Time Tax:

Which Groups Wait in the Longest Lines on Election Day?

Stephen Pettigrew
pettigrew@fas.harvard.edu
March 22, 2013

Abstract

In the last few election cycles, long lines to vote seem to have become the new normal in some areas. Anecdotal evidence suggests that long waiting times disproportionately affect minority and lower-income neighborhoods. Previous work, however, has been plagued with data limitations and concerns with omitted variable bias. I overcome these problems by using large surveys from two recent elections to test the hypothesis that minority and low-income individuals are subject to a "time tax" when voting, which white and high-income voters are not. Using a hierarchical Bayesian regression model, I demonstrate that the racial composition of a neighborhood has a tremendous effect on the average amount of time voters in that neighborhood wait in line. Voters in areas that are just 30% black will wait, on average, three times as long to vote as voters in a 5% black area. The results on whether there is a similar effect for income are slightly less conclusive, and will require further investigation. This paper begins to answer questions which have become the source of attention in the political science and legal communities, as well as the media.

1 Introduction

In the November 2012 general election, approximately one in every ten voters waited in line to vote for more than a half hour. About 3.5 million voters waited in excess of an hour, with some standing in line for as many as three or four hours. Long lines at the polls became such a hot topic in the media that President Obama acknowledged in his victory speech that the issue was one that needs to be fixed.[1] The news media reported countless horror stories about multiple-hour lines to vote and voter attrition due to waiting times. The problem was particularly pronounced in Florida, where lines stretched so long during the early voting period that one person waited nine hours to cast a ballot and live-blogged the whole experience.[2] Long lines at the end of the early voting period prompted election officials in three Florida counties to provide for in-person absentee balloting in the two days after the end of early voting but before Election Day.[3] Other states, including Texas, Virginia, and Wisconsin, had lines during early voting and on Election Day which were longer than most would deem acceptable.[4,5]

Despite to growing attention in the media over the past decade to problem of lengthy lines at polling stations, very few political scientists have weighed in on the subject. Most of the academic work in this area makes policy recommendations about election administration. This area of research tends to be in the format of a comparative cost-benefit analysis of two or more voting technologies, weighing the impact each technology has on waiting time and

[1] "Transcript of President Obama's Election Night Speech." Available at: http://www.nytimes.com/2012/11/07/us/politics/transcript-of-president-obamas-election-night-speech.html?pagewanted=all. Accessed Jan. 21, 2013.

[2] "Photo of My 9 Hour Wait to Vote in Florida." http://www.dailykos.com/story/2012/11/03/1155089/-Photo-of-My-9-Hour-Wait-to-Vote-in-Florida. Accessed Jan. 21, 2013.

[3] "Long Lines for Early Voting before Tuesday's Election" http://www.cnn.com/2012/11/05/politics/election-voting/index.html. Accessed Jan. 21, 2013.

[4] "Some Glitches, Long Lines, and Voter ID Confusion at Polls." http://www.jsonline.com/news/statepolitics/five-hours-into-election-long-lines-some-problems-surface-ur7hbv1-177502851.html. Accessed Jan. 21, 2013.

[5] "Long Lines but Few Snags in US Election." http://www.cnn.com/2012/11/06/politics/election-voting/index.html. Accessed Jan. 21, 2013.

the costs of decreasing line length by adding more machines. Empirical case studies which draw their inferences from the data of a handful of counties in a limited geographic region are also common in research about line length. Legal scholars have also weighed in on the subject and tend to put forth the argument that long lines for voting function as a quasi-poll tax. They suggest that lines tend to be longer in minority and poor areas and that election administration policies of the states, whether in a deliberate or unintended manner, unfairly discriminate against voters in these areas.

This current body of work about line length in elections suffers from two important shortcomings. Much of the work, particularly from law review journals, lacks a high degree of statistical or methodological rigor. Correlation tables which show that African-Americans and Hispanic wait in longer lines than whites provide initial evidence that there may be some form of discrimination going on. But without simultaneously taking into account important factors like shifts in turnout, we cannot draw strong conclusions about what types of solutions may be required to alleviate these differences. The second shortcoming, which typically characterizes the more methodologically sophisticated literature, is one of limited data, derived from their use of either case-study or simulation data.

The focus of most of the work to date is on explaining the variation in waiting times, based on factors such as voting technologies and poll worker efficiency. To my knowledge, there has been no empirical study with the primary goal of identifying differences in waiting times based on demographic characteristics. In this paper I set forth with the purpose of filling this gap in the literature. Specifically, I test the hypothesis that racial and ethnic minority groups are subject to longer lines to vote than the white electorate and that these lengthier waits are not easily explained by other factors like shifts in turnout.

