

Electronic Companion Supplementary Information



The supporting material in the electronic companion is divided into the following sections: Figures (EC.1); Tables (EC.2); Description of Data Sources (EC.3); Multiple Imputation Methods (EC.4); Instrument Justifications (EC.5); Extension of Famighetti et al. (2014) (EC.6); and Average Waiting Time Simulation Methods (EC.7).

EC.1. Figures



Figure EC.1 Percent of non-white citizens who registered to vote by state. Data taken from U.S. Census Bureau reports on the voting and registration for states by race (see <https://www.census.gov/topics/public-sector/voting/data/tables.html>).

Voting Equipment Usage

-  Paper ballots only
-  DREs and paper ballots

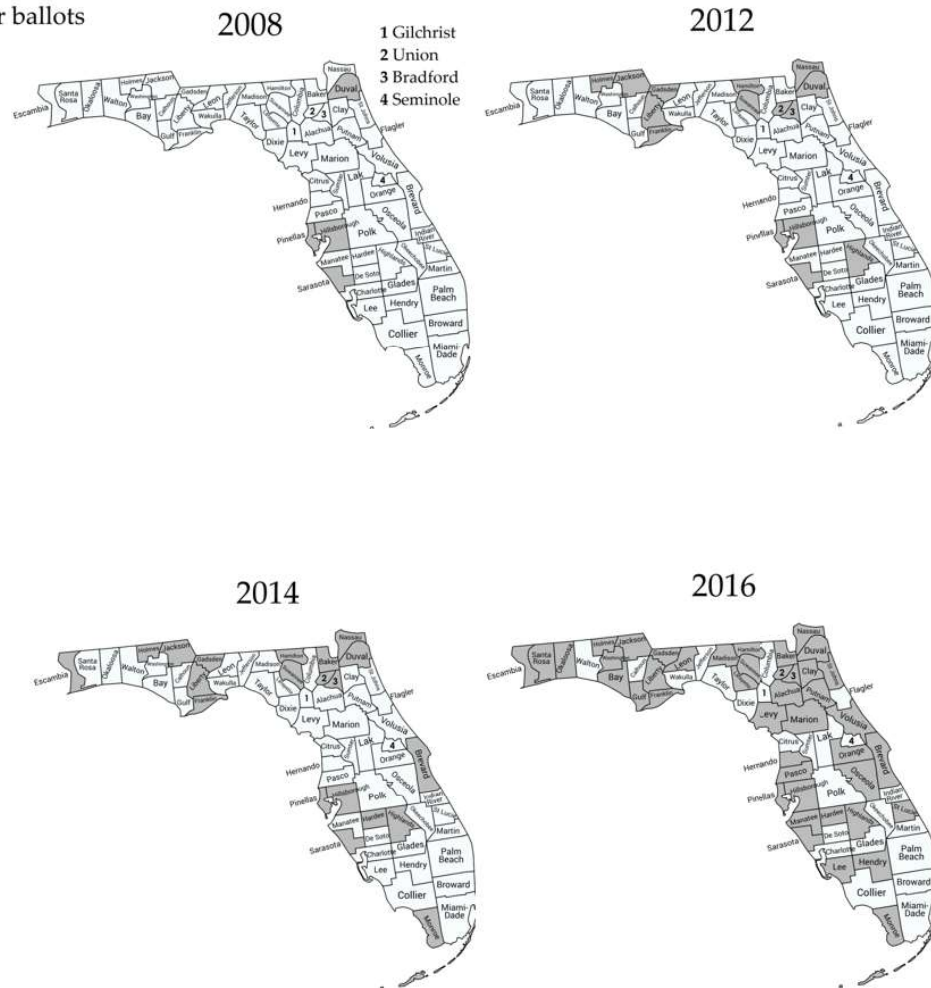


Figure EC.2 Change in Florida's voting equipment usage over time courtesy of the Verified Voting Foundation (figure created courtesy of <https://mapchart.net/>)

EC.2. Tables**Table EC.1 Summary statistics for key variables across Florida's 67 counties and the 2008 to 2016 general elections, including the number of poll workers per polling place (PWPerPP)**

