Maintaining the Machinery of Democracy: Assessing the Quality of Voting Equipment in the 21st Century

Nadine Suzanne Gibson
The University of Texas at Austin

July 26, 2018

Abstract

Underlying the concerns of local election officials about the state of their voting equipment is the notion that voting technology has a substantial impact on the quality of American elections. Voting equipment can quite literally be considered the machinery of democracy. Therefore it is important to understand the relationship between quantity and quality of voting equipment on voting behavior. This paper will examine the impact of differences in resource allocation on turnout. I use the term "resource allocation" loosely to describe financial resources spent and labor mobilized to maintain or improve a jurisdiction's voting equipment. Now that HAVA funding is no longer provided to localities at the intended capacity, there is substantial variability in the amount of resources available to localities to spend on voting equipment. I find empirical evidence in support of turnout improvements through specific types of investments such as vendor services in the form voter education programs and with newer equipment.

1 Introduction

Despite the obvious relevance of the election experience for subsequent democratic participation, the voting and elections literature has overwhelmingly focused on voting behavior rather than the actual act of voting itself. Since the controversial 2000 presidential election, there has been an increasing demand for information about improving the conduct of American elections. With only a decade-and-a-half of sustained attention by political scientists, our understanding of election administration has grown greatly. Most notably, research has focused on politically salient issues like turnout, residual vote rates, voter identification, and voter suppression. Although these issues are important and contemporaneous, persistent less visible problems plague the system and attract scant scholarly attention. It takes major election mishaps to garner attention to issues that have been of utmost concern to local election officials all along, such as creating foolproof ballots. It can be argued that crisis in 2000 could have been avoided if punch card voting machines allowed for longer and bigger ballots.

This study hopes to use datasets not fully leveraged to shed light on widespread persistent problems facing election administrators. What we know about these sorts of issues is for the most part anecdotal and descriptive. Some research has helped us better understand elements of election administration, but they are (by and large) from the perspective of the voter. The data used for this analysis is derived from county, municipal, and state contracts for the acquisition of voting equipment by election equipment vendors. This study will establish a relationship between turnout and resource allocated to elections. In particular, financial resources spent on voting equipment by county governments.

2 Election Administration in the 21st Century

The decade of 2000-2010 saw unprecedented election reform (Montjoy 2010). One of the most substantial was the Help America Vote Act of 2002 (HAVA; P.L. 107-252). In a nutshell, "HAVA created the Election Assistance Commission (EAC), established a set of election administration requirements, and provided federal funding, but did not supplant state and local control over election administration" (*The Help America Vote Act and Election Administration* 2015). As a result, the election administrative system has become increasingly complex, leaving the burden on local election administrators to navigate and implement changes (Montjoy, 2008).

In a survey of local election officials, Kimball et al. (2013, p.567) report that election officials interpret their policy environment as administratively burdensome due to "an ongoing set of unfamiliar requirements that have made their life more difficult". Currently, election administrators feel increased pressure to find "quick-fixes" and act as problem-solvers. Many administrators find it difficult to keep up with financial and labor costs associated with the requirements set by HAVA. Recent research shows that there are many non-trivial additional costs (new ballot forms, additional hours worked, rental space, etc.) associated with upgrading voting equipment. Furthermore, these additional costs vary year to year based on market prices. These rising costs and decreasing budgets are of utmost concern in an era where administrators have to oversee elections where partisan suspicions are high. As a result, the administrators have become handicapped in their ability to provide high quality elections. Like a straw that breaks the camel's back, these less visible but persistent problems can become catastrophic.

Our current knowledge of how much localities spend on voting equipment is speculative and anecdotal. According to a national survey of local election officials themselves, voting technology is the most daunting issue in upcoming elections. The results of the 2013 Survey of US Local Election Officials (Ansolabehere, Shaw, and Stewart, 2013) show about one fourth of LEOs believe that voting technology is the area in most need of significant improvement or an upgrade. Without a sense of the costs of elections, it will become increasingly difficult for localities to take preventative measures before an Election Day catastrophe.

Underlying the concerns of local election officials about the state of their voting equipment is the notion that voting technology has a substantial impact on the quality of American elections. Voting equipment can quite literally be considered the machinery of democracy. Therefore it is important to understand the relationship between quantity and quality of voting equipment on voting behavior.

This paper will examine the impact of differences in resource allocation on turnout. I will use the term "resource allocation" loosely to describe financial resources spent and labor mobilized to maintain or improve a jurisdiction's voting equipment. Now that HAVA funding is no longer provided to localities at the intended capacity, there is substantial variability in the amount of resources available to localities to spend on voting equipment. In other words, some localities are resorting to austerity while others are free to make necessary purchases on a regular basis.

3 Hypotheses

Of most interest in this study is the degree to which county expenditures on voting equipment impact county-level turnout. Turnout is arguably the most important empirical metric for judging the health of democracy. As a starting point, it is reasonable to investigate the state of voting equipment in the United States as it relates to turnout.

H1: Higher turnout rates will be associated with higher levels of resource expenditure on voting equipment.

More specifically,

H1a: Higher turnout rates will be associated with more dollars spent per registered voter.

