

Voter Registration List Quality Pilot Studies: Report on Methodology*

Stephen Ansolabehere
Eitan Hersh
Department of Government
Harvard University

Alan Gerber
David Doherty
Department of Political Science
Yale University

June 8, 2010

*This research was funded by The Pew Charitable Trusts. Any opinions or estimates expressed herein are those of the authors and do not necessarily reflect the views of The Trusts. Joe Williams and Anthony Fowler aided with compiling these data. Viridiana Rios aided with Spanish translation. The LA County Registrar and Recorder provided invaluable guidance and support implementing the study. In particular, we thank Dean Logan, Michael Petrucello and Ray Ching in the LA election office for their cooperation. Rod Petrey and Mark Pritchett at the Collins Institute provided helpful advice as well.

Key Results

- Initial analysis of state and county voter files revealed some problems in the compilation of county voter lists. Most notably, lists from two of Florida's sixty-seven counties were incomplete.
- The rate of undeliverable mail was 9% in the Florida random sample and 5% in the Los Angeles random sample. These rates compare favorably to the Census 2000 undeliverable mail rates of 10% in Florida and 6% in California. Relative to other government lists, registration lists in both locations are quite current.
- Undeliverable mail may be indicative of various registration problems and should not be used to purge voters. Thirty to fifty percent of those for whom two pieces of mail came back as undeliverable actually voted in the 2008 general election.
- Based on a statistical method that combines undeliverable mail rates with self-reported discrepancies by survey respondents, 10-25% of Florida registration records and 8-17% of Los Angeles records are estimated to be invalid entries.
- The costs associated with conducting a registration list audit are enumerated. A statewide mail-based list audit runs approximately \$19,000 plus research costs. Using this research design, a national study would cost approximately \$1,000,000 in addition to costs for design implementation and analysis.

1 Introduction

Between August 2008 and July 2009, audits were conducted to assess the quality of voter registration lists in two areas of the United States. These audits, conducted by Professors Stephen Ansolabehere and Alan Gerber with research assistance from David Doherty and Eitan Hersh, and funded by the Pew Center on the States, represent an initial effort to evaluate the quality of voting records nationwide. According the National Commission on Federal Election Reform, inaccurate voting lists present a significant obstacle to the integrity of U.S. elections, adding to costs of election administration, exposing elections to the possibility of fraud, and penalizing poor and less educated citizens (2001). This pilot study offers a first look at the quality of registration lists in two jurisdictions and develops a methodology for assessing list quality generally.

This memo serves to describe, in detail, the methodology for conducting audits of voter lists. To begin, we will review the specific goals of our approach. Second, we will describe the sampling procedure. Third, we will share some problems and surprises we encountered in our research. Fourth, we will review statistics on the undeliverable rate and response rate of mailed surveys. Finally, we will explain how to measure overall list quality based on the undeliverable mail rate and registrant-reported list inaccuracies. An accompanying report provides a detailed analysis of data gathered in the two pilot audits.

2 The Approach

Los Angeles County, California and the state of Florida were selected as pilot study locations. Voter registration in each state is limited to U.S. citizens who reside in the state and are at least 18 years of age as of the day of the next election. Individuals deemed mentally incapacitated by the courts are ineligible to vote in both states as are felons who are either in prison or on parole.

Los Angeles County was selected in part because we had prior experience working with election administrators there and so we could conduct a study of voters in cooperation with the county government. L.A. County is also of interest because in California, counties manage voter roles, periodically sending updates to the state. The state stores the voter roll information separately for each county, but does not play any role in updating or maintaining the lists. Only in a few states is this “bottom-up” structure in place, and we are interested to understand how list maintenance functions in a jurisdiction where the state does not put resources into cleaning or updating registration records (Electionline 2005).

Conversely, the state of Florida has a top-down structure, meaning that the state and counties share responsibility for the maintenance of a centralized statewide voter list which county administrators can update directly. As records are centralized in Florida, academic researchers can obtain a list of all voters in the state directly from the state’s Division of Elections for a nominal fee. As a large and diverse state in a different region of the country from Los Angeles, and as the focal point for the election problems associated with the 2000 Presidential contest, Florida presented an excellent second jurisdiction for this pilot study.

With locations selected, we drew several random samples from the two populations of voters (the sampling technique is described in the next section). Two identical pieces of first-class mail were sent to each randomly selected voter. The two mailings were sent approximately one week apart from each other. The first part of the audit method consists of analyzing the mail that was returned by the U.S. Post Office as undeliverable. Sending two pieces of mail rather than one enables researchers to distinguish mail that was undelivered due to Post Office mishaps and mail that was undelivered because of invalid addresses.

Undeliverable mail can signal several kinds of problems with voter lists. Mail might be returned because of typos or incomplete address data, because a voter is no longer at the address listed by the registrar, or perhaps because the voter has died. We analyze the rate of undeliverable registration addresses as it compares with other undeliverable mail studies,

such as those associated with the U.S. Census and jury summonses. We also analyze how undeliverable mail varies across jurisdictions.

2.1 The Survey

Contained in the mailing was a brief survey that voters were asked to fill out and to return in a business-reply envelope. This survey gauged the accuracy of the information listed on the voter file. Voters were asked to confirm or correct their name, address, birth date, political party, vote history, and in the case of Florida, racial identification. Unlike California, Florida keeps a record of each registrant's race on the voter file. Registrants were asked if they moved residences in the last two years and if they changed their voter registration record in the last two years.

The questionnaire listed the registrant's vote history record from the 2006 general election, the 2008 Presidential primary election, and in the case of Florida, the 2008 general election. Since the Los Angeles study was fielded before the 2008 general election, Los Angeles registrants were asked if they intended to vote in that election rather than if they had voted. For each election, registrants were asked to report their method of voting, whether by mail, in person, or in Florida, by early voting.

Some of the respondents who received both mailers completed and returned both surveys. However, most did not do this as the second letter mailed contained the following line in bold letters on the top of the page: "If you have already completed this survey, Thank You! IF NOT, please fill out and return the questionnaire."

The second part of the audit method consists of studying the incidence of voter-reported inaccuracies on the voter rolls. A detailed analysis of the survey responses is included as a separate report. The questions in the survey reveal discrepancies between voters' reported names, addresses, vote history and other information and these data as they are recorded by the election registrar.

The letters and questionnaires that were sent to voters are included in Appendix B. Notice two features of these materials. First, the letter and questionnaire were sent to voters in English and in Spanish. This provides an unobtrusive way to measure language choice among registrants. Second, the Los Angeles study was a cooperative effort with the L.A. County Registrar while the Florida effort was conducted solely by the researchers. As such, in L.A. the mailers had the official insignia of the Registrar and mail was sent from the Registrar's office. Before conducting the study in Florida, researchers placed several phone calls and emails to the Florida Division of Elections in an effort to work cooperatively with their office. We sent drafts of the letter and questionnaire to the Division of Elections ahead of the study. However, we received no response. To allay concerns from registrants in the sample regarding the legitimacy of the study, a toll-free number was set up and advertised in the Florida mailing to provide registrants with a method of contacting the researchers should they have questions. We received and returned approximately 25 calls.