Statistic	Voters Per PW*	Pct Democrat	Pct White	AbsenteeBallots*	Early Ballots*	Elect-ionDay Ballots*	Provi-sional Ballots*	Poll Diff†	Person PerSq Mile*	House Price	Med Inc*	Use DRE†	Pct 65Plus	PW Per PP
Mean	235.64	44.46	78.51	320.24	406.69	540.69	3.64	3.36	346.18	142.23	44243	0.73	18.51	8.19
Std. Dev.	93.11	14.60	14.09	229.48	295.14	223.86	4.02	1.06	531.99	23.53	7648	0.45	6.84	2.74
Minimum	79.70	19.42	18.37	30.88	16.00	142.69	0.00	1	9.76	97.55	29642	0	7.90	3.09
Maximum	581.74	88.38	95.80	1251.83	1591.49	1162.39	32.09	5	3486.37	214.61	69523	1	53.10	20.39
25th%ile	168.58	33.20	72.42	148.80	187.94	342.29	1.00	2.7	46.12	124.44	37778	0	13.80	6.00
50th%ile	220.01	42.28	82.84	257.12	315.44	543.90	2.22	3	166.57	137.22	43787	1	17.10	7.64
75th%ile	280.40	54.22	87.29	428.34	549.50	707.63	5.24	4	413.12	159.29	49135	1	21.70	9.74

*Non-logged values shown

†Statistics calculated by averaging all imputations for each variable. See Section EC.4.

Table EC.2 Main regression specification checks

Model Specification Checks (Note: $\Delta\hat{\varepsilon}_{i,t} = \Delta\varepsilon_{i,t} - \Delta\varepsilon_{i,t-1}$)	1-Step	2-Step
Number of instruments:		61
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\varepsilon_{i,t-1}$ Max p-value among imputations:	0.001	0.027
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\varepsilon_{i,t-2}$ Min p-value among imputations:	0.372	0.722
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\varepsilon_{i,t-3}$ Min p-value among imputations:	0.748	0.636
Hansen test of overidentifying restrictions Min p-value among imputations:		0.168

Table EC.3 F-statistics resulting from the regression of the first differenced covariate (pooled across county and election years) on only the lagged instrument candidate(s) for all endogenous covariates.

General Election Years: 2008 to 2016

First differenced covariate	Lags used as Instruments	F-Statistic
PctDemocrat _t	t-2	59.92
PctWhite _t	t-2	52.00
Voters/PollWkr _{t-1}	t-2/t-3	41.72
Voters/PollWkr _{t-2}	t-2/t-3	33.77*
AbsenteeBallots _{t-1}	t-2/t-3	576.83
AbsenteeBallots _{t-2}	t-2/t-3	42.17*
EarlyBallots _{t-1}	t-2/t-3	681.21
EarlyBallots _{t-2}	t-2/t-3	104.35*
ElectionDayBallots _{t-1}	t-2/t-3	56.77
ElectionDayBallots _{t-2}	t-2/t-3	19.94*
ProvisionalBallots _{t-1}	t-2/t-3	196.88
ProvisionalBallots _{t-2}	t-2/t-3	96.33*
PollDiff [†] _t	t-1	73.53

*The first difference of the second lag of the variable is fully identified by the two lags used as instruments so we report the minimum F-statistic resulting from the first-stage, pooled county regression of the first difference of the second lag on each individual lagged instrument candidate.

†We conducted the F-test for all 10 imputed cases for PollDiff and report the minimum F-statistic.

Table EC.4 Robustness check results for coefficients of interest with descriptions below

General Election Years: 2008 to 2016 — Counties: 67 — First Differences: 268

DV: Voters/PollWkr	(1)AltProxy [†]	(2)NoLag [†]	(3)PctNoParty [†]	(4)Shelby [†]	(5)Neighbors [†]
PctDemocrat	0.029** (0.012)	0.032** (0.013)	0.028*** (0.010)	0.033** (0.014)	0.038*** (0.012)
PctWhite	0.024 (0.035)	0.056 (0.037)	0.039 (0.032)	0.055 (0.042)	0.034 (0.039)
AbsenteeBallots _t	0.356** (0.159)				
EarlyBallots _t	-0.065 (0.109)				
ElectionDayBallots _t	0.394*** (0.133)				
ProvisionalBallots _t	-0.048* (0.028)				
PctNoParty			0.086** (0.034)		
ShelbyCounties				-0.003 (0.076)	
NeighborAbsentee _{t-1}					0.303 (0.212)
NeighborAbsentee _{t-2}					0.022 (0.144)
NeighborEarly _{t-1}					0.013 (0.173)
NeighborEarly _{t-2}					-0.156* (0.083)
NeighborElectionDay _{t-1}					0.160 (0.187)
NeighborElectionDay _{t-2}					0.235 (0.154)
NeighborProvisional _{t-1}					0.017 (0.033)
NeighborProvisional _{t-2}					-0.008 (0.031)
Other Variables in Regression:					
Voters/PollWkr Lags	t-1/t-2	No	t-1/t-2	t-1/t-2	t-1/t-2
Absentee-, Early-, ElectionDay-, Provisional- Ballots Lags	No	t-1/t-2	t-1/t-2	t-1/t-2	No
PollDiff, PersonPerSqMile, HousePrice, MedInc, UseDRE, Pct65Plus	Yes	Yes	Yes	Yes	Yes
Election Year Fixed Effects	Yes	Yes	No	Yes	Yes
Presidential, Time Trend	No	No	Yes	No	No

