Revisiting Roll-Off in Alerted Optical Scan Precincts: Evidence From Illinois General Elections.

Michael G. Miller
Barnard College, Columbia University
Department of Political Science
mgmiller@barnard.edu

Michelle D. Tuma
Fordham University
School of Law
mtuma@fordham.edu

Logan Woods
University of Michigan
Department of Political Science
ltwoods@umich.edu

Abstract

In November of 2007, Illinois passed a law requiring optical scan voting machines to provide voters with an audible warning if they failed to register a vote in elections to any one of five statewide offices. The policy took effect in 2010, but 12 counties failed to upgrade their equipment in time to implement the policy. Miller (2013) exploited this opportunity to determine whether the alert reduced undervoting in the precincts where it was implemented, and found no significant effects in primary elections. We extend the analysis to precinct-level returns from the 2010 Illinois General Election, and find little evidence for reduced aggregate under-voting resulting from the audible undervote alert. We do, however, find some evidence that the alert was more effective for races that appeared on the lower portion of the ballot.

For the first time ever during the 2010 Illinois General Election, optical scanning machines notified voters when they cast an incomplete ballot. Under the provisions of a 2007 law, upon inserting their ballot into the scanner, Illinois voters heard an audible alert and saw their ballot “kick back” to them if it failed to register a vote in elections for any one of the five statewide, “constitutional offices” of Governor/Lieutenant Governor, Attorney General, Secretary of State, Treasurer, or Comptroller. This alert was often accompanied by a verbal warning from an election judge. The policy was intended to curb erroneous under-voting in the five elections for which it was applicable. The spirit of the law was therefore in-line with Section 301a of the Help America Vote Act, which sought to reduce erroneous “residual” votes, and which had mandated alerts for over-voted ballots nationwide.

Yet, in contrast to over-votes, the vast majority of which are resultant of voter error, it is unclear whether a law designed to reduce rates of *under*-voting should be effective, as in any given election a (frequently large) percentage of voters choose not to express a preference. Indeed, Miller’s (2013) analysis of the Illinois General Primary Election found practically no effect of the alert on undervote rates in that state. However, it is at least plausible that—as Miller concedes—the undervote alert might be more effective in a general election as opposed to a primary. One might reasonably assume that those who vote in primary elections are more dedicated and informed than the average voter who participates in a general election. In other words, the composition of the primary electorate might itself be conducive to very low rates of unintentional under-voting, leaving few instances for the alert to do its job. Unintentional under-voting may occur more frequently among the less committed and informed voters participating in a general election, however. If so, then even though the policy did not reduce under-voting in the 2010 Illinois General Primary Election, it may have done so in the 2010 Illinois General Election. In this paper, we investigate that possibility.

Undervote Alerts and Determinants of Roll-Off

An omnibus election bill passed by the Illinois General Assembly in 2007 mandated that beginning in 2010, "No Precinct Tabulation Optical Scan Technology voting system shall be approved unless it…will identify when a voter has not voted for all statewide constitutional offices" (10 ILCS 5/24B-16(e-5)).[1] In practice, this alert is delivered via an audible "beep" when the ballot is presented to the optical scanner, which occurs after the voter has left the voting booth but before he or she exits the polling place. Upon issuing the audible beep, the machine "kicks back" the ballot to an election judge, who examines the ballot and informs the voter of the reason for the alert. The voter is then afforded the opportunity to return to a voting booth and correct the omission if desired. The policy made Illinois the first state to mandate that voters receive a warning from an election judge about *under*-voting prior to leaving an election place, and was ostensibly intended to curb unintentional under-voting (see: Miller 2013).

Yet, if the alert policy is to diminish the undervote rate, one must accept the premise that a relatively large proportion of undervotes are errant, and it is not clear whether such expectations are well-founded. For instance, Knack and Kropf (2003) report evidence that a significant percentage of undervotes in presidential elections is intentional, while experimental evaluations have determined the erroneous undervote rate to be as low as 0.2% of ballots cast (Herrnson et al. 2008). That said, Tomz and Van Houweling (2003) find evidence that some undervoting (or "roll-off") is in fact unintentional, resulting from confusion and inexperience among voters. Indeed, changes to such factors as ballot design, equipment, or election administration can potentially reduce undervote rates (e.g., Herrnson, Hanmer, and Niemi 2012;

[1] The policy, including legislative background and subsequent reaction, is described at length in Miller (2013).