The letter accompanying the questionnaire impressed upon recipients that the information gathered would not be used to change their registration records, but rather it would be used only to measure the accuracy of voter lists. We emphasized this point not only to reinforce the confidentiality of the survey but also to inform voters that if their registration record contained inaccuracies, they would have to contact the election administration directly to resolve the problem.

2.2 Limitations of the Audits

Analysis of undeliverable mail and survey responses demonstrate both the scale of inaccuracies as well as the variation of inaccuracies on two voting lists. The analysis helps answer questions about how many records contain voting problems and about the kinds of problems the lists contain. Answers to these questions can guide policy makers who are interested in determining the appropriate level of priority for voter registration reform and the kinds of

reform ideas that will best address the current shortcomings of registration list maintenance.

By way of acknowledging the limitations of this study, two points must be made especially clear. First, while the method of analysis described here should be easily replicable in other jurisdictions, the results of two small pilot studies are not indicative of nationwide list quality. Voter registration list quality is likely to vary significantly across states and, to some extent, across counties within states. Florida and Los Angeles are important cases to study, but they are not to be treated as necessarily representative of the nation.

Second, while these pilot studies document the rate of inaccurate and erroneous records on voter lists, this report is not intended to assign blame for negligent administration to election officials. Given the simple facts that voters move, die, change surnames and political parties, and sometimes submit incorrect information to election officials, there will inevitably be inaccuracies on voter lists. We leave it to policymakers and advocates to determine the level of inaccuracies that they deem to be acceptable.

3 Sampling Method

To sample from the populations of registered voters in Los Angeles and in Florida, voter files were acquired from each jurisdiction. As the Los Angeles study was a cooperative effort between researchers and the County Registrar, the Registrar supplied a list of all voters in the county, including information on each voter's name, address, age, party, and vote history for the 2006 general election and the 2008 Presidential primary. The list was acquired in September 2008. Following the November 2008 general election, the Registrar provided an updated voter list noting vote history for the 2008 general election. The Los Angeles mailing was sent out in October 2008.

The Florida Secretary of State's Division of Elections assembles sixty-seven county files into CD-ROM and makes the full state voter list available to researchers for a \$10.00 fee.

The state keeps two files for each county: one file that contains the name, address, and other identifying information of registered voters and one that contains vote histories for each registered voter. The Florida file was acquired in May 2009, and the study was conducted in July 2009.

In Los Angeles, two samples were drawn. We drew a random sample of 7,500 from the full voter file. Following the sampling strategy of Alvarez, Ansolabehere, and Sikora (2003), we then selected five precincts that had very low turnout in prior elections and sampled an additional 100 registrants from each of these precincts.¹

Table 1: Sample Sizes

	Los Angeles	Florida
Random Sample	7,500	10,000
Low Turnout Sample	500	N/A
Inactive Sample	N/A	1,000
Total	8,000	11,000

Prior to drawing these samples, we experimented with clustered sampling as an alternative to simple random sampling. Clustering by precinct would involve first randomly sampling a certain number of precincts, and then sampling voters within those precincts. An advantage of this form of sampling is that by ensuring that different kinds of precincts are represented, the sample might more accurately reflect the variation in the population. However, we found that this approach actually introduced bias in the sample. Experimenting with a voter file from Florida that contained race identifiers, we found that, due to residential racial segregation, clustering caused severe under-representation or over-representation of minorities depending on whether a concentrated minority precinct happened to be in the sample in any given draw. As a result of these preliminary sampling exercises, we opted for simple random sampling.

¹The low turnout precincts are identified by precinct numbers 3850010D, 38550360A, 9001046M, 9001780C, 9003922A. In selecting low turnout precincts, we ignored precincts with very small populations such that a sample of 100 registrants was not possible.

Table 2: Sample vs. Population in Turnout and Partisanship

	Official		Los Angeles		Florida	
	LA	FL	Full Sample	Low Turnout	Active	Inactive
<i>Turnout</i>						
2008 General	81.9%	75.2	77.5	52.7	73.6	3.8
2008 Pres. Primary	55.3	41.8	52.2	26.4	35.2	1.3
2006 General	52.1	46.8	47.2	22.8	40.2	4.4
<i>Partisanship</i>						
Pct. Democrat	51.8	42.1	50.8	59.2	42.9	38.4
Pct. Republican	24.1	35.8	26.1	14.0	35.4	30.9

L.A. turnout figures: County of Los Angeles Registrar-Recorder/County Clerk (2009); L.A. party registration figures: California Secretary of State (2008). Florida turnout figures: Florida Department of State (2009a); Florida party registration figures: Florida Department of State (2009b).

In Florida, two samples were also drawn: one sample of 10,000 active registrants and a second sample of 1,000 inactive registrants.² Each sample was drawn from counties in proportion to the county populations. For example, since Miami-Dade County includes 14.1% of the state population, the active sample included 1,410 registrants from Miami-Dade and the inactive sample included 141 registrants from Miami-Dade. Because of rounding, the active voter sample actually contained 9,999 individuals and the inactive sample included 1,002 individuals. The larger sample in Florida is helpful in answering questions about list quality in sparsely populated counties of the state.

Table 1 summarizes the sample sizes in the two studies. Note that the Los Angeles voter file supplied by the county office did not include a designation of active-inactive status. Table 2 compares the turnout rates among registrants in each of the four samples for the 2008 general election, the 2008 Presidential primary election and the 2006 general election. The table also includes the proportion of registered Democrats and registered Republicans

²In Florida, election officials periodically send voters address confirmation notices. If a voter does not respond to the notice within 30 days, he or she is marked as inactive. Voters remain inactive until they vote, update their registration record, or request an absentee ballot. After two federal general elections without any activity, inactive voters are removed from the voter list.

in the samples. These numbers can be compared to the official data reported on state and county websites.

Notice that whereas the party registration data in the samples match the official tallies closely, the sample turnout estimates are slightly at odds with the official tallies. There are two explanations for this. First, voter registration systems are updated on a continuous basis. When researchers acquire a list of registered voters, they see only one snapshot in time. Official statistics on turnout appear soon after an election. A snapshot taken a few months later will contain a slightly different set of voters due to updates.

The second reason for inconsistencies between official turnout figures and those in the samples is that official statistics are generally reported based on the count of ballots on Election Day. Counting ballots is, for better or worse, an entirely separate activity from checking off the names of voters as they enter the polling place or submit an absentee ballot. Due to mishaps in indicating registrants who vote in a given election, and machine inaccuracies in tabulating all ballots effectively, the sum of voters who are listed as having voted will almost always differ from the sum of ballots counted.

4 Data Problems

Apart from the audit of individual records, we encountered three surprising problems when working with the voter lists, each of which speaks to issues of data quality. First, an initial diagnostic of Florida county voter files acquired from the the state identified unusual data from Hillsborough County, a populous county that includes the city of Tampa. The list of Hillsborough registrants included only a small fraction of the number officially reported as registered. Moreover, the registrants who were listed had a vote rate of only 50% in the 2008 general election, far lower than the correct turnout rate.