*p<0.1, **p<0.05, ***p<0.01

[†] Windmeijer (2005) corrected standard errors reported in parentheses

Robustness Check Descriptions

(1) AltProxy: We replace the lagged proxies for forecasted voter demand with contemporaneous proxies. These contemporaneous variables represent clairvoyant forecasts that approximate the true, unobserved forecasted voter demand.

(2) NoLag: We remove both lags of the dependent variable.

(3) PctNoParty: To fully account for variation in political party composition, we include the control *PctNoParty*, the percentage of active registered voters who had no political party affiliation. To prevent the number of instruments from exceeding the panel size (Roodman (2009)), we replace the election-year fixed effects with *Presidential*, a dummy variable indicating whether the election year was presidential, and a linear time trend.

(4) Shelby: We include an exogenous variable, *ShelbyCounties* which equals 1 in the 2014 and 2016 elections and zero otherwise for those counties that were effected by the 2013 Supreme Court decision *Shelby County v. Holder*.

(5) Neighbors: Instead of creating proxies for the forecasted turnout based on lags within each county i , we create new proxies based on the turnout of neighboring counties that share a geographic boundary with county i in election year $t - j$. NeighborAbsentee/NeighborEarly/NeighborElectionDay/NeighborProvisional is the log of the average of the number of absentee/early/Election Day/provisional ballots cast per polling place across all counties neighboring county i in election year $t - j$.

Table EC.5 Robustness specification checks

Model Specification Checks (Note: $\Delta\hat{\varepsilon}_{i,t} = \Delta\varepsilon_{i,t} - \Delta\varepsilon_{i,t-1}$)	(1)AltProxy	(2)NoLag	(3)PctNoParty	(4)Shelby	(5)Neighbors
Number of instruments:	61	53	67	62	61
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\hat{\varepsilon}_{i,t-1}$ Max p-value among imputations:	0.011	0.010	0.023	0.026	0.032
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\hat{\varepsilon}_{i,t-2}$ Min p-value among imputations:	0.565	0.891	0.930	0.726	0.494
H_0 : No correlation between $\Delta\hat{\varepsilon}_{i,t}$ and $\Delta\hat{\varepsilon}_{i,t-3}$ Min p-value among imputations:	0.873	0.407	0.924	0.629	0.613
Hansen test of overidentifying restrictions Min p-value among imputations:	0.207	0.193	0.315	0.158	0.247

Table EC.6 F-statistics resulting from the regression of the first differenced covariate (pooled across county and election years) on only the lagged instrument candidate(s) for robustness checks

General Election Years: 2008 to 2016

First differenced covariate	Lags used as Instruments	F-Statistic
AltProxy Robustness Check		
AbsenteeBallots _t	t-2/t-3	371.49
EarlyBallots _t	t-2/t-3	538.32
ElectionDayBallotst	t-2/t-3	23.15
ProvisionalBallots _t	t-2/t-3	59.80
PctNoParty Robustness Check		
PctNoParty	t-2/t-3	13.22
Neighbors Robustness Check		
NeighborAbsentee _{t-1}	t-2/t-3	2702.50
NeighborAbsentee _{t-2}	t-2/t-3	94.67*
NeighborEarly _{t-1}	t-2/t-3	1655.11
NeighborEarly _{t-2}	t-2/t-3	251.97*
NeighborElectionDay _{t-1}	t-2/t-3	61.07
NeighborElectionDay _{t-2}	t-2/t-3	32.50*
NeighborProvisional _{t-1}	t-2/t-3	489.12
NeighborProvisional _{t-2}	t-2/t-3	214.00*

*The first difference of the second lag of the variable is fully identified by the two lags used as instruments so we report the minimum F-statistic resulting from the pooled county regression of the first difference of the second lag on each individual lagged instrument candidate.