I accomplish this goal by addressing the key shortcomings of previous work. Using data from the 2008 and 2012 Survey of the Performance of American Elections (SPAE) and the 2008 Cooperative Congressional Election Study (CCES), I study individual wait times

from 85% of all the counties in the US. This broad cross-section allows me to make inferences which are more generalizable than those generated from case studies. I analyze survey responses, which total about 35,000 in number, in a hierarchical linear Bayesian model. As I will explain in more detail later, the Bayesian approach appropriately takes into account differences in the data generation process between the surveys. Such an analysis could be done using mixture modeling in the frequentist setting, but at the cost of interpretational ease and model parsimony.

The layout of the remainder of this paper is as follows. I begin with a more thorough discussion of the current state of the literature. I begin this section with a discussion of why this topic of one more deserving of attention of researchers. I then review the existing literature and develop my specific theoretical expectations. In the next section I describe the data and the set of the model. After that I report the findings of my analysis, showing that, even when I account for the appropriate covariates, there is strong evidence to support the hypothesis that minority groups wait longer to vote than whites. I then conclude with a discussion of potential policy implications of this work, as well as several avenues for future academic exploration.

2 Why Do Wait Times Matter?

I begin my exploration of the previous literature about line length by enumerating several reasons why political scientists, as well as ordinary citizens in a representative democracy, should be concerned with the topic. There are three major reasons for why the topic is an important one: long lines cause voter attrition, long wait times may change the outcome of elections, and long lines diminish voter confidence in the electoral process as a whole.

2.1 Voter Attrition

In reading news reports about long lines in the 2008 and 2012 presidential elections, there are countless quotes from prospective voters who failed to cast a ballot because of the length of the line to vote.[6] Non-participation of this type could be considered a form of disenfranchisement, given that many voters cannot commit multiple hours to voting because of other obligations like taking care of their children or getting to work. The prospect of waiting in a long line deters two different types of would-be voters from participating: those who never show up to their precinct to vote (referred to in the literature as "balkers") and those who show up, but leave before voting ("renegers").

Several studies have sought to quantify the number of registered voters who fit into either category. The Final Report for the 2008 SPAE indicates that 11% of non-voters in 2008 failed to vote because of lengthy wait times at the polls. In the 2012 SPAE, this figure is estimated to have risen to 14.5%. If we take these figures to be true, and assume that without any wait time these balkers would have voted, the turnout in 2012 would have increased by 16 million voters, a 12% increase in participation. A separate study indicates that in 2012 as many as 49,000 people in Central Florida alone balked at voting because of long lines.[7]

The best attempt to quantify the number of people who reneged on voting was done by Spencer and Markovits (2010). During the 2008 presidential primary election in California, the researchers conducted a field study at 30 precincts in three different San Francisco-area counties. Data collectors at each precinct recorded arrival rates as well as waiting and voting times for everybody that showed up at the polling station. This research provided, for the first time, data about the rate at which people get into line but leave without attempting to

[6] "Long Lines extend Voting Past 7 p.m. in Central Florida." http://articles.orlandosentinel.com/2012-11-06/news/os-election-day-today-20121106_1_mitt-romney-electoral-votes-polls-show. Accessed Jan. 23, 2013.

[7] "Researcher: Long Lines at Polls Caused 49,000 Not to Vote." http://www.orlandosentinel.com/news/politics/os-discouraged-voters-20121229,0,215136.story. Accessed Jan. 23, 2013.

cast a ballot. They find that almost two percent of people who show up to vote eventually renege. The authors also show convincingly that the rate of reneging increases as line length grows. The probability of an individual reneging quadruples when there goes from being no line to just having five people in front of them. Although this study has limitations in its scope, it does tell us that long lines may be deterring millions of voters from casting ballots.

2.2 Electoral Implications of Long Lines

The second reason why we should be concerned about long lines at polling places is that there is the potential for voter attrition to alter election results. Long lines became a hot-topic following the 2004 presidential election, particularly after the publication of the controversial report by the House Judiciary Committee Democratic Staff, "Preserving Democracy: What Went Wrong in Ohio." This report made the case that long lines created from the misallocation of resources had turned away Kerry voters from the polls, costing him enough votes to lose the state of Ohio, and with it the presidential election. A followup study shows that the relationship between voting machines and turnout was not strong enough to have an the electoral impact as suggested (Highton, 2006). The shortage of machines did seem to cost Kerry a few thousand votes, but this was not enough to have swung the result of the election.