Stein et al. 2008; Ansolabehere and Stewart 2005; Kimball and Kropf 2005; Shocket, Heighberger, and Brown 1992; Darcy and Schneider 1989). Moreover, Nichols and Strizek (1995) find that electronic voting machines featuring lights that blink until a vote is cast in a given race can reduce under-voting; Nichols (1998) presents similar evidence. The Illinois alert differs significantly from the lights examined in Nichols and Strizek (1995); while the latter stimulus was experienced *during* the act of voting, the former occurs *after* the voter has completed the ballot. Thus, correcting an errant undervote would be more costly for an Illinois voter, who would have to return to a voting booth, possibly after waiting in line, identify the location of the undervote, and fix it.

That said, it is at least plausible that Illinois' optical scan alert would reduce the overall undervote rate. However, in the only examination of the relationship between the alert and the undervote rate, Miller (2013) reported practically no effect of the policy. However, Miller's analysis is limited to primary elections, whose voter composition is likely to differ markedly from that of a general election. Relative to general elections, much of the time primaries are low-salience, low-information affairs, drawing a smaller, more committed slice of the electorate. Turnout in the 2010 Illinois General Election, for instance, was more than 20 points higher than in the primary election that year, so it stands to reason that the electorate in the general election was less informed and/or dedicated on average. Indeed, primary election voters are likely to be more informed (Burden 2004; Norrander 1996), more active (Norrander 1986), and older (Geer 1988) than the electorate in a general election—all factors that could conceivably affect errant voting rates. As such, Miller (2013) concedes that "the possibility remains that the alert has stronger effects on aggregate undervoting in general elections (than in primaries)."

In short, while previous studies have demonstrated that election reforms can potentially reduce undervoting rates, the lone examination of the Illinois policy found no evidence of such an effect. By extending evaluation of the audible alert to general elections, we hope to discover whether the alert reduced undervotes among the less committed, presumably less informed general electorate—which is comprised of people who we expect to be more likely to cast erroneous undervotes.

Data and Method

We adhere generally to the analytic strategy employed by Miller (2013), who examined the effect of the undervote alert in *primary* elections using a panel analysis of same-precinct roll-off through time from 2006 to 2010, controlling for time-variant factors such as turnout. To that end, we obtained from the Illinois State Board of Elections (SBE) precinct election returns from the 98 counties in Illinois that employed the same optical scanning technology for both early voting and Election Day voting in each election.[2] Although the format in which Illinois counties submit their returns varies, all returns include information about the number of ballots cast in each precinct, as well as the number of registered voters. The majority of systems also explicitly report the number of undervotes in a given contest.[3]

Of the 98 counties included in our analysis, twelve refused to implement the undervote alert in time for the 2010 primary election, which occurred in February. Miller (2013) exploited this uneven implementation to provide a counterfactual comparison group in a panel analysis of undervotes in the 2006 and 2010 primary elections. Despite the fact that there was an

[2] There are 102 counties in Illinois, but the policy affects only ballots cast on optical-scan machines. Kane and Peoria Counties employ DRE technology for Election Day voting, while Cook County utilizes DRE machines for early voting in Chicago precincts as well as in many suburban precincts on Election Day. As there is no way in the record to distinguish between DRE and optical scan ballots, we exclude Cook County, as did Miller's (2011) analysis. Sangamon County changed election technology between 2006 and 2010, so its precincts cannot serve as a panel for the present study.

[3] We utilized this information where it was available. In counties where the systems did not report an undervote number, we calculated undervotes as the total votes cast subtracted from the total ballots cast.

exceptionally long period of time between the 2010 primary and the November general election, according to the SBE none of the twelve abstaining counties executed the change before the general election, either.[4] As such, as in Miller (2013) we treat the implementation of the undervote alert as a natural experiment, designating a "treatment" group of counties that upgraded their optical scan systems to provide the alert, and a "control" group of counties that did not add the alert to their optical scan systems.[5] Assignment to the "treatment" group was non-random, which is a point we return to below. However, considering that most voters likely had little knowledge of this policy change, and that they received the alert (or not) based solely upon the precinct in which they were assigned to vote, we are comfortable describing this design in quasi-experimental terms.

We therefore proceed by examining precinct-level trends in undervotes from 2006—the last general election in which Illinois' constitutional offices were on the ballot before the alert was implemented—to 2010, the first one after the law took effect. We calculate the difference in undervote rates between 2006 and 2010 within each precinct, and subtract the mean rate of change in the "treated" precincts from that of the "control" group. The mean difference-in-differences allows us to derive the effect of the alert on undervote rates. This claim is grounded on two assumptions. First, since no counties alerted voters to under-voted ballots in 2006, the trend in under-voting from 2006 to 2010 in the control group establishes a counterfactual slope, allowing us to surmise how the rate of under-voting would have changed in the treatment group had the alert not been implemented. Second, employing a panel design eliminates as possible

[4] As noted in Miller (2013), some county clerks exhibited concerns about voter privacy and/or ballot secrecy in refusing to implement the policy. Others, however, were no doubt concerned about cost, as the larger counties were saddled with the higher price of upgrading many more machines—at the height of a deep economic recession that hit Illinois particularly hard.

