Whose Vote is Lost by Mail? Evidence from Philadelphia in the 2020 General Election*

Daniel J. Hopkins
University of Pennsylvania
danhop@sas.upenn.edu

Marc Meredith
University of Pennsylvania
marcmere@sas.upenn.edu

Anjali Chainani
Anavi Strategies
anjaliachainani@gmail.com

Nathaniel Olin
Independent Researcher
nathaniel.olin@gmail.com

October 28, 2022

Abstract

The 2020 elections raised the salience of procedural frictions that prevent people who intend to vote by mail from casting a counted ballot. Here, we focus on the incidence of such lost mail votes in Philadelphia County during the 2020 general election. There were heightened concerns about lost mail votes in that election because a judicial ruling issued just before the election stated that mail ballots returned outside of a secrecy envelope—commonly referred to as "naked ballots"—would not count. Our combination of observational and experimental analyses show who was at the greatest risk of casting a lost mail vote and how voter education efforts reduced their incidence. We find that voters from racial/ethnic minority backgrounds were more likely to return a naked ballot than White voters, partly due to language barriers. Younger voters were less likely to return a naked ballot than older voters, a finding which may be linked to widespread communication efforts seeking to educate citizens about how to successfully cast a mail ballot, often via social media. Our analysis of a randomized experiment exploring the effects of a postcard sent to almost 18,000 mail-ballot recipients further demonstrates how communication can reduce the risk of lost mail votes, especially by getting people to return their mail ballots more quickly.

1   Introduction

Since Gerber and Green (2000), hundreds of experiments have been conducted to understand which communication efforts most effectively mobilize people to cast ballots (see Green and Gerber 2019 for a summary). Less is known about what can be done to increase the share of mobilized citizens who cast ballots that are counted. Some would-be voters leave a polling place without casting a ballot, particularly when waiting times are long (Spencer and Markovits, 2010; Stein et al., 2020), and others cast ballots that are ultimately rejected. This rejection could occur for many reasons, including because the potential voter failed to present required documentation, cast the ballot at an incorrect polling place, or had their registration record canceled (Merivaki and Smith, 2020). Stewart (2020) argues that mail voters are particularly sensitive to the specifics of election administration, noting that about 20 percent of mail ballots distributed in the 2016 presidential election were never returned and that 1 percent of the returned mail ballots were ultimately rejected. Stewart refers to mail ballots that do not ultimately count because of imperfect election administration as "lost votes by mail."

This paper focuses on lost votes by mail in the 2020 presidential election in Philadelphia County, Pennsylvania.[1] Concerns about lost votes by mail were especially salient in this context. Philadelphia is a majority-minority city in a highly competitive swing state, and previous research shows that mail ballots cast by minority voters are more likely to be rejected than those cast by White voters (Baringer, Herron and Smith, 2020; Shino, Suttmann-Lea and Smith, 2022). Moreover, the combination of a recent law change and increased demand for mail ballots due to the COVID-19 pandemic meant that many voters were casting mail ballots for the first time, which increases the chances of that their mail ballot gets rejected (Cottrell, Herron and Smith, 2021). Finally, there were widespread concerns about naked ballots, a phrase used to describe mail ballots returned outside of a secrecy envelope. The Pennsylvania Supreme Court ruled just before the election that naked ballots were not to be

---

[1]Philadelphia County is coterminous with the City of Philadelphia.

counted, generating substantial apprehension that voters unaware of this requirement would cast tens of thousands of mail ballots that must be rejected (Deeley, 2020). Ultimately, about 4,000, or 1 percent, of mail ballots returned in Philadelphia were naked ballots, which was similar to the estimated share statewide (Hopkins, Meredith and Wang, 2021).

A combination of observational and experimental analyses document which groups of voters were especially likely to have lost mail votes—and also suggest why there may have been fewer lost votes by mail in Philadelphia than many had feared. Using novel, individual-level administrative data compiled from multiple sources, we first document which registrants failed to return a mail ballot that they requested. Next, we conduct an observational analysis of what predicted whether a mail ballot was rejected because it was returned outside of a secrecy envelope. We then evaluate a block-randomized field experiment, in which government officials sent a subset of registrants an instructional postcard after they were sent their mail ballots. This postcard specifically emphasized how ballots and associated documentation should be filled out to ensure that a returned mail ballot is counted.

Several of our findings are consistent with prior research. For instance, our observational analysis shows that people with more voting experience, and particularly previous experience voting by mail, were more likely to return their requested mail ballot and less likely to cast a naked ballot. Also in keeping with previous work, we find that voters from racial/ethnic minority backgrounds were more likely to cast a naked ballot than White voters. Additional analyses that incorporate information on the Census demographics of a voter's neighborhood suggest that this is partially caused by minority voters being more likely to face language barriers than White voters.