Despite the fact that in this high-profile case, long lines do not appear to have altered the result of the election, we should not discount the possibility of such an event ever occurring. If long lines afflicted polling places in an area that strongly favored one party or candidate other the other, then the resulting voter attrition could potentially swing the election result. In a presidential election such partisan disparities are unlikely to determine the outcome of the race. In local, state, or even congressional races however one precinct with an unusually long line may alter the outcome of an election decided by just a few dozen or hundred votes.

Such an event would require that precincts have a different partisan balance than the district as a whole and that long waiting times systematically afflicted the precincts of one party over the other. We know from decades of research that there tends of be geographic clustering of political opinions by neighborhoods, rather than a random distribution of voters across large geographic regions. Thus we should expect that voting precincts organized geographically will produce election results that mirror the opinions of its voters. Likewise, there is strong evidence that the amount of time a voter waits in line is strongly correlated with their partisan opinion. In 2008, Obama supporters waited on average 43% longer to vote than McCain supporters–19:04 minutes versus 13:21 minutes, respectively. By 2012 the gap had narrowed slightly to 35%, with an average wait of 10:10 for Romney supporters and 13:41 for Obama supporters.

2.3 Confidence in the Electoral System

The final reason why we should be concerned with this topic is that long lines generate dissatisfaction. A study by Claassen, Magleby, Monson, and Patterson (2008) finds that voters who wait in line longer tend to evaluate the whole voting experience more negatively. Long lines are correlated with voters having less belief that the electoral system is working. Although the drop off in confidence is small, it has the potential to have important downstream impacts on things like confidence that one's vote was counted or confidence that all ballots were accurately counted. These considerations should motivate researchers and policymakers alike to seek solutions to this problem.

3 What Increases Line Length?

I now turn to considering what the prior research tells us about the factors that increase waiting times in American elections. This section is divided into to parts. First I describe the

on-the-ground factors which increase line length at a particular precinct. Second I describe the broad-based factors, such as race, socioeconomic status, and large shifts in turnout which may have an effect on line length. Although I make this division, these two areas are not mutually exclusive and do not operate independent of one another. Ultimately, each explanation comes back to a discussion of the allocation of resources like poll workers or voting machines. Longer ballots create longer lines unless more voting stations are available. Higher turnout can strain the system if more resources are not available. And race becomes an issue if these resources appear to be allocated in an unequal manner. Assessing the extent to which racial inequalities exist is the main focus of this paper. But in order to draw any conclusions about this relationship, it is necessary to take into account all the other correlates of line length.

3.1 Factors at the Precinct

Most of the work on waiting times on Election Day pulls from queuing theory, which is an area of practical research about systems in which there are "arrivals" "queuing," and "service." In the context of an election, arrivals refers to the process of voters showing up at their precinct to vote (Allen and Bernshteyn, 2006). Queuing occurs as the voters wait in line, or as Spencer and Markovits (2010) put it, "Lines form when the rate of service being provided cannot keep pace with arrivals." In the case of voting, service is the act of filling out one's ballot.

Much of work that directly applies queuing theory comes in the form of comparative policy analyses, often considering which voting technology would yield the most efficient and cost-effective election administration. One such article compares Direct Recording Electronic (DRE) voting systems to paper ballot optical scan (PBOE) machines (Edelstein, 2006). In addition to emphasizing that voting technology can affect wait time, the authors also contribute a few general ideas to the study of wait times in elections. Most important of

these ideas is the author's emphasis on the need to account for non-uniform arrival rates of voters. Another important takeaway is that counties ought to employ a voting technology which gives flexibility to poll workers to add or take away voting stations, depending on the number of voters in line.

A similar policy analysis emphasizes this final point, while also accounting for the variable of ballot length. Edelstein and Edelstein (2010) demonstrate, using simulation data, that with a moderately high turnout and a ballot that takes more than five minutes to fill out on average, election officials would need to double the number of voting machines in each precinct to be able to keep waiting times at a reasonable level throughout the day. Such flexibility is easy when using a paper ballot system, in which additional voting stations require only a table and privacy walls. Precincts that use touch-screen computer ballots do not have such flexibility, since poll workers cannot add new machines on Election Day and purchasing new machines before an election is expensive.