Table EC.7 Base simulation inputs

Simulation parameter	Selection/Value
Polling hours (from Florida's Division of Elections website)	7am to 7pm
Arrival pattern	mid-morning peak
Voting technology	paper ballot + optical scanner
Time to check in (minutes) (Stewart III (2015))	2
Time to complete a ballot (minutes) (Stewart III (2015))	5
Time to scan ballot (minutes) (Stewart III (2015))	0.5
% voters unable to check-in (observed % of provisional ballots)	1%
% of voters arriving before polls open (default for simulation)	1.20%
Same day voter registration	no
Expected number of voters	660
Number of check-in stations	2
Number of voting stations	4

Table EC.8 Number of active registered voters per poll worker for Florida and South Carolina across the 2008 to 2016 elections

Voters/PollWkr*	2008	2010	2012	2014	2016
Florida	162.54	228.97	229.13	284.75	276.74
South Carolina†	171.43	196.71	217.60	213.60	214.33

*Non-logged values shown

†Based on EAVS questions A3a and D3 or equivalent

EC.3. Data Sources

Voters/PollWkr

Florida Division of Elections provides information on the number of active registered voters by county for each election (see <https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/bookclosing/>). Election Administration Voting Survey provides county-level information on the number of poll workers by county from question D3 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

PctWhite/PctDemocrat/PctNoParty/PctRepublican

Florida Division of Elections provides information on the number of active registered voters by county for each election (see <https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/bookclosing/>). Florida Division of Elections provides information on the racial and political party composition of active registered voters by county for each election (see <https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/bookclosing/>).

AbsenteeBallots

Election Administration Voting Survey provides county-level information on the number of Uniformed and Overseas Citizens Absentee Voters Act (UOCAVA) voters who voted absentee, Federal Write-in Absentee Ballots (FWAB), and domestic civilian absentee voters from question F1 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>). Election Administration Voting Survey provides county-level information on the number of polling places in an election from question D2 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

EarlyBallots

Election Administration Voting Survey provides county-level information on the number of ballots cast at an early vote center from question F1 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>). Election Administration Voting Survey provides county-level information on the number of polling places in an election from question D2 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

ElectionDayBallots

Election Administration Voting Survey provides county-level information on the number of ballots cast on Election Day from question F1 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>). Election Administration Voting Survey provides county-level information on the number of polling places in an election from question D2 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

ProvisionalBallots

Election Administration Voting Survey provides county-level information on the number of provisional ballots cast in an election from question E1 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>). Election Administration Voting Survey provides county-level information on the number of polling places in an election from question D2 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

PollDiff

Election Administration Voting Survey provides county-level information on poll worker recruitment difficulty in an election from question D5 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

HousePrice

Federal Reserve Bank of St. Louis collects county-level house price data through its “All-Transactions House Price Index” in the FRED database (see <https://fred.stlouisfed.org/>). The All-Transactions House Price Index was only available for 56 out of 67 counties. For the 11 counties missing data, we imputed the index based on the average of the housing indexes from adjacent counties.

MedInc

U.S. Census Bureau collects county-level information on household income in the past 12 months and publishes its 5-year estimates in report ID S1901 (see <https://www.census.gov/programs-surveys/acs/>). The U.S. Census only reports household median income for all counties in Florida from 2009 and onward. We use the 2009 median incomes as a proxy for 2008.

PersonPerSqMile

Florida Office of Economic and Demographic Research publishes county-level information on population estimates used for calculating revenue sharing (see <http://www.edr.state.fl.us/Content/population-demographics/data/index-floridaproducts.cfm>). U.S. Census Bureau

published information on the total square mileage in a county in 2010 (see <https://www.census.gov/quickfacts/fact/note/US/LND110210>).

UseDRE

Verified Voting Foundation publishes polling place equipment information by county (see <https://www.verifiedvotingfoundation.org/about-vvf/>). See below for how we impute missing data for the 2010 election.

Pct65Plus

U.S. Census Bureau collects county-level information on age and publishes its 5-year estimates in report ID S0101 (see <https://www.census.gov/programs-surveys/acs/>). The U.S. Census only reports age statistics for all counties in Florida from 2009 and onward. We use the 2009 age metrics as a proxy for 2008.

PWPerPP (for Table EC.1)

Election Administration Voting Survey provides county-level information on the number of poll workers by county from question D3 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>). Election Administration Voting Survey provides county-level information on the number of polling places in an election from question D2 or equivalent (see <https://www.eac.gov/research-and-data/election-administration-voting-survey/>).

