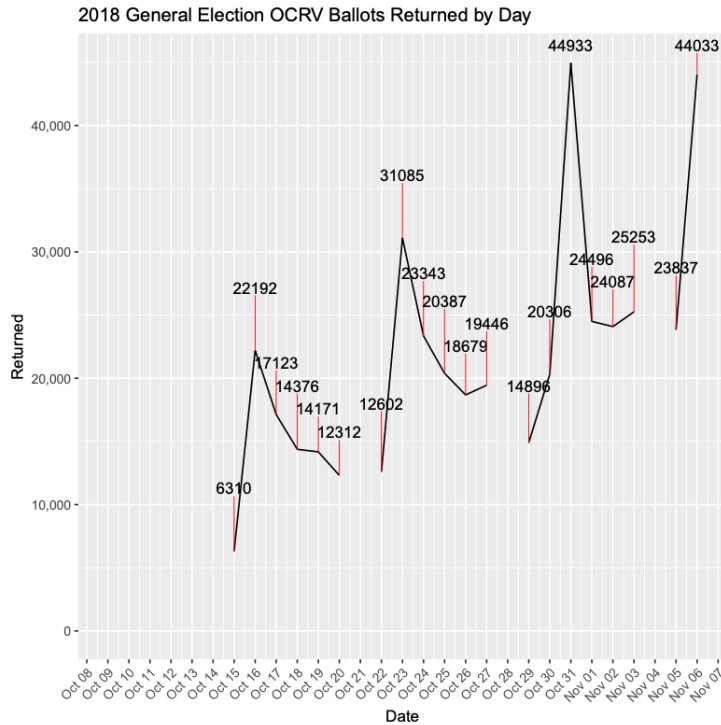
Figure 1: Monthly Changes in Oregon Total Percent Registered

tant part of the elections ecosystem in California. Illustrative results are displayed in Figure 2. Mail ballot flows would be an important component of a full performance audit in Oregon, with the same kind of forensic tools in place to identify unusual patterns in ballot returns, signature rejections, and other similar features of the system.

Our team recommends that a performance audit for Oregon should monitor vote by mail issuance and receipt. A more fine-grained auditing system could monitor additional elements of the VBM ballot flow (see the next section for more detail).

2.1.3 Public Opinion About Oregon’s Elections System

The second aim of the project was to design, implement, and analyze a public opinion survey in Oregon and establish baseline estimates of public knowledge and information about Oregon elections, public evaluations of the performance of Oregon elections, and public opinion about the security and integrity of the state’s elections. Questions are often raised after elections, particularly close and competitive ones, about the integrity and security of the system. Often these claims are illusory, based on misinformation, or originate in baseless complaints made by the losing side. Nonetheless, it important to understand at the outset the current levels of public knowledge about and confidence in this most vital part of the democratic process.



Daily data from the Orange County Registrar of Voters. These figures are estimates, and may not reflect the final totals.

Figure 2: Orange County Daily Mail Ballot Return, 2018 General Election

Knowing these levels will help us identify avenues for improvement—whether this be informing the public about things they may not know, helping them understand security processes that are not well understood (or inaccurately described by the news media), and informing them about new features in the system that increase integrity and security.

2.1.4 Pilot Project Outcomes

It is important to stress that this project was a pilot study, seeking to build a proof of concept. The short time frame for the work (one month), with an intervening election in May, meant that we were unable to collect as much data as we would have liked to in order to create forensic tools for the State of Oregon. In addition, as the report makes clear, the data access system in place at the state level severely limited our ability to obtain sufficient data in a timely manner. This is a consequence of an outdated database architecture and data access portal that could benefit from a reevaluation and refresh to bring them up to modern standards. We speak to this more in our recommendations section.

During the project, we also conducted one county site visit, to Multnomah County, and conducted extended personal interviews with two systems experts, the current Application Administrator for the Oregon Secretary of State’s Elections Division, and a Systems Engineer and

Technical Liaison at the Electronic Registration Information Center, Inc. who previously served as a Business Analyst for the Oregon Secretary of State. We used these qualitative data collection opportunities to supplement the results from our quantitative research and gain important knowledge about the history and current usage of the system.

Based on the preliminary election performance pilot project and the public opinion survey, we have studied how to provide an independent and transparent process to evaluate the current structure and procedures of the Oregon election database system. Our report provides recommendations about how to establish a comprehensive election performance methodology, and a process for evaluation of the methodology, for implementation in the 2020 election cycle. The report also includes recommendations for data access and dissemination to improve upon current processes and procedures at the state.

3 Voter Files and Election Workspace Auditing Pilot Project Results

The goal of large database auditing is to be able to confirm the integrity of the data—that the data is accurate, does not contain large numbers of duplicate records, and has not been tampered with or manipulated. In large administrative datasets, these are likely to be low-incidence issues. We also want to make sure that any forensic tools for auditing these datasets do not identify an overly large number of false positive events for investigation. These goals require that the auditors have access to frequent instances of the database, so that the normal behavior of the dataset can be estimated or modeled, allowing identification of anomalies.

An ideal situation for auditing the voter database and election workspace for Oregon is for the audit team to receive *daily snapshots* of both these databases, which could then be analyzed for duplicates and linked to previous daily instances for analysis of changes in the data. By conducting these analyses with daily snapshots, rather than weekly, monthly, or quarterly snapshots, the auditing team can construct the necessary time-series that can help determine normal database changes so that only true anomalies are flagged for investigation. This is the scenario implemented in Orange County, and it represents the standard for administrative database auditing that we wished to implement in this Oregon pilot project. A daily time-series also allows for more timely analytics, which can flag anomalies for investigation in near real-time.

Thus, the first requirement for a performance auditing system is to have sufficient snapshots of key portions of the OCVR and EW so as to allow us to capture regularized data flows and estimate key statistical parameters which are part of the forensic toolkit.

The frequency and content of the snapshots are partially a function of a state or county elections ecosystem. For example, in a state with very low absentee ballot usage, it may be less vital to monitor mail ballot issuance and returns, while states that allow for same-day and Election-Day registration would require close attention to how same-day registration (SDR) and Election Day registration (EDR) data flows come in from polling locations and counties.

Our best understanding of the data flows that are theoretically collected by the counties in Oregon, but potentially not all sent to the state, are represented in the Oregon Ballot Return Pathways, shown in 3. This is a first cut at identifying key points in the ballot flow that could be captured (represented by squares colored in blue). The figure is based on information gathered during a visit to the Multnomah County Elections Division on May 20, 2019. Our thanks to Tim Scott, Director of Elections for Multnomah County, and his staff who helped us understand their processes. It is important to recognize that the steps followed in other counties may be somewhat different because of variations in election technology across counties. This would be documented in future reports. Ideally, each of the steps coded in blue, and additional or alternative paths that may not be reflected in the figure can be monitored in future auditing studies.

Based on our analysis of Oregon’s election ecosystem and the unique features of the state, a fully