[5] We include in our initial dataset information from the 5,472 Illinois precincts in our 98-county frame that existed with the same geographic boundaries in both 2006 and 2010, of which 3,986 were in the "treated" condition.

confounders covariates—observable or not—that we expect remained relatively constant within each precinct between the 2006 and 2010 elections, such as racial composition, mean household income, or population.[6] Accordingly, the bivariate relationship between the alert's implementation and undervote rates would be distorted only by omitted variables that changed by both precinct *and* time.[7] This feature of the difference-in-difference approach makes it particularly attractive for studying the effects of policy, and it has been well-utilized in studies of other election policies, including voter ID laws (Erickson and Minnite 2009) and changes in voter registration policies (Hanmer 2009).

That said, the law was not implemented in random fashion, which leads to additional considerations. A number of factors appear to have influenced whether counties upgraded their systems, including the manufacturer of their optical scanning equipment and the number of machines for which an upgrade was necessary. Miller (2013) found that more populous counties were less likely to have completed the upgrade before the election. This difference carries through to the precincts included in our analysis, as those in the "control" condition contained roughly 100 more registered voters than precincts in which the upgrade was implemented.[8] Such a difference serves as a reminder that while the piecemeal implementation of the alert is fortuitous for analysts seeking a counterfactual, the covariate balance between groups should be established prior to conducting an analysis of undervote trends (see: Ho *et al.* 2007a; Sekhon 2009).

The leftmost columns of Table 1 contain the mean values of relevant covariates for precincts in the treatment and control condition, and confirm that control precincts were both

[6] The correlation coefficient for the number of registered voters in a given precinct in 2006 and 2010 is 0.9.
[7] Illinois state law ensures that ballot presentation is not among these potential confounders, as the state mandates that all five of the constitutional elections appear in predetermined uniform order.
[8] One-tailed tests indicate that this difference is statistically significant (p<.0001).

more populous and saw more ballots cast in the 2010 election. In addition, control precincts appear to have a slightly higher percentage of Hispanic voters (as defined in Ansolabehere and Rodden 2011) than treated precincts, and to have been more supportive of Barack Obama in the 2008 presidential election.[9] We correct this imbalance by pre-processing our data with a genetic matching algorithm, from which we derived weights for our subsequent analysis (On pre-processing, see Ho *et al.* 2007a; 2007b. On genetic matching and search algorithms, see: Sekhon 2008; Sekhon 2006; Diamond and Sekhon 2006; Mebane and Sekhon 1998; Sekhon and Mebane 1998). The algorithm seeks optimal balance via the construction of matched pairs consisting of one precinct that implemented the alert and one precinct that did not, and should ideally result in a dataset displaying superior balance without parsing treated precincts.

Table 1: Mean Values of Observable Covariates and Balance Assessment, Precinct-Level Data

	Means: Unmatched Data		Means: Matched Data		Percent Improvement			
	Treated	Control	Treated	Control	Mean Diff.	eQQ Med	eQQ Mean	eQQ Max
Propensity Score	0.77	0.62	0.77	0.77	99.83	57.10	37.75	6.87
2010 Registered Voters	682	780	682	676	94.17	3.26	9.07	0.00
2010 Ballots Cast	341	374	341	337	86.42	9.76	6.98	0.00
2006 Residual Vote Perc.	3.63	3.07	3.63	3.60	95.38	27.87	30.11	-44.90
Perc. Obama Vote, 2008	51.44	55.35	51.44	51.40	99.07	23.99	22.65	0.00
Perc. African-American	4.48	5.91	4.48	3.90	59.66	4.31	9.71	13.15
Perc. Hispanic	5.29	7.17	5.29	4.80	73.38	35.73	29.48	1.67
Perc. in "Collar County"	23.11	61.98	23.11	22.80	99.23	0.00	39.78	0.00

The matching algorithm seeks balance on eight covariates observed before voters cast ballots during 2010 general election. These include measures of precinct size (including the number of registered voters and ballots cast in 2010), historical under-voting rates, racial

[9] In both cases, one-tailed tests indicate that the differences are statistically significant, $p<.0001$.

demographics, and the percentage of the vote received by Barack Obama in the 2008 election.[10] We also include a dummy variable denoting a precinct's presence in the suburban "collar counties" that ring Cook County in the Northeastern corner of the state. These counties are part of the Chicago media market, and are generally understood to be distinct from much of the remainder of downstate Illinois in terms of partisan preference and socio-economics.[11] To aid the matching exercise, we also include the propensity score in the algorithm, which we calculated with a logistic regression model employing all observed covariates (see: Diamond and Sekhon 2006).