Another important factor which affects wait times at precincts is the arrival rate of voters. In their 2006 article, Allen and Bernshteyn provide a new, more sophisticated calculus for counties to use when allocating resources, such as poll workers and voting machines. Their analysis, which uses simulated data, assumes a Poisson rate of arrival for voters and an exponential distribution for the amount of time taken by voters to out their ballots. With a Poisson arrival rate, there is an average number of voters who show up within a time window, but there is also variation in the rates, as was emphasized by Edelstein (2006). In other words, a precinct may average 50 voters per hour, but it may have 150 show up when people get out of work at 5:00 and only 15 at 10:00 in the morning. Allen and Bernshteyn's work highlights to need to focus on the arrival rate at peak points during Election Day, rather than just the average rate. They demonstrate that small fluctuations in the average amount of time it takes to fill out the ballot can have large impacts on the wait time for other voters: "A ballot initiative adding 30 seconds on average to an otherwise three-minute ballot

could add one half hour to the precinct wait times with 2 [voting] machines." This finding suggests the need to account for ballot length, when attempting to explain line length.

Spencer and Markovits's (2010) field study in California yielded the first empirical data regarding arrival rates at individual precincts. Arrival rates at the thirty precincts in their data are not uniform, but there are similar patterns across precincts, suggesting that the rates may be predictable. In the first hour there was a high number of arrivals, a dip around lunchtime, and a large spike at 5:00 p.m. Such a finding suggests that an individual's wait time may be significantly correlated with the time they show up at the polling location.

3.2 Broad-Based Factors

The main hypothesis that I test in this paper is that racial minority groups wait in line longer to vote than whites, even when taking into account the factors described above. The idea that certain groups might be subject to longer lines has been called a "time tax," a term firmed used in print by Berkeley Law School Dean Christopher Edley, Jr. in 2008. Elora Mukherjee (2009) explains why such a tax is problematic, "Like the poll tax, the time tax burdens a citizen's fundamental right to vote. It is a government policy or practice that forces one citizen to pay more in time to vote compared to her neighbor." In this article, Mukherjee uses 2008 CCES data to show that blacks and Hispanics were disproportionately subject to the time tax. She also shows that such a racial effect was most pronounced in the South.

A big problem with drawing any conclusions based on her analysis, however, is that she does not account of increased turnout among minority groups in the 2008 election because of the presence of Obama on the ballot. The 2008 SPAE Final Report implies that part of the reason that African-Americans waited longer was because participation rates in these communities were higher than has traditionally been the case. If such racial disparities persist even when accounting for changes in turnout, we would then have stronger evidence

of the existence of the time tax. I account for turnout shifts by including a covariate in my model for the percent change in vote since the previous presidential election.

Prior work has also suggested that income may have an effect on line length. Spencer and Markovits's study finds that high income neighborhoods had significantly more efficient poll workers. A $10,000 increase in median household income for a census black group translated into eleven fewer seconds in the check-in process for each voter. Such an effect, multiplied over hundreds or thousands of voters in a day, could turn a five minute wait into a 45 minute wait. Using data on individual income level and zipcode-level median income I test this socioeconomic hypothesis as well.

The final broad factor that has been suggested to increase line length is voter identification laws. The story here is simple: requiring poll workers to ask voters to provide photo identification adds precious second to the check-in time, potentially lengthening the wait time for those waiting to check in. I test this hypothesis using a dummy variable for whether the state requires photo ID to be allowed to vote.

4 Data and Model

To test my hypotheses, I estimate a hierarchical Bayesian model using three different sources of survey data. The 2008 and 2012 Survey of the Performance of American Elections provide data regarding the administration of elections and voter opinion regarding their experience on Election Day. The 2008 SPAE provides 8250 observations for my analysis; the 2012 SPAE adds an additional 6409. The third data source is the 2008 Cooperative Congressional Election Study, which contributes 19564 observations to the analysis.

Because of the differences in the way in which these surveys are conducted, I use Bayesian analysis to facilitate my inference. Using a hierarchical Bayesian model allows me to circumvent the issue of different sampling and survey methodologies by allowing the

results from one of the studies to serve as the prior for the likelihood function for the other studies, and vice versa. I do not need to designate one dataset as the prior and the other as the data, since such a decision is arbitrary and computationally irrelevant. Similarly, I could estimate the posterior distribution for one survey, then use some approximation of its shape as the prior for the second survey. Doing so, however, would introduce unnecessary noise to the model. Instead, I simultaneously estimate the parameters for both the SPAE and CCES. This provides the most statistically efficient estimates possible.