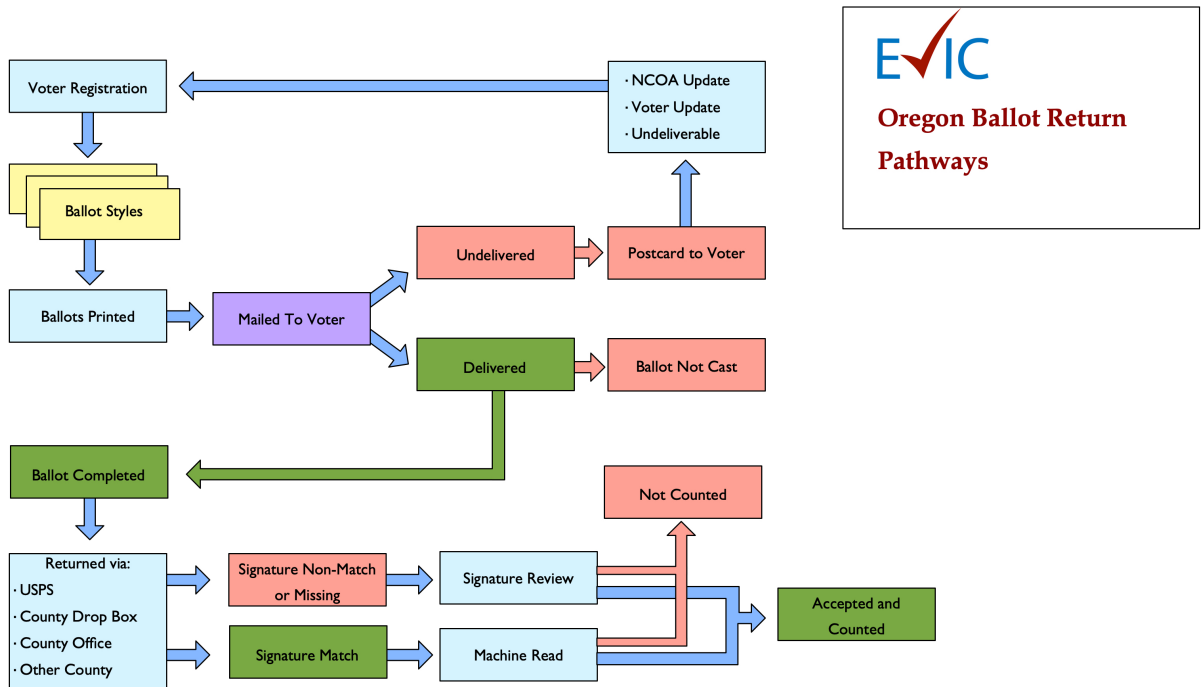


Figure 3: Pathways for Oregon Ballot Delivery and Return

functional performance auditing system would monitor data flows on:

1. Transactions and activity in the statewide voter registration system (OCVR), including:
 - New registrations.
 - Changes to existing registrations.
 - Online voter registrations.
 - Automatic voter registrations.
 - Automatic address changes.
 - Removals due to voter inactivity or ERIC reporting.
2. Transactions and activity in the Elections Workspace, including
 - Mail ballots issued.
 - Mail ballots returned (by date).
 - Signatures verified and signatures rejected.
 - Ballots accepted for counting.
3. Post-Election Reporting, including:
 - Voting turnout by precinct.
 - Election totals by precinct.
 - Split ticket voting by precinct.

As already noted, ideally we would receive daily snapshots, but these do not necessarily need to be exchanged on a daily basis. Reports could be accumulated and transmitted less frequently (weekly, biweekly), and it may be feasible to work from audit and change logs to reconstruct daily changes to the database. However, a failure to obtain regular, granular data snapshots renders a performance audit nearly impossible.

3.1 Best Case Scenario: The Orange County Example

In the recently completed pilot project conducted in Orange County, CA, Dr. Michael Alvarez reported on the data flows that were made available to their research team and the forensic tools that they developed as part of that project (Alvarez, 2019). We have already noted the differences working with a single county, and a single registrar, versus a statewide system and, even in a top-down system like Oregon's, a diverse set of counties with different election technology and different capabilities. Nonetheless, it is illustrative to recognize the analytical power that resulted from regularized data flows that were made available in the Orange County project.

Orange County was able to supply, on a weekly basis, daily snapshots of the complete voter registration database. The snapshots were created via an automated script (often referred to "cron" process—a utility that allows a user to schedule a job to run at a particular time). The snapshots were then bundled into a single compressed archive, encrypted, and stored on a secure cloud-server. The research team was able to access these archived snapshots each week, allowing for a detailed and nuanced evaluation of the Orange County system. <https://monitoringtheelection.us/2018-orange-county-election-performance-project> contains the final report and other project materials.

In the following sections, we describe what we were able to learn, given the limited data snapshots that we had to work with, and include recommendations for data acquisition and dissemination if a decision is made to continue a performance audit project through the November 2020 general election.

3.2 Results: Data Access and Retrieval

Data access was provided through manual query exports by the Secretary of State's Election Division office. The team started with common data queries that were easy to run and which could be easily provided to the auditing team via a secure FTP server. After initial exploration of data elements, the team requested a larger set of variables for the complete voter file and election workspace.

This pilot project was severely hampered by an inability to obtain regular data snapshots from the OCVR or Elections Workspace. It became quickly apparent that the OCVR was not designed to allow for comprehensive snapshots. Our understanding, elaborated below in personal interviews with the original architect of the system and the current system administrator, is that these

problems stem from the fact that the OCVR is 15 years old, and that it was not designed to accommodate the kind of data retrieval tasks that a modern auditing and performance and data access requires. It is clear to our team that if and until the OCVR can be more easily accessed by the Elections Division, assuring the integrity and security of the system may be a difficult task.

Custom database queries, data exports, and data sharing all have to be done manually unless they have been developed by the vendor. This requires using limited time and resources under the contract with the vendor, therefore these requests have to be weighed against other future support needs.

This is not an uncommon problem for states with top-down systems. In some cases we learned that states have successfully moved operations in-house, but that can require a sizeable investment. Other states have not started this process and are not even able to read their own data. This places Oregon in the middle of the pack in terms of modernizing the state's HAVA-era registration system, and makes it a timely moment to consider modernizing and updating the system.

Finally, the system requires that any data files are posted on a secure FTP site maintained by the Division. Secure FTP is an awkward method to access data and may require specialized skills. There are easier to use cloud storage options available (Amazon Web Services, Google Cloud, Microsoft Cloud, Box, Dropbox) that are typically used for these kinds of data exchanges. Our understanding from our interviews is that the SFTP policy is driven by Oregon's Department of Administrative Services, so changing this data exchange system may be difficult.

In response to these challenges, we used *five OCVR snapshots* and *two EW snapshots* for analysis in this project, a number that only allowed us an initial exploratory investigation.

We strongly recommend that the Oregon election management system switches to a database product that facilitates in-house query development and deployment along with tools for the export of data, as it can affect the administration practice as well as research outputs as is this report. Not only will this allow for an outside auditor to assure the integrity of the system, but, after initial investment for upgrades, will save resources and free up time for the system administrator.

In summary, the OCVR system was designed in response to Help America Vote Act (2002) requirements. When developed, many different features were designed into the system based on the state of technology, and the information needs of the time. However, since 2002 election administration and data management have changed in Oregon. Today, efforts are underway to update the system, including moving more of the infrastructure directly under SoS control. Currently about half of the OCVR system is under the direct control, with the remaining half relying on outside vendor support.

Resolving these issues is **an important longer term step** in moving forward for to a more exten-

sive performance auditing project.

3.3 Results: Data Schema and Database Design

The first task of the study, once data access was established, was to understand the data model used in the Oregon Centralized Voter Registration (OCVR) system, which contains all of the voter registration and election administration data in Oregon, and the Elections Workspace (EW), a subset of the OCVR that contains the complete ballot history and voting history in OCVR since its inception.

In the rest of this section, we describe the structure of this database, constructed from the perspective of the outside auditor.

3.3.1 Oregon Centralized Voter Registration System (OCVR) and Elections Workspace (EW)

The **OCVR** is a centralized, top-down, registration system maintained by the Oregon Secretary of State, Division of Elections. This system allows county election officials to access and manage voter registration records. Outside of an active election period, the system allows for updates from online voter registration and Oregon's automatic voter registration to be loaded into the database and reviewed by local election officials.

Several data flows emerge from the voter registration and ballot processing data. Prior to elections, National Change of Address (NCOA) transactions, Oregon Motor Voter (OMV), and online voter registration (OVR) flow into staging tables. These staging tables are available for Counties to review and then accept into the OCVR voter table as official registration changes or updates. These updates are performed through a user interface that virtually accesses the OCVR database from the County level. These updates from NCOA, OMV, and OVR also trigger a postcard to be sent to voters – which may result in an additional update if the voter responds with a request to change party affiliation or opt out of registration.

From an election data audit perspective, the OCVR system has two key areas to monitor for changes. The first is the voter registration table. The second is the Election Workspace.

Within the active election period, starting 61 and 45 days prior to an election, the **Election Workspace (EW)** is updated with current registered voters to generate ballot data for each voter and election. The EW contains a record for every ballot created in Oregon since the inception of OCVR. It is appended before each election with a current list of voters in each jurisdiction. This data is then available to edit and be updated by counties with voter, ballot, and other information. This record can be edited by counties during the election, and at any point after.