H1b: Higher turnout rates will be associated with the purchase of supplemental vendor services.

To date, there are no published studies in any political science journal regarding the impact of vendors on any aspect of the electoral process. Election equipment in the United States is almost exclusively purchased from private-sector vendors. When a jurisdiction purchases voting equipment, they are actually purchasing the hardware and software along with a variety of services for the initial implementation and long-term service and support of the system. In other words, not only is voting equipment purchased, but so are services provided by the vendors to maintain the equipment. Unlike other industries, customers cannot "substitute away" from voting equipment when vendors increase their prices. Because voting equipment uses proprietary software, LEOs also cannot mix and match products from different companies. Therefore, firms with large product catalogues are desirable.

In the 10 states used in this analysis, there are five main vendors[1]:

1. Hart Intercivic, Inc.(“Hart")

2. Election Systems & Software, LLC.(ES&S)

3. Dominion Voting Systems, Inc. (“Dominion") [2]

4. Danaher Corporation

5. Unisyn

The implementation of HAVA was short sighted in that it forced states and localities to purchase higher quality equipment than they were predisposed to purchasing, but it created barriers for entry of new vendors through rigorous a certification process. The resulting voting technology industry is “dysfunctional” and “malformed” (Interview with Greg Miller). Choices in voting technology have dramatically decreased following HAVA. Given this, LEOs have not been able to purchase new equipment and/or make the necessary upgrades.

H2: Lower turnout rates will be associated with older voting equipment.

One of the main contributions of this proposed chapter is the incorporation of independent variables related to voting equipment quality. Unlike many other types of goods, computers are primarily subject to “technological depreciation” (Cho 2011). Physical depreciation, on the other hand, is a less prevalent reason for replacement of computers. The average working lifespan of computing equipment observed by Cho (2011) from 1989 to 1999 decreased from six to five years due to an acceleration in technological depreciation. The cost of maintaining legacy computers increases with age. If we equate the life-span of a voting machine to a typical computer, the voting machines purchased by states through HAVA funds in 2002-2006 are in need of replacement. Additionally, when machines are taken out of service for repairs, overall productivity is handicapped.

[1]Some localities use regional value-added resellers.

[2]Dominion Voting Systems purchased Premier Election Solutions, formerly Diebold Election Systems, Inc. and Sequoia Voting Systems, Inc. in 2010

4 Turnout in a Hyper-Federalized System

Unlike other aspects of American politics, elections are inherently geographic. Article 1, Section 4, Clause 1 of the U.S. Constitution explicitly grants states the responsibility of running elections. This responsibility however has devolved to the local level. Consequentially, leaving many localities footing the bill of major election expenditures. This study is motivated by the relationship between election expenditures and the quality of elections in the eyes of the voters. In other words, can localities buy high quality elections?

As a result of elections being conducted by the states and more precisely the localities, there is not widespread uniformity in American electoral procedure. Theoretically, this lack of uniformity should lead to policy improvement because jurisdictions can act as laboratories of democracy. In reality, there is little sharing of empirical data and scientific research between jurisdictions. This makes sense when we consider that localities (often with little resources) are responsible for the vast majority of election conduct. Those primarily in charge of elections, local election officials, are faced with many institutional and budgetary restrictions with little concrete evidence for what will actually work the best.

Given the nature of the data, it is appropriate to use spatial methodology to generate inferential statistics. Due to the nature of federalism, where regional subunits of government jointly share authority with the national government, location in the United States determines a great deal about the manner in which citizens are represented. This is consistent with the Madisonian conception of American popular sovereignty in that power is distributed throughout the system. Moreover, Ewald (2009, p.97) argues, "uniformity is actually not a central value of American elections". Even during the Founding, the notion of a decentralized election system was not a controversial topic. Given this, it would be naive to consider measures gauging electoral participation, such as turnout, as taking on uniform values across the country.

The traditional way of calculating turnout is by dividing the total number of voters by the voting age population (Burden and Neiheisel 2013). For the purposes of this study, turnout is calculated by dividing the total number of voters registered on Election Day. Since this study is primarily focused on measuring the impact of election administration variables on facilitating voters in casting their ballots, it is not necessary to include those who are not eligible to vote or do not wish to participate.[3]

We should expect to find “neighborhood effects" in county-level turnout for two reasons in particular:

1. Not all electoral districts are nested within county boundaries.
2. County boundaries were not constructed (nor redefined) to accommodate socio-political communities.

As an example of this phenomenon, Figure 1 is a map of the City of Austin, TX. Highly competitive races for Austin city government will drive turnout not only in Travis County (which contains the majority of the city), but also in Bastrop, Hayes, and Williamson counties (which also contain a some of the city). As a result, this scenario would lead to a clustering pattern in turnout across the spatial units, y_{Travis} and $y_{Williamson}$.

