2010

Long Lines at Polling Stations? Observations from an Election Day Field Study

Douglas M. Spencer
University of Colorado Law School

Zachary S. Markovits
Election Initiatives, Pew Center on the States

Citation Information

Douglas M. Spencer & Zachary S. Markovits, *Long Lines at Polling Stations? Observations from an Election Day Field Study*, 9 Election L.J. 3 (2010), https://doi.org/10.1089/elj.2009.0046, available at https://scholar.law.colorado.edu/articles/1384.

ELECTION LAW JOURNAL
Volume 9, Number 1, 2010

DOI: 10.1089/elj.2009.0046

Original Article

Long Lines at Polling Stations? Observations from an Election Day Field Study

Douglas M. Spencer and Zachary S. Markovits*

ABSTRACT

This pilot study represents the first systematic attempt to determine how common lines are on Election Day, at what times of day lines are most likely to form, what are the bottlenecks in the voting process, and how long it takes an average citizen to cast his or her ballot. This study highlights the importance of evaluating polling station operations as a three-step process: arrival, check-in, and casting a ballot. We collected data during the 2008 presidential primary election in California, measuring the efficiency of the operational components of 30 polling stations across three counties. We found statistically significant, and meaningful, variation in the service rates of poll workers and voting technology. Our findings should better help election officials make critical decisions about the allocation of critical resources.

INTRODUCTION

THE SUCCESS OF America's electoral process hinges on the performance of hundreds of thousands of polling places housed in schools, churches, municipal buildings, and neighbors' garages. Local elections officials manage the poll workers whose task is to administer an efficient voting process. When this process breaks down, it is critical for these local officials to both identify and solve whatever is wrong. Too often, election officials lack the data they need to effectively measure both what has gone wrong and how a policy change will lead to better performance. As an example, long lines at polling stations are a visible indication that something is wrong, yet little reliable information is available to explain the operational inefficiencies that lead to lines at polling stations or to guide policymakers in their choice of remedy.

Voters have long complained about standing in line to vote, but in 2004 the issue became the subject of popular concern. In the presidential election of 2004, long lines were reported in Ohio, Pennsylvania, and Florida—some as long as ten hours.[1] Several states responded to these reports with new early voting rules. For example, in 2005, the Illinois General Assembly approved a bill that provided for in-person early voting at newly established vote centers.[2] Later that year, Ohio's General Assembly

* Douglas M. Spencer is a researcher at the Survey Research Center, University of California, Berkeley and a doctoral candidate at the UC Berkeley School of Law's Jurisprudence and Social Policy Program. Zachary S. Markovits is a Senior Associate, Election Initiatives, Pew Center on the States.

The authors presented a version of this paper at the 2009 Annual Meetings of the Midwest Political Science Association on April 2, 2009 in Chicago, IL. Portions of this research appeared in "Dispatches from the Polls: A Case Study," *Electionline.org Briefing*, Washington, D.C. (2008), pp. 4–7.

[1] M. Powell and P. Slevin, "Several Factors Contributed to 'Lost' Voters in Ohio," *Washington Post*, Washington, D.C. (December 15, 2004), p. A01.

[2] 10 Ill. Comp. Stat. § 5/19A-10. A 2006 Illinois State Board of Elections pamphlet states that people vote early because "they want to avoid long lines at the polling place on Election Day." See "Early Voting: No Reason Needed to Vote Early," http://www.elections.state.il.us/Downloads/ElectionInformation/PDF/EarlyVoting.pdf.

passed a law allowing no-fault absentee voting.[3] The Utah legislature followed suit the next year with an early voting law of their own.[4] All of these laws aimed to spread out the voting time in order to relieve polling place congestion on Election Day.

Despite these and other attempts to reduce the waiting time for voters, long lines remained one of the major problems reported by voters during the 2008 election cycle. According to a recent survey, 11% of respondents who did not vote in the 2008 general election cited long lines as a major factor in their decision to stay home.[5] That result suggests that long lines may have discouraged up to 2.6 million people from voting.[6] Among those who braved long lines to vote, African Americans were twice as likely to report standing in line for more than half an hour than white or Hispanic voters.[7] These reports suggest that consequences of long Election Day lines are more serious than the mere inconvenience of standing on one's feet for several hours; there are strong reasons to seek a solution to this problem.

Unfortunately, there are few data on which to base an explanation of the prevalence of lines at polling stations. The data that do exist come from post-election, self-reported interviews.[8] To date, nobody has physically monitored the inner-workings of polling stations to observe how their location, personnel, and organization affect the flow of voter traffic, to measure the time it takes people to vote, or to determine how these two variables are related.

On February 5, 2008 we dispatched 119 data collectors to 30 California polling stations to collect such data. This study represents the first systematic attempt to determine how common lines are, at what times of day lines are most likely to form, at what point in the voting process lines form, and how long it takes the average citizen to cast his or her ballot. In addition to uncovering data about the voting process, our project also aims to establish a simple research design to guide future large-scale data collection projects.

Much of the field research on elections examines activities that take place before elections, such as voter registration and poll worker training, and relies on information that is available after the election, such as turnout and election results, but treats the polling station as a black box. We capture data related to voter arrivals, poll worker availability, voting machine operation, and the length of lines. Our targeted independent variable is voting technology. We also analyze how various poll worker characteristics affect the time it takes people to get served at a polling station. This study departs from prior research on elections by peering inside the black box and focusing on the internal operation of polling stations.

One major reason why polling place inefficiency has yet to be adequately studied is that the administration of elections in the United States is extremely complicated. Each state creates its own rules, budgets its own money, and constructs its own election processes. In some states, such as Wisconsin and Michigan, local jurisdictions have primary autonomy over election administration. In others, such as Oklahoma and Delaware, all election officials are state employees. Still others share administrative duties between state and local election officials. For example, in California, counties have significant authority, yet they operate within a broad framework established by the Secretary of State. On the federal level, the United States Constitution pre-

---

[3] Sub. H.B. No. 234. During the 2006 general election, a then-record 15.3% of Ohio voters cast absentee ballots (up from 10.6% in 2004 under the old law). The Columbus Dispatch quoted Matthew Damschroder, director of the Board of Elections in Franklin County, where nearly 25% of voters cast absentee ballots, as attributing the increase in absentee voting to "lingering memories of the long lines at many county polling stations in . . . 2004." See M. Niquette, "Voting law made few big waves," *The Columbus Dispatch*, Columbus, Ohio (December 17, 2006).

[4] Utah Election Code 20A-3-601, amended by Chapter 256, 2007 General Session.

[5] M. Alvarez, S. Ansolabehere, A. Berinsky, T. Hall, G. Lenz, and C. Stewart, "2008 Survey of the Performance of American Elections—Final Report" (Boston 2009), p. 36 (see Table III-3), http://vote.caltech.edu/drupal/node/231.

[6] *Id*., at 59.

[7] *Id*., at 38. 29% of black voters reported standing in line for more than thirty minutes compared to 14% of white voters and 15% of Hispanic voters.