The EW creates a new set of data flows that can be monitored to track changes to ballot information. These flows result in updates to OCVR as counties process ballots. Ballot data is updated throughout the election cycle by county election offices. The primary sources of new data up-

dates stem from the mailing of ballots (and any returned or updated via US Postal Service), and through the return of ballots. Ballots returned as undeliverable or with forwarding information are used to update voter registration data, but not necessarily ballot data in the Election Workspace. There is some variation county by county on whether the ballot data is updated in the EW. Some of these updates may happen during the election cycle, or even after the election has been certified. Changes in the EW do not automatically update data in the voter registration table.

Counties are not required to update the current Election Workspace, though some do. In these counties, data are uploaded that reflect some of the steps shown in Figure 3. When ballots are received, they are initially evaluated in a *first scan*. In the *first scan*, ballots are run through an optical scanner capturing the signature block, ballot identification data, and election data encoded on the envelope. Ballots are sorted at this stage to remove those that need to be sent to another county or that are for the wrong election. The data from this scan is also used to verify the voter signature matches those on file. For this pilot project, we did not identify whether or not this first scan is conducted and information uploaded to the EW in all Oregon counties.

After the first scan is completed, data in OCVR is updated to indicate the ballot has been received and its status in review. A *fine scan* follows the first scan and indicates that ballots are being accepted for opening, signature verification, and counting. Ideally, after the *fine scan*, data are transacted again with OCVR. However, in some cases it was reported to the research team that counties may batch the first and fine scan data in one final OCVR update.

It remains unclear to the research team how many of these EW transactions can be monitored during the election period. For this pilot project, we did not identify whether the results of multiple scans are uploaded to the EW in all Oregon counties, particularly smaller counties that may not have the same level of automation as Multnomah County, where we conducted a site visit.

Finally, because the EW represents the basic data for historical analysis of elections. While reports are generated after each election that create a snapshot at that time – the ability of counties to edit the EW at anytime may cause the data to shift even after an election is certified. This strikes our research team as a potentially serious issue, if results generated from analyses of the EW do not correspond to certified election results (perhaps we misunderstand how the databases operate).

3.4 Results: Data Quality (Static)

Past research has recommended the use of eleven data quality indicators to evaluate the integrity of voter registration data (Ansolabehere and Hersh, 2010). Of the eleven indicators, many require the use of third-party data, such as commercial data vendor files, which we are not utilizing for this project.

In our analysis of the Oregon voter registration data, we can examine some of the same indicators recommended in Ansolabehere and Hersh (2010), such as duplicate records, missing address information, missing birth date information, and the distribution of birth dates in the data file. Because date of birth is not provided to our research team but just the year of birth, we were unable to accurately produce some of these measures. In addition, because we did not use third-party data, some address validation measures and deadwood estimations are also unavailable.

Here are some insights using the available metrics.

- *Missing Address Information:* The missing address information was straightforward to calculate—by the June 13th snapshot, excluding the confidential voters, there were 7,008 observations with missing residential information among 3,248,394 voter records, which is 0.22% of the data. This is slightly higher than the national estimate of 2010, which is 0.13%. If we include uninformative entries such as “00 0”, the number goes up very slightly. Because Oregon is an all-mail election state and there are confidential voters whose residential addresses are intentionally redacted, the mailing address matters more than the residential address in Oregon. Only 1,432 mailing addresses were missing (0.04%), either in zip code, city, or street address.
- *Birthdate Coverage:* 1,597 records (0.05%) had birth years missing. This is much better than the national estimate of 2010 which was 13.2%.
- *Duplicated Records:* excluding confidential voters and voters whose first names were missing (89 voters), we have a rough estimate of 11,806 duplicated records (0.36%). Compared to the national estimate of 2010 (1.5%), this is a low percentage.
- *Deceased Voters:* Although we do not have external validation measures, we could reasonably assume that voters over age of 110 are unlikely. It is plausible that many of these cases are voters who are deceased but not removed from the voter roll, and in many cases simply turned “inactive.” Just 119 voters were recorded to have birth years before 1909. This is much better than the national estimate of 2010, which was 1% of voters. Note however that this is a conservative estimate, and likely a lower bound on the true rate of deceased voters.
- *Registration Date Coverage:* The initial registration date was not provided—instead, we only had access to the most recently updated registration date. Instead of the number for registered voters for whom their initial date of registration is unknown, we calculate the same for the most recent date of registration—there were no such cases. However, in 722 cases,

the registration date was before January 1, 1909, such as “1111-11-11”. These are likely erroneous entries.

- *Registration Distribution:* The percent of registrants listed with most recent registration date of January 1st is 0.24% (7,855 voters). This is also much better than the national estimate of 2010, in which 1 in 50 registrants were listed as having registered at January 1st.

Overall, the static quality of Oregon’s voter data was good. However, we were only able to investigate a few metrics. We recommend that full birth date is provided for auditing purposes, as it will provide more indicators for quality analysis, as well as improve the analysis of the dynamic quality changes discussed in the next Subsection. We also recommend that the administrators contact the few voters with missing first names to ensure accuracy.

3.5 Results: Data Quality (Dynamic)

As noted at the outset of this report, a performance auditing system requires sufficient snapshots of key data flows so that we can create statistical parameters around expected rates of change in key data elements, such as the new and changed voter registrations, anomalous addresses, removals, and so on.

Our methodology also lets us look for other types of discrepancies, in particular anomalous rates of change in important fields in the voter registration data, like names, party affiliations, addresses, and the administrative status of the voter.

We were unable to secure enough snapshots of the voter data to present a full analysis due to the issues illustrated in Section 3.2. Therefore, here we present preliminary results using the data that we have been provided. We were so far given five snapshots of the voter file: April 5, April 9, June 6, June 10, and June 13 (5 snapshots during a period when we would ideally have obtained 50-90 snapshots).

Figure 4 shows the number of changes in the voter file by specific fields such as names, party affiliation, address, and active status. Note that because there is a large temporal gap between April 9 and June 6, the observed changes can be misleading, especially the continuous, linear trend between April 9 and June 6. Rather, the accumulated difference of two months are being shown as the peak in the Figure.

Because the snapshots were not yet regularly automated, with a large interval between April 9 and June 6 due to elections, there is little information available—the Figure only contains four dates. At the moment, the Figure is thus purely illustrative of the possibilities of the analysis. A longer and more detailed time series of these measures of data quality would allow us to discern trends and anomalies in changes more precisely.

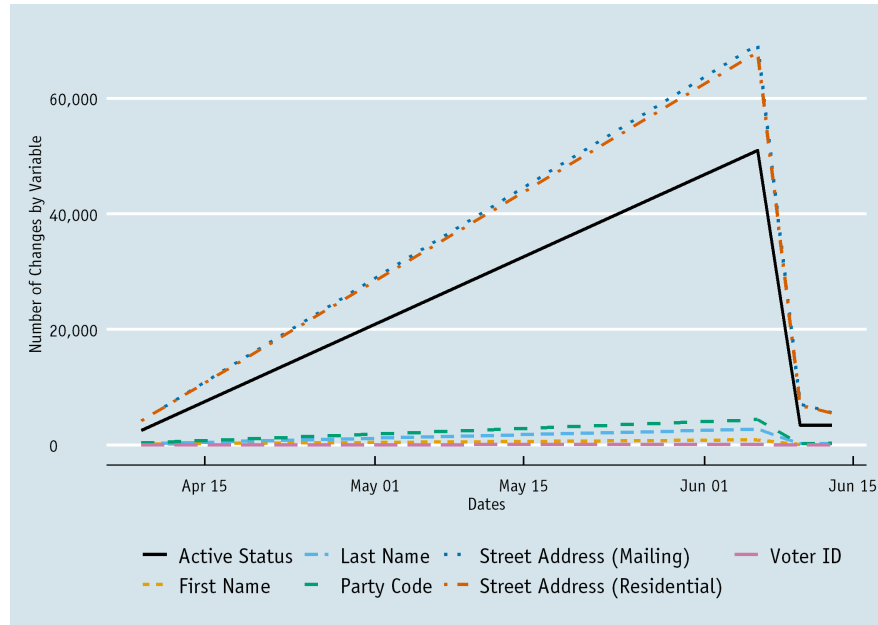


Figure 4: Number of Changes by Variable, Voter Registration Data, April 5 to June 13, 2019 (See Main Text for Accurate Interpretation)

3.6 Results: Election Workspace (EW)

Our team has received only two snapshots of the Election Workspace, first on April 8, 2019, and second on June 10, 2019. Between those snapshots, there was a special election on May 21, and the June 10 snapshot reflected the ballots sent and returned to the OSS, and both data exports showed only ballots for this election. The first snapshot was exploratory, and helped the team understand the role of the EW in the Oregon system. The second snapshot was a testbed for data export.