However, both the observational and experimental analyses also suggest that the extensive informational campaigns that arose to educate voters about mail balloting in the 2020 general election may have reduced the incidence of lost votes by mail. Contrary to previous work, we find that *younger* mail-ballot voters were less likely to cast naked ballots than older mail-ballot voters in the 2020 general election. We suggest that at least some of this

difference—as well the higher rate of naked ballots among Republican registrants than Democratic registrants—was potentially a consequence of younger voters and Democrats having more exposure to information about naked ballots than older voters and Republicans.

Our experimental analyses highlight three different ways in which being sent an instructional postcard increased the likelihood that a registrant successfully cast a mail ballot. First, mailers caused registrants to return their mail ballots more quickly, even though it lacked any reference to when mail ballots should be returned. This suggests that the postcard increased the salience of an unreturned mail ballot. Second, postcards reduced the likelihood that a mail ballot was canceled. Some of the potential mechanisms that could cause this to occur include the aforementioned speedier return of mail ballots among postcard recipients, the postcard reducing the incidence of clerical errors on mail ballots, and the postcard increasing understanding of—and confidence in—the mail-balloting process. Finally, registrants sent a postcard were more likely to cast a mail ballot than registrants who were not, although this was partially offset by differences between the two groups in in-person voting. The combination of increased follow-through on returning requested mail ballots and increased speed of mail-ballot return suggests that sending a registrant an instructional postcard after sending their mail ballot may increase the number of counted ballots and reduce the administrative burden on elections officials. Postcard recipients were also about 0.1 percentage-points less likely to return a naked ballot than people in the control group, although this difference is not statistically significant ($p = 0.171$, one-tailed). Considered together, these results reinforce the conclusion that election administration policies and procedures do not have static effects. Instead, their impacts depend on the actions of election administrators, campaigns, and other actors (see also Hopkins et al., 2017).

This paper proceeds as follows. Section 2 situates this study within the broader academic literature on unreturned and rejected mail ballots. Section 3 describes why unreturned and rejected mail ballots were particularly salient in the 2020 presidential election in Philadelphia County. Sections 4 and 5 detail the postcard experiment and the data employed to analyze it,

respectively. Section 6 presents our findings. Section 7 concludes by discussing the broader implications of our findings and suggests directions for future research.

2   Related Research

In this section, we detail prior research on lost votes by mail and develop expectations for their incidence in Philadelphia in 2020. Stewart (2020) defines a lost vote as a case in which imperfect election administration prevents someone from casting a vote for the candidate of their choice. Features of election administration that could generate lost votes include faulty equipment, confusing ballots, registration problems, and poor polling place operations (Caltech/MIT Voting Technology Project, 2001). Stewart highlights some aspects of election administration that are unique or especially challenging for mail balloting relative to in-person balloting, including processing mail ballot requests, distributing mail ballots to registrants who request them, and receiving, verifying, and tabulating received mail ballots in a timely way. As a consequence, people casting mail ballots are at a greater risk of having their vote lost than people casting an in-person ballot.

Estimating the incidence of lost mail ballots has proven challenging, in part because of data limitations. Prior to the development of the Election Assistance Commission's Election Administration and Voting Survey (EAVS), there was no systematic documentation of quantities like the number of mail ballots that were requested, distributed, received, or counted in the United States. Stewart (2010) used the measures of these quantities by county from the 2008 EAVS to estimate that about 7.6 of the 35.5 million mail ballots requested in the 2008 presidential election (21%) ultimately did not count because they were either not transmitted to voters, not returned, or not counted. However, two features of EAVS data limit what can be learned about lost votes from such analyses. First, while the EAVS continues to improve its data integrity, any administrative survey of its size and scope includes measurement error and missing data. Second, and more problematic, an unknown number of the mail ballot

requests documented in EAVS data are to registrants who have no intention of voting by mail, particularly in states that utilize vote-by-mail or permanent absentee lists. To account for this, Stewart (2020) uses administrative data from North Carolina's 2016 general election to identify cases in which a mail-ballot recipient decided to vote in-person or purposively abstain. He concludes that only about half of the unreturned mail ballots in that election could represent lost votes.