[8] While exit polling usually is limited to turnout and voter preference, some interviewers ask voters how long they stood in line. See R. Claassen, D. Magleby, Q. Monson, K. Patterson, "'At Your Service': Voter Evaluations of Poll Worker Performance," *American Politics Research* Vol. 36, No. 4 (2008), pp. 612-34; M. Alvarez, et al., *supra* note 5; R. M. Stein and G. Vonnahme, M. Byrne, and D. Wallach, "Voting Technology, Election Administration and Voter Performance," *Election Law Journal: Rules, Politics, and Policy* Vol. 7, No. 2 (June 2008), pp. 123–35.

serves the right of Congress to supersede state laws regulating congressional elections.[9] The result is a complex web of overlapping jurisdictions and 10,071 government units that administer elections.[10] To complicate matters further, authority in all jurisdictions is ceded to two million poll workers who control the success or failure of each election.[11]

Our study contributes to the literature on elections administration in two ways. First, we break down the voting experience into a three-step Election Day precinct place voting process that is common to all jurisdictions regardless of local rules. Second, we create reliable measures for the arrival rate of voters and the service time of the poll workers and voting machines, respectively. Using these measures, we are better able to predict when voters will arrive at the polls and to evaluate the efficiency of various poll worker protocols and the effects of different voting technology.

PREVIOUS RESEARCH

Lines form when the supply of a service cannot meet the demand to be served. By increasing the efficiency of voting services, elections administrators can decrease the likelihood of lines forming at polling stations during an election. Unfortunately, elections administrators do not often think of polling stations as operation centers and much of the research on elections treats the polling station as a black box. Operations research is not prominent in the election literature, although it is extensively employed in the private sector when businesses are looking for ways to boost efficiency or cut costs. More specifically, firms apply the methods of queuing theory to study the efficiency of assembly line production, computer processors, and customer-oriented services.[12]

This study makes a case for applying operations research to elections and illustrates how to use the framework of queuing theory to improve the voting process.

Lines and operations

Our study relies on the basic assumption of queuing theory that lines form when the rate of service being provided cannot keep pace with arrivals. In every polling station, two services are provided: identity verification and balloting. Thus, we analyze the voting process in three steps: 1) arrival at the polling station, 2) check-in and identity verification (what we call Service One), and 3) balloting (what we call Service Two). Figure 1 illustrates this process.

Congestion at a polling station will occur for either of two reasons. In the first case, a line will form when Service One cannot keep pace with arrivals (Service One $<$ Arrivals). For example, if there is only one poll worker checking a voter's identification and registration, and she can only serve three voters per minute, then a line will form if four or more people arrive each minute. In the second case, a line will form when Service Two cannot keep pace with the flow of traffic into the voting booth (Service Two $<$ Arrivals $<$ Service One). For example, if three voters arrive and have their registration verified each minute, but upon leaving the check-in table find only one available voting machine that requires four minutes to use, a voter backup will ensue.

To date, the research that examines polling station operation and congestion has not evaluated how the distribution of voter arrivals interacts with the rate of poll worker service and the time a voter stands in the voting booth. In the two most prominent articles on polling station congestion, the au-

---

[9] Art. I, sec. 4 of the United States Constitution reads: "The Times, Places, and Manner of holding Elections for Senators and Representatives, shall be prescribed in each State by the Legislature thereof; but the Congress may at any time by Law make or alter such Regulations . . . "

[10] Election Data Services, "2008 Voting Equipment Report, Washington, D.C. (2008), http://www.electiondataservices.com/images/File/NR_VoteEquip_Nov-2008wAppendix2.pdf.

[11] See "Helping Americans Vote: Poll Workers," *Election line.org Briefing*, Pew Center on the States (September 2007), http://www.pewcenteronthestates.org/uploadedFiles/ERIP Brief19_final.pdf.

[12] For example, in 1997, IBM teamed with Air Canada to apply queuing models to decrease customer wait times for preflight services like check-in, security checks, and boarding. They developed a model that allowed Air Canada to run various "what if" simulations. Thus, Air Canada could project the performance outcomes of different resource allocations, such as changing the number of customer service agents or introducing electronic check-in machines. See J. L. Snowdon et al., "IBM Journey Management Library: An Arena System for Airport Simulations," *The Journal of the Operational Research Society* 51, no. 4 (2000), pp. 449–56.

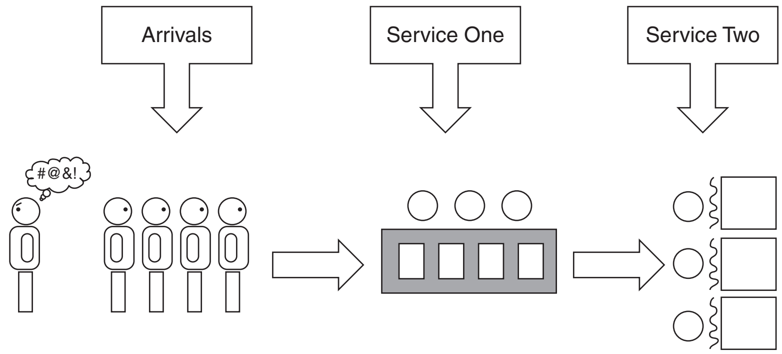

FIG. 1. Three steps of the voting process.

thors focus on the 2004 presidential election in Franklin County, Ohio, where people reported waiting in line up to three hours to vote. In both articles, the authors rely on assumptions about polling place activity to estimate line lengths using proxies related solely to voting equipment. Highton relies on voting machine allocation as a proxy for variation in line length.[13] Allen and Bernshteyn use both machine allocation and ballot length as proxies for line length and also point to after-hours service at polling stations as evidence that long lines persisted throughout the day.[14] These articles illustrate how various administrative decisions can negatively affect voter participation. Indeed, both articles estimate that more than 20,000 people failed to vote in Franklin County because of long lines.

A more recent report by Edelstein employs a simulative queuing model, again based on voting equipment, to estimate the distribution of average wait times during an election.[15] Edelstein artificially manipulates the input of voter arrivals, assuming that the rate of arrivals doubles for a one hour period in the morning, during the lunch hour, and at the end of the day. Edelstein simulates 10,000 elections to generate a series of graphs of peak wait times based on various scenarios of voting machine allocation and ballot completion rates. These models are a welcome contribution to the field of election operations, and the graphs are a good example of how information about the process might help voters to make an informed decision about when to go to the polls.

Like the two previous articles, however, Edelstein was unable to link his theoretical model to actual Election Day circumstances and voter behavior. For example, the focus on voting machines and ballot lengths fails to account for variation in the first service provided to voters at the check-in table and its effect on line length. In addition, the distribution of voter arrivals is based on an assumption of randomness or other untested hypotheses about voter behavior. In contrast, our study aims to both broaden the scope of observation and to use observational data to test general assumptions about voter behavior.