The forensic analysis of data from the Election Workspace is extremely preliminary as a result. However, for each election, using the patterns in the ballot sent and received, we can start checking for common patterns and anomalies.

One example of the analysis that would be possible is *the proportion of ballots that were sent late and received late*. This could be a critical indicator to the state and to individual counties if some portions of their electorate is in danger of being disenfranchised due to poor postal delivery or lack of understanding of how to correctly return ballots on a timely basis.

Figure 5 shows countywide proportions of ballots sent late compared to all ballots, and 6 shows countywide proportions of ballots received late compared to all ballots. The proportion is estimated through a partially pooled Bayesian model, also called hierarchical models. For more on hierarchical and multilevel models, see Gelman and Hill (2012).

While smaller counties have much wider confidence intervals, in Figure 5, *Hood River County*

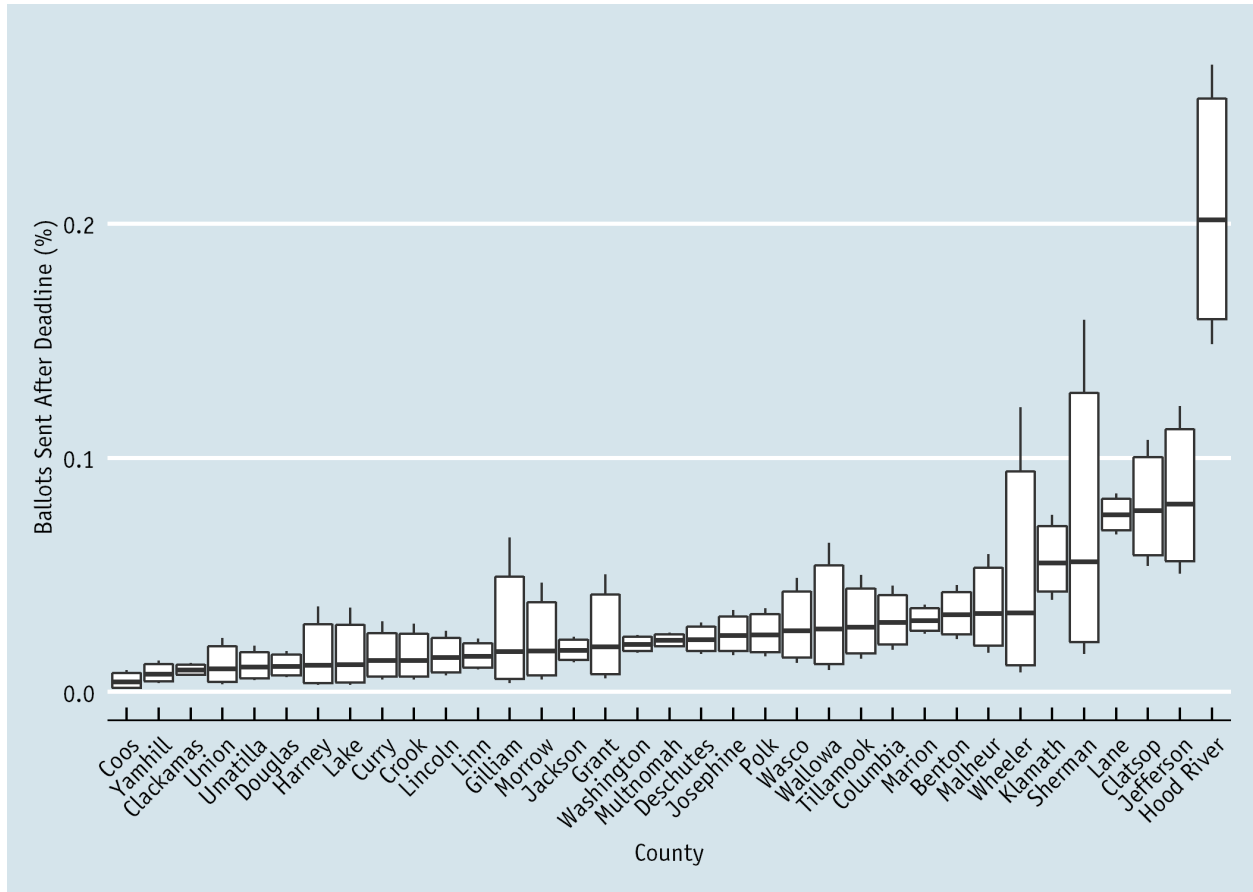


Figure 5: Estimated Proportion of Ballots Sent After May 16 Deadline

stands out as having had more ballots sent out after the deadline, compared to other counties. This is an excellent demonstration of how a proactive performance auditing system can alert local and state administrators to potential issues in their ballot delivery system.

In general, there are few ballots sent out later than designated. In Figure 6, smaller counties such as Lake, Umatilla, and Klamath have had more late-returned ballots than other counties, more than 1% of ballots. This is presumed to be an issue of geography and the postal system. In these large rural counties, more voters live farther away from the centralized postal facility that all mail passes through in the state.

While we should not be overly concerned about these preliminary estimates, because they are conditional on ballots sent/received dates being reported correctly, it may be advisable to check with the aforementioned counties to check if there were any obstructions in election administration for the May 21, 2019 election.

Finally, in the future, provision of Election Workspace data from previous election cycles could provide important additional data. Data from the EW could be used to build sophisticated