We contacted the Hillsborough list manager directly to obtain a correct county list. The

manager was not aware of the problem with the state's version of the list, nor was he aware of how such a problem could have occurred. As the county had a much more plausible version of the registration list than the one acquired from the state, this seems to have been a merging problem between the two systems. This problem suggests a vulnerability when counties and states are jointly responsible for list management: the more data are moved around, the more opportunities there are for merging problems.

The second problem, more worrisome than the first, involved the list of voters in Duval County, Florida, another large county and home to Jacksonville. The Duval County list looked generally normal, except for data from the 2008 Presidential primary election. In the Duval list obtained from the state, only 1,140 people voted in that election in contrast to the roughly 200,000 reported by the county. Several phone calls to the county office did not yield an explanation for this oddity.

The Duval county problem seems to be due to the county either not properly collecting or not properly reporting which registrants did and did not vote in the Presidential primary election. We will not speculate as to why this problem occurred, but it certainly suggests that states should follow the recommendations of Atkeson, Alvarez and Hall (2008) and develop tools for precincts to reconcile differences between the number of voters who sign in to the polls and the number of ballots counted. Failure to label registrants as having voted is a serious problem if it occurs in federal general elections, as this could result in a voter improperly being removed from the voter rolls under the purging provisions of the National Voter Registration Act of 1993.

The third problem with the voter lists was found in merging the first list supplied by Los Angeles County in September 2008 with an updated list supplied in December 2008. The voters who were sampled for the mail study were drawn from the first list. The updated list contained vote history from the 2008 general election and so the two files were merged. In so doing, we noticed that 344 of 8,000 individuals in the sample, or 4.3%, were no longer on

the voter file.

This discrepancy could be attributed to coding problems that caused voter identification numbers to be different in the two lists. The discrepancy could also be the result of purged entries. Los Angeles County officials were not able to give us a determination of what happened to these records.

The characteristics of the 344 missing registrants differ from those in the sample who were listed in both versions of the voter file. For example, the missing registrants were about four years younger and five percentage points more likely to be affiliated as Democratic than the non-missing registrants. Thirty-three percent of the missing registrants had voted in the 2008 primary, compared to 51% of non-missing registrants. The response rate to the questionnaire was half as high among the missing registrants as the non-missing registrants.

While the missing registrants look substantially different from non-missing registrants on every variable of interest in the data, a significant percentage of missing registrants were voters in the 2008 primary and responded to the questionnaire at the address listed in the voter file. Thus, the reasons are unclear for why these voters no longer appear on the voter list.

5 Undeliverable Mail and Survey Response

In Los Angeles County, for 5% of registrants in the random sample the Post Office returned at least one of the two pieces of mail. In Florida, for 9% of registrants in the active voter sample the Post Office returned mail as undeliverable. Table 3 reports the undeliverable rate as well as the response rate to the survey, breaking down the Florida data by individual county.

Table 3: Active Voter Sample Size, Response Rate, and Undeliverable Rate by County

County	Sample Size	Undeliverable Rate	Response Rate
Los Angeles	7,500	5%	26%
Florida (All)	9,979	9	21
<i>Florida Counties</i>			
Alachua County	135	13	21
Baker County	14	0	29
Bay County	93	12	23
Bradford County	16	0	13
Brevard County	298	7	25
Broward County	1,014	9	20
Calhoun County	8	0	25
Charlotte County	89	8	27
Citrus County	74	11	36
Clay County	88	6	29
Collier County	157	9	21
Columbia County	35	3	18
DeSoto County	20	15	18
Dixie County	9	44	20
Duval County	487	7	20
Escambia County	184	13	21
Flagler County	31	3	30
Franklin County	7	29	0
Gadsden County	28	7	31
Gilchrist County	9	0	11
Glades County	7	14	17
Gulf County	8	0	0
Hamilton County	8	13	29
Hardee County	17	18	7
Hendry County	23	13	15
Hernando County	82	4	29
Highlands County	55	7	31
Hillsborough County	624	16	18
Holmes County	12	0	0
Indian River County	70	6	23
Jackson County	29	3	11
Jefferson County	8	0	13

Continued on Next Page...

Table 3 – Continued

County	Sample Size	Undeliverable Rate	Response Rate
Lafayette County	4	0	0
Lake County	132	4	18
Lee County	275	8	29
Leon County	150	15	29
Levy County	22	5	14
Liberty County	4	0	0
Madison County	12	0	8
Manatee County	164	7	25
Marion County	162	11	24
Martin County	79	8	22
Miami-Dade County	1,401	9	16
Monroe County	50	2	20
Nassau County	36	3	11
Okaloosa County	107	17	27
Okeechobee County	22	9	25
Orange County	561	10	18
Osceola County	108	8	14
Palm Beach County	707	13	24
Pasco County	216	8	20
Pinellas County	574	7	22
Polk County	303	8	19
Putnam County	44	9	15
St. Johns County	77	4	28
St. Lucie County	121	9	21
Santa Rosa County	74	9	18
Sarasota County	203	8	32
Seminole County	228	8	19
Sumter County	33	9	33
Suwannee County	22	18	11
Taylor County	12	0	17
Union County	8	25	17
Volusia County	277	7	25
Wakulla County	14	7	31
Walton County	25	4	33
Washington County	13	8	33

Note: The undeliverable rate includes respondents for whom one or two pieces of mail was returned as undeliverable. The survey response rate is calculated based on voters for whom

neither mailing was returned as undeliverable. These cases are excluded from the response rate calculation because they reflect inaccuracies in the voter roll rather than a refusal to complete the survey. Statistics are calculated only from the full random samples and exclude the low-turnout precinct and inactive voter samples.

In assessing the rate of undeliverable mail from these pilot studies, the most important question to ask is whether these figures, 5% for Los Angeles and 9% for Florida, are high or low. What is the baseline for evaluation?

The Census undeliverable rate is the optimal available comparative metric. In the decennial Census, the Census Bureau generates a list of all mailable addresses in the United States. It does so by use of government sources and by purchasing lists from commercial vendors (Mihm 1999). This list generation mechanism is quite different from the voluntary process of the voter registration system. In spite of the superior data and resources available to the Census Bureau, the Census nevertheless faces undeliverable mail problems just as the voter registration systems do.

The national undeliverable rate for the 2000 Census was 7.2% (Treat 2001, Kohn 2003). That is, of all the Census forms distributed by mail, 7.2% were returned to the Census Bureau by the Post Office as undeliverable as addressed. However, there is quite a lot of variation in the undeliverable rate across states. California has one of the lowest undeliverable rates, at 6.3%, and Florida has a much higher rate of undeliverable mail, at 10.2% (Kohn 2003). Thus, if one treats the Census statewide undeliverable rates as a baseline, Florida and Los Angeles both have undeliverable rates lower than expected and, though Florida's rate is higher than Los Angeles', these rates are comparable, relative to the Census baseline (each is 1 percentage point below the state Census rate).³

Table 4 includes some additional statistics on undeliverable mail comparisons. The 2005 American Communities Survey, another study conducted by the Census Bureau, witnessed

³We note that Census form undeliverable rates in California are an imperfect proxy for undeliverable rates in Los Angeles County. Unfortunately Census undeliverable rates are unavailable at the county level.