Because of the challenges of observing lost votes in EAVS data, a number of recent papers instead use individual-level administrative data to identify characteristics that are associated with casting mail ballots that do not count, either because they are received too late or contain a clerical error (Baringer, Herron and Smith, 2020; Cottrell, Herron and Smith, 2021; Shino, Suttmann-Lea and Smith, 2022). Some common patterns emerge from this work. Arguably the most consistent finding is that young voters, especially those who recently became eligible to vote, are more likely than older voters to cast mail ballots that are rejected. It is also generally the case that racial or ethnic minority voters are more likely to cast a rejected mail ballot than White voters. In addition, Cottrell, Herron and Smith establish that both past turnout and previous experience with mail balloting are associated with a reduced likelihood of casting a mail ballot that gets rejected. They also find that the share of mail ballots rejected for being received too late or containing a clerical error in Florida was substantially lower in the 2020 general election than in either the 2016 or 2018 general elections.

This paper builds upon this previous work in several ways. First, we look not just at the set of voters who cast a mail ballot, but also at the broader universe of registrants who requested mail ballots. We show that just over 50 percent of the unreturned mail ballots in Philadelphia were generated by registrants who substituted returning their ballot for voting in person. This share increasing to almost 60 percent among Republican registrants. This suggests that a smaller share of unreturned mail ballots in the 2020 presidential election represented lost votes than in other recent presidential elections, particularly in

heavily-Republican areas. Second, we examine how neighborhood-level characteristics, in addition to individual-level characteristics, relate to mail-ballot rejection. Doing so provides suggestive evidence about some of the mechanisms that produce the relationships between individual-level characteristics and ballot rejection documented in previous work. Finally, our combination of observational and experimental results establishes the role that informational campaigns can play in reducing the number of lost votes by mail, which potentially contributed to the reduced share of returned mail ballots that were rejected in the 2020 general election.

3   Philadelphia Context

Here, we explain why Philadelphia's 2020 general election is an especially valuable case in which to explore lost votes by mail. Concerns about such lost votes were very salient during the 2020 general election in Philadelphia and in Pennsylvania more broadly. Given that Donald Trump's 44,000 vote margin in Pennsylvania helped to propel him to the presidency in 2016, Pennsylvania was once again hotly contested in the 2020 presidential race.

According to U.S. Election Assistance Commission (2017), at least 2,534 of the 266,208 mail ballots returned in Pennsylvania in the 2016 general election (about 0.9 percent) were rejected. There were several reasons why people anticipated substantially more lost votes by mail in the 2020 general election. Pennsylvania's Act 77, passed in 2019, ended the requirement that registered voters must have a qualified excuse to request a mail ballot. More people took advantage of their newfound ability to request a mail ballot than expected due to the COVID-19 pandemic and higher-than-normal turnout in the 2020 elections, meaning that many Pennsylvanians were voting by mail for the first time.[2]

Alarm about lost votes by mail grew when the Pennsylvania Supreme Court ordered that

---

[2]The Pennsylvania Secretary of State reported that 2,704,147 of the 6,945,045 ballots cast in the 2020 general election (about 39 percent) were cast by mail (Pennsylvania Department of State, 2021). As a point of comparison, only 266,208, or about 4 percent, of the 6,223,150 ballots cast in the 2016 general election were cast by mail (U.S. Election Assistance Commission, 2017).

election officials disqualify mail ballots not returned within a secrecy envelope on September 17, 2020 (see *Pa. Democratic Party v. Boockvar*, No. 133 MM 2020). Pennsylvania is one of 16 states that provides voters with a smaller secrecy envelope that gets placed within a larger mail-ballot return envelope to protect the privacy of voters' choices.[3] The 2020 primary election revealed that county elections officials lacked a common understanding about whether state law mandated the rejection of mail ballots returned outside of a secrecy envelope, which were commonly referred to as naked ballots. In this election, almost 6 percent of the 7,643 of the absentee ballots cast in Lawrence County and almost 5 percent of the 8,548 absentee and provisional ballots cast in Mercer County were rejected for being outside of a secrecy envelope (Wachter, 2020). In contrast, a Philadelphia elections official estimated that they counted 15,000 to 20,000 naked ballots in the 2020 primary (Deeley, 2020). Extrapolating to the 2020 general election, Commissioner Deeley estimated that 30,000 to 40,000 naked ballots might be disqualified in Philadelphia County alone.[4]

Although less salient than naked ballots, there was also apprehension that improperly filled out declarations on mail-ballot return envelopes could generate lost votes by mail. Below the declaration, voters are asked to sign, date, and print their names and addresses. It was already well established that a mail ballot is disqualified if the voter did not sign the declaration. Less clear was what happened if a voter failed to date, print their name, or print their address on the declaration. It wasn't until after the election that the Pennsylvania Supreme Court clarified that mail ballots should not be disqualified in the 2020 general election solely because a voter did not date or print their name and address on the declaration (see *In re: Canvass of Absentee and Mail-In Ballots of November 3, 2020 General Election*,

---

[3]See Table 13 of National Conference of State Legislatures (2020) for list of states that use secrecy envelopes.