EC.4. Multiple Imputation Methodology

Predicting *PollDiff*

We use the following as predictors of *PollDiff* in counties: total active registered voters, male-to-female ratio, percentage of the population with a Bachelor’s degree or higher, median age, median income, unemployment rate, percentage Democrats, urban sprawl (people per square mile), political activeness (Election Day turnout percent) (Kimball et al. (2009), Burden and Milyo (2015)). We also use past values of *PollDiff* within a county as a predictor. Gender and age data is from the U.S. Census Bureau report ID S0101 5-year estimates (see <https://www.census.gov/programs-surveys/acs/>). Educational attainment data is from the U.S. Census Bureau report ID S1501 5-year estimates (see <https://www.census.gov/programs-surveys/acs/>). Unemployment data is from the U.S. Census Bureau report ID S2301 5-year estimates for the population 16 years and over (see <https://www.census.gov/programs-surveys/acs/>). Election Day turnout data is calculated from the number of ballots cast on Election Day divided by the total number of active registered voters.

Predicting *UseDRE*

If a county used (or did not use) DREs during both the 2008 and 2012 elections, then we assume they would have used (or not used) DREs during the 2010 election. There are 56 counties for which we infer their 2010 usage in that manner. For the remaining 11 counties, we perform multiple imputation in which we randomly assigned each county to either use DREs or not use DREs across 10 imputations.

Imputing Results

We impute the regression coefficients and standard errors per Rubin (2004) with adjusted degrees of freedom suggested in Barnard and Rubin (1999) and 10 imputations based on guidance from White et al. (2011).

EC.5. Instrument Choices for the First Differences

In this section, we explain the logic for the lags we chose as instruments for the first differences and qualitatively justify both exogeneity and relevance. Quantitative checks for relevance can be found in Table EC.3. Recall that we also aimed to be parsimonious in the number of lags we chose as instruments due to the findings in Bowsher (2002).

First Differences for *PctWhite* and *PctDemocrat*

Minimum time period lag needed for exogeneity?

t-2: The percentage of white and Democrat registered voters at least two elections prior should be unrelated to voter or poll worker shocks in the current election or the last election.

Lags chosen as relevant instruments?

t-2: As shown in Table 1, the change in *PctWhite* and *PctDemocrat* across subsequent elections is not as large as it is for the forecasted voting demand variables and also does not appear to depend on whether the election goes from presidential to midterm or midterm to presidential. Thus, we assumed the most recent election lag qualifying as an instrument (regardless of whether it is a presidential or midterm election) would suffice as a relevant instrument for each first difference. Note that we confirm relevance of all our instrument choices described in this section in Table EC.3.

First Differences for *AbsenteeBallots*, *EarlyBallots*, *ElectionDayBallots*, and *ProvisionalBallots* Lags

Minimum time period lag needed for exogeneity?

t-2: Ballots cast per polling place two elections prior should be unrelated to voter or poll worker shocks in the current election or the last election.

Lags chosen as relevant instruments?

t-2/t-3: As shown in Table 1, the change in ballots cast per polling place across subsequent elections can be large, and the direction of the change depends highly on whether the election goes from presidential to midterm or midterm to presidential. Thus, we assumed the two most recent election lags qualifying as instruments (which include both a presidential and midterm election) would be needed to create relevant instruments for each first difference.

First Differences for *Voters/PollWkr* Lags

Minimum time period lag needed for exogeneity?

t-2: See Arellano and Bond (1991) and StataCorp (2013).

Lags chosen as relevant instruments?

t-2/t-3: As shown in Table 1, the change in voters per poll worker across subsequent elections can be large, and the direction of the change depends highly on whether the election goes from

presidential to midterm or midterm to presidential. Thus, we assumed the two most recent election lags qualifying as instruments (which include both a presidential and midterm election) would be needed to create relevant instruments for each first difference.

First Differences for *PollDiff*

Minimum time period lag needed for exogeneity?

t-1: Poll worker recruitment difficulty in the last election should be unrelated to voter or poll worker shocks in the current election or the last election since recruitment occurs before the election and poll workers need to go through training prior to the election.