The results of the matching exercise are displayed in Table 1, which in addition to the pre-match means described above also depicts the means in the post-match control group after poorly matched precincts were parsed from the dataset (Column 4).[12] The four rightmost columns of Table 1 contain percentages that indicate how much matching improved balance between the treatment and control groups on a range of metrics. On the whole, Table 1 indicates that matching resulted in a dataset with improved balance. Specifically, matching reduced the mean difference between groups for all variables included in the algorithm, and the extent of the improvement exceeds 70% for all but one variable. Similar improvement is apparent on the majority of other metrics in Table 1 as well. As such, we believe that matching results in a well-balanced dataset, which enhances our confidence in unbiased inference about the effect of the undervote alert in the results we report below.

[10] We obtained racial demographics and Obama's vote percentage in each precinct from Ansolabehere and Rodden (2011). We employ 2010 information for registered voting and ballots cast because a voter must necessarily both register *and* receive a ballot at his/her precinct prior to voting (or casting a ballot with a non-vote) in a given race.
[11] Among counties included in our initial dataset, these are: DuPage, Lake, McHenry, and Will. Barack Obama performed better in these counties by about six percentage points, relative to other downstate counties.
[12] All "treated" precincts were successfully matched. After conducting the match, 4,874 precincts remain in our dataset. Effectively, the matching exercise parsed 598 control precincts from the dataset that did not pair well with any treated precinct, and supplied weights for each surviving precinct to be used in subsequent analysis.

In the analysis below we present straightforward calculations of mean difference-in-differences in roll-off patterns, with relevant one-tailed tests. However, we also calculate the difference-in-differences with OLS regression models of same-precinct roll-off. This approach affords the advantage of holding time-variant covariates constant, allowing the models to return the treatment effect on the treated, or the independent effect of implementing the alert on the percentage of ballots cast with undervotes.

We calculate an undervote percentage (or "roll-off percentage") as the proportion of ballots cast in a given race on which no vote was recorded. This quantity ranges from 0 to 100, with higher values indicating a larger percentage of voters submitting ballots with no preference registered for a given contest. We report results for each of the five elections to constitutional offices (Governor/Lt. Governor, Attorney General, Secretary of State, Comptroller, and Treasurer) separately, but we also compute the percentage of undervotes in constitutional races for a given precinct overall in an effort to analyze whether the alert affected undervoting in the aggregate. This percentage is calculated as:

$$R = 100\left(\frac{\sum U_{IP}}{5B_P}\right)$$

Where $\sum U_{IP}$ is the total number of undervotes cast for each *I* of the five statewide races in precinct *P*, and *B* is the total number of ballots cast. Accordingly, *R* represents the proportion of all possible votes in a given precinct that were recorded as undervotes, ranging from 0 to 100. The dependent variable in our analysis is the first-differenced quantity of the undervote percentage in a given precinct in a given race for 2006 subtracted from the undervote percentage in the same precinct and race in 2010. For the overall measure *R* in precinct *P* described above, this quantity is therefore calculated as:

$$\Delta R = R_{P_{2010}} - R_{P_{2006}}$$

For all models, the independent variable of interest is a dichotomous indicator of "treatment" coded 1 if the precinct was within a county that implemented the alert in time for the election, and 0 otherwise. Given the first-differenced nature of the dependent variable, the regression coefficient for this indicator therefore reflects the difference-in-differences in undervote percentage, or the effect on ΔR of implementing the undervote alert. Considering the manner in which ΔR is calculated, if the alert reduced instances of under-voting, then the coefficient for the treatment indicator would be negative and statistically significant.

We add to all models one additional control variable to account for the change in turnout in each precinct from 2006 to 2010. Turnout is theoretically related to roll-off in its own right; as more people turn out to vote for races on the top of the ballot (such as federal elections or governor), the average knowledge of and/or interest in the various races farther down the should be expected to decrease. In other words, as turnout rises, so too should the percentage of undervotes. However, we also believe that precinct-level turnout can serve as a proxy for less observable characteristics of an election, such as campaign spending and/or intensity. This is important, since mobilization activities have been shown to affect turnout in the context of gubernatorial elections (Nickerson, Friedrichs, and King 2006).[13] We therefore calculate turnout in each precinct as the percentage of registered voters who cast a ballot, and include as a control a first-differenced variable reflecting that figure (ranging between 0 and 100) for 2006 subtracted from that for 2010.