[4]Categorizing naked ballots as "lost votes" does not mean that elections officials engaged in imperfect election administration by following the Pennsylvania Supreme Court's order and rejecting mail ballots submitted outside of a secrecy envelope. Rather, we believe that the continued existence of the law requiring mail ballots to be submitted in a secrecy envelope is imperfect election administration. As Deeley (2020) notes, the secrecy envelope helped maintain the anonymity of mail ballots when Pennsylvania had poll workers count the mail ballots cast by voters from their precinct. However, once counties started counting mail ballots centrally, usually using a machine, the secrecy envelope was no longer necessary to maintain their anonymity.

No. 31 EAP 2020, No.32 EAP 2020, No.33 EAP 2020, No.34 EAP 2020, No.35 EAP 2020, No. 29 WAP 2020).

Not all mail ballots returned outside of a secrecy envelope or with a disqualifying clerical error on the declaration become lost mail votes. While the Pennsylvania Supreme Court ruled that counties were not required to contact voters who submitted potentially problematic mail ballots (see *Pa. Democratic Party v. Boockvar*, No. 133 MM 2020), election administrators in some counties, including Philadelphia, canceled some returned mail ballots before Election Day that they knew contained a disqualifying clerical error (Jones, 2020). When election administrators canceled a mail ballot, information about the cancelation then appeared in the Pennsylvania Secretary of State's ballot tracker. Elections officials in some counties also attempted to contact voters directly to make them aware of potential problems with their mail ballot (Couloumbis and Martine, 2020). Registrants who, for whatever reason, were concerned that their mail ballot had a disqualifying error could ask for their mail ballot to be canceled, and then either request a replacement mail ballot or cast an in-person ballot on Election Day.[5]

A Pennsylvania Supreme Court decision partially mitigated additional concerns that a substantial number of mail votes would be lost because they would not be received by election officials on time. These concerns partly grew out of voters' experiences in the 2020 primary election, in which several counties, including Philadelphia, struggled to distribute mail ballots in a timely way to registrants who requested them (Lai, 2020*b*). While Pennsylvania statue requires mail ballots to be received by election officials before polls close on Election Day, tens of thousands of registrants did not receive their mail ballots with sufficient time for this to happen in that election (Hopkins et al., 2021). Thus, seven counties had the Election Day receipt deadline changed to an Election Day postmark deadline for the 2020 primary election through a combination of gubernatorial executive order and judicial order

---

[5]Legal challenges brought by Republican candidates attempting to invalidate the ballots cured through this process were dismissed, because courts ultimately concluded that counties were permitted to do so (see *Barnette v. Lawrence*, No. 2:20-cv-05477, *Donald J. Trump for President v. Montgomery Cty. Bd. of Elections*, No. 2020-18680).

(Commonwealth of Pennsylvania Governor's Office, 2020; Lai, 2020a). Further, the general counsel of the United States Postal Service (USPS) warned the Pennsylvania Secretary of State that Pennsylvania's deadline to request mail ballots of seven days before Election Day was inconsistent with USPS' delivery standards (Lai and Rushing, 2020). Consequentially, the Pennsylvania Supreme Court ruled that only for the 2020 general election, mail ballots received by county election officials within three days of Election Day should count unless the preponderance of the evidence demonstrates that they were mailed after Election Day (see *Pa. Democratic Party v. Boockvar*, No. 133 MM 2020).[6]

While voters nationwide were exposed to extensive informational campaigns about mail balloting in the lead-up to the 2020 general election, this was particularly true in Pennsylvania because of the state-specific concerns about lost mail votes and its central role in the Electoral College. In the wake of the Pennsylvania Supreme Court ruling on naked ballots, the Pennsylvania Department of State (2020) issued guidance strongly encouraging counties to develop instructional inserts describing how voters should mark and return their mail ballots that included a warning about the necessity of returning their ballot in a secrecy envelope. This guidance also encouraged counties to publicize locations that may assist and educate voters, such as county websites and ballot collection sites. Pennsylvania-based media frequently discussed naked ballots, sometimes highlighting one of the many celebrity public service announcements that emphasized the important of returning a mail ballot in a secrecy envelope (e.g., Hatmaker, 2020; Kiner, 2020).