I chose to use a Bayesian model for a few reasons. Utilizing frequentist methods, even through mixture modeling, creates problems when one attempts to take into account the different data-generation processes of each study. A true frequentist could proceed in one of two ways. She could estimate the coefficients of interest in separate likelihood estimations for each study, or she could naively append the datasets together and draw inferences from that. The downside of the first approach is that we could not directly compare the magnitude and significance of the analogous coefficients in the two regressions because of differences in the data generation process and dependent variables. The second approach makes strong–and untestable–assumptions regarding the similarities in the data generation processes.

The dependent variable in my analysis is how long each voter waited in line prior to casting their ballot. Luckily, from the perspective of model simplicity, this question is constructed identically in the CCES and SPAE surveys. In each, the respondent is asked, "Approximately how long did you wait in line to vote?" They are then presented with five options to choose as their response: "not at all", "less than 10 minutes", "10 to 30 minutes", "31 minutes to an hour", and "more than an hour." Those who said they waited longer than an hour were then prompted to specify their wait time in an open ended followup question. The responses from this question were recoded to be on a time scale, in minutes. Respondents who fall into the first four categories were recoded to be the midpoint of their

response category (i.e. 0, 5, 20, and 45 minutes).[8] Respondents who waited longer than one hour had their wait time coded as their response to the open ended question. In the handful of instances in which the respondent did not answer the open ended question, their wait time was coded to be the mean of the wait times for all other 1+ hour respondents in their survey. Because the dependent variable is mostly continuous in nature, I estimate the model with a normal-distribution likelihood function.[9]

My model includes covariates at four different hierarchical levels: individual, zip code, county, and state. The main covariates, which test the race time-tax hypothesis, are included at the individual and zip code levels. I include dummy variables for each respondents' race– black, hispanic, asian, and other, with white left as the reference category. At the zip code level, I include the proportion of people in the zip code in each of these four racial groups (again using white as the reference category), which I calculated using data from the 2010 Census.

To account for the possibility that differences in wait times by race are driven by turnout shifts, I include a county-level covariate for change in turnout, calculated as the percent change in turnout between presidential elections. I also include a zip code level covariate for logged median income, to take into account neighborhood socioeconomics. At the state level, I include a dummy variable for whether the state requires photo identification as a prerequisite for voting. Aside from an indicator variable for year, the remainder of my covariates serve as controls and are at the individual-level of the hierarchy.[10]

[8]Alternatively, for each observation, I could have drawn a value from a uniform distribution for that observation's category. Because of the large sample size, however, doing so does not change the substantive results from the model.

[9]Because the dependent variable is truncated at zero, a more complicated likelihood function, such as one based on the half-normal, log-normal, or negative binomial distribution, may perhaps be more appropriate. I chose the normal likelihood because of its ease of interpretation, and because it yields results that are substantively similar to the more complicated models.

[10]These covariates are: logged age, four-category ordinal scales for both education and income, and dummy variables for democratic and republican party identification (with independents serving as the reference group).

Because I used a normal distribution in my likelihood function, I use a multivariate normal prior distribution for my covariates.[11] Each dimension of this multivariate surface is assigned a mean of zero and a variance of three. I assume independence between the parameters, so the off-diagonal values of the covariance matrix are set at zero.[12] Because of the large number of datapoints, the results of my estimation are robust to changes in the prior distributions.

5 Analysis

I estimated my model using the RStan package in R . Rather than using the more common Gibbs Sampler, RStan utilizes Hamiltonian Monte Carlo (HMC) to sample from the posterior distribution. Although there are costs to computational efficiency, HMC has the distinct advantage of being able to more easily explore multimodal posterior surfaces. It accomplishes this by including an ancillary momentum parameter which accounts for the gradient of the posterior surface multiple times during each iteration of the algorithm. I ran three chains of 10,000 iterations each and used a 2,000 iteration burn-in period to arrive at my results. I assessed convergence using several different tests, yielding sufficient evidence that the chains had converged.[13]

The results from my model are presented in Table 1. The coefficient sizes are approximated by the median estimated coefficient, as well as the 95% credible intervals. The

[11] The one slight exception is the prior for the ancillary variance parameter. Because its value is constrained to be positive, I log transform the parameter and assign a normal distribution, with mean 0 and variance 3, for the prior of the logged variance. I then exponentiate the parameter when I use it in the likelihood function.

[12] In essence, I have assigned an independent normal(0,3) prior distribution to each parameter in the model.