[Insert Figure 1 here]

4.1 Spatial Autocorrelation

In political science research, especially research involving elections, there is important geographic variation. With more attention being paid to election administration since the 2000 election, there is an increasing need for political scientists to be aware of how to account for geographic dependencies in the regression analysis. Controlling for spatial autocorrelation is also common practice in quantitative geographic, medical, and demographic research.[4]

The term spatial refers to “how areal units are arranged on a planar map” (Griffith 1987, p. 10). Autocorrelation occurs when the ordering of observations produces a relationship between pairs of

[3]Since voters must be registered well in advance of Election Day, the exact number of registered voters known on Election Day.

[4]Concerns over spatial autocorrelation is particularly evident in research on communicable diseases in medical statistics.

individual observations. Formally, autocorrelation means

$$h_i = f(h_j), i \neq j \tag{1}$$

where an individual observation h_i is a function of other observations.

There are two types of spatial autocorrelation that researchers should investigate when using spatial data. The first type of spatial autocorrelation is in the dependent variable. The second type is in the regression error term. Substantively, the difference between these two types of spatial autocorrelation relates to the functional form of the spatial processes (Griffith 1987).

When spatial autocorrelation is observed in the dependent variable, it is because the data is organized in such a way that observations' placement and proximity to one another are non-random. Equation 2 presents the functional form of autocorrelation in the dependent variable,

$$\begin{aligned} y_i &= f(y_1, y_2, ..y_n), i \notin N \\ N &= \{1, 2, .., n\} \end{aligned} \tag{2}$$

where y_i is a function of the values of other observations of the random variable Y at other locations, $y_1, y_2, ..y_n$. When this is the case, there is clustering of similar (positive spatial autocorrelation) or dissimilar (negative spatial autocorrelation) observations.

Odland (1988, p. 53) defines spatial autocorrelation in the error term as instances where "the error at each location depends on the errors at other locations". This generally occurs when the spatial process generating autocorrelation is caused by some unobserved variable. Consider the linear regression model in Equation 3 where ϵ_i is correlated with ϵ_j and $i \neq j$.

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{3}$$

When the errors in one point in space, *i*, are dependent on another location, *j*, the errors of the regression model are no longer independent. The model in Equation 3 would then be in violation of the Guass-Markov Theorem condition of non-autocorrelation, indicating that OLS is a suboptimal estimator relative to the models outlined in the next section of this report.

A statistically significant result would imply that we can reject the null of randomness and independence of observations. It would then be appropriate to consider the dependent variable as being systematically organized across space. In other words, the pattern in the dependent variable observed is unlikely to have occurred if it was truly randomly distributed across space. Global spatial statistics estimate the degree to which the dataset is spatially organized in clusters of like-values. The most common statistic for testing spatial autocorrelation in continuous data is Moran's I.[5]

4.2 Moran's *I*

Moran's I is a correlation coefficient between observations which are nearest neighbors (Moran 1950). Moran's I is asymptotically normal with an expected value of $-\frac{1}{(n-1)}$ under the null hypothesis of independently distributed observations (Griffith 1988; Cliff and Ord 1981; Moran 1950; Moran 1948). The theoretical sampling distribution can then be used to generate a confidence interval and a Z-statistic to test whether it is appropriate to reject the null hypothesis of independence of observations. Statistical significance of Moran's I can also be determined through non-parametric means. Exact p-values of Moran's I can be obtained from a random permutation test with a permutation distribution composed of $n!$ Moran's I statistics.[6]

Table 1 presents the results of a two-sided Moran's I test. The weights matrix uses queen contiguity which is when the weighting scheme includes all neighbors sharing at least one border and all neighbors sharing at least one vertex.[7] The results indicate that there is indeed spatial autocorrelation present in the data. The results are nearly identical across approaches. The positive statistics across Moran's I specifications and deviation from the expected value indicate that there

[5]It should be noted that Moran's I is one of many statistics measuring spatial dependency.

[6]Due to large number of permutations necessary for most geographic analyses, it is prudent to use a Monte Carlo approach to approximating the permutation distribution (Anselin 2009).

[7]This is in contrast to Rook contiguity where only neighbors sharing borders are included in the weighting scheme.

are instances of clustering of *high-high* and *low-low* values among neighboring units. The Lagrange multiplier (LM) test presented in following section will determine the appropriate spatial autoregressive regression model to use on the data.

[Insert Table 1 here]

4.3 Lagrange Multiplier Test for Spatial Dependence

When determining the appropriate model to use for analysis, Anselin et al. (1996) suggest using the LM test for spatial dependence. The LM test is used for testing hypotheses about parameters in the likelihood framework. More precisely, the LM test tests the hypothesis of a simpler model by maximizing the log likelihood subject to restrictions. When the LM statistic is large, the null hypothesis of a simpler model should be rejected. Compared to the Wald test and the Likelihood-ratio test, the LM test is the least stringent and most appropriate for testing model specifications (Engle 1980).