Lags chosen as relevant instruments?

t-1: We assumed poll worker recruitment difficulty in an election to be similar to poll worker recruitment difficulty in the last election since Burden and Milyo (2015) suggest that poll workers tend to be motivated by civic duty. We also find evidence for this assumption in Table 1 where we see that the change in *PollDiff* is not as large as it is for the forecasted voting demand variables and also does not appear to depend on whether the election goes from presidential to midterm or midterm to presidential. Thus, we assumed the most recent election lag qualifying as an instrument (regardless of whether it is a presidential or midterm election) would suffice as a relevant instrument for each first difference.

EC.6. Famighetti et al. (2014) Poll Worker Result Calculation with Percentage of White Voters

Famighetti et al. (2014) do not explicitly examine the relationship between the percentage of white registered voters and the number of Election Day eligible voters per poll worker in their study (i.e., they look at black, Latino, and other minority registered voters). However, using their data and regression methodology, we were able to examine this relationship by regressing the number of Election Day eligible voters per poll worker on the percentage of white registered voters in each precinct and included county fixed effects. To convert this regression estimate into a percentage (0.26%, as reported in our results), we divided the regression estimate by the mean number of Election Day eligible voters per poll worker in the data.

EC.7. Average Waiting Time Simulation

We use three different methods to measure the impact of a change in the percentage of Democrats on the average wait time. The first adjusts the average number of check-in stations in response to changes in *PctDemocrat*, leaving all other parameters constant. Let Q be the initial number of check-in stations. Assuming that the average change in check-in stations per polling place is proportional to the average change in voters per poll worker, the average number of check-in stations after the change is

$$Q/e^{\hat{\beta}_1 P}$$

where $\hat{\beta}_1$ is our estimate for the change in voters per poll worker ($\hat{\beta}_1 = 0.028$), and P is the percentage change in Democrats in a county. To overcome the integer constraint on the number of voting resources (e.g., it is not possible to reduce the number of check-in stations across all polling locations by 0.1, but it is possible to reduce the number of check-in-stations in 10% of polling locations), we presume that the fraction ρ of polling places has $Q - \delta$ check-in stations and the remaining fraction, $1 - \rho$, has $Q - \delta + 1$ stations, where

$$\delta = Q - \left\lfloor \frac{Q}{e^{\hat{\beta}_1 P}} \right\rfloor$$

and

$$\rho = Q \left(1 - \frac{1}{e^{\hat{\beta}_1 P}} \right) + 1 - \delta$$

For most changes $\delta = 1$, but it is possible that a larger change in voting resources is needed. With our first measure we let W_1 be the new average waiting time due to a P percent change in Democrats

$$W_1 = \rho \Omega(660, 2 - \delta, 4) + (1 - \rho) \Omega(660, 2 - \delta + 1, 4)$$

where $\Omega(v, c, s)$ is the simulated average wait time for a polling place with v expected voters, c check-in stations, and s voting stations.

The second method is analogous to the first except now the number of voting stations is reduced:

$$W_2 = \rho \Omega(660, 2, 4 - \delta) + (1 - \rho) \Omega(660, 2, 4 - \delta + 1)$$

The third method increases the number of voters per polling place holding voting resources constant:

$$W_3 = \Omega\left(660e^{\hat{\beta}_1 P}, 2, 4\right)$$

References

- Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58(2):277–297.
- Barnard J, Rubin D (1999) Small-sample degrees of freedom with multiple imputation. *Biometrika* 86(4):948–955.
- Burden B, Milyo J (2015) The quantities and qualities of poll workers. *Election Law Journal* 14(1):38–46.
- Famighetti C, Melillo A, Pérez M (2014) Election day long lines: Resource allocation. *Brennan Center for Justice* at New York University, <https://www.brennancenter.org/sites/default/files/publications/ElectionDayLongLines-ResourceAllocation.pdf>.
- Kimball D, Baybeck B, Gross C, Wiedlocher L (2009) Poll workers and election administration: The view from local election officials. Unpublished doctoral dissertation, University of Missouri, St. Louis.
- Roodman D (2009) How to do xtabond2: An introduction to difference and system gmm in stata. *The Stata Journal* 9(1):86–136.
- Rubin D (2004) *Multiple imputation for nonresponse in surveys*, John Wiley & Sons, Hoboken, NJ.
- StataCorp L (2013) Stata linear dynamic panel-data estimation. *College Station, TX: StataCorp LP* .
- Stewart III C (2015) Managing polling place resources. *Caltech/MIT Voting Technology Project*, <http://web.mit.edu/vtp/Managing%20Polling%20Place%20Resources.pdf>.
- White I, Royston P, Wood A (2011) Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* 30(4):377–399.
- Windmeijer F (2005) A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126(1):25–51.