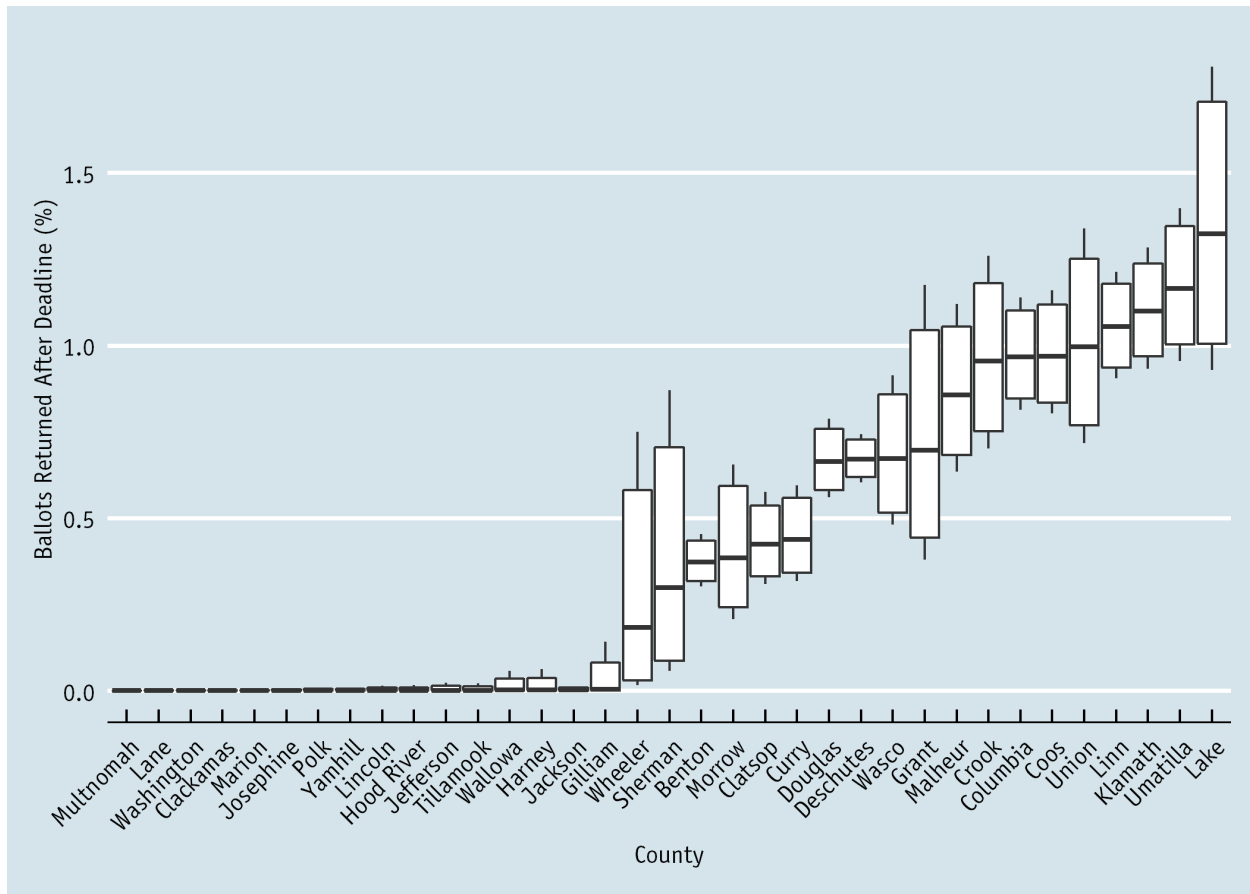


Figure 6: Estimated Proportion of Ballots Returned After May 21 Deadline

statistical anomaly detection tools that could be deployed in the 2020 and future elections.² EW data from past elections could also be used to engage in forensic analysis of past elections, which could provide both a useful baseline for future analysis as well as important insight regarding the integrity of this critical component of Oregon’s election infrastructure.

²See Chandola, Banerjee and Kumar (2009) for further discussion of statistical anomaly detection.

4 June 2019 Oregon Election Audit and Voter Confidence Survey

The second key aim of this pilot project was to design and implement a large-scale survey of Oregon registered voters. The purpose of this survey was to set a baseline to understand how much Oregonians know about our elections system, how much trust and confidence they express in the system, and to identify potential avenues to improve the knowledge and confidence levels that citizens express.

4.1 Public Opinion and Voter-Centric Election Administration

Traditionally, election officials have rightfully focused on *administering* the election—making sure that the registration rolls are accurate and complete, that registered voters are assigned to the correct precincts, that secure and accessible voting machines and polling places are made available, and that results are counted accurately and on a timely basis. Obviously, as a fully vote by mail state, the county clerks and state officials in Oregon have a somewhat different set of tasks, for example, building close relationships with the US Postal Service and assuring that ballot drop boxes are appropriately deployed, but in general, the administration of elections has not historically prioritized the *customer experience* of the voter.

This focus has changed in the past twenty years, one of the positive responses to the events surrounding the 2000 presidential election. “Voter-centric” election administration prioritizes policies and processes meant to make voting more accessible and enjoyable. Election officials nationwide strongly endorse voter-centric election administration (Adona and Gronke, 2019b).

Citizen trust and confidence in the American elections system is at a low point when examined over a 20 year years, since the 2000 election. As shown in Figure 7, reproduced from Adona and Gronke (2019a), national surveys of American voters show a slight decline in the level of confidence that their own ballots have been counted as cast, and a deep drop in confidence in the integrity of the national count. Charges of “hacking and rigging” certainly hurt voter confidence in 2016, but voter confidence had already been in decline.

Can election administrators stem the decline in confidence in the “local” vote? Previous work has shown a significant impact of the quality of local election administration on voter confidence (Atkeson, Alvarez and Hall (2015), Atkeson and Saunders (2007)). We are encouraged that the Oregon Division of Elections wants to understand how Oregonians view our state elections system, and how to best convey to the public that the system is reliable, accessible, and secure.

4.2 Survey Research Methodology

The survey was designed to establish a baseline of public opinion, including:

- Public knowledge levels of key aspects of Oregon elections.
- Trust and confidence in election outcomes and elections officials.

While voters remain confident that their own votes are being counted as intended, confidence in national results has declined.

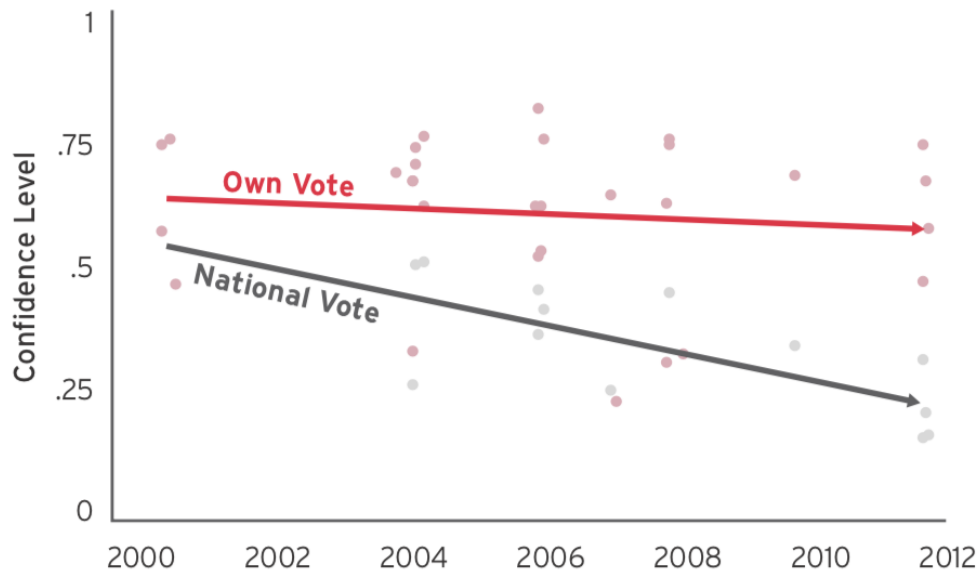


Figure 7: Voter Confidence in Own and National Votes

- Perceived levels of election integrity problems.

The survey questionnaire was designed by the principal investigators, Dr. Michael Alvarez and Dr. Paul Gronke, in close consultation with the client, and in collaboration with DHM Research (<http://dhmresearch.com>), a non-partisan opinion research firm located in Portland, OR. From June 15 to 19, 2019 DHM Research conducted a survey of Oregon registered voters. The purpose of the survey was to assess Oregonians experience with voting and their confidence in the state's election system. The survey consisted of 831 registered voters and took approximately 5 minutes to complete. This is a sufficient sample size to assess voter opinions generally and to review findings by multiple subgroups.

The mixed-mode approach was used because using either approach alone misses key demographics of interest, and thus produce a less representative survey.³ On one hand, using only online interviews misses registered voters who only use a landline phone, and risks missing those who are less tech savvy. These characteristics tend to be clustered among older and some categories of Hispanic citizens. Using only a telephone contact approach risks missing younger voters (who do not have a landline), and consequently, it may miss registered voters with lower

³For additional discussion of this, see Lonna Rae Atkeson, Alex N. Adams, and R. Michael Alvarez, "Nonresponse and Mode Effects in Self- and Interviewer-Administered Surveys", *Political Analysis*, 2014

incomes and among other demographic groups. A mixed approach was judged most effective in providing the most robust and generalizable survey results.

To ensure a representative sample, quotas were set by age, gender, area of the state, and political party.

4.3 Questionnaire, Top Line Results, and PowerPoint Summary

The survey questionnaire, topline survey responses, and a PowerPoint presentation are included with this report. Contents of the PowerPoint are also reflected in the writeup below.

4.3.1 Results: Voter Confidence in Oregon is High

The good news from the survey is that very few voters in Oregon reported any problems receiving (4%) or marking (3%) their ballots.

Some voters experienced problems receiving or marking their ballots in the November 2018 election.