Anselin et al. (1996) note that the is no theoretical basis for the assumption of $W_1 \neq W_2$ in applied research. Since there is no substantial reason for this assumption in terms of this analysis, the LM statistic used for hypothesis testing will be simplified such that $W_1 = W_2 = W$. [8]

Model selection in this paper will rely on two LM tests presented in Anselin et al. (1996). Equation 4 tests the hypothesis of a spatial lag ($H_o : \rho = 0$) in the presence of spatial disturbances. Equation 5 tests the hypothesis of spatial disturbances ($H_o : \lambda = 0$) in the presence of a spatial lag.

$$LM_\rho = \frac{[\frac{\tilde{\mu}'Wy}{\tilde{\sigma^2}} - \frac{\tilde{\mu}'W\tilde{\mu}}{\tilde{\sigma^2}}]^2}{N\tilde{J}_{\rho\cdot\beta} - T} \tag{4}$$

$$LM_\lambda = \frac{[\frac{\tilde{\mu}'W\tilde{\mu}}{\tilde{\sigma^2}} - T(N\tilde{J}_{\rho\cdot\beta})^{-1}\frac{\tilde{\mu}'W\tilde{\mu}}{\tilde{\sigma^2}}]^2}{T[1 - T(N\tilde{J}_{\rho\cdot\beta})]^{-1}} \tag{5}$$

Both tests allow for the parameter not of interest to be unrestricted.[9]

[8] The only caveat to making this simplification is that the null of simultaneously testing for ρ and λ (the general spatial process regression model) cannot be tested using the LM test due to identification issues (Anselin et al. 1996).

[9] In practice, there is a two-step procedure for implementing LM tests for spatial dependence. Depending on the

[Insert Table 2 here]

4.4 Spatial Autoregressive Error Regression Model (SAER)

When the regression error in one location is dependent on the error in another location, it is necessary to use the spatial autoregressive error regression (SAER) model. In the SAER model, the autocorrelated error term, $\boldsymbol{\mu}$ is a function of the autocorrelation parameter, λ, a matrix of spatial weights for paired observations, $\mathbf{W}$, the autocorrelated error term of another observation, and an identically and independently distributed error term, $\boldsymbol{\epsilon}$ (Odland 1988).

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu} \\ \boldsymbol{\mu} &= \lambda \boldsymbol{W} \boldsymbol{\mu} + \boldsymbol{\epsilon} \end{aligned} \tag{6}$$

with

$$\begin{aligned} \epsilon &\sim N(0, \sigma_{\epsilon}^2 I^2) \\ -1 &< \lambda < 1 \end{aligned} \tag{7}$$

To find a consistent estimate for $\boldsymbol{\beta}$ it is necessary to use spatially weighted least squares (sometimes referred to as spatial Cochrane-Orcutt) to estimate λ (Anselin and Rey 2014).

5 Data and Variables

The focus of this analysis is the examination of the impact of the cost of voting equipment as well as the impact of purchasing vendor services on turnout and the Election Day experience. When localities purchase vendor services, it is for the most part optional. These services are purchased with the intent to improve efficiency and accuracy in conducting elections. There is, however, no academic research on whether these services are actually producing better high quality elections.

results of the typical LM for spatial autocorrelation in the dependent variable or the error term, it may be necessary to use a more robust LM test. The results of the robust LM test should identify whether the spatial process generating autocorrelation is in the dependent variable or the error term. The robust LM test for spatial dependence can be implemented in R using the package “spdep".

Funds from the MEDSL were used to construct a unique dataset was created to measure the cost of voting equipment using data from state and county-level contracts for the acquisition of voting equipment. The states included in this analysis are California, Delaware, Illinois, Nebraska, New Mexico, Nevada, Rhode Island, Texas, Utah, and Vermont. Future iterations of this study will include a more representative sample of states (or a sample of counties within a representative sample of states).

Open records requests to state and local election officials served as the basis of this new dataset. An open records request, also known colloquially as a Freedom of Information Act (FOIA) request, is the process by which a citizen may ask to obtain a copy or inspect documents that are considered to be public information, but are not made publicly available.[10] Municipal contracts are considered public information (the bidding process, however, is not). County and municipal level contracts for the acquisition of voting equipment will provide data on:

1. Cost per voting equipment unit

2. Geographic variability in cost

3. Number of units currently in use

4. Services subcontracted to vendors

Variables that can be derived from these contracts include age of voting equipment, number of units per registered voter, and dollars spent per registered voter on voting equipment.

Data collected from the contracts, was then merged with turnout data from the chief election officer of each state, the Verified Voting Foundation, Inc., U.S. Election Assistance Commission, and demographic data from the United States Census Bureau. This is possible through the inclusion of geodesic place codes like Federal Information Processing Standards (FIPS) code.

[10]Although both terms are identical in terms of the type of request, the Freedom of Information Act is a federal law. Governmental transparency laws are referred to by different names depending on the state. For example, in Texas the Texas Public Information Act governs open records requests made to the state and local governments.

5.1 Independent Variables

Investment in voting equipment per registered voter. Voting equipment is considered any piece of hardware used to facilitate the counting and casting of ballots.[11] The total investment in voting equipment is the dollar amount spent on the hardware of most current system of voting equipment in use in 2016 purchased by counties from voting equipment vendors.