Table 4: Comparison of Undeliverable Rates

	Los Angeles	Florida
Undeliverable Rate (current study)	4.9%	9.3%
Census 2000 Undeliv. Rate (by state)	6.3	10.2
Census 2000 Undeliv. (national)	7.2	7.2
2005 ACS Undeliv. (national)	12.0	12.0
Jury Undeliv. Rate (by state)	17.4	13.0
Alvarez et al. (2003) study	9.8	N/A

an undeliverable mail rate of 12% (Citro and Kalton 2007). A national study of jury selection found that 14.6% of summonses were returned as undeliverable, with actually a higher undeliverable rate in California than in Florida (Mize, Hannaford-Agor, and Waters 2007); however, the jury rates are in either case higher than the registration rates. In Alvarez, Ansolabehere, and Sikora’s (2003) study with a similar research design as the present study, the undeliverable mail rate was 9.8%, double the rate of undeliverable mail in this study.

When compared with undeliverable mail in other contexts, the voter registration systems in Los Angeles and Florida seem to perform quite well. No compilation of addresses can possibly keep up with the constant population fluctuations attributable to deaths and moves. Even the Internal Revenue Service recently reported that it has not been able to return \$123.5 million to taxpayers in refund checks because nearly 108,000 refunds were returned by the Post Office as undeliverable (IRS 2009). Nevertheless, maintaining current lists and, to the greatest extent reasonable, reducing undeliverable mail rates remains a desirable goal. Los Angeles and Florida, two jurisdictions with very different registration administrations, appear to be doing better than expected at keeping undeliverable mail below levels typical of government mailing lists.

Undeliverable Mail and Ghost Voters

The key to the methodology for undeliverable mail is to use returned mail to estimate the fraction of addresses on the lists that are no longer valid. The extreme assumption holds that

when an address yields returned mail, that registration may be treated as invalid because the person is no longer at that address or the address is incorrect. Of course, that may not be entirely true. There may be a typographical discrepancy on the list and the person may still vote even though there is an discrepancy. Also, when people move residences but stay in the general vicinity of their old address they sometimes leave their old registration in place and go back to that precinct to vote. It is unclear how to treat either of these examples. Both produce returned mail and neither is technically a valid registrant, but election officials let people vote in these instances, in the first case because the problem just stems from a clerical mistake and in the second case because the person looks like a valid registrant at the polls on Election Day.

Table 5: Turnout Rates of Undeliverables

Vote Rate	Letters Returned		
	0	1	2
<i>Los Angeles</i>			
2008 General	79%	44	48
2008 Pres. Prim.	54	18	29
2006 General	49	19	22
<i>Florida</i>			
2008 General	78	43	31
2008 Pres. Prim.	38	11	10
2006 General	43	16	18

Numbers are calculated from the random samples in each location and ignore the inactive and low-turnout precinct samples.

One of the most surprising findings of this study is that a substantial percentage of registrants whose mail is undeliverable continue to vote at the invalid address. Statistics on these “ghost voters” are reported in Table 5. In Los Angeles, nearly half of the voters for whom both pieces of mail were returned as undeliverable voted in the 2008 election. In Florida, 31% of these individuals voted. To the extent that ghost voters are movers who return to their old precincts or are true residents with typographical inaccuracies on

their registration forms, their voting habits suggest that although their records are incorrect election offices should use caution in purging names from lists because mail is returned. An alternative explanation is that these invalid names are suggestive of fraud. Some of the votes that occur with the “invalid” addresses may be cast by people other than the registrants listed, or perhaps by registrants voting in multiple jurisdictions. The results here do not allow us to distinguish between these explanations. However, it should be emphasized that we have neither sought nor found any evidence of fraud in this study.

6 Inferences about the Registration Population

This pilot study contains two different measures of list quality: the rate of undeliverable mail and the rate of discrepancies reported by respondents to a short survey (the survey is shown in Appendix B). In this section, we develop a method of combining these two measures to produce an overall rate of inaccuracy (ROI) for Florida and for Los Angeles. Estimating an inaccuracy rate requires that we make a number of assumptions about how to infer inaccuracies from undeliverable mail and discrepancies reported by registrants. No assumption is perfect or indisputable, so the best approach is to articulate the assumptions that seem to us to be most reasonable as clearly as possible. Alternative assumptions will yield different rates of inaccurate records, and so after deriving one estimate, we will alter some of the assumptions and offer three additional estimates.

For these estimations, we use data from the random samples and ignore the inactive and low-turnout precinct samples. The measurement strategy can be divided into three parts: the ROI for registrants for whom two pieces of mail were returned as undeliverable, the ROI for registrants for whom no mail was returned, and the ROI for registrants for whom one piece of mail was returned. We will estimate these rates and then sum them together to arrive at a final estimate.

A. Rate of Inaccuracy for Registrants with Two Pieces of Undeliverable Mail

The first assumption we make is that if a registrant has two pieces of mail returned as undeliverable, then his/her registration is inaccurate. That is,

$$Pr(I|BothLettersReturned) = 1. \tag{1}$$

where I indicates that a listed registration is inaccurate. The probability of the Post Office making a mistake in mailing both letters to the same valid address is extraordinarily small. Thus, the assumption that having both mailers returned as undeliverable implies an invalid address is reasonable. We have shown that many registrants who fall under this category continue to vote at their listed addresses. Nevertheless, these registrants are either not voting where they are living or have an inaccuracy on their registration record, or else they are involved, as a victim or as a perpetrator, in voter fraud.

To estimate the proportion of registrations that are invalid due to two returned mailings, we multiply the probability of an invalid address (under this assumption, 1) by the probability of having two returned mailings. In Florida, 6.4% of active voters had both mailers returned, so we begin to aggregate invalid records by assuming that **6.4%** of records are invalid. The equivalent figure for L.A. County is **2.0%**.

B. Rate of Inaccuracy for Registrants with Zero Pieces of Undeliverable Mail

We now add inaccuracies ascertained from the questionnaire. Details on questionnaire responses are included in the accompanying report. Here, we focus on inaccuracies reported by registrants that could reasonably affect one's ability to vote on Election Day. These are included in a measure labeled "reported correction," defined in the accompanying report. To estimate the ROI for the full registration list, we cannot apply the self-reported discrepancies directly because of selection bias. Whereas our sample of voters is random, the individuals

who took the time to fill out the survey and mail it back are not chosen randomly. As we detail in the accompanying report, respondents tended to be older and were more likely to be voters than non-respondents. Because discrepancies are correlated with age, vote history and other key variables, we must account for the selection bias in the sample of questionnaire respondents. We do so by modeling self-reported inaccuracies as a function of age, voter turnout, party, and in the case of Florida where the data exists, race. We use a logistic regression model and generate predicted probabilities for all registrants. These predicted probabilities provide a way for us to estimate the inaccuracy rate among registrants who did not respond to the survey by allowing us to calculate an inaccuracy rate that accounts for the observed differences between individuals who did and did not respond to the survey. Coefficients from the logit models are displayed in Table 6.