Our goal in designing this study is to bring actual data related to the operation of polling stations to bear on these models and to identify not just when a line may form, but exactly where and why the process breaks down. Every election jurisdiction runs its elections differently; thus, our study is by no means a comprehensive analysis of this issue, nor can it be universally applied. However, we believe that our dataset is valuable in that it describes actual, observed processes that underlie how polling places operate.

[13] B. Highton, "Long Lines, Voting Machine Availability, and Turnout: The Case of Franklin County, Ohio in the 2004 Presidential Election," *PS: Political Science and Politics* Vol. 39, No. 1 (2006), pp. 65–68. See also J. Knapp, "Effect of Voting-Machine Allocations on the 2004 Election—Franklin County, Ohio," (2009), http://coolohio.com/machinery/.
[14] T. Allen and M. Bernshteyn, "Mitigating Voter Waiting Times," *Chance* Vol. 19, no. 4 (2006), pp. 25–36.
[15] W. A. Edelstein, "New Voting Systems for NY—Long Lines and High Cost " *New Yorkers for Verified Voting* (2006), available at: http://www.nyvv.org/voterlines.shtml. See also W. A. Edelstein and A. D. Edelstein, "Touchscreen Voting Machines Cause Long Lines and Disenfranchise Voters," (2008), http://arxiv.org/abs/0810.5577.

STUDY DESIGN

Setting

California is home to 12% of the country's registered voters. Each of California's 58 counties trains its own poll workers and selects its own voting system. In 2007, California Secretary of State Debra Bowen conducted a "top-to-bottom review" of the state's voting systems and withdrew her approval for the use of several voting machine models, most notably touch-screen direct recording electronic (DRE) machines, whose manufacturers had failed to meet the Secretary's published standards.[16] As a result, four voting systems were used in California during the 2008 primary election: (1) non-touch-screen DRE machines manufactured by Hart Intercivic; (2) paper ballots with precinct-based optical scanners; (3) paper ballots with centrally located optical scanners; and (4) the "InkaVote Plus" hybrid precinct-central scanner with security ink marker.[17] The Hart Intercivic DRE machine, called an eSlate, was used by Orange and San Mateo counties and required voters to spin a rotary "SELECT Wheel"—much like an iPod track wheel—to scroll through the on-screen selections. Twenty-eight counties used precinct-based optical scanners, twenty-four counties used centrally-located optical scanners, and Los Angeles county used the one-of-a-kind InkaVote hybrid scanner system where ballots are scanned at the precinct to verify they have been filled out correctly, then re-scanned at a central location where the individual votes are tallied.[18]

Any registered voter in California may alternatively request to vote absentee. California is one of just five states that allow permanent no-excuse absentee voting, meaning any registered voter may request to receive an absentee ballot automatically for all future elections. Absentee voters, whether permanent or for the particular election only, receive a ballot in the mail several weeks before an election and have the option of either mailing their ballot to the registrar's office—postmarked on or before Election Day—or dropping their absentee ballot off at any polling station in their county. In the 2008 primary election, nearly 42% of all votes were cast by absentee ballot.[19]

The variation in voting systems between counties provides a natural setting for our study. On February 5, 2008, we dispatched 119 data collectors to monitor the flow of voter traffic at 30 polling stations—10 polling stations in each of three neighboring Bay Area counties. Each county employed different voting technology: Alameda County voters submitted paper ballots into optical scanners at each polling station, San Mateo County voters used the Hart Intercivic eSlate DRE machine, and Napa County voters cast paper ballots that were collected and later scanned at a central location. We did not study voting locations using the InkaVote system.

Selection

California's 58 counties are comprised of 25,090 precincts. Our sample of 10 polling stations in each county allows us to observe small effects among the variables across counties.[20] The 30 polling stations were selected in geographic clusters that were stratified by income. The decision to organize our sample into clusters rather than strict randomization was purely economical. The clusters were arranged such that the support van could make a stop at each polling station in less than an hour. We stratified our precincts by income to produce an adequate sample of lower income communities in Napa and San Mateo counties and then randomly chose precincts within these strata that were representative of different income lev-

---

[16] "Top-to-Bottom Review," California Secretary of State website, http://www.sos.ca.gov/elections/elections_vsr.htm. Despite their decertification, Secretary Bowen allowed 14 counties to use Premier (formerly Diebold) AccuVote touch-screen DRE machines as their state-mandated disability-accessible voting machine.

[17] "Voting Systems Used by Counties," California Secretary of State website, http://www.sos.ca.gov/elections/voting_systems/systemsinuse_110606.pdf.

[18] As the nation's largest voting jurisdiction, Los Angeles has a special contract with Election Systems & Software, Inc. (ES&S) as the only consumer of the InkaVote Plus system, serving Los Angeles county voters since 2003.

[19] See "Voter Participation Statistics by County," in *Statement of Vote—February 2008 Primary Election*, California Secretary of State's Office, 2008, http://www.sos.ca.gov/elections/sov/2008_primary/contents.htm.

[20] The assumptions that motivated our power analysis—using Cohen's effect size conventions—proved to be conservative. We underestimated the overwhelming turnout that occurred on February 5th, an error that strengthened the robustness of our data.

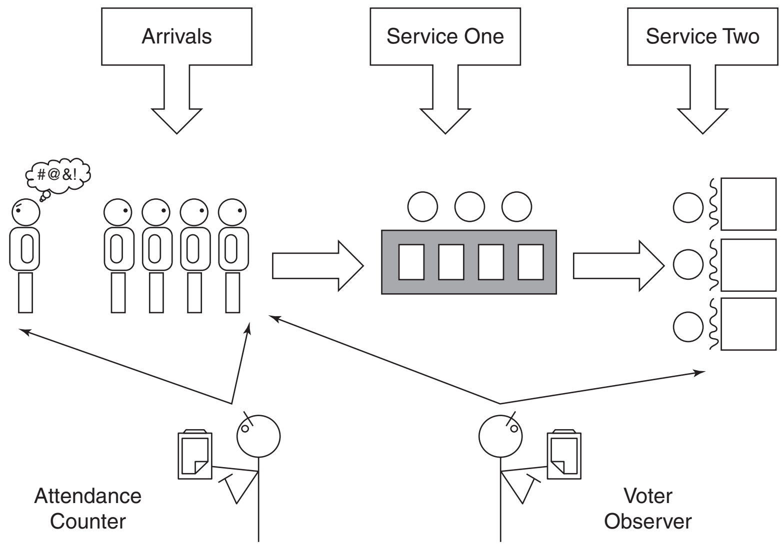

FIG. 2. Data collection.

els.[21] The result is a binary variable—"high income" and "low income"—where "high income" is defined as median household income greater than or equal to California's three-year-average median household income of $55,864.[22]

The geographic clusters and the income strata did not perfectly overlap. Thus, each geographic cluster contained at least one "high income" and at least one "low income" polling station. As far as feasible, we selected polling stations within each geographic cluster that represented the heterogeneity that existed in the county as a whole.

Data collection

The study was structured to record: 1) the arrival rate of voters, 2) the attrition rate of the line, 3) the rate of Service One, 4) the rate of Service Two, and 5) general poll worker characteristics.