Regardless of whether the state or county is responsible for the acquisition of voting equipment, there are unobserved state-level variables that determine expenditures (i.e. regional vendor market, tax revenue structure, state laws and guidelines). The within-between approach to regression is a hybrid of the traditional fixed and random-effects estimators. The within-between approach is sometimes referred to as the "Mundlak approach" or the "hybrid fixed-effects model". In this specification, the coefficient estimates of the state-level means and the county-level deviations (from the state mean) are uncorrelated (Dieleman and Templin 2014.

$$y_{jn} = \beta_1(x_{ij} - \bar{x}_j) + \beta_2 x_j + \omega_{jn}$$

$$where,$$

$$\omega_{jn} = \mu_j + \epsilon_{jn}$$

Vendor Services. Depending on the vendor, counties may elect to purchase additional election services to facilitate the planning and conduct of elections. Across most vendors, services for training, Election Day support, ballot production, and project management, are available for purchase.

Training. As part of equipment acquisition agreements, vendors offer courses for local election officials and poll workers on how to use the equipment. In general, these courses are in-person with hourly rates.

[11] Software was not included in the calculation of investment in voting equipment due to the disparities in licensing agreements across vendors. For instance, some vendors offer perpetual licenses, while others may only provide annual licenses. In addition, in the FOIA request made to counties for voting equipment contracts did not make an explicit request to obtain copies of contracts for software licenses. For many counties, this additional request would have complicated the data gathering task substantially.

Voter Outreach. Localities may purchase vendor created materials aimed at increasing the public's understanding of how to use voting equipment. Examples of materials are posters with instructions, social media campaigns, pocket guides, among others.

Project Management. When localities are unable to independently run elections in their jurisdiction, vendors offer services to help plan and conduct elections on behalf of jurisdictions.

Election Support. In order to make sure that equipment related issues are kept at a minimum and solved in a timely manner, jurisdictions may choose to purchase easily accessible mechanical and technological support directly from the vendor. Usually the time-frame of this type of support is Election Day.

Mode of counting and casting of ballots. There are five possible manners in which ballots may be cast and counted in the United States:

1. Direct-recording electronic voting machine (DRE)
2. Paper-based system using Optical Scanners
3. Both DRE and paper-based systems made available to voters (Mixed)
4. Hand-counted paper ballots
5. Vote-by-Mail

The model includes a sole dummy variable labeled "DRE Only" to indicate counties using DRE only systems. In this voting system configuration, both voters with disabilities and voters without disabilities use the same device. The choice to include only one indicator in the model is due to identification issues with state-level variables in voting equipment.[12]

Age of voting system. The age of a voting system is determined from the date on the first county contract for the acquisition of voting equipment of the model currently in use in 2016. For coding purposes, the start of a voting system's lifespan would commence following the first June (the start of the annual election cycle) the county was in possession of the equipment. For instance, if a

[12] The vendor Hart does not offer a paper-based system with ADA compliant ballot marking devices.

county purchased equipment in September of 2006, then the equipment was coded as nine years old in 2016. It should be noted that many counties made subsequent minor purchases to supplement and replace devices to meet state and federal guidelines.

Average number of machines per registered voter.The total number of voting machines is the sum of either all DRE machines or Ballot Marking Devices depending on the county's chosen mode of ADA compliance.[13] The total number of voting machines was provided by the U.S. Assistance Commission's 2016 Election Administration and Voting Survey. The number of machines is divided by the number of registered voters as a manner of standardization across counties.

Average number of machines per precinct. The total number of machines is then divided by the total number of precincts in each county. The building block of all electoral districts is the precinct. In every precinct, all voters receive the same ballot. Alternatively, all voters in a precinct vote for all the same offices. This number is an approximation of the number of machines the average voter will see in the polling place. The number of machines is divided by the number of precincts in each county.

Demographic variables. In addition to variables measuring elections, demographic controls are included in the model. County level demographic values was obtained from the 2016 American Community Survey 5-year estimates data compiled by the United States Census Bureau.

6 Results

6.1 Descriptive Data Analysis

Table 3 presents the state averages of county-level variables related to the acquisition and investment in voting equipment. The statistics which are unique to this study are the average equipment investment per registered voter and the average price per voting equipment unit. Both statistics illustrate the high cost of acquiring and maintaining voting equipment in the United States.

What is most notable about these results is the high cost of voting equipment. This statistic is

[13] HAVA requires all counties to have ADA compliant voting equipment available to voters with disabilities.

calculated by dividing the total investment in voting equipment by the number of units used in the 2016 election as reported in the EAVS 2016 survey.

[Insert Table 3 here]

6.2 Spatial Regression Analysis

The model in the first column of Table 4 presents the results of spatially autocorrelated error model with county investment in voting equipment per registered voter expressed as the raw value. In contrast, a within-between specification for investment in voting equipment per registered voter is used estimate both state-level and county-level effects in the second column. Despite the different specifications of county investment in voting equipment per registered voter, there is no statistical evidence in support of the central hypothesis.

[Insert Table 4 here]

On the other hand, the coefficient for “Age of System" is negative and statistically significant in both model specifications. As hypothesized (H2), older systems are associated with lower the turnout. The coefficient estimate of the variable “State Contract" is negative and statistically significant across both models. This suggests that the hyper-federalized state election systems promote turnout all else being equal.