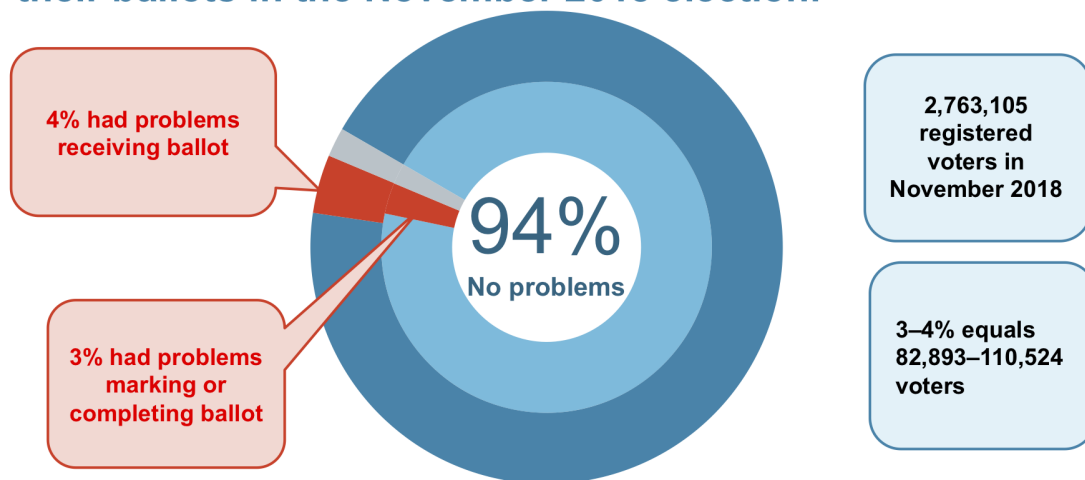


Figure 8: Relative Few Face Ballot Problems

As shown in Figure 8, 94% said they had “no problems” receiving or marking their ballot. This is not meant to minimize the nearly 100,000 voters who may have encountered some kind of problem, but overall, a system where 94% of your customers have no problems using your service strikes us as one that is functioning well. These are figures that we would expect to be very low given a mature vote by mail system with accurate addresses and ballots that have been designed to work well within that system.

In Figure 9, we compare voter confidence levels from the 2019 Oregon Election Audit Survey

to results obtained from the 2018 Cooperative Congressional Election Study (CCES). While the comparison is not precise because the CCES was conducted in the few weeks following the 2018 election, while the Oregon survey was conducted in June, 2019, we highlight two notable findings from these surveys.

Oregon voters express a **higher level of confidence** in the integrity of the county level ballot counting (the first bar) when compared to national voter confidence after the 2018 election (the second bar below). Confidence levels in Oregon in 2019 also exceeded state level confidence from the three vote by mail states (Colorado, Oregon, and Washington) (the bottom bar).

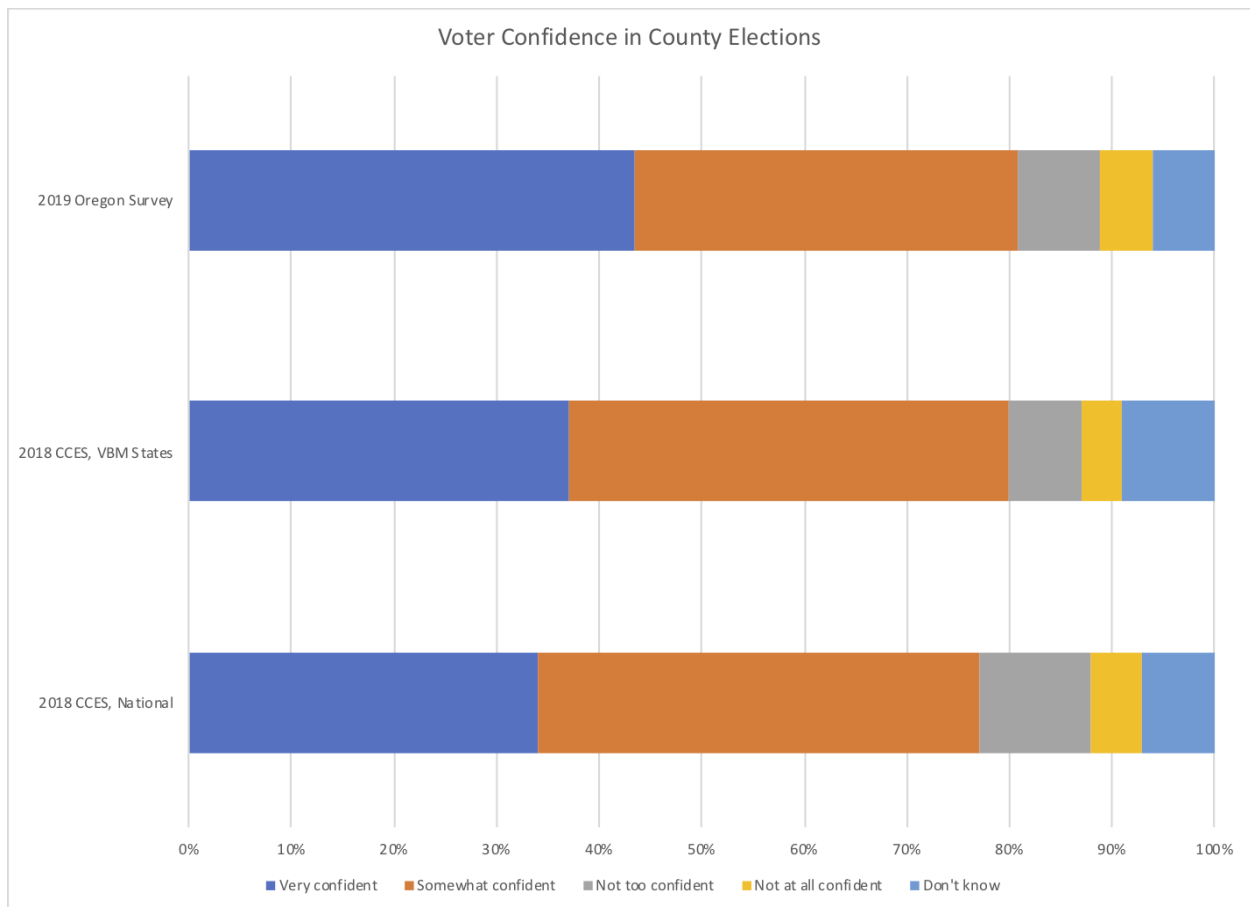


Figure 9: Voter Confidence in Oregon and the Nation

In the past, voter confidence in VBM states has lagged behind voter confidence nationally. The data from 2019 shows that voter confidence in Oregon, Washington, and Colorado exceeded the national totals by four percentage points. And, as already mentioned, a larger proportion the Oregon electorate expresses the highest level of confidence (“very confident”) in the integrity of the vote count in 2018, while the overall confidence level (“very confident” plus “somewhat confident”) is approximately equal to the level expressed by the respondents in all three VBM states.

These differences are not enormous, but may be an indicator that the public in VBM states no longer expresses somewhat lower confidence in the VBM ballot process, as has been found in previous work. More comparisons are necessary to see if this is a lasting pattern, or something that only occurred in 2018 and lasted through 2019.

With respect to important subgroups, Democrats and college graduates express higher confidence while non-affiliated voters and those with a high school or lower education express lower levels of confidence.

Voter confidence in the integrity of the county election count is high. These figures are not particularly surprising given the political and demographic makeup of the state—there is almost certainly some aspects of the “winner’s effect” where Democrats express higher levels of confidence after an election in which Democrats did well both locally and nationally. Another reason could be the low number of voters who reported problems with their ballot. Nonetheless, these indicate a well-functioning system that is viewed positively by the citizenry.

4.3.2 Results: Oregonians Are Somewhat Concerned About Fraud and Worry About Hacking

The survey revealed that Oregonians are somewhat concerned about specific types of election fraud. From 35-47% say that fraud almost never occurs, compared to just 5-14% who think it is common. Over two-thirds told us that it “almost never” or “infrequently” occurs that:

1. People vote more than once.
2. People steal or tamper with ballots.
3. Officials falsify election results

Only 7% of fewer of our respondents think that these forms are “very common.”

Votes express higher levels of concern about two other types of fraud. Fewer than half think that these types of fraud are infrequent or never happen:

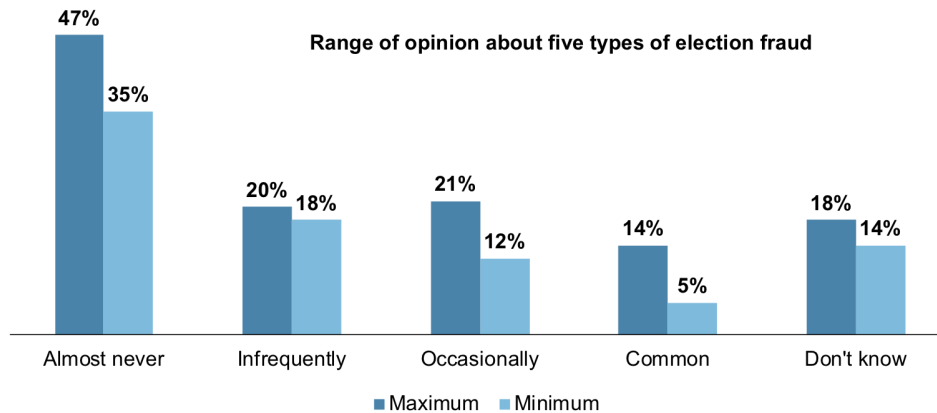
1. People voting with a ballot intended for another person.
2. People voting who are not US citizens.