Our design called for two data collectors. See Figure 2. The first data collector recorded the arrival of voters in ten-minute intervals from the opening of the polls at 7:00 a.m. to their close at 8:00 p.m. We noted the arrival of 11,858 people. In addition, we recorded the number of people that reneged. "Reneging" is the technical term for a person who arrived, stood in line, and then left without immediately voting. The data collectors were instructed to mark all those people who got out of line as a "renege" unless they were absolutely sure that they got back into line. In our sample, 225 people reneged (1.89% of all arrivals).

The second data collector noted the exact time in seconds that every fifth voter made contact with the poll workers at the check-in table, physically left the check-in table, physically arrived at the voting booth or station, and physically left the voting booth or station. In Alameda County, the time was noted when the voter scanned his or her ballot. Thus, we were able to calculate the service times for 2,160 voters.

Our data gatherers also administered a short questionnaire to 153 poll workers, or 89% of all observed poll workers, with simple demographic and poll worker history information.[23] This allowed us to link the characteristics of each poll worker to the voters they assisted and, thus, to measure the effects of a poll worker's age, race, sex, income level, em-

[21] Because Napa County has both a significantly higher median income and, more interestingly, a significantly more polarized income distribution than either of the other two counties, we were careful to prevent this from driving our results.

[22] The 3-year-average was calculated from 2005–2007, in current 2007 dollars, as reported by the U.S. Census Bureau. "Three-Year-Average Median Household Income by State: 2005–2007," (U.S. Census Bureau, 2007).

[23] We asked the following eight multiple choice questions in the survey: (1) Are you male or female?; (2) Which of the following age groups do you fit into?; (3) Including today, how many times have you worked as an election official?; (4) What is your current occupational status?; (5) What is the highest level of education you have completed?; (6) Do you live in the precinct of this polling place you are stationed at?; (7) What is your current marital status?; (8) What is your race?

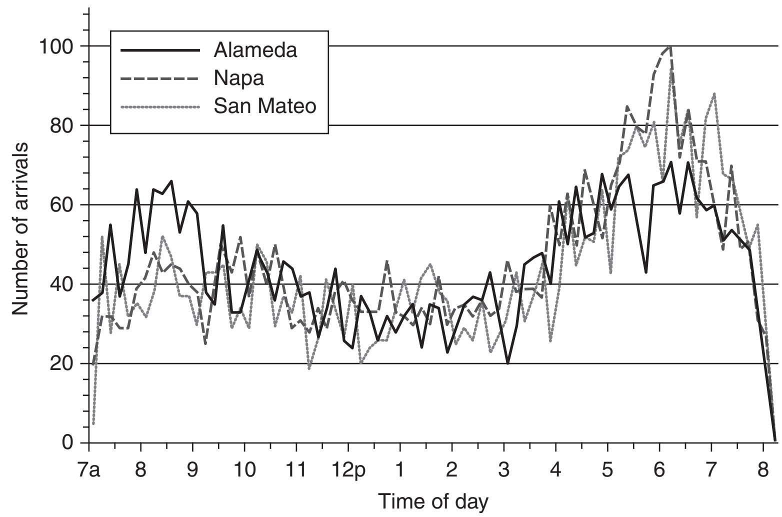

FIG. 3. Distribution of arrivals. This figure previously appeared in "Data for Democracy: Improving Elections Through Metrics and Measurement," *Pew Center on the States*, Washington, D.C. (2008), p. 5, used with permission.

ployment status, education level, and experience on his rate of service.

FINDINGS

Our study yielded data on the arrival rate of voters, the number of minutes voters spent at Service One and the number of minutes voters spent at Service Two. We also observed actual line lengths and attrition. These data shed some light on the processes inside the walls of a polling place.

Voter arrivals

Unsurprisingly, we found that voter arrivals are not randomly distributed across the day. Knowledge about the flow of traffic is significant for local election officials attempting to make allocation decisions. Admittedly, the supply of some resources, such as voting machines, cannot realistically be adjusted during the day, but other resources, such as poll workers, could be. Most officials look to general data on turnout when allocating resources for Election Day, but there are few data quantifying when voters actually arrive throughout the day.[24] Figure 3 shows the density function for the number of arrivals in our sample.

The uniformity we observed both within and between counties suggests that the arrival rate of voters may be predictable. This is not to say that all jurisdictions would see this same double-humped pattern; however, some discernible pattern is likely.[25] In our sample, voters arrived in about the same distribution: a small surge during the first hour of operation, a noticeable dip during the lunchtime hours of 12:00–2:00 p.m., and then a wave of voters ($\approx$150% increase) between 5:00–7:00 p.m. Arrivals sharply declined after 7:00 p.m. and only 3% of voters cast a ballot in the final 30 minutes of the day.

Although consistent with expectations, our more precise information is useful for a number of rea-

[24] In 2006, Maryland used electronic poll books to track the time people checked in at the registration table. These data are somewhat different from an arrival rate as they do not account for voters who arrive when there is a queue, only tallying voters once they reach the check-in table. See Dan Seligson, "How Data Has Improved Election Management," in *Data For Democracy, Pew Center on the States*, Washington, DC (2008), p. 5.

[25] We are cautious about the external validity of our findings; anecdotal reports from Florida during the November 2008 general election suggest that lines were consistently longer in the early morning hours, but consistently very short in the late afternoon and evening. Parts of Florida, of course, contain large numbers of retirees. There is evidence that voters in other jurisdictions arrive in a pattern quite similar to what we observed. See, for example, K. Dow, "Study of Voter Flow at the 2006 General Election, Columbia County, NY," New Yorkers for Verified Voting (2007), http://tinyurl.com/Dow-Voter-Flow-pdf; D. Magleby and H. Christensen. "When People Vote and What Difference It Makes." Prepared for presentation at the Annual Meetings of the American Political Science Association, New York, New York, September 1-4, 1994 (Figures 1 and 2); D. Fuchs and J. Becker, "A Brief Report on the Time of Day When People Vote," *Public Opinion Quarterly*, Vol. 32, No. 3 (1968), pp. 437–40 (Table 1).

sons.[26] For example, armed with the knowledge that more than 25% of all voters arrive between 5:00–7:00 p.m., local election officials might establish a policy that encourages poll workers to take their dinner break before or after the evening rush or allocate more poll workers to stressed precincts in those evening hours. In polling stations that use paper ballots or in small locations, such as someone's garage, election officials might also instruct poll workers to reduce the number of privacy booths during the middle of the day, when the arrival rate is lowest. These privacy booths are relatively simple to collapse and in polling stations where voters are standing shoulder to shoulder, creating more space would allow them to move around more freely and allow for greater voter privacy.

Data on voter arrivals are simple to capture and highly useful. We hope our example will prompt scholars and elections officials alike to gather such data.