Table 6: Logit Models for Predicted Probabilities of Inaccuracies in Registrations with No Returned Mail

<i>Dep Var.: Any Discrepancy</i> Ind. Vars.	Florida			Los Angeles		
	Coef.	Std. Err.	$P > z $	Coef.	Std. Err.	$P > z $
Voted Gen. 08	-1.04	0.20	0.00	-0.63	0.26	0.02
Voted Pres. Prim. 08	-0.86	0.18	0.00	-0.43	0.20	0.03
Voted Gen. 06	-0.44	0.18	0.01	-0.41	0.20	0.04
Age	0.00	0.00	0.55	-0.01	0.00	0.02
Democrat	-0.60	0.34	0.08	0.98	0.61	0.11
No Party	-0.29	0.36	0.43	0.69	0.63	0.27
Republican	-0.49	0.34	0.16	0.75	0.62	0.23
White	-0.10	0.29	0.72			
Black	0.24	0.38	0.52			
Hispanic	-0.26	0.36	0.47			
Constant	0.24	0.49	0.62	-1.38	0.67	0.04
Observations	1911				1722	
Log Likelihood	-624				-529	

All variables aside from age are indicator variables. The dependent variable includes discrepancies in name, address, birthdate, vote history and reported deaths as described in the text.

Using the predictions from this logit model, we arrive at an average predicted probability of a registrant having an inaccuracy given that no mail was returned as undeliverable: 18.1%

for Florida registrants. We multiply this number by the percent of the sample for whom no mail was returned, 90.7%, and arrive at an estimate of **16.4%** of the total sample of registrations invalid due to discrepancies in names, addresses, birthdates, and vote history. The equivalent estimate for Los Angeles is 13.3% of registrations with no mail returned having discrepancies, times 95.1% of the sample with no returned mail = **12.7%**.

C. Rate of Inaccuracy for Registrants with One Piece of Undeliverable Mail

The final part of the estimation exercise is to account for those registrants for whom one piece of mail was returned as undeliverable. We cannot treat these individuals like those for whom both pieces of mail were returned because some of them did receive the other piece of mail. We know this because a small number of them responded to the questionnaire. However, the response rate for this group is so low (5.5%) that it would be unreasonable to expect that all of these individuals received the non-returned piece of mail.

We thus divide this group of registrants into two subgroups. One subgroup will be treated as if both mailers were returned as undeliverable, and we will thus assume a 100% invalid rate for this group. The other subgroup will be treated as if no mail was returned as undeliverable, and so we will assign them the invalid probability devised in the previous section.

To divide these registrants, consider the response rate to the survey of those for whom one piece of mail was returned: 16 out of 291 registrants in this category responded to the questionnaire in the Florida sample. Some set, x , of these 291 registrants, where $x \geq 16$, did actually receive the other survey. We might suppose these x voters would respond at the same rate as those who received both mailers, 21.1%. We solve for x by assuming that 16 is $.216 \times x$ and so $x = 74$. Thus, 74 of these 291 registrants (25.4%) will be given the invalid rate of 18.1% and the rest will be given the invalid rate of 100%.⁴ In mathematical terms,

⁴To the extent that receiving two mailers rather than one increases the response rate, we might sup-

$$Pr(I|OneLetterReturned) = .254(.181) + .746(1) = .792. \quad (2)$$

Among Florida registrants, we estimate that 79.2% of those for whom one piece of mail was returned as undeliverable have invalid records. We multiply this by the percentage of the sample for whom one piece was returned, and arrive at $.792 \times .029 = .023$. That is, **2.3%** of the registration list in Florida is invalid due to inaccuracies among those with one piece of undeliverable mail. The equivalent percentage in Los Angeles is **2.6%**.

D. Total Rate of Inaccuracy in Florida and Los Angeles

To arrive at an estimate of the total rate of invalid records, we simply add together the invalid rates of the three mutually exclusive and collectively exhaustive groups of registrants. The results are displayed in Table 7. We estimate that the probability of a record being invalid either because the address is undeliverable or because there is a discrepancy on the registration record is 25.5% in Florida and 17.3% in Los Angeles.

Table 7: Estimated Total Rate of Inaccuracy in Florida and Los Angeles

	0 Pieces Returned		1 Piece Returned		2 Pieces Returned	Total
	% Invalid	+	% Invalid	+	% Invalid	
Florida	(0.181×0.907)	+	(0.792×0.029)	+	(1×0.064)	= 0.251
Los Angeles	(0.133×0.951)	+	(0.893×0.029)	+	(1×0.020)	= 0.173

For each subgroup, the invalid rate is equal to the rate of inaccuracy for the subgroup times the proportion of the sample population in that subgroup. Note that these estimates require a number of assumptions that, though justifiable, are subject to reasonable disagreement.

pose that using the 21.1% response rate for these registrants is too high. This assumption is therefore a conservative choice in that it may deflate the expected overall invalid rate.

Estimation under Alternative Assumptions

We now offer several alternative measures under slightly different assumptions from our first estimate. Each of these alternative specifications treats undeliverable mail in the same fashion as described above but alters inferences made about the survey responses.

In the first alternative estimate, we ignore address discrepancies reported by survey respondents. The rationale for doing so is that most people who said their address was incorrect were recent movers. Since it would be infeasible for a voter registration database to keep up with movers on an instantaneous basis, these recent movers may very well have their records corrected by the next election.

Registrants who change their permanent address with the U.S. Post Office can have their mail forwarded for the period of twelve months. After that time, their mail will be returned to the sender as undeliverable. Recent movers will thus not show up in the undeliverable mail unless they did not leave a forwarding address. Undeliverable mail for movers can signal that election administrations are not keeping records up to date; registrant-reported recent moves do not signal as serious a problem, since no time is allowed for the administrators to catch up with the movers. In this first alternative estimate, then, we alter the discrepancy rate for registrants by ignoring address discrepancies.

In the second alternative estimate, we remove an additional factor from the discrepancies identified in the questionnaires. In this estimate, we ignore vote history inaccuracies. Similar to the rationale in address changes, it may be the case that election administrators take some time to update vote histories. Updating vote histories requires assembling Election Day precinct sign-in sheets as well as collecting vote histories from registrants who voted by mail or participated in early voting. This process can take time, and it may not be an administrator's highest priority. Furthermore, ignoring vote history inaccuracies obviates the need to account for any over- or under-reporting of turnout in the surveys (see the companion report for a more detailed discussion of this issue). Of course, vote history discrepancies can

pose a serious problem if they are basis for removing voters from the registration system, but perhaps these are less serious discrepancies than those associated with a voter’s name, birthdate, or address. In the second alternative, we ignore *both* vote history discrepancies *and* address discrepancies from our measures.

In yet another estimation, we assume that survey non-response is ignorable. In the first estimate, we argued that responding to the questionnaire was correlated with registrant characteristics, namely their age and their voting patterns. On this basis, we estimated the rate of inaccuracies by generating predicted probabilities from a logistic regression model. Here, we ignore any response bias and use the rate of inaccuracies reported by registrants to estimate the rate of inaccuracies for the full sample of registrants for whom no mail was returned as undeliverable. This assumption we apply to the initial estimate as well as the first two alternative estimates.