Results

We begin by examining overall undervote percentages in control precincts in 2006 and 2010. We do this in an effort to assess the key assumption of our panel difference-in-differences design:

[13] Turnout is an exogenous predictor in this case, as a voter travels to a polling place prior to voting (or not voting).

that the control group provides a reliable counterfactual trend. If the roll-off percentage in precincts for which there was no intervening treatment displays high correlation between the two elections, we can begin with higher confidence that our findings are not affected somehow by unobserved variables. Figure 1 therefore depicts mean overall roll-off rates from the precincts included in our matched dataset; rates for 2006 are plotted against those from 2010. As is evident in Figure 1, control precincts demonstrate a strong, positive correlation (*r*=.57), indicating stability through time. As such, we are confident that the control group yields a reliable counterfactual basis for inference about the effect of the undervote alert.

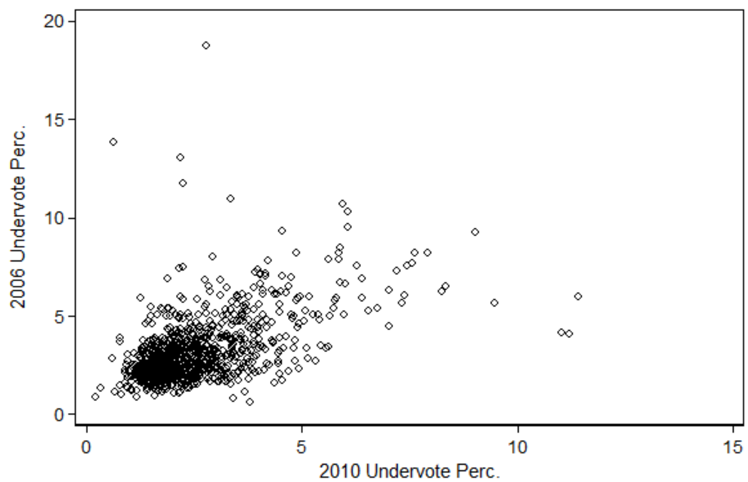

Figure 1: Correlation of Undervote Rate in Illinois Control Counties, 2006 and 2010. Note: Rate ranges between 0 and 100.

Table 2 contains mean same-precinct undervote rates overall (*R* above), and also for each of the five contests for constitutional office. Importantly, the rates are calculated using the full, unmatched sample of precincts described above. We present rates separately for the treatment and control group, for each election year. Table 2 also contains the first differences for each group (2010 rate-2006 rate), and the difference-in-differences (Treatment Difference- Control Difference), which can be interpreted as the effect of the alert—albeit holding no time-variant factors constant. Two trends are worth mentioning. First, for both the treatment and control groups, undervote rates declined both overall and in elections to four of the five constitutional offices from 2006 to 2010. The lone exception is the election for Secretary of State, in which the roll-off rate increased by roughly .4 and .75 points in the control and treatment groups, respectively.

Second, the difference-in-differences (rightmost column) yields mixed findings regarding the effect of the alert. Notably, the effect of the alert on undervote rates overall appears to be negligible. The difference does not achieve statistical significance, and is substantively small, at seven-hundredths of one percent. However, the differences in Table 2 are consistent with lower roll-off in treated precincts in the elections for two offices: Comptroller and Treasurer. The undervote rate trended significantly lower in treated counties by about 0.5 and 0.6 percentage points in these elections, respectively. In contrast, roll-off trended significantly *higher* in the election for Governor and Secretary of State in precincts that implemented the alert.

Notably, regression models on the unmatched data controlling for same-precinct turnout fluctuations return the same findings with regard to both substantive size and statistical significance. These results are reported in Appendix Table A1. With such varied findings—and especially considering the apparent inability of the alert to affect overall undervote rates—it is difficult to conclude anything other than that there is little evidence in Table 2 for a uniform effect of the undervote alert. There is, however, evidence in the unmatched data that the alert was effective in the races for Comptroller and Treasurer, which are generally less-visible than the other elections and which are found in the last two rows of the ballot for constitutional offices.

Table 2: Panel Difference-in-Differences, Undervote Rate, Unmatched Data

	Means: Matched Control		2010-2006: Control	Means: Matched Treated		2010-2006: Treated	Difference-in-Differences
	2006	2010		2006	2010		
Overall	3.07	2.45	-0.62*	3.63	2.94	-0.69*	-.07
Governor	2.41	1.35	-1.06*	2.34	1.60	-0.74*	.315*
Attorney General	2.45	2.06	-0.39*	3.06	2.67	-0.39*	-0.004
Secretary of State	1.73	2.14	0.41*	1.76	2.54	0.78*	0.37*
Comptroller	4.17	3.22	-.95*	5.13	3.71	-1.42*	-0.47*
Treasurer	4.56	3.47	-1.09*	5.84	4.18	-1.66*	-0.57*

* $p<.05$. One-tailed tests. Undervote rate ranges from 0 to 100.