Service One

Service One includes all of the activities that occur at the check-in table such as checking ID, looking up the voter's name in the voter registration book, and producing a ballot or voting card. These tasks aim to verify the voter's identity and provide him or her with either a voting card for a direct recording electronic (DRE) voting machine or a blank paper ballot. The most important independent variable in Service One is the poll worker. Whatever affects the performance of a poll worker should have a measurable effect on the efficiency of Service One and thus the voting process at large.

Table 1 shows the results of a multivariate regression that models the Check-in Service Time of voters in our sample on a host of operational and social variables. The reported standard errors are robust standard errors, clustered by polling station. The first variable in the table is a dummy variable that shows if the poll worker that served the voter at the check-in table was over 55 years old. The second variable is a dummy variable that shows if the poll worker that serves the voter is an experienced poll worker. In order to show robust significance, we also include an interaction variable between age and experience. Because earlier studies have shown that older poll workers are generally less comfortable with DRE machines than younger poll workers, we include an interaction variable between age and the San Mateo County dummy variable to account for the possibility that this discomfort would be driving a resultant slower service time.[27] We do not include an interaction between the San Mateo dummy and experience. This election was the first time San Mateo County used these voting machines, so any interaction between these dummies would be coincidental.[28]

Several of our early assumptions were confirmed by our data. With a check-in baseline of 1 minute 50 seconds, voters who later cast their ballots using a voting machine reserved for disabled voters re-

[26] Information about voter arrivals may be important for political reasons. For example Republican strategist Karl Rove complained that 15,000 voters were deterred from voting in Florida's panhandle in 2000 after several network news programs called the election for Al Gore ten minutes before the polls closed. See K. Jamieson and P. Waldman, eds., *Electing the President, 2000: The Insiders' View*, Philadelphia: Univ. of Pennsylvania Press, (2001), p. 232. Rove was convinced that a large percentage of voters procrastinated until the final hour before going to the polls. Our data suggest that that may not have been the case. See, however, Crespin, M. and R. Vander Wielen. "The Influence of Media Projections on Voter Turnout In Presidential Elections from 1980–2000," Prepared for presentation at the Annual Meetings of the Midwest Political Science Association, Chicago, April 25–28, 2002, crespin.myweb.uga.edu/Mediaresubmission.pdf; M. Delli Carpini, "Scooping the Voters? The Consequences of the Networks Early Call of the 1980 Presidential Race," *Journal of Politics* Vol. 46, No. 3 (1984), pp. 866–85, for evidence of depressed turnout when media projections are substantially early, i.e. two or more hours before the polls close.

[27] Initial assessments have found that poll workers are not particularly confident in their ability to avoid mistakes. A post-primary survey of poll workers in Cuyahoga County, Ohio found that only 42 percent were satisfied with their own performance on election day. See K. Frakas, "Study: Poll Workers, Not Machines, Build Voter Confidence," *The Plain Dealer*, Cleveland (February 21 2007). Hall, Monson and Patterson observe that poll workers over the age of 55 were significantly more likely to have start-up or shut-down problems with electronic machines than younger poll workers. T. Hall, Q. Monson, and K. Patterson, "Poll Workers and the Vitality of Democracy: An Early Assessment," *PS: Political Science and Politics* Vol. 40, No. 4 (2007), pp. 647–54. See also M. Alvarez, M. Llewellyn, and T. Hall, "Are Americans Confident Their Ballots Are Counted?," *Working Paper* No. 49, Caltech/MIT Voting Technology Project (2006).

[28] The county dummy variables act as proxies for voting machine technology. It is certainly possible that this could be picking up other variation between the counties, but each regression includes local demographic characteristics, such as population density, education, and family income, as well as election characteristics, such as arrival rate, provisional balloting, and absentee drop-offs to control for alternative explanations for the variation.

TABLE 1. MODEL OF FACTORS AFFECTING SERVICE ONE

Variable	*Coeff.*	*Robust SE, clustered by polling station*	p > |t|
Poll worker is over 55 years old	−0.108	0.187	0.570
Poll worker has experience	0.518	0.256	0.054
Disabled voter	0.711	0.169	0.000
Absentee	−0.902	0.371	0.023
Provisional	1.778	0.447	0.001
% of voting age population			
African American	−0.038	0.012	0.003
Hispanic or Asian	0.008	0.006	0.743
Median household income (per $10k)	−0.185	0.031	0.000
Alameda County dummy	0.959	0.420	0.032
San Mateo County dummy	0.587	0.385	0.140
Interaction of age and experience	−0.122	0.273	0.657
Interaction of age and San Mateo dummy	−0.466	0.253	0.078
Constant	1.839	0.399	0.000
R^2	0.178		
N	1,420		

Other non-statistically significant control variables in the model:
- Voter experienced ballot error later
- San Mateo voter opted for paper ballot
- Percent of population with bachelor's degree or higher
- Population density
- Number of people arriving
- Number of people reneging
- Number of poll workers at the check-in table
- Number of people standing in line at the check-in table
- Number of voting stations/privacy booth set up for use.

quired 43 more seconds of the poll worker's time.[29] Voters who cast provisional ballots were at the table twice as long as traditional voters—an additional one minute and forty-seven seconds—while absentee voters stood at the table for half the time of traditional voters—a total of fifty-six seconds.[30] There was a significant increase in poll worker efficiency at polling stations in high-income neighborhoods. For each additional $10,000 of median household income in the Census block group most closely aligned to the precinct, voters could expect to stand at the check-in table for eleven fewer seconds. Thus, some voters in neighborhoods of Napa County where median household income is $46,000 experienced an average service time at the check-in table that was two full minutes longer than voters in other neighborhoods of Napa County where median household income exceeds $200,000.

Two findings surprised us and were completely contrary to our *a priori* hypotheses. First, experienced poll workers were not more efficient, even when we controlled for age. To the contrary, the service time of poll workers with experience during at least one prior election was correlated with an extra thirty-one seconds of voter check-in time.[31] With all of the changing rules and regulations between elections, it may be the case that poll workers with experience are more likely to become confused about applying a rule that has changed several times. It may also be the case that experienced poll workers are more relaxed and likely to visit with voters whom they have served over the past several

[29] Our data collectors were instructed to check a box when a voter used the voting machine reserved for disabled voters. We did not record whether (or how) individual voters were disabled. Because some non-disabled voters also used the voting equipment reserved for the disabled, our results likely underestimate the effect of disabled voters on the process.

[30] In California, absentee voters are able to fill out their ballot at home and drop it off at any polling place in the county where they are registered.

[31] This finding fell just outside the conventional 0.05 alpha level (p = 0.054). Thus, at the very least, we cannot reject the counterintuitive hypothesis that experienced poll workers are less efficient.

years.[32] Both of these explanations appear plausible and further research is warranted to determine what factors may be driving our finding.