We suspect that some voters were more likely than others to be exposed to this messaging about how to cast a mail ballot. Because substantially more Democrats were voting by mail than Republicans, the Democratic Party was more aggressively working to inform its supporters about how to successfully cast a mail ballot. A spokesman for the Pennsylvania Democratic Party described its education campaign about naked ballots as an "all-hands-on-deck effort from us and the Biden team and the DNC all kind of working hand in hand to

---

[6]While Republican candidates continued to appeal this decision through Election Day, these appeals ultimately failed.

make sure that we spread the word and educate voters"(Meyer, 2020). Thus, we hypothesize that Democrats were more likely to be exposed to messaging about naked ballots than Republicans. Social media was one of the primary avenues for voter education, with Instagram accounts such as @nakedballot using memes and infographics to emphasize the importance of putting mail ballots in their secrecy envelopes. For example, millions of viewers engaged with social media campaigns featuring celebrities or election administrators wearing no clothes to raise the salience of naked ballots (Kurtz, 2020; Walters, 2020). Accordingly, groups with higher levels of social media consumption may have greater exposure to messaging about naked ballots.

4   Experimental Design

We helped to design and implement a field experiment to gain causal leverage on how voter education efforts affected the incidence of lost votes by mail in Philadelphia in the 2020 general election. Figure 1 displays the informational postcard sent by the City of Philadelphia to a randomized subset of Philadelphia registrants after they were sent their mail ballot.[7] This mailer draws attention to the importance of placing the mail ballot in the secrecy envelope before placing it in the mailing envelope. It also highlights the need to sign the declaration on the back of the mailing envelope. Given the uncertainty highlighted previously about the date by which mail ballots had to be received by county election officials and whether declarations needed to include a date or address, this postcard contains no reference to the necessity of other parts of the declaration or when to return the mail ballot.

Even though the information presented in Figure 1 also appears in the original instructions accompanying the mail ballot, there are several reasons why reiterating this information via an informational postcard may increase the likelihood that the voter complies with the instructions. First, it may increase the likelihood that these instructions are ever seen. And
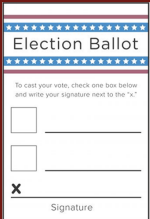
---

[7]Special thanks to Jessica Olarsch and the rest of the Penn Leads the Vote team for designing this mailer.

Figure 1: Mailer Sent to Subset of Philadelphia Registrants Who Requested Mail Ballots

even if the mailing is only reinforcing something that a voter already read, it still may increase that voter's ability to recall this information when casting their ballot. This could happen, for example, if seeing the postcard causes the voter to initiate the process of filling out their ballot. Moreover, receiving this mailing could highlight the importance of complying with the instructions necessary for the ballot to count. This information might be particularly important and novel for voters with limited previous experience voting and casting a mail ballot.

The informational postcard displayed in Figure 1 was similar to a lot of other direct mail being distributed in Pennsylvania in the lead up to the 2020 general election by campaigns, interest groups, and election officials. Speaking about the focus of the Pennsylvania Department of State's office, Kathy Boockvvar, the Secretary of State who oversaw the 2020 general election, mentioned "raising awareness through social and paid media, emails and direct mail that ballots need to be clothed (Murphy, 2020)." Similarly, the Biden campaign emphasized how "its direct mail to voters in Pennsylvania features detailed steps about the mail-in process as well as images based on the state's secrecy envelopes" (Otterbein, 2020). Thus, one should not interpret our experiment as estimating a treatment effect of registrants receiving an informational postcard relative to not receiving a single informational postcard. Rather, the treatment effect that we estimate reflects the marginal effect of registrants receiving one additional mailer beyond those already received, albeit a mailer directly from a government source (see also Hopkins, Schwarz and Chainani, Forthcoming).

The universe for the experiment was a subset of the 338,155 Philadelphia County registrants with an approved mail-ballot application as of October 15, 2020. The goal when constructing the experimental universe was to select a subset of mail-ballot applicants who remained interested in casting a mail ballot but had not already done so before receiving the mailer. To this end, several factors determined which registrants with an approved mail-ballot application were included. First, the experimental universe only included registrants who had recently been sent their mail ballot or had their mail-ballot application approved

but who had not yet been sent their mail ballot as of October 15. Specifically, it included registrants who were sent a mail ballot on October 10, October 14, or October 15, or who had their mail ballot application approved on October 13, October 14, or October 15 but who had not yet been sent a mail ballot. Among this subset, registrants were excluded if they requested that a mail ballot be sent to a ZIP code outside of