We now turn to the results of our regression models on the matched sample, which calculate the independent effect of the alert on undervote rates, holding fluctuations in voter turnout constant. We fit these models with a parsed dataset resulting from the matching exercise above, and employ weights derived from the matching algorithm. The models are intended to calculate more accurately the treatment effect of the alert, employing a balanced dataset while holding a known time-variant confounder constant. Coefficients and standard errors from the models are contained in Table 3, which depicts results both for overall undervote rates and for elections to each constitutional office.

Table 3: OLS Regression Coefficients: Panel Difference-in-Differences in Overall Roll-Off, Optical Scan Precincts, Matched Data

	(1) Overall	(2) Governor	(3) AG	(4) Sec. State	(5) Comptroller	(6) Treasurer
Treatment Indicator	-0.061	-0.077	0.208	0.068	-0.185	-0.321*
	(0.093)	(0.095)	(0.106)	(0.091)	(0.135)	(0.139)
Change in Turnout	0.032*	0.011*	0.042*	0.026*	0.035*	0.047*
	(0.005)	(0.005)	(0.007)	(0.004)	(0.008)	(0.009)
Constant	-0.676*	-0.679*	-0.668*	0.675*	-1.295*	-1.414*
	(0.089)	(0.090)	(0.101)	(0.087)	(0.129)	(0.127)
Observations	4,874	4,874	4,874	4,874	4,874	4,874
R-squared	0.016	0.002	0.021	0.011	0.010	0.010
Root Mean Sq. Error	1.763	1.798	2.084	1.715	2.517	3.432

* p<0.05
Robust standard errors in parentheses. Weights derived from genetic matching algorithm.
Dependent variable is 2006 undervote percentage subtracted from 2010 undervote percentage.

Table 3 offers relatively little evidence of reduced undervote rates in the precincts that implemented the alert. The coefficient for overall undervote rates—arguably the most relevant outcome in terms of judging the efficacy of the policy—is substantively small and does not achieve statistical significance. The same can be said of the coefficients for undervote rates in the elections for Governor and Secretary of State. The coefficient for roll-off rate in the Attorney General elections is relatively large but is incorrectly signed. Results from the unmatched data indicated significant, negative effects of the alert for the Treasurer and Comptroller elections, and the matched models also return negative values for the treatment indicator in those races. However, only in the Treasure's election is the coefficient statistically significant (*p*=.021) in the matched models. That said, the coefficient for the undervote trends in the Treasurer election does suggest a meaningful effect: The 0.32-point reduction in Treasurer roll-off amounts to a decrease in under-voting of about 7% in that election.

We stress, however, that in no other election do the models of matched data suggest that the alert reduced undervote rates, and the results of the matched and unmatched data agree in all elections except that for Comptroller. There is therefore strong evidence for undervote reductions in the election for Treasurer, and weaker support for the notion that the alert reduced under-voting in the election for Comptroller. Thus, we are left with a nuanced conclusion: the alert might have diminished instances of under-voting in *some* races (especially those lower on the ballot), but with regard to the rate of undervotes overall, there is scant evidence for a reduction in precincts that implemented the alert.

Table 4: 90% Confidence Intervals About the Simulated Treatment Indicator Coefficients

	Lower Bound	Mean	Upper Bound
Overall	-0.066	-0.061	-0.057
Governor	-0.081	-0.076	-0.070
AG	0.202	0.207	0.213
Sec. State	0.070	0.075	0.080
Comptroller	-0.188	-0.181	-0.174
Treasurer	-0.336	-0.328	-0.321

Following the recommendations in Rainey (2014), we conclude by investigating the nature of the apparently negligible effects that our models suggest. To that end, we conducted Monte Carlo simulations in which we drew a sample of 1,000 coefficients from the treatment indicators for each of the six models presented in Table 3.[14] We then calculated 90% confidence intervals about the means of the resulting sample. The means and confidence interval bounds are contained in Table 4. The confidence intervals for both the Comptroller and Treasurer elections do contain values that would denote non-negligible negative effects of the alert on undervote rates in elections to those offices. However, for the "overall" coefficient, which is the one with greatest relevance to the question of how the alert affects undervote rates across all elections, the

[14] We ran these simulations with the Clarify package for Stata (Tomz,Wittenberg, and King 2003)

confidence interval does not contain non-negligible effects. Again, we are left with a somewhat split decision: While we are comfortable concluding that our analysis yields little evidence that the alert reduced undervote rates in the aggregate, there is evidence for some effect on races relatively low on the ballot.