Second, we found that the number of poll workers at the service table is positively correlated to the length of the lines that we observed. In other words, when there were more poll workers at the check-in table, lines were longer. There is almost certainly some level of simultaneity between the variables and so we are uncertain whether additional poll workers get in each other's way or, if upon seeing a long line forming, additional poll workers run to the table to help their co-workers.[33] In our model, additional poll workers did not make the process any more efficient. Thus, regardless of why additional poll workers are positively correlated to a longer line, extra human capital may not be the most efficient solution to the operational problem of long lines. One possible explanation is that poll workers are poorly trained; researchers and policymakers have recently evaluated poll worker training programs as a way to assess poll worker efficiency. For example, the Election Assistance Commission created a briefing book of "best practices" for poll worker training based on surveys they conducted of local election officials throughout the United States.[34]

Perhaps a more realistic explanation of the surprising reverse correlation between the number of workers and quick check-in is that poll worker inefficiency results from a system design failure. In a polling station, the tasks at the check-in table are very simple and not well suited to many operators. For example, it only takes one person to scan for a voter's name on an alphabetized list of registered voters. Thus, an additional poll worker may actually hinder the process.

In essence, local officials face two distinct poll worker problems: a managerial separation between supervisor and employee and an institutional knowledge gap between skilled and unskilled poll workers. Election officials oversee an institution with history and rules governing its operation, but lack the time to devote to long-term training. Poll workers, on the other hand, are not polling place specialists ready to complete discrete and specialized tasks, but are often untrained or poorly-trained workers. Election officials and academics are aware that this is the case, yet they lack the data necessary to evaluate its effect on poll worker consistency and efficiency. Our study provides an introductory analysis of these effects. Nonetheless, additional research should be conducted to explore the full extent of this difficult problem.

Service Two

The second stage in the voting process is a voting platform that allows voters to interact with the ballot. Service Two incorporates the length and design of ballots, the design and model of a voting machine, and the number of voting machines or privacy booths in use. For every voter, a unique ballot is marked, verified, and tabulated. Different voting technologies combine these processes in different ways. For example, most voters in the United States mark a paper ballot that they feed into an electronic optical scanner. This scanner combines the process of verification and tabulation. In a growing number of jurisdictions, voters use direct recording electronic (DRE) machines that combine ballot marking, verification, and tabulation all into one act.[35]

---

[32] Poll workers in Utah and Ohio reported that "being with people I enjoy" was the third most important motivating factor for their service (behind "doing my share" and "it is my civic duty"). See D. Magleby, Q. Monson and K. Patterson, "Evaluating the Quality of the Voting Experience: A Cross Panel Pilot Study of the November 7, 2006 Election in Franklin County, OH, Summit County, OH and the State of Utah." *Center for the Study of Elections and Democracy*, Brigham Young University (2008), http://csed.byu.edu/ (follow "Publications" hyperlink; then follow "By Year" hyperlink).

[33] We regressed line length on the number of poll workers at the check in table, controlling for service time, county effects, poll worker demographic effects, arrival and reneging rates only to find a 0.095 positive correlation ($p < 0.05$). Yet when we reversed the two variables of interest—regressing number of poll workers on line length and the host of other variables—we saw a similar 0.036 positive correlation ($p < 0.025$).

[34] See U.S. Election Assistance Commission, "Successful Practices for Poll Worker Recruitment, Training and Retention", Washington, D.C. (2007); See also S. Mockabee, Q. Monson, and K. Patterson, "Evaluating on-Line Training: A Study of Poll Worker Training in Butler and Delaware Counties, Ohio for the March 4, 2008 Presidential Primary Election," *Center for the Study of Elections and Democracy*, Brigham Young University, forthcoming, Pew Center on the States (2009).

[35] In the aftermath of the 2000 presidential election in Florida, the United States Congress enacted the Help America Vote Act which, *inter alia*, requires states to replace old voting systems (specifically the problematic punch card system) with new, electronic voting machines that allow voters to verify their ballot before it is officially counted. The result has been a nationwide trend toward optical scanners and DRE machines. In 2000, 41.9% of registered voters lived in a county that used either an optical scanner or DRE machine. See "Voting Equipment Summary by Type as of 11/07/2000," *Election Data Services*, (2004). In 2008, that number increased to 88.8%. See "2008 Voting Equipment Study," *supra* note 11.

Table 2. Difference Between Percent of Ballots Cast

	Total Ballots Cast	*In Precinct on Election Day*	*Major Party Presidential Candidate*	*Prop 91*	*Prop 92*	*Prop 93*	*Prop 94*	*Prop 95*	*Prop 96*	*Prop 97*
San Mateo	211,697	54.24	82.24	89.24	92.02	93.65	94.63	94.61	94.26	94.26
Alameda	428,930	51.48	81.21	89.57	92.24	93.12	94.28	94.18	94.12	93.70
Difference			−0.97%	−0.33%	−0.22%	0.53%	0.35%	0.43%	0.14%	−0.56%

Source: Alameda and San Mateo Registrars of Voters' websites, 2008.

Voting technology is our key independent variable on waiting times. Each county we observed employed a different voting technology. In Alameda County voters used paper ballots ("complete the arrow") that were scanned at each polling station. Napa County also used paper ballots, but stored them to scan at a central location at the end of Election Day. San Mateo County used DRE machines manufactured by Hart Intercivic.

Voters in Napa County spent the fewest minutes casting their ballots—about three minutes in the voting booth. Voters in Alameda County spent three minutes and twenty-five seconds in the voting booth while voters in San Mateo County spent approximately four minutes and thirty seconds each to complete their ballot, controlling for ballot length.[36]

How much is the technology itself the cause of these differences? Possible explanations include the tendency to cast an incomplete paper ballot, the length of the ballot, and the various socio-economic characteristics that differ by precinct and by county.

Incomplete Ballots. One possible explanation for why DRE machines could take longer to use is that DRE users voted for more ballot issues than paper ballot voters. When a voter using a paper ballot decides to skip one or two measures, he or she can still submit the ballot to the scanner without incident. On a DRE machine, if a voter decides to skip a measure, he or she will be reminded that the ballot is incomplete and given the opportunity (with the press of a button) to go back and complete the ballot. Therefore, casting a complete ballot on a DRE is simpler for the voter and takes less time. One possible explanation for longer vote times at DRE machines, therefore, may be that a larger ratio of voters cast a *complete* ballot than those who vote via paper ballots. If true, the observed inefficiency of the DRE machines may be justified to the extent one regards avoiding drop-off as desirable. However, we did not observe a disparity in ballot completion rates between counties.

The biggest difference between precincts is the percent of ballots cast for major party presidential candidates. This entire difference can be explained by the fact that voters in Alameda County were more than four times as likely to vote for Green Party candidate Cynthia McKinney and twice as likely to vote for a Libertarian candidate than voters in the other two counties. Among the seven state ballot initiatives, there was less than 1% difference in votes cast between the counties and, contrary to the above hypothesis, three of the seven ballot propositions saw a higher overall percent of votes among paper ballot voters in Alameda County than among voters using a DRE machine in San Mateo County. See Table 2.

Thus, it does not appear to be the case that ballot completion rates are driving the difference in service times.