Table 8: Alternative Estimates for Total Rate of Inaccuracy

	Response Bias Not Ignorable		Response Bias Ignorable	
	Florida	Los Angeles	Florida	Los Angeles
Primary Estimate	25.1%	17.3	19.3	13.7
Alternative Estimate 1 (Address discrepancies excluded)	20.2	12.8	16.4	10.7
Alternative Estimate 2 (Address/vote history discrepancies excluded)	10.8	9.6	10.0	8.1

Table 8 shows a total of six total discrepancy estimates for each location. The table shows the primary estimate that we described in detail, in addition to the estimates ignoring address and vote history corrections. For each of these three estimates, we present versions that do and do not take into account response bias. For Florida, the inaccuracy ranges from a 10% to 25% probability than an individual record will contain an inaccuracy. In Los

Angeles, the estimated probability ranges from 8% to 17%.

The rate of invalid records is evidently higher in Florida than in Los Angeles not only because there was a much higher undeliverable mail rate in Florida, but also because Florida respondents were more likely to identify discrepancies on their vote records than Los Angeles respondents. The quality of the Los Angeles list may be surprising, given that the state is uninvolved in list maintenance in California. Evidence from Los Angeles seems to contradict the 2005 claim by the Commission on Federal Election Reform that “bottom-up systems are not capable of providing a complete, accurate, current, and valid registration list (11).” As a large county with greater resources than most local jurisdictions, Los Angeles may be an exception rather than a rule. Nonetheless, relative to Florida’s top-down approach, Los Angeles seems to be maintaining a very high quality list without state intervention.

7 Cost of Conducting Audits

In this section we review the costs associated with conducting mail-based registration audits, in the event that other scholars or practitioners wish to replicate this process. Table 9 reports the itemized cost of the audits. The roughly \$15,000 spent for two mailings to 8,000 Los Angeles registrants compares to a cost of approximately \$19,000 for two mailings to 11,000 Floridians. In Los Angeles, there was no cost to purchase the voter list, though unless done in coordination with list administrators, there would generally be a cost associated with the data acquisition as well (see Gronke and Schreiber 2008).

Having conducted a statewide audit in Florida, we will use the Florida cost estimate to provide a approximate cost for auditing the entire United States. The 11,000-person sample size in Florida does not allow for a detailed examination of less populous counties, but it includes enough cases across the state so that like counties can be grouped together for analysis. Gronke and Schreiber (2008) report that to collect a full national list of voters from

Table 9: Printing Costs for Audits

Item	Los Angeles		Florida	
	First Mailing	Second Mailing	First Mailing	Second Mailing
Data Processing	\$170.00	42.50	340.00	31.25
Print Surveys	1419.32	1419.32	1416.74	1416.74
Print Business Reply Envelopes	446.94	435.69	592.79	570.67
Outgoing Envelopes	152.28	152.28	231.79	231.79
Labor (Addressing and Stuffing)	1400.00	1400.00	1921.68	1921.68
Outgoing Metered Postage	3424.00	3424.00	4699.87	4699.87
Business Reply Charges		846.41		875.60
Total		\$14,732.74		\$18,950.47

the 49 states with voter registration, plus Washington DC, costs \$102,826. Thus conducting an 11,000-person survey in 50 statewide jurisdictions would cost $\$102,826 + 50 \times \$19,000 = \$1,052,826$. Note that this estimate of approximately one million dollars does not include costs associated with survey design and data analysis.

Rather than purchasing data from individual states and counties, which is quite expensive, one alternative for researchers is to purchase lists from political data vendors. Several companies, mostly based in Washington DC, serve political campaigns and interest groups by providing cleaned-up versions of the voter files in every state. These companies, such as Aristotle and Catalist, invest substantial resources in studying and improving the voter data they acquire from the states. As a result, the vendors not only have voter lists to sell but they also possess a wealth of knowledge about the quality of data coming from individual states and counties. Ansolabehere and Hersh are currently conducting a national study of voter registration systems utilizing this approach, findings forthcoming.

8 Conclusions

In this memo, we confronted several surprising findings. First, whereas we expected to find problems with the voting registration system through studying undeliverable mail and self-reported discrepancies, we did not foresee the data merging and data completeness prob-

lems associated with voter lists in Hillsborough, Duval, and Los Angeles counties. Indeed, these problems, identified in Section 4 above, are quite concerning since they have the potential to cause problems for large numbers of voters. A typo on an individual record seems far less serious than the potential for an entire precinct or county file being corrupted on account of a faulty data merge. Initial evidence here suggests that more research must be conducted to study a.) the number of counties where these problems occur, b.) whether or not these data problems are permanent or temporary, and c.) how such problems might affect the ability of citizens to vote.

The second surprise was that in spite of the fact that information on the voter registration list is provided by individuals and is manually entered into voter files by administrators, the quality of the lists, measured by undeliverable mail, is apparently superior to the the 2000 U.S. Census mailing list and state jury lists, which utilize a variety of tools to identify individual addresses, including commercial data.

A third surprise was that undeliverable mail does not necessarily imply an obsolete registration address. The fact that 30-50% of people whose addresses are not mailable still vote at these addresses is quite a puzzle. It certainly suggests that election administrators ought to be very cautious before purging registrants on the basis of undeliverable mail. Moving forward from here, election administrators will need more research to figure out why this phenomenon exists. Who are these individuals who are voting at unmailable addresses, and how should they be dealt with in the voter registration system?

Aside from identifying these three surprises, this memo has demonstrated a method for conducting registration list audits and a method for estimating the rate of inaccuracy on a list by utilizing self-reported discrepancies as well as undeliverable mail. These methods, as well as the data from the two pilot studies described in our reports, should pave the way for future research assessing the quality of voter registration lists in the United States.

References

- Atkeson, Lonna Rae, R. Michael Alvarez and Thad E. Hall. September 22, 2008. “The New Mexico 2006 Post Election Audit Report.” University of New Mexico and University of Utah.
- California Secretary of State. 2008. “October 20, 2008 Report of Registration.” <http://www.sos.ca.gov/elections/ror/ror-pages/15day-presgen-08/county.pdf>.
- Citro, Constance F. and Graham Kalton, eds. 2007. *Using the American Community Survey: Benefits and Challenges*. Washington, DC: National Academies Press.
- County of Los Angeles Registrar-Recorder/County Clerk. 2009. “Los Angeles County’s - Past Election Results.” <http://rrcc.co.la.ca.us/elect/pastrslt.html-ssi>.
- Electionline.org. June 2005. “Assorted Rolls: Statewide Voter Registration Databases Under HAVA.” http://www.pewcenteronthestates.org/report_detail.aspx?id=35430. Election Reform Information Project.
- Florida Department of State. 2009a. “Election Reporting System.” <https://doe.dos.state.fl.us/elections/resultsarchive/index.asp>.
- Florida Department of State. July 2009b. “Voter Registration by Party Affiliation.” Florida Division of Elections, Voter Registration Statistics. <http://election.dos.state.fl.us/nvra/affiliation.asp>.
- Gronke, Paul and Bailey Schreiber. December 2008. A 50-State Assessment of Data Availability and Data Reporting. In *Data for Democracy*. Pew Center on the States pp. 48–53.
- Internal Revenue Service. November 5, 2009. “IRS Seeks to Return \$123.5 Million in Undeliverable Refunds to Taxpayers; IRS Reminds Taxpayers to Use E-file and Direct Deposit.” <http://www.irs.gov/newsroom/article/0,,id=215117,00.html>. Press Release.