Conclusion

In total, our analysis is strongly suggestive that undervote rates trended significantly lower in the 2010 election for Treasurer in precincts that implemented the alert, and there is also some weaker evidence of negative effects on roll-off in the Comptroller election. However, there are no negative, non-negligible effects apparent in elections to the other three constitutional offices in Illinois, and most important, we find practically no evidence that the alert reduced rates of under-voting overall in the precincts where it was implemented. This conclusion is somewhat more nuanced than that in Miller (2013), who found essentially no effects of the alert in primary elections. We believe that our finding leaves one obvious unanswered question that might be addressed in future research: Why might the undervote work in only the Treasurer's race?

There are at least two possible answers to this question. One explanation may be that the Treasurer's race appears at the bottom of the ballot relative to the other elections we analyze, and therefore unsurprisingly features higher instances of roll-off in all cross-tabulations depicted in Table 2 above. It is possible that given its placement on the ballot, the election for Treasurer was more likely to be accidentally overlooked, and that the alert therefore did correct instances of mistaken under-voting. The Treasurer's race is also unique in that there was no incumbent running in either 2006 or 2010, nor did either election feature a candidate who would have been considered to be well-known during the election in question. It is not clear how this anomaly might translate to the effectiveness of the alert, but we believe it is worth mentioning both of these possibilities to guide further research and to inform policymakers in subsequent elections. At a minimum, future research should consider whether efforts to reduce erroneous voting should be expected to have disparate effects with regard to an election's position on the ballot.

With regard to the overall inefficacy of the policy in reducing undervotes, our findings have implications both for judging the alert as policy and for extant research on the manner in which external stimuli affects voting behavior. For one, given the apparent failure of the alert to affect undervote rates overall, our analysis is consistent with the long-held belief that under-voting is a mostly intentional act. That said, we reiterate that since our analysis is conducted at the precinct level, we will refrain from speculating about the alert's impact on any aspect of mass voting behavior, including propensity to unintentionally "roll off" in a given contest. There are in fact multiple factors that could contribute to the alert's ineffectiveness. Voter apathy about certain races and/or lack of information about the candidates may simply not be overcome by the audible warning. The relatively costly act of returning to the voting booth to complete the ballot and re-submit may also reduce the impact of the alert. As such, this paper should be read as analysis of the alert's overall effect, and not an analysis of individual voting behavior.

To that end, our analysis calls into question the wisdom of enacting such an alert as policy, as doing so has not been without cost to the Illinois counties that upgraded their machines. For instance, the 66 counties utilizing Diebold Accuvote machines spent about $135 per machine in order to implement the undervote alert (Guidroz 2010). Moreover, as noted in Miller (2013), it is possible that the policy may also generate other, unforeseen costs to voters and counties alike, including longer wait times and more poll workers to oversee the process. This is particularly problematic given the well-documented budget issues in Illinois in the early half of the 2010s. That said, our findings should give pause to any reformer—however well-meaning—looking to enact a policy to curb unintentional under-voting.

References

Ansolabehere, Stephen, and Jonathan Rodden. 2011. "Illinois Data Files, Version 2." Retrieved October 31, 2011 from: http://hdl.handle.net/1902.1/15845

Ansolabehere, Stephen, and Charles S. Stewart III. 2005. "Residual Votes Attributable to Technology." *The Journal of Politics* 67(2): 365-389.

Burden, Barry C. 2004. "Candidate Positioning in US Congressional Elections." *British Journal of Political Science* 34(2): 211-227.

Darcy, R., and Anne Schneider. 1989. "Confusing Ballots, Roll-Off and the Black Vote." *Western Political Quarterly*

Diamond, Alexis, and Jasjeet S. Sekhon. 2006. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." Available at: http://sekhon. berkeley. edu/papers/GenMatch. pdf [Accessed March 16, 2009].

Erikson, Robert S., and Lorraine C. Minnite. 2009. "Modeling Problems in theVoter Identification-Voter Turnout Debate," *Election Law Journal* 8(2): 85-101.

Geer, John G. 1988. "Assessing the Representativeness of Electorates in Presidential Primaries." *American Journal of Political Science* 32(4): 929-945.

Hanmer, Michael J. 2009. *Discount Voting: Voter Registration Reforms and Their Effects.* Cambridge University Press.

Herrnson, Paul S., Richard G. Niemi, Michael J. Hammer, Benjamin B. Bederson, Frederick Conrad, and Michael Traugott. 2008. *Voting Technology: The Not-So-Simple Act of Casting a Ballot.* Brookings Institution Press.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007(a). "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(1):199-236.