Socio-Economic Status Effect. In addition to mechanical differences between counties, demographic dissimilarities may be a good predictor of how long it takes to vote and provide a better sense of the contributing effect of voting technology. For each of our 30 monitored precincts, we gathered Census data on precinct-level median household income, percent with a bachelor's degree or more, population density, and the percent of voting age population that is African American, Hispanic, and Asian. Furthermore, we know the sex of each voter in our sample and whether he or she used voting

[36] We controlled for ballot length in two ways. First, we held constant the number of race and ballot propositions. Second, we held constant the number of words on each ballot (including instructions). Our findings were the same using either method.

TABLE 3. MODEL OF FACTORS AFFECTING SERVICE TWO

Variable	*Coeff.*	*Robust SE clustered by polling station*	p > |t|
Alameda dummy	(dropped)		
San Mateo dummy	1.668	0.471	0.002
Ballot length	−0.130	0.166	0.442
Arrival rate	0.002	0.014	0.903
Number of poll workers	0.022	0.066	0.740
Number of privacy booths (or DRE machines)	−0.040	0.062	0.523
Voter experienced machine error	0.702	0.946	0.465
Male or female	−0.025	0.149	0.870
Voter used disabled voting equipment	1.835	3.117	0.561
Voter cast a provisional ballot	1.269	0.330	0.001
Voter cast a paper ballot in San Mateo	−2.799	0.201	0.000
Median household income ($10k)	−0.048	0.052	0.360
% of precinct with BA degree or more	0.013	0.007	0.071
% voting age population			
African American	0.035	0.019	0.073
Hispanic or Asian	−0.005	0.010	0.596
Population density (per sq. km.)	−0.027	0.074	0.714
Constant	3.691	0.660	0.000
R^2	0.111		
N	1,485		

equipment reserved for the disabled. We added several mechanical control variables such as machine errors, provisional ballots, arrival rate, number of poll workers, number of voting booths, San Mateo voters that used paper, and controlled for ballot length in order to analyze the effect of these operational and socio-economic factors on voting time. See Table 3. In order to account for possible variation caused by a mixture of data from distinct units of analysis, we report robust standard errors that are clustered by each polling station.

None of the socio-economic characteristics themselves have a significant effect. However, the cluster of polling stations in neighborhoods that exceeded California's median household income saw voters cast their ballots on average 32 seconds faster than those who live in neighborhoods that fell below California's median household income of $55,864, regardless of the voting technology.

After controlling for 14 operational and social variables, including ballot length, voters who used a DRE took on average one minute forty seconds or 45% longer to vote than voters using paper ballots. Thus, it is likely that the technology of the voting machine *itself* explains much of the variation in voting times. To test this hypothesis, we compared DRE voters in San Mateo with their in-county peers who opted to use paper ballots. On average, DRE voters spent two minutes forty-eight seconds or 76% longer to cast their ballots than their paper ballot counterparts. The difference is so distinct that even with a small sample ($n = 34$), the significance was very high ($p < .001$).

This finding is important inasmuch as DRE machines are significantly more expensive than paper ballot printing costs and privacy booths, a fact that may lead to fewer stations being available for voters in DRE jurisdictions.[37] Thus, not only might it take longer to use a DRE machine, but there may be fewer of them available, compared to privacy booths for paper ballot voters. In this case, a polling station that employs DRE voting machines would be more likely to see a line form on Election Day.

[37] In extreme cases, the cost of DRE machines is so high that states are forced to consolidate precincts and yet still face a shortage of voting stations on Election Day. The implications of this downsizing was apparent in Utah County during the 2008 primary election where 107 precincts were consolidated into 59. See M. Rigert, "Long Waits at the Polls Mark Election," *Daily Herald*, February 7, 2008.

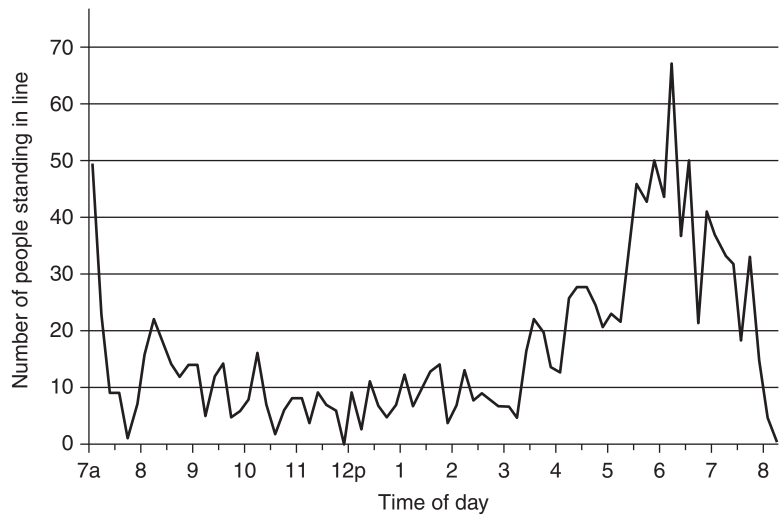

FIG. 4. Aggregate length of lines.

Line length

In addition to tallying arrivals, we counted the number of people who were standing in line at the end of each ten-minute interval throughout Election Day. In queuing theory, lines are a product of incongruence between arrivals and service times. As we noted above, the arrival of voters we observed was not equally distributed across the day. The rates of both Service One and Service Two, however, were constant across the day; poll workers did not pick up the pace as arrivals increased, nor did voters fill out their ballots more quickly when lines started to form.

These findings suggest that there is a threshold of arrivals beyond which the constant service rate in our sample was inadequate. Although the individual lines that we observed were not particularly long, our data reveal that lines were longest when the rate of arrivals was at its peak, namely, between 5:00–7:00 p.m. Figure 4 represents the aggregate line length of every polling station in our sample.[38]

Elections officials are likely to view this data favorably, and we applaud the registrars of voters and poll workers in our sample. As concerned citizens, we were quite relieved that we did not encounter hours-long lines. Yet as researchers, this lack of long lines prevents us from making any concrete normative claims about individual voter behavior. Despite this, our data is important inasmuch as they confirm the hypothesized correlation between voter arrivals and line length, and because they provide baseline information against which the existing research on lines can be tested. Furthermore, analyzing line length at the aggregate level is useful because we can measure correlations between line length and other variables, such as attrition.

Attrition

Our final data relate to the deterrent effect of lines. In our sample of 11,858 arrivals, 225 people (1.89%) "reneged."[39] Figure 5 illustrates the percentage who reneged, depending on the line length. At the end of each 79 ten-minute interval during the thirteen-hour Election Day, we noted the number of people standing in line—anybody waiting to be served by a poll worker, but not yet served. We matched these 2,181 data points with our data on reneging for each corresponding ten-minute interval.[40] We then created a four-by-two frequency table to calculate the rate of observed reneging in

[38] The median line length in our sample was zero and the longest line in any one polling station reached thirteen people. Only fifty-one voters (of nearly 12,000) waited in line for more than ten minutes.