Among the possible election services offered by vendors, the coefficient for voter outreach/education programs is positive statistically significant. Counties are able to increase turnout by a substantial amount by purchasing materials aimed at educating the public in the mechanics of properly casting a ballot. In the model with the standard specification, the coefficient for “Project Management" is negative and statistically significant.

6.3 Individual State Analyses

In this study's current iteration, there are two states which have been fully coded as well as contain a sufficient number of counties to run individual state-wide multivariate regression analysis.

Table 5 presents the results of spatially autocorrelated error regression models for Texas (n=245) and Nebraska (n=92). There is one important difference between the states' manner of acquiring voting equipment. Texas devolves the responsibility of purchasing voting equipment to the counties, while the Secretary of State of Nebraska as responsible for acquiring voting equipment. As a result, the models are specified slightly different. Due to the lack of variation across counties in Nebraska in services purchased, vendor, age of equipment, and type of equipment, those variables were omitted from the model. In contrast, Texas exhibits a fair bit of variability as counties had little uniformity requirements imposed by the state.

[Insert Table 5 here]

The results of both models is supportive of the central hypothesis of higher levels of investment in voting equipment is associated with higher levels of turnout. In the regression on turnout in Texas, the coefficient estimate for investment in election equipment per registered voter is positive and statistically significant. Similarly, the same coefficient estimate in the Nebraska model is weakly statistically significant (p-value = 0.0534). Although these results contradict the previous findings in Table 4, the "within" estimates for investment in equipment per registered voter is more precise in Table 5. On the other hand, we are less able to generalize these findings to other states.

7 Limitations

It should be noted that the results presented in the current iteration of this paper are preliminary. As of the time of this conference, I have gathered contracts from Alaska (complete), Arizona (incomplete), California (incomplete), Delaware (complete), Illinois (incomplete), Kentucky (incomplete), Maine (complete), Nebraska (complete), Nevada (complete), Pennsylvania (incomplete), Rhode Island (complete), Texas (complete), Utah (complete), Vermont (complete), and Virginia (incomplete), Wisconsin (incomplete).[14]

[14]The information provided in the contracts from Alaska and Maine was not sufficient for inclusion in this analysis. Contracts from Arizona, Kentucky, Pennsylvania, Virginia, and Wisconsin have not yet been coded.

This incomplete nature of the data is pertinent to the interpretation of the variable "State Contract". Given the labor intensive nature of obtaining county-level contracts, it is reasonable to gather state-wide contracts prior to embarking on the mission of contacting thousands of county election officials. As a result, the majority of states in this analysis have made state-wide contracts with vendors for the acquisition of voting equipment.

Although there is statistically significant coefficient estimates presented in this analysis, caution should be taken in generalizing these findings to the entire country. Without data from a more diverse sample of states, the statistical results for "State Contract" may, in fact, be capturing the difference overall state trends in turnout relative to Texas (contains the majority of counties in this analysis). Due to the collinearity between state fixed-effects and an indicator variable for state-level equipment contracts it is not possible to tease out the true independent effect of the choice of states to make contracts on behalf counties or to give counties that responsibility. Fortunately, this issue can be settled with more data and will be addressed in future iterations of this study.

8 Conclusion and Discussion

Methodologically, the results of this study suggest that modeling spatial data appropriately does have substantive implications on regression analysis. By including specifications for spatial autocorrelation in the error term the fit of the model was improved. Considering the recent deluge of publicly accessible *big data* produced by governmental entities, it is imperative for researchers to understand how to recognize and model spatial data. Political scientists studying local government, in particular, should be cognizant of the difficulties in dealing with spatial data. As there are roughly 3,000 counties in the United States, there exists the possibility of a complex scheme of spatial dependencies that must be taken into account in any county-level analysis.

Using county-level turnout data from Texas and Nebraska, the results presented in Table 5 suggest a positive relationship between spending on voting equipment and turnout. Substantively, these results are non-trivial. Counties that invest more resources into elections appear to have higher levels of turnout than those who invest less, all else being equal. It should be noted that these results suggest a causal relationship. Unfortunately, this results cannot confirm this conclusion. Analogous to the age old question of the *chicken or the egg*, this analysis cannot detect whether increasing resources invested in elections is the definitive causal mechanism explaining increases turnout.

The age of voting equipment also seems to influence turnout in the hypothesized direction. These results confirm worries of those who believe that voting equipment in the United States is woefully outdated. At the end of the day, voting machines are like any other piece of computerized technology. Unlike punch-card and lever voting systems which are mechanical in nature, DRE and optical scanning voting technology cannot last for decades-upon-decades. Despite the flaws of mechanical systems, they were long-lasting and did not require proprietary knowledge in their maintenance. If the future of voting technologies lies in computerized systems, it is prudent for localities to prepare to update these systems at regular and short intervals.

The empirical evidence in support county purchases of voter outreach efforts should be reassuring for local election officials. Voter education programs are a concrete and practical solution for localities in improving the turnout and the Election Day experience for their constituents. Voter education programs can include elements such as posters, pocket cards, social media, and in-person demonstrations on how to use the equipment. These results suggest that by providing voters with information on what they can expect at the polling place, voters are more able to quickly and confidently cast their ballots.[15] For voters who have impediments to voting, or uncertain of how to use computer technology more broadly, may find the information from these efforts useful in overcoming obstacles in casting a ballot.