[13] The traceplots for my second chain indicate that it took a bit longer to converge than the first and third chains. This is also reflected in the Geweke diagnostic statistics for the second chain, which suggests that this chain had not converged for a few of the variables. The Gelman diagnostic however calculates values of almost exactly one for all parameters in the model, indicating that the parameters in all three chains yield the same substantive interpretation. Any potential problems generated by chain two are irrelevant though, because all results I present are based on chain one.

Table 1: Estimated Coefficient Sizes

Dependent variable: wait time (in minutes)

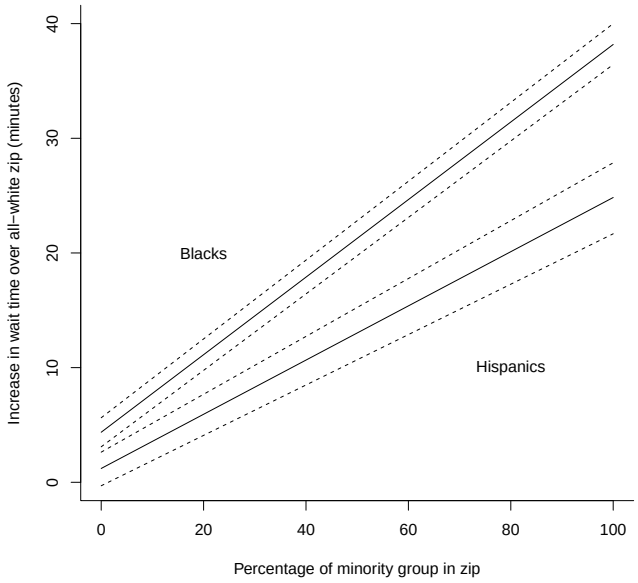
	2.5%	Median	97.5%
Individual			
Black	3.09	4.36	5.63
Hispanic	-0.30	1.21	2.63
Asian	-4.70	-1.91	0.78
Other Race	-0.68	0.77	2.33
log(Age)	-5.71	-4.89	-4.07
Democrat	-0.74	0.33	1.42
Republican	-2.52	-1.42	-0.34
Income-$30-60k	0.79	1.64	2.45
Income-$60-100k	1.70	2.59	3.46
Income-$100k+	2.31	3.26	4.19
Education-HS Degree	-0.04	0.70	1.39
Education-Bachelors	1.10	1.92	2.73
Education-Grad School	0.70	1.74	2.76
Zip code			
Pct Black	0.32	0.34	0.36
Pct Hispanic	0.21	0.24	0.27
Pct Asian	-0.02	0.03	0.09
Pct Other Race	-0.42	-0.36	-0.30
log(Median Income)	3.26	3.79	4.35
County			
Pct Change Turnout	6.56	10.33	14.20
State			
Voter ID Law	6.12	6.96	7.83
2012	-4.73	-3.98	-3.26
Intercept	-19.69	-14.38	-8.98

Linear Hierarchical Bayesian Model
N=34223

covariates are partitioned based on their level in the hierarchy. Because of the setup of the likelihood function, coefficient values can be easily interpreted as increases or decreases in the dependent variable, wait time, which is measured in minutes. What is immediately noticeable is that, even when controlling for shifts in turnout, there still appears to be a strong relationship between race and wait time, particularly for African Americans.

Figure 1 presents the estimated marginal effects for the black and Hispanic covariates. The upper solid line shows median difference (bounded by 95% credible intervals) in the number of minutes an African American waits in line to vote compared to a white person,

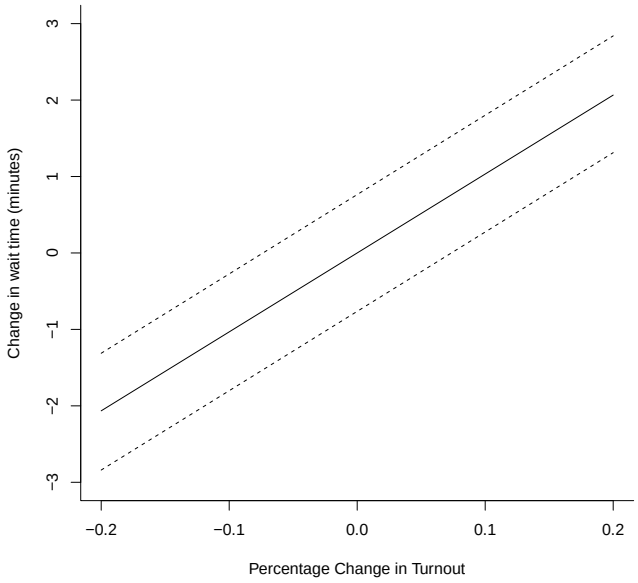
Figure 1: Difference in wait time for racial minorities compared to whites