One-third think they “commonly” or “occasionally” occur, with younger, Republican, non-affiliated, and voters of color expressing higher levels of concern.

Overall, as shown in Figure 10, not very many Oregonians think that election fraud is “common,” when summarized across all five types that were asked about in the survey.

There are important subgroup differences that are an area of concern. Younger, non-affiliated, Republican, and citizens of color are significantly more likely to think that some kinds of fraud occurs with greater frequency.

Voters are more likely to think election fraud almost never occurs (range of opinion 35–47% for five specific types of fraud) than that it occurs commonly (5–14%).



DHM RESEARCH | REED COLLEGE OREGON ELECTION AUDIT SURVEY | JUNE 2019

17

Figure 10: Perceptions of Election Fraud, Across Five Types

Election fraud demographic tables

Common/occasional	Total	18–29	65+	Dem	Rep	NAV	White	POC
Steal/tamper with ballots	18%	21%	9%	14%	21%	19%	17%	16%
Vote more than once	19%	27%	10%	10%	31%	19%	15%	34%
Officials changing reported vote count	18%	22%	14%	9%	28%	19%	18%	20%
Non-US Citizens voting	31%	32%	13%	12%	51%	34%	29%	49%
Ballot intended for someone else	28%	37%	16%	23%	28%	34%	26%	45%

Almost never	Total	18–29	65+	Dem	Rep	NAV	White	POC
Steal/tamper with ballots	44%	38%	56%	54%	43%	35%	46%	30%
Vote more than once	47%	39%	62%	58%	42%	42%	51%	35%
Officials changing reported vote count	45%	45%	49%	51%	44%	41%	48%	33%
Non-US Citizens voting	35%	34%	39%	49%	23%	32%	38%	17%
Ballot intended for someone else	36%	14%	49%	44%	37%	27%	39%	16%

DHM RESEARCH | REED COLLEGE OREGON ELECTION AUDIT SURVEY | JUNE 2019

24

Figure 11: Subgroup Differences in Perceptions About Election Fraud

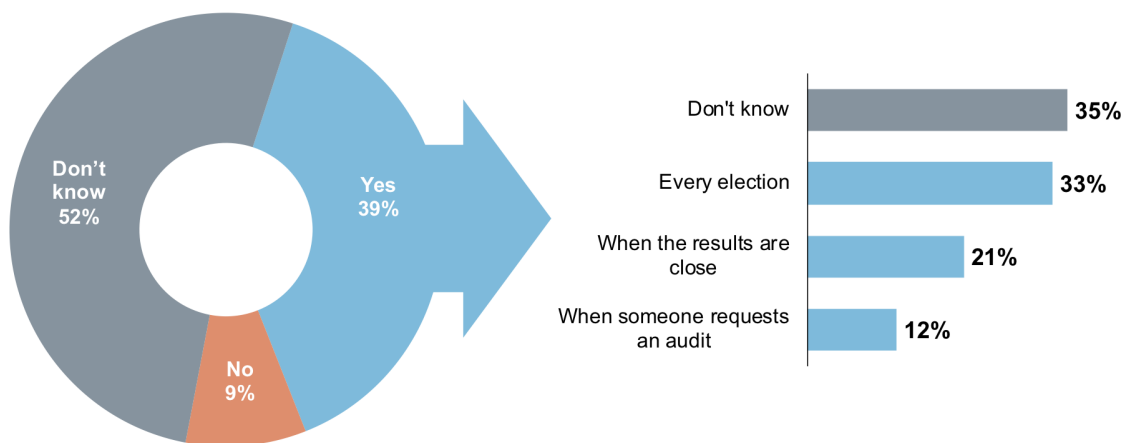
These subgroup differences will be important to monitor over the next year, and specialized outreach may be advisable to address the concerns of these groups.

Finally, respondents express varying levels of concern about “hacking,” a topic of active conversation in national political dialogue. Over half of the respondents (51%) said that they were at least somewhat concerned that Oregon’s statewide voter registration system was vulnerable to hacking. It is interesting that this is one area where younger voters expressed more confidence that the state’s system is secure, perhaps because of higher levels of comfort with new technologies.

4.3.3 Results: Election Audits May Increase Confidence, Only If Voters Know Audits Occur

One way to improve voter confidence is to assure that final results are accurate. Post-election audits are sometimes suggested as one method to improve voter confidence, and Oregon has a longstanding post-election audit system. The problem is that a majority of voters do not even know that audits are being conducted (see Figure 12), and among those who do know they take place, not many know the specific details of the process.

52% of voters don’t know whether state officials audit election results. Of the 39% who said officials do audit results, one-third didn’t know when such audits take place.



DHM RESEARCH | REED COLLEGE OREGON ELECTION AUDIT SURVEY | JUNE 2019

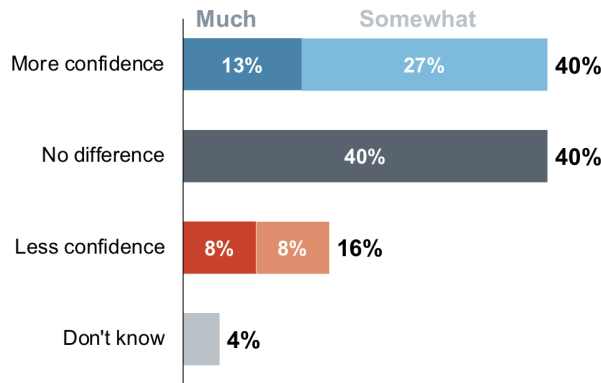
13

Figure 12: Voter Information About Oregon Election Audits

Furthermore, even if this information is conveyed to voters, it is not clear that this will move the needle much in terms of public opinion. While prospective judgments are notoriously prone to error, at least when asked, 40% of our respondents told us that knowing there was a hand-count, post-election audit would increase their confidence in the integrity of the count, but only slightly. The rest of the respondents said it would make no difference, or that they did not know.

In summary, the good news is that Oregonians generally view their election system positively and are confident that ballots are being counted as cast. The less good news is that some of

Knowing that election officials perform audits by hand counting a small percentage of ballots increased confidence in the election system among 40% of voters—in most cases moderately—and made no difference to another 40%.



DHM RESEARCH | REED COLLEGE OREGON ELECTION AUDIT SURVEY | JUNE 2019

14

Figure 13: Voter Confidence and Election Audits

the features of the system that should increase confidence are not known to the majority of Oregonians, and it is not clear that this knowledge would do much to change their views.

4.4 Summary and Recommendations from the Public Opinion Survey

Oregonians overall express a high level of confidence in the security and integrity of ballot counting at the county level, but there are important variations among subgroups that should be monitored.

Obviously, election officials would like to drive these numbers as close to zero as possible, but there is a lower limit to all of these kinds of questions, given the state of the national political conversation and lack of clear information about the real (and imaginary) security threats faced by local and state election systems.

This is an area where a future performance report could test to see if some of these figures can be moved by directed messaging from the state.

For example, the survey showed a soft impact of election audits on perceptions of voter confidence—knowing about audits made no different or a modest positive difference for two-thirds of the respondents. However, less than half (48%) of the respondents knew that election audits took place at all. If Oregonians were more informed about the existence of election audits, and how and when they were conducted, this could move the needle on voter confidence.

Future versions of a survey should also ask a more detailed set of information items about the

Oregon elections system. It is not clear to us how well most Oregonians understand the path of the vote by mail ballot and what security protocols are in place to assure that ballots are delivered to and cast by the intended registered voter.

Generally, this first baseline survey provided important data that can help the Secretary of State, the Division of Elections, and county election officials understand how the voting public evaluates the voting services that are provided. In order to study any change in the evaluations of the voting public, we recommend that surveys like these be conducted on an ongoing basis, in particular in years of statewide elections and when important changes in election administration are implemented.