In contrast, the results of the coefficient estimates for project management services from vendors is more troubling. These results could, however, be a proxy for localities that are least able to manage elections due to the size the election office. In many counties, the official responsible for running elections is also responsible for many other essential government tasks. As vendors are not in fact

[15]This study serves as the basis for my first empirical in my dissertation. The claim that voter education programs increase confidence voters' ability to cast their ballot is investigated in the second empirical chapter of my dissertation using individual level data from the 2018 SPAE and the 2018 CCES.

public employees, their objectives may not be as in-sync with the interests of the public we would hope.

One of the secondary conclusions that can be inferred from the results of this study is that hyper-federalized election systems are beneficial to maintaining a healthy American democracy. The evidence presented supports the notion that counties are more in-tune with the needs of their voters. Local election officials are more likely to understand the traffic patterns of individual polling places and the types of machines that will best accommodate their constituency. When counties take over the reins of acquiring voting equipment, they may in fact be able to directly increase turnout.

References

Anselin, L., and Sergio Joseph Rey. 2014. *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*. Chicago, IL: GeoDa Press LLC.

Anselin, Luc. 2009. “Spatial Regression”. In *The SAGE Handbook of Spatial Analysis*, 255–275. London: SAGE Publications, Ltd.

Anselin, Luc, et al. 1996. “Simple diagnostic tests for spatial dependence”. *Regional Science and Urban Economics* 26 (1996): 77–104.

Burden, Barry C., and Jacob R. Neiheisel. 2013. “Election Administration and the Pure Effect of Voter Registration on Turnout”. *Political Research Quarterly* 66 (1): 77–90.

Cho, Sung-Jin. 2011. “An empirical model of mainframe computer investment”. *Journal of Applied Econometrics* 26 (1): 122–150.

Cliff, Andrew David, and J. K. Ord. 1981. *Spatial processes: models & applications*. London: Pion.

Dieleman, Joseph L., and Tara Templin. 2014. “Random-Effects, Fixed-Effects and the within-between Specification for Clustered Data in Observational Health Studies: A Simulation Study”. Ed. by Andrew R. Dalby. *PLoS ONE* 9 (10).

Engle, Robert F. 1980. “Hypothesis Testing in Spectral Regression; the Lagrange Multiplier Test as a Regression Diagnostic”. In *Evaluation of Econometric Models*, 309–321. Academic Press.

Ewald, Alec C. 2009. *The Way We Vote: The Local Dimension of American Suffrage*. Nashville, TN: Vanderbilt University Press.

Griffith, Daniel A. 1988. *Advanced spatial statistics: special topics in the exploration of quantitative spatial data series*. Boston, MA: Kluwer Academic Publishers.

— . 1987. *Spatial autocorrelation: a primer*. Washington, D.C: Association of American Geographers.

Kimball, David C., et al. 2013. “Policy Views of Partisan Election Officials”. *The UC Irvine Law Review* 3 (3): 551–574.

Montjoy, Robert S. 2010. "The Changing Nature… and Costs… of Election Administration". *Public Administration Review* 70 (6): 867–875.

Moran, Patrick A.P. 1950. "Notes on Continuous Stochastic Phenomena". *Biometrika* 37 (1/2): 17–23.

— . 1948. "The Interpretation of Statistical Maps". *Journal of the Royal Statistical Society. Series B (Methodological)* 10 (2): 243–251.

Odland, John. 1988. *Spatial autocorrelation*. Newbury Park, CA: Sage Publications.

The Help America Vote Act and Election Administration. 2015. Congressional Research Service.

Figure 1
Boundary Map of Austin, TX

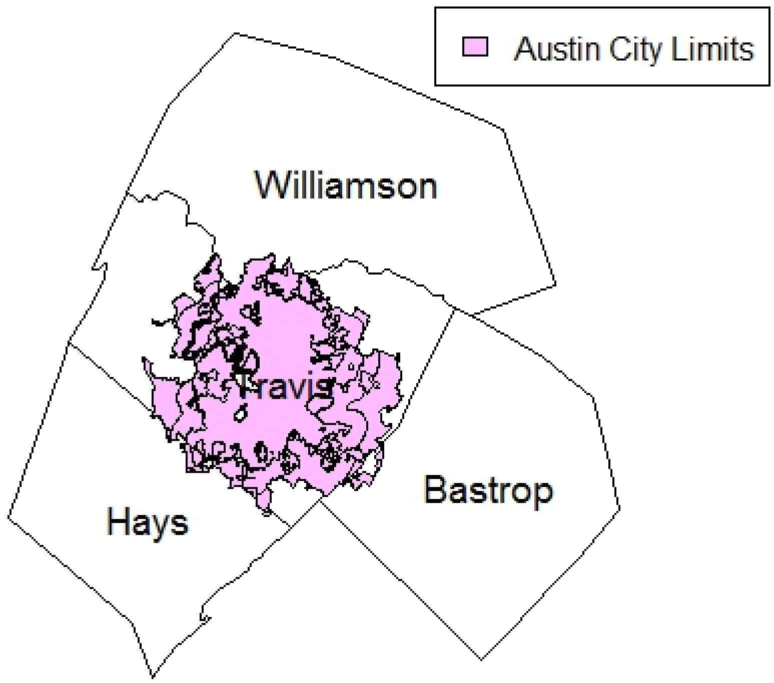

Table 1
Moran's I statistics

	Moran's I Statistic	$E(I)$	Deviation from $E(I)$	P-value
Parametric approach:				
Dependent variable	0.147	-0.002	0.149	0.000***
Error term	0.052	-0.016	0.068	0.025*
Monte Carlo approach:				
Dependent variable	0.147	-	-	0.001***