based on the racial composition of the zip code in which the African American lives. What we find is that black neighborhoods have wait times that are substantially larger than white neighborhoods. A three percent increase in the number of blacks in a zip code roughly corresponds to an average wait time that is a full minute longer. To put this into context of the actual demographic makeup of the country, approximately 54% of zip codes are composed of fewer than 5% of black residents while 11% of zip codes are at least 30% black. A black person in a 30% black zip code will wait approximately 12:49 longer in line than a white person in a 5% black zip code.

The model also provides strong evidence that Hispanics experience lines on Election Day that are a lot longer than whites. A four percent increase in the number of Hispanics in a zip code adds a minute to the wait time of voters in the neighborhood. The model also suggests that there is a 5.7% probability that Hispanics will experience wait times that are

Figure 2: Marginal effect of turnout changes on wait times

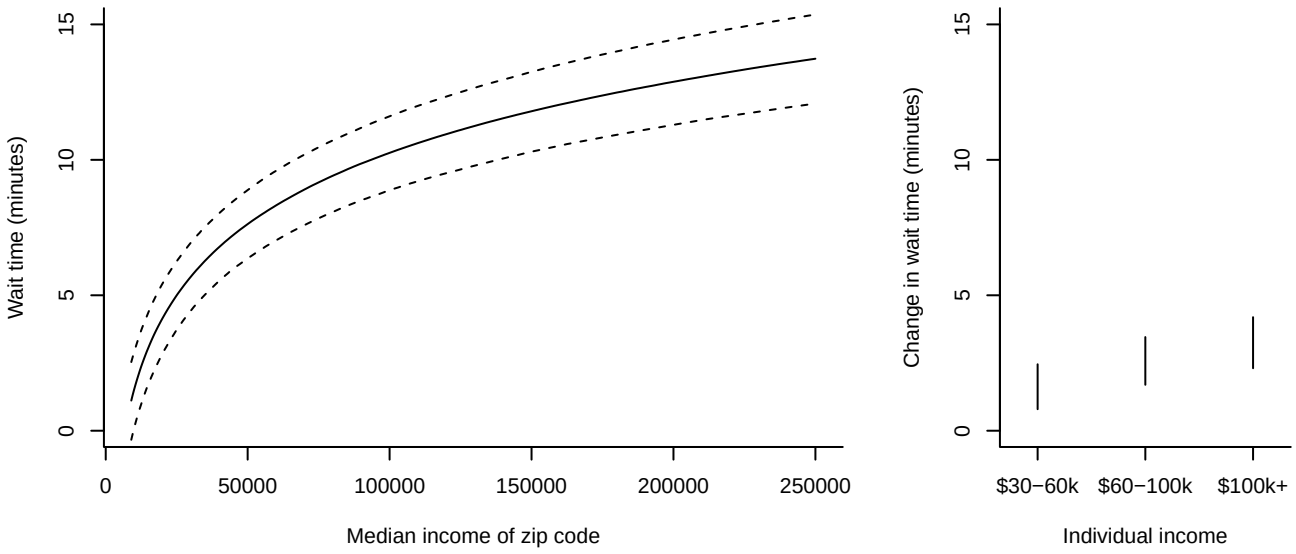


longer than whites, ceteris paribus.[14] The lower set of bars in figure 1 show the differences in wait times for Hispanics compared to whites. Approximately 9% of zip codes have more than 30% of residents who identify as Hispanic, while about 49% are less than 5% Hispanic. The median difference in wait time between a Hispanic person in a 30% Hispanic neighborhood and a white person in a 5% Hispanic zip is 6:07.

These results provide strong evidence to support the time tax hypothesis–that certain groups of citizens are required to spend more resources (in this case, time) to exercise one of their civic rights. What makes this finding even more strong is the fact that these results are robust to the inclusion of shifts in turnout. Figure 2 shows the marginal effects of inter-election changes in turnout. Changes in turnout rates, even shifts as sizable as 20%, do not have nearly as large a magnitude as the effects for race. These results mitigate the argument that any racial differences that exist in the 2008 or 2012 resulted from Obama's presence on

[14]The analogous percentage for blacks is estimated to be zero. None of the 8,000 draws of the individual-level black covariate were negative.