5 A Path Forward For A Comprehensive Performance Audit of the Oregon Elections System

This pilot project yielded a number of important insights regarding the current architecture of the voter registration and voting-by-mail transaction databases used by the Oregon Secretary of State, as well as providing a baseline for understanding Oregon voter awareness of and confidence in the state's election procedures and administration.

5.1 Auditing Oregon's Voter Registration Database: Immediate Next Steps

In our efforts to work with the Oregon Secretary of State's team to obtain voter file extracts for auditing, it became clear that obtaining the daily voter file snapshots that would be ideal for our auditing techniques was more difficult than anticipated. Despite the Division's best efforts, they had difficulty producing daily snapshots for various technical reasons apparently related to the original design of the database's architecture. Additionally, in our examination of the database structure, we determined that the current structure of the database appeared overly complicated, with many empty tables. The consensus in our conversations with Elections Division, one set of county officials, and an outside expert on the Oregon system is that the state's voter registration database is that it is difficult to use and to audit. The voter registration database was built in the early 2000's, and at this point is antiquated and in need of updating.

As we noted, because of these database issues, we had difficulty obtaining the daily voter file snapshots that would provide finely granular auditing data. This means that we have not been able to audit the database thoroughly at this point, though we remain optimistic that in the near future we can work with the Division to develop a process to obtain the requisite data to audit the database well.

The fact that we had trouble obtaining the data we needed, along with our general observation that the current voter registration database is complicated, difficult to use, and may be in need of redesign and redevelopment, led us to note that the database might be vulnerable to malfunction and a target for the types of issues raised in the various recent reports on the vulnerabilities of voter databases in the United States.

In our opinion, this raises the stakes for the need to move forward to develop a sophisticated auditing process and methodology immediately, so that we can produce analytical auditing reports for the 2020 election cycle that will identify emerging issues, and alleviate concerns about database errors and vulnerabilities.

Our recommendations are divided into three parts. First, we have a set of immediate recommendations, detailed below, related to the current architecture and data reporting capabilities of the OCVR and ES systems. Second, we propose two additional projects that we argue should be part of a comprehensive performance audit review which could inform system modernization after

November, 2020.

1. In the immediate term, we recommend:

- *Auditing analyses should be initiated using available data.* For example, Our work indicated that the Elections Office has a historical archive of weekly voter database extracts, albeit with a limited number of database fields. We recommend that *the weekly extracts be used for an initial auditing efforts*, with the historic snapshots analyzed and used as a baseline from which to audit contemporary weekly snapshots.
- Secondly, *weekly database transactions logs* should be produced and provided to the auditing team. The transactions logs can be audited, and possibly used to reconstruct daily snapshots for deeper analysis.
- Thirdly, the auditing team should continue to work closely with Elections Office (and the voter registration system vendor) to develop a process to *obtain detailed daily voter file snapshots* as soon as feasible.

2. In the longer term, we recommend:

- First, we recommend that the audit team develop *a partnership with one or more of Oregon's counties*, so as to track the data flow from county to state, and back, as well as to explore how Oregon's *ballot tracking system* can be taken advantage of to assure system integrity.
- Second, the Elections Division should start a process to determine how to *modernize the state's antiquated voter registration database process and architecture*, so as to have a system that is easily auditable.

5.2 Vote-By-Mail Transactions Data Auditing: A Pilot Project for January 2020

There is also a pressing need for the development of quantitative daily auditing of vote-by-mail transactions data in Oregon. Detailed and real-time forensic auditing of the vote-by-mail process would identify potential issues in the processing of ballots in Oregon, and help establish safeguards that will provide better assurance that the process is operating optimally.

However, due to the timing and scope of this pilot project, the auditing team was unable to obtain vote-by-mail transactions data that could be used to attempt this type of auditing. Thus our recommendation is that a similar pilot project be launched during the next major statewide election cycle, either working with transactions data provided by OSS, or with data from a set of representative Oregon counties. This subsequent pilot project would work to build a secure data pipeline, to build and implement data auditing procedures, and to establish a process of rapid reporting of the auditing results to the appropriate election officials and to the public.

5.3 Public Awareness and Engagement on Election Auditing: Monitoring the Public Pulse Through November 2020

The baseline survey of Oregon voters we conducted as part of this project yielded important information. In our assessment, there are two important results from the survey. First, that Oregon registered voters are not well informed about the procedures used to audit election administration in the state. Second, it's not clear that there currently is much of a connection in the minds of Oregon's voters between the election auditing that is done, and their confidence in the integrity of the state's elections.

Our recommendations are two-fold. First, we recommend that the Elections Division develop and implement a public communications plan to better inform Oregon's electorate and election's stakeholders about the many steps that state and county election officials in Oregon take to secure and to audit the state's elections. This public communications effort should also seek to make it clear to Oregon's voters and stakeholders that by securing and auditing the state's elections data and administration, that there should be strong confidence in the security and integrity of the state's elections.

Second, in order to evaluate the effectiveness of this public communications strategy, that periodic public opinion surveys, like the one we conducted, should be fielded during and after the initial public communications plan. The Division should continue to measure voter awareness about auditing and election security, to study voter confidence in Oregon's elections, and to obtain detailed feedback from voters about their experiences with election administration in future statewide elections.

References

- Adona, Natalie and Paul Gronke. 2019a. Understanding the Voter Experience: The Public's View of Election Administration and Reform. Technical report Democracy Fund Washington DC: .
- Adona, Natalie and Paul Gronke. 2019b. Stewards of Democracy: The Views of American Local Election Officials. Technical report Democracy Fund Washington, DC: .
- Alvarez, R. Michael. 2019. Assuring Election Integrity: A Comprehensive Ecological Framework for Evaluating Elections in Southern California 2017-2019 Project Final Report. Technical report California Institute of Technology.
URL: <https://monitoringtheelection.us/2018-orange-county-election-performance-project>
- Alvarez, R. Michael, Lonna Rae Atkeson and Thad E. Hall. 2012. *Evaluating Electins: A Handbook of Methods and Standards*. Cambridge University Press.
- Ansolabehere, Stephen and Eitan Hersh. 2010. "The quality of voter registration records: A state-by-state analysis." *Report, Caltech/MIT Voting Technology Project* .
- Atkeson, Lonna Rae and Kyle L. Saunders. 2007. "The Effect of Election Administration on Voter Confidence: A Local Matter?" *PS: Political Science and Politics* 40:655–660.
- Atkeson, Lonna Rae, R. Michael Alvarez and Thad E. Hall. 2015. "Voter Confidence: How to Measure It and How It Differs from Government Support." *Election Law Journal: Rules, Politics, and Policy* 14(3):207–219.
- Chandola, Varun, Arindam Banerjee and Vipin Kumar. 2009. "Anomaly Detection: A Survey." *ACM Comput. Surv.* 41.
- Gelman, Andrew and Jennifer Hill. 2012. *Data analysis using regression and multilevelhierarchical models*. Vol. 1 Cambridge University Press New York, NY, USA.
- Hall, Thad E., Lonna Rae Atkeson and R. Michael Alvarez. 2012. *Confirming Elections: Creating Confidence and Integrity Through Election Auditing*. New York: Palgrave Macmillan.
- Huefner, Steven F., Daniel P. Tokaji, Edward B. Foley, Nathan A. Cemenska, Moritz College of Law and Election Law @ Moritz. 2007. *From Registration to Recounts: The Election Ecosystems of Five Midwestern States*. Columbus, Ohio: Ohio State University Michael E. Moritz College of Law.
- Kim, Seo-Young Silvia, Spencer Schneider and R. Michael Alvarez. 2019. "Evaluating the Quality of Changes in Voter Registration Databases." *American Politics Research* .
URL: <http://vote.caltech.edu/working-papers/134>
- McGhee, Eric, Paul Gronke, Mindy Romero and Rob Griffin. 2018. Voter Registration and Turnout under "Oregon Motor Voter": A Second Look. In *Annual Meeting of the American Political Science Association*. Boston, MA: p. 26.

Orange County Registrar of Voters. 2018. "2004-2018 Voter Data Trends." <https://www.ocvote.com/election-library/docs/Voter%20Data%20Trends%20Report.pdf>.

Oregon Secretary of State. 2019. "Voter Registration Comparison by County July 2019."