Table 2
Lagrange Multiplier Test for Spatial Dependence

Model	**Lagrange Multiplier Test**
Autoregressive Lag	1.380
Autoregressive Error	2.917.

Table 3
State Averages of Voting Equipment Variables

State	Age of Equip.	Machines Per Reg. Voter	Machines Per Precinct	Equip. Investment Per Reg. Voter	Price of 1 Unit
CA	9.063	0.001	0.732	$15.33	$17,383.24
DE	10.000	0.002	3.138	$ 4.68	$ 2,254.72
IL	11.176	0.001	1.031	$ 4.35	$ 5,364.19
NE	11.000	0.002	0.797	$17.65	$ 9,498.45
NM	2.000	0.002	0.953	$15.30	$ 4,995.99
NV	12.000	0.007	3.184	$17.16	$ 2,523.95
RI	10.000	0.001	1.000	$ 4.86	$ 3,984.08
TX	10.037	0.005	2.912	$19.79	$ 5,378.43
UT	10.000	0.006	2.748	$26.26	$ 4,526.49
VT	14.286	0.004	2.143	$ 3.01	$ 6,355.73
Overall	9.956	0.004	2.143	$17.84	$6781.65

Table 4

Spatially Autocorrelated Error Regression Model for County-Level Turnout in 10 States in 2016

	Standard		Within-Between	
(Intercept)	88.305	***	96.682	**
	(33.950)		(35.384)	
Election Administration:				
DRE Only	-0.661		0.189	
	(5.145)		(5.240)	
Machines/ Registered Voters	-63.395		-66.847	
	(223.900)		(223.630)	
Machines/ Precinct	-0.420		-0.372	
	(1.260)		(1.261)	
Investment in Equip./ Reg. Voters	0.043		–	
	(0.086)			
State Average Investment in Equip.	–		-0.379	
			(0.515)	
Deviation in Investment in Equip.	–		0.061	
			(0.089)	
Age of System	-2.621	**	-2.640	**
	(0.862)		(0.864)	
State Contract	-18.713	***	-19.310	***
	(5.672)		(5.731)	
Vendor Services:				
Training	5.320		4.316	
	(5.943)		(6.044)	
Voter Outreach	11.803	*	12.723	*
	(5.309)		(5.447)	
Election Day Support	5.947		6.197	
	(4.843)		(4.850)	
Project Management	-9.567	*	-8.230	
	(4.566)		(4.828)	
Demographics:				
Median Age	-0.305		-0.360	
	(0.371)		(0.377)	
Per Capita Income	0.000		0.000	
	(0.000)		(0.000)	
% College	0.000		0.000	
	(0.000)		(0.000)	
% White	0.161		0.184	
	(0.324)		(0.326)	
% Black	-0.629		-0.615	
	(0.489)		(0.491)	
% Hispanic	-0.172		-0.145	
	(0.321)		(0.323)	
Total Population	0.000		0.000	
	(0.000)		(0.000)	
Lambda	0.214	*	0.221	*

Table 5
Spatially Autocorrelated Error Regression Model for Texas and Nebraska County-Level Turnout in 2016

	Texas		Nebraska	
(Intercept)	-9.581		7.766	
	(15.288)		(18.170)	
Election Administration:				
DRE Only	-0.135		–	
	(0.704)			
Machines/ Registered Voters	-17.151		328.150	
	(27.341)		(232.250)	
Machines/ Precinct	-0.039		0.198	
	(0.173)		(3.845)	
Investment in Equip./ Reg. Voters	0.035	**	0.349	.
	(0.011)		(0.181)	
Age of System	0.537		–	
	(0.373)			
Vendor Services:				
Training	0.500		–	
	(1.018)			
Voter Outreach	-1.117		–	
	(0.792)			
Election Day Support	-0.009		–	
	(0.642)			
Project Management	0.019		–	
	(0.787)			
Demographics:				
Median Age	0.107		-0.106	
	(0.070)		(0.273)	
Per Capita Income	0.001	***	0.000	
	(0.000)		(0.000)	
% College	0.000		0.002	
	(0.000)		(0.002)	
% White	0.516	***	0.369	.
	(0.156)		(0.190)	
% Black	0.387		-0.098	
	(0.169)		(1.131)	
% Hispanic	0.348		0.096	
	(0.153)		(0.250)	
Total Population	0.000		0.000	
	(0.000)		(0.000)	
Lambda	0.345	***	0.200	
	N=245		N=92	