[39] We were unable to track people once they left the polling station. Our data collectors were instructed to count an instance of reneging if the person who left the line did not immediately return. If a person left the line and was gone for several hours, our data collectors likely did not remember him. Thus, such people were tallied as having arrived twice and reneged once.

[40] We excluded data for one polling station where several different lines formed (and overlapped) for access to different precinct tables because we were not confident in the accuracy of the data. We also excluded 14 data points from another polling station where our observers failed to record these data.

each of three categories—from less crowded to more crowded—where 0 = nobody got out of line, and 1 = at least one person got out of line and did not immediately vote. Each data point in Figure 5 represents the percentage of cases in each "crowdedness" category where reneging = 1. Because there were no systemic problems causing long lines in our sample, the number of observations is heavily skewed towards zero. Three-fourths (or 1,616) of our data points, in fact, occur where line length = 0. Of the remaining one-fourth of our data points, 24.5% occur when lines were between one and five people long. There are 31 data points (1.49%) when lines were between 6 and 10 people long, and only three ten-minute intervals (0.1%) where line lengths exceeded 10 people (with a maximum of 13).

Figure 5 plots the observed rate of reneging in our sample. We see a positive correlation between line length and the likelihood of reneging. This unsurprising observation confirms the conventional wisdom about voter behavior in these circumstances.

Because of our unevenly distributed sample, the data points in the first two categories are statistically significantly different—marked by black dots and a solid line—while the data points in the third category is not—marked by a white dot and a dotted line. More research on this topic is warranted. With more information from a sample with a wider range of line lengths, we would be able to evaluate this observed correlation between line length and reneging in more detail and with more confidence.

Whether or not the rate of reneging is related to polling station operation variables, the fact that 1.89% of arrivals reneged is, in itself, noteworthy. A handful of recent, closely watched elections were decided by a fraction of a percent. For example, the 2008 Minnesota Senate race, the 2004 Washington gubernatorial race, and the 2000 presidential race in Florida were all decided by 0.01% or less. Presuming that some, or even many of the 1.89% of our sample that reneged returned to vote later in the day, and presuming that the rate of reneging is equally distributed between political parties (there is no reason to believe otherwise), our observation is still cause for concern. Reneging is important both as a measure of the internalized cost of standing in line as well as because it may affect turnout and election results.

DISCUSSION AND FUTURE RESEARCH

Our study suggests that there is predictable variation in each of the three stages of voting and our models identify key variables for improving the efficiency of polling station operations. We recognize the limitations on our data and caution against extrapolating from our findings. Elections vary from jurisdiction to jurisdiction and from election year to election year and our study is limited to three counties in one single state during one single primary election; we do not measure the effect of voter ID laws on the check-in process; we do not measure

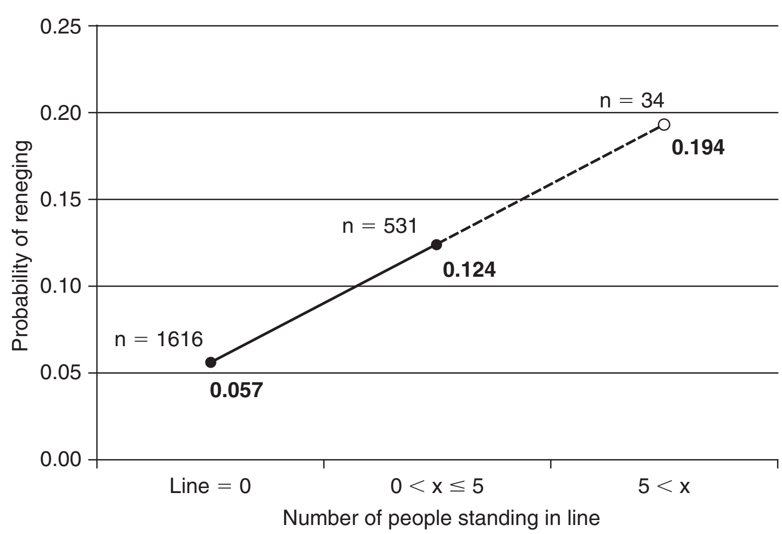

FIG. 5. Probability that voters reneged.

the effect of early voting on turnout; and our sample is limited to a model of DRE machine that represents just 20% of the DRE market.

However this article does paint an introductory picture of operation inside a polling station's black box and illustrates the need for further and more extensive data collection if we are going to have any idea how the process actually works. Our initial systematic look at the inner workings of the polling place highlights the importance of evaluating polling station operations as a three-step process: voter arrivals and a series of two services. The bulk of research on lines focuses on the voting machine, and how their poor allocation may be correlated with depressed turnout. This research fails to account for the very important step—and potential bottleneck—of "checking in" and any systemic problems that may exist there. Our study finds that it is a critical component of any such research.

In addition, we have collected the most comprehensive data to date on the operations and inner workings of the polling place. Post election survey questions related to voting are typically unspecific about the process and in any event rely on voter recall. They therefore lead to unreliable responses. For example, the question "how long did it take you to vote" may elicit a response that relates to the time a voter spent in the voting booth (Service Two), or the time spent inside the polling station (Service One + Service Two), or the time spent away from work (Service One + Service Two + commute to and from the polling place). Our goal was to improve on earlier studies that either collected partial information or made assumptions about unobserved voter behavior by filling in data gaps.

This study aims to be a bridge between the problem of long voting lines and cost-effective solutions. Despite recent advances in the field of election administration, from centralized voter registration lists to electronic voting machines, the process of voting is still largely evaluated against anecdotes and assumptions. Our study brings new data to bear on this process and also presents a model for future data collection and analysis on the operation of polling stations—America's retail outlets for democracy.

ACKNOWLEDGMENTS

The authors are grateful to Henry Brady and the Survey Research Center at the University of California, Berkeley who funded this research, and to the Pew Center on the States' electionline.org for their generous assistance. We are especially indebted to Morgan Hanger and Jonny Morris, our co-investigators, for their inimitable contributions to the overall project, and to the 119 data collectors, most of them UC Berkeley students, who diligently worked from 5:00 a.m. to 10:00 p.m. on Election Day to collect this dataset. Special thanks to Madonna Camel, Cindy Palwick, Eva Seto, and Jon Stiles of the University of California, Berkeley Survey Research Center for their help in organizing and implementing this field study, and to Sean Gailmard for his generosity in helping us recruit data collectors. Barry Burden, Doug Chapin, Bill Edelstein, Sean Greene, Paul Gronke, Ed Johnson, John Koomey, Lauren Lambie-Hanson, Brent Nakamura, and Mike Urbancic provided very useful comments and suggestions for which we are thankful. We wish to acknowledge the cooperation of the three registrars of voters in whose counties this study was conducted: Dave Macdonald of Alameda County, John Tuteur of Napa County, and Warren Slocum of San Mateo County.

Address correspondence to:
Douglas M. Spencer
Survey Research Center
University of California, Berkeley
2538 Channing Way
Berkeley, CA 94720

E-mail: dougspencer@gmail.com