Kohn, Felipe. April 10, 2003. The United States Postal Service Undeliverable Rates for Census 2000 MailOut Questionnaires Final Report. Technical report Decennial Statistical Studies Division, U.S. Bureau of the Census.

Mihm, J. Christopher. November 19, 1999. “Decennial Census: Information on the Accuracy of Address Coverage.” Letter to Rep. Dan Miller, Chairman, Subcommittee on the Census, House Committee on Government Reform. <http://www.gao.gov/archive/2000/gg00029r.pdf>.

Mize, Gregory E., Paula Hannaford-Agor and Nicole L. Waters. April 2007. “The State-of-the-States Survey of Jury Improvement Efforts.” National Center of State Courts and the State Justice Institute. Local Court Surveys.

National Commission on Federal Election Reform. August 2001. “To Assure Pride and Confidence in the Electoral Process.” Miller Center of Public Affairs and The Century Foundation.

The National Commission on Federal Election Reform. September 2005. “Building Confidence in U.S. Elections.” Center for Democracy and Election Management, American University.

Treat, James B. August 5-9, 2001. Analysis of the Effect Undeliverable Questionnaires have on Mail Response Rates. United States Department of Commerce, Bureau of the Census Proceedings of the Annual Meeting of the American Statistics Association.

Appendix B



**Massachusetts
Institute of
Technology**

Stephen Ansolabehere

Professor, Department of Political Science

<date>

Dear <first> <last>:

You have been selected at random as part of a study by researchers at the Massachusetts Institute of Technology and Yale University on the accuracy of voter registration lists in the United States. The voter registration lists contain information about your registration and participation in elections and are public record. It is important to have accurate lists to make sure elections work properly.

This study is an effort to verify the accuracy of the information in these public records. This study is just a check on the information on the voter lists. It will not be used to change your registration in any way, but it will help researchers and the local elections office better understand the quality of the registration lists. Your participation in this study is voluntary, and any information gathered in this study will be kept entirely confidential.

Please complete the information on the reverse side of this letter and return it in the enclosed envelope.

This research is funded by the Pew and JEHT Foundations. Any inquiries should be directed to Professor Stephen Ansolabehere (1-877-278-2981) at MIT.

Thank you for your cooperation.

A handwritten signature in black ink that reads 'Stephen Ansolabehere'.

Professor Stephen Ansolabehere
Massachusetts Institute of Technology

PLEASE FILL OUT AND RETURN IN THE ENVELOPE ENCLOSED.

Have you moved within the past two years? Yes___ No___

Have you changed your registration within the past two years? Yes___ No___

We have you listed as follows:

Name: <first> <middle> <last> <suffix>

Address: <address1> <address2>

<city>, <state>, <zipcode>

Date of Birth: <birthdate>

Is this information correct? Yes ___ No ___

If not, please write in the correct information next to the information above.

Your race is listed as <race_full>.

Is that correct? Yes ___ No ___ .

If not, please write in the correct information next to the information above.

Your party registration is listed as <party_full>.

Is that correct? Yes ___ No ___

If not, what should your party registration be? _____

According to the registration rolls, you <voted_gen06> in the November 2006 General Election.

Is that correct? Yes___ No___

If you voted, were you registered at the above address at the time? Yes___ No___

If you voted in the 2006 general election, what method did you use? (circle one)

Vote By Mail Vote At Polling Place Voted Early in Person

According to the registration rolls, you <voted_pri08> in the 2008 Presidential Primary Election.

Is that correct? Yes___ No___

If you voted in the 2008 Presidential Primary Election, what method did you use? (circle one)

Vote By Mail Vote At Polling Place Voted Early in Person

According to the registration rolls, you <voted_gen08> in the 2008 General Election. Is that correct? Yes___ No___

If you voted in the 2008 General Election, what method did you use? (circle one)

Vote By Mail Vote At Polling Place Voted Early in Person

THANK YOU!

<survey_id>



**Massachusetts
Institute of
Technology**

Stephen Ansolabehere

Professor, Department of Political Science

<date>

Estimado(a) Sr(a). <first> <last>:

Usted a sido seleccionado al azar para forma parte de un estudio realizado por el Massachussets Institute of Technology y Yale University, sobre la calidad de la información contenida en el registro electoral en los Estados Unidos. El registro electoral contiene información sobre el registro y la participación de los electores en las elecciones y es del dominio público. Es importante que la calidad de la información contenida en el registro electoral sea adecuada para asegurar el buen desempeño de las jornadas electorales.

El principal objetivo de este estudio es verificar la exactitud de la información que está contenida en el registro electoral. El presente estudio no será usado para cambiar su registro de ninguna manera; sólo será útil para entender la calidad del registro electoral. Su particiación en este estudio es voluntaria y cualquier información colectada durante el estudio será mantenida en estricta confidencialidad.

Por favor, complete la información contenida en el reverso de esta carta y regrésela en el sobre adjunto.

La presente investigación es patrocinada por la Fundación Pew y JEHT. Por favor, dirija cualquier pregunta al Prof. Stephen Ansolabehere (1-877-278-2981) en MIT.

Gracias por su cooperación.

A handwritten signature in black ink, appearing to read 'Stephen Ansolabehere', written in a cursive style.

Professor Stephen Ansolabehere
Massachusetts Institute of Technology

POR FAVOR COMPLETE LA SIGUIENTE FORMA Y REGRÉSELA EN EL SOBRE ADJUNTO.

¿Ha cambiado de domicilio los últimos dos años? Sí__ No__

¿Ha cambiado su registro de votante durante de los últimos dos años? Sí__ No__

Su información muestra lo siguiente:

Nombre: **<first>** **<middle>** **<last>** **<suffix>**

Dirección: **<address1>** **<address2>**
<city>, **<state>**, **<zipcode>**

Fecha de nacimiento: **<birthdate>**

¿Esta información es correcta? Sí __ No__

Si no es correcta, por favor escriba la información correcta al costado de la información proveída.

Su raza está registrada como **<race_span>**.

¿Es correcto? Sí __ No __

Si no es correcta, por favor escriba la información correcta al costado de la información proveída.

Su afiliación partidista está clasificada como **<party_span>**. ¿Es correcto? Sí __ No __

Si su afiliación partidista no es correcta, ¿Cuál es su afiliación partidista? _____

Según el registro, usted **<voted_gen06sp>** en la elección General en noviembre del 2006

¿Es correcto? Sí__ No__

Si usted votó, ¿Estaba registrado en ese tiempo en la misma dirección proveída arriba?

Sí__ No__

Si usted votó en la elección general 2006, ¿Cómo lo hizo? (Por favor, circule la respuesta correcta)

Por Correo Temprano en el Área de Votar Votó Temprano Personalmente

Según el registro, usted **<voted_pri08sp>** en las elecciones primarias presidenciales del 2008.