Figure 3: Effect of Income on Wait Times



the ballot driving up turnout. This should not be entirely surprising. Although there were big spikes in voter participation in minority neighborhoods in 2008, such spikes were easily anticipated by election officials, who would have been able to provide adequate resources for those precincts.

Another suggestion of previous research has been that income has a significant relationship with line length. Recall that the Spencer/Markovits article found that check-in times decreased as average income in the neighborhood grew. Figure 3 seems to tell a different story. The left half of the graph shows the estimated average wait time by zip code median income levels. The relationship is clearly a positive one, and is slightly difficult to explain. It may be that more wealthy people tend to be better educated, and that more educated people tend to take longer filling out their ballots because they consider each ballot measure more carefully. The right half of figure 3, however, does not seem to provide strong evidence to support this theory. While there is a significant difference in the wait time for

the lowest income group ($0-30k) compared to all three higher income brackets, changes in income after the $30,000 do not dramatically change wait times. The relationship uncovered here, between income and wait time, is one that merits further, more in-depth exploration.

Voter identification laws have been accused of having the detrimental effect of increasing waiting times for voters. The results in table 1 support this argument. States that require photo ID to vote experience an average wait time that is six to eight minutes longer than states that do not have such laws. Six to eight minutes may seem like too long of an added wait for such a task as seemingly quick as producing a drivers license. But it is important to remember that the thirty second task of each voter producing an ID has to be multiplied by every person in line ahead of you. In addition, this large six to eight minute increase may perhaps also be evidence that lots of voters could not produce valid ID and took more time at the check-in table as a result. Such a large increase in wait may diminish over time, as voters become more accustomed to producing ID before voting. With the evidence at hand however, it is impossible to weigh in definitively on any of these possibilities.

6 Conclusion and Future Work

In this paper, I have provided evidence that African Americans and Hispanics are subject to longer lines and lengthier waits when trying to cast a ballot in American elections. These results persist even after accounting for key factors like income and turnout shifts. The paper is the first piece of research which provides more evidence from cross-sectional data than just bivariate correlations between race and waiting times.

There are several avenues in which this research can be further developed. In 2008 and 2012, there seemed to be some reports by the media that early voters tended to stand in lines that were *longer* than the lines on Election Day. This would be an interesting finding, given that one of the most common justifications for implementing early voting is that it

alleviates line length on Election Day. Inclusion of an early voting dummy variable, or even subsetting the data between early voters and Election Day voters could provide insight for policymakers into the possible effects early voting have on wait times.

Another possibility to expand this research is to more thoroughly take into account differences in voting technologies. Several of the previous studies about wait times have focused on comparing the effects of voting machine types on line length. None of this research though has used a broad cross-section of data, as is employed in this paper. Utilizing vote technology data from the Election Administration and Voting Survey, I could provide a more rigorous test of whether any particular voting method tends to produce shorter waiting times for voters.

Other methodological changes, which were not possible in this version of the paper because of computing limitations, could be made in an effort to capture all the major factors which affect line length. State fixed effects could account of variation in election administration laws, as well as differences in administrative efficiency at the state level. County fixed effects may perhaps also be used to factor in the key variable of ballot length.

References

Allen, Theodore, and Mikhail Bernshteyn. 2006. "Mitigating Voter Waiting Times." *Chance* 19.

Alvarez, Michael R., Stephen Ansolabehere, Adam Berinsky, Gabriel Lenz, Charles Stewart III, and Thad Hall. 2009. *2008 Survey of the Performance of American Elections - Final Report*.

Claassen, Ryan L., David B. Magleby, J. Quin Monson, and Kelly D. Patterson. 2008. "At Your Service: Voter Evaluations of Poll Worker Performance." *American Politics Research* 36(4): 612–634.

Edelstein, William A. 2006. "New Voting Systems for New York–Long Lines and High Cost." *New Yorkers for Verified Voting* 10.

Edelstein, William A., and Arthur D. Edelstein. 2010. "Queuing and Elections: Long Lines, DREs and Paper Ballots.".

Highton, Benjamin. 2006. "Long Lines, Voting Machine Availability and Turnout: The Case of Franklin County, Ohio in the 2004 Presidential Election." *PS: Political Science* .

Mukherjee, Elora. 2009. "Abolishing the Time Tax on Voting." *Notre Dame Law Review* 85.

Spencer, Douglas M., and Zachary S. Markovits. 2010. "Long Lines at Polling Stations? Observations from an Election Day Field Study." *Election Law Journal* 9.

Stan Development Team. 2012. *Stan Modeling Language: User's Guide and Reference Manual*.