Kimball, David C., and Martha Kropf. 2005. "Ballot Design and Unrecorded Votes on Paper-Based Ballots." *Public Opinion Quarterly* 69(4): 508-529.

Knack, Stephen, and Martha Kropf. 2003. "Roll-Off at the Top of the Ballot: Intentional Undervoting in American Presidential Elections." *Politics and Policy* 31(4): 575-594.

Miller, Michael G. 2013. "Do Audible Alerts Reduce Undervotes? Evidence from Illinois." *Election Law Journal: Rules, Politics, and Policy* 12(2): 162-178.

Mebane, Walter R., and Jasjeet S. Sekhon. 1998. "Genetic Optimization Using Derivatives (GENOUD)." Available at http://sekhon.polisci.berkeley.edu/rgenoud/. Accessed March 16, 2009.

Nichols, Stephen M., and Gregory A. Strizek. 1995. "Electronic Voting Machines and Ballot Roll-Off." *American Politics Quarterly* 23(3): 300-318.

Nichols, Stephen M. 1998. "State Referendum Voting, Ballot Roll-Off, and the Effect of New Electoral Technology. *State & Local Government Review* 30(2): 106-117.

Nickerson, David W., Ryan D. Friedrichs, and David C. King. 2006. "Partisan Mobilization Campaigns in the Field: Results from a Statewide Turnout Experiment in Michigan." *Political Research Quarterly* 59(1): 85-97.

Norrander, Barbara. 1986. "Selective Participation: Presidential Primary Voters as a Subset of General Election Voters." *American Politics Quarterly* 14(1-2): 35-53.

Norrander, Barbara. 1996. "Presidential Nomination Politics in the Post-Reform Era." *Political Research Quarterly* 49(4): 875-915.

Rainey, Carlisle. 2014. "Arguing for a Neglible Effect." *American Journal of Political Science* 58(4): 1083-1091.

Sekhon, Jasjeet S. 2006. "Alternative Balance Metrics in Matching Methods for Causal Inference." Available at: http://sekhon.berkeley.edu/papers/SekhonBalanceMetrics.pdf [Accessed March 16, 2009].

Sekhon, Jasjeet S. 2008. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R.." Available at: http://sekhon.berkeley.edu/papers/MatchingJSS.pdf [Accessed March 16, 2009].

Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." Available at: sekhon.berkeley.edu/papers/opiates.pdf [Accessed March 16, 2009].

Sekhon, Jasjeet S., and Walter R. Mebane. 1998. "Genetic Optimization Using Derivatives." *Political Analysis* 7(1): 187-210.

Shocket, Peter A., Neil R. Heighberger, and Clyde Brown. 1992. "The Effect of Voting Technology on Voter Behavior in a Simulated Multi-Candidate City Council Election: A political Experiment of Ballot Transparency." *Western Political Quarterly* 57(1): 15-25.

Stein, Robert M., Greg Vonnahme, Michael Byrne, and Daniel Wallach. 2008. "Voting Technology, Election Administration, and Voter Performance." *Election Law Journal* 7(2): 123-135.

Tomz, Michael, and Robert P. Van Houweling. 2003. "How Does Voting Equipment Affect the Racial Gap in Voided Ballots?" *American Journal of Political Science* 47(1): 46-60.

Appendix: Regression Results from Unmatched Data

Table A1: OLS Regression Coefficients: Panel Difference-in-Differences in Overall Roll-Off, Optical Scan Precincts, Unmatched Data

	(1) Overall	(2) Governor	(3) AG	(4) Sec. State	(5) Comptroller	(6) Treasurer
Treatment Indicator	-0.075	0.314*	-0.009	0.371*	-0.476*	-0.576*
	(0.050)	(0.054)	(0.059)	(0.050)	(0.074)	(0.082)
Change in Turnout	0.026*	0.010*	0.035*	0.021*	0.026*	0.039*
	(0.005)	(0.004)	(0.006)	(0.004)	(0.007)	(0.008)
Constant	-0.654*	-1.070*	-0.440*	0.378*	-0.989*	-1.147*
	(0.041)	(0.046)	(0.049)	(0.042)	(0.063)	(0.059)
Observations	5,472	5,472	5,472	5,472	5,472	5,472
R-squared	0.011	0.007	0.014	0.017	0.012	0.012
Root Mean Sq. Error	1.730	1.800	2.039	1.685	2.492	3.301

* $p<0.05$. Robust standard errors in parentheses. Dependent variable is 2006 undervote percentage subtracted from 2010 undervote percentage.