¿Es correcto? Sí__ No__

Si usted votó en las elecciones primarias 2008, ¿Cómo lo hizo? (Por favor, circule la respuesta correcta)

Por Correo Temprano en el Área de Votar Votó Temprano Personalmente

Según el registro, usted **<voted_gen08sp>** en la elección General en noviembre del 2008.

¿Es correcto? Sí__ No__

Si usted votó en la elección general 2008, ¿Cómo lo hizo? (Por favor, circule la respuesta correcta)

Por Correo Temprano en el Área de Votar Votó Temprano Personalmente

GRACIAS!

<survey_id>



DEAN C. LOGAN
Registrar-Recorder/County Clerk



**Massachusetts
Institute of
Technology**

STEPHEN D. ANSOLABEHRE
Professor, Department of Political Science

<date>

Dear <name of voter>,

You have been selected at random as part of a study by researchers at the Massachusetts Institute of Technology and Yale University on the accuracy of Los Angeles County's voter registration lists. The voter registration lists contain information about your registration and participation in elections.

This study is an effort to verify the accuracy of the information in these public records. It is being conducted with the cooperation of the county election office. This study is just a check on the information on the voter lists. It will not be used to change your registration in any way, but it will help researchers and the Los Angeles County elections office better understand the quality of the registration lists. Your participation in this study is voluntary, and any information gathered in this study will be kept entirely confidential.

Please complete the information on the reverse side of this letter and return it in the enclosed self addressed envelope.

This research is funded by the Pew and JEHT Foundations. Any inquiries should be directed to Professor Stephen Ansolabehere, John Lovett, or Eitan Hersh at MIT.

Thank you for your cooperation.

Dean Logan
Registrar-Recorder/County Clerk

Professor Stephen Ansolabehere
Massachusetts Institute of Technology

Have you moved within the past two years? Yes___ No___

Have you changed your registration within the past two years? Yes___ No___

We have you listed as follows:

Name:_____

Address:_____

Date of Birth:_____

Is this information Correct? Yes ___ No ___

If no, please write in the correct information next to the information above.

Your party registration is listed as <Democrat/Republican/Other Party Name/Unaffiliated>.

Is that correct? Yes ___ No ___

If not correct, what should your party registration be ? _____

According to the registration rolls, you <VOTED/DID NOT VOTE> in the November 2006 general election.

Is that correct? Yes___ No___

If you voted were you registered at the above address at the time? Yes___ No___

If no, where? _____

If you voted in the 2006 general election, what method did you use? (circle one)

Vote By Mail Vote At Polling Place

According to the registration rolls, you <VOTED/DID NOT VOTE> in the 2008 presidential primary election. Is that correct? Yes___ No___

If you voted in the 2008 primary election, what method did you use ? (circle one)

Vote By Mail Vote At Polling Place

Do you intend to vote in the 2008 presidential general election on November 4 this year?

Yes___ No___ Not Sure___

PLEASE RETURN IN THE ENVELOPE ENCLOSED.



DEAN C. LOGAN
Registrar-Recorder/County Clerk



**Massachusetts
Institute of
Technology**

STEPHEN D. ANSOLABEHERE
Professor, Department of Political Science

Estimado <Nombre del Votador>

Usted a sido escogido al azar como parte de un estudio por la Oficina de Elección del condado de Los Angeles y los investigadores del Massachusetts Institute of Technology y Yale University para la certitud de la lista de registro de votadores en Los Angeles. Las listas del registro de votantes contienen la información sobre su registro y participación en elecciones.

Este estudio es un esfuerzo para verificar la exactitud de la información en estos expedientes públicos. Este estudio es nada mas una averiguación de la información sobre las listas del votante. No será utilizado para cambiar su registro de ninguna manera, sino que ayudará a la oficina de la elección en Los Angeles a entender mejor la calidad de las listas del registro. Su participación en este estudio es voluntaria, y cualquier información personal obtenida en este estudio será mantenida enteramente confidencial.

Termine por favor la información al reverso de esta carta y regrese a la oficina de elección de Los Angeles en el sobre incluido.

Esta investigación es financiada por el banco y las fundaciones de Pew y JEHT. Cualquier pregunta se debe dirigir al profesor Stephen Ansolabehere de MIT.

Gracias por su cooperación.

Dean Logan
Registrar-Recorder/County Clerk

Professor Stephen Ansolabehere
Massachusetts Institute of Technology

¿Usted se ha movido en el plazo de los últimos dos años? Si__ No__

¿Usted ha cambiado su registro de voto durante de los últimos dos años? Si__ No__

Su información enseña lo siguiente:

Nombre: _____

Dirección: _____

Fecha de nacimiento: _____

¿Está esta información correcta? Sí __ No__

Si no, escriba por favor la información correcta al lado de la información arriba.

Se enumera su registro del partido como <Demócrata/Republicano/Otro Partido/Sin Afiliación.> ¿Está eso correcto? Sí __ No __

Si no, corrija. ¿Qué debe ser su registro del partido? _____

Según los rodillos de registro, usted <Voto/No Voto.> en la elección General en noviembre del 2006

¿Está eso correcto? Si__ No__

¿Si usted votó, estaba registrado en ese tiempo en la dirección arriba? Si__ No__

¿Si no, dónde? _____

¿Si usted votó en la elección general 2006, cual manera uso usted? (círculo uno)

Por Correo Temprano en Área de Votar

Según los rodillos de registro, usted <Votó/No Votó> en las elecciones primarias presidenciales del 2008.

¿Está eso correcto? Si__ No__

¿Si usted votó en las elecciones primarias 2008, cual manera uso usted? (círculo uno)

Por Correo Temprano en Área de Votar

¿Usted se propone votar en la elección general presidencial 2008 el 4 de noviembre este año? Si__

No__ No Estoy Seguro__



Prof. Stephen Ansolabehere
 Department of Political Science
 Massachusetts Institute of Technology
 77 Massachusetts Avenue
 Building E53, Room 449
 Cambridge, MA 02139-4307

PRESORTED
 FIRST CLASS



LAKE WORTH FL 33467

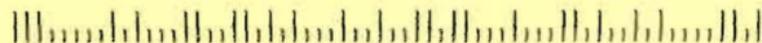
334 DFE 1 909C 00 09/11/09
 RETURN TO SENDER



MOVED LEFT NO ADDRESS
 UNABLE TO FORWARD
 RETURN TO SENDER

BC: 02139430799 *1697-03336-11-34

38





NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 21 NEW HAVEN CT

POSTAGE WILL BE PAID BY ADDRESSEE



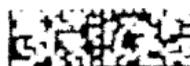
MIT/ YALE FLORIDA REGISTRATION PROJECT
PO BOX 208331
NEW HAVEN CT 06520-9917



ATTN TECHNICAL SERVICES BUREAU
REGISTRAR-RECORDER/COUNTY CLERK
PO BOX 3024
SANTA FE SPRINGS, CA 90670-9908

PRESORTED
FIRST CLASS

US POSTAGE
00 32



LOS ANGELES CA 90024-8304



900 DE 1 00 10/25/C

RETURN TO SENDER
ATTEMPTED - NOT KNOWN
UNABLE TO FORWARD

BC: 90670002424 *1062 05313 25 2

90670@0024
HALF 551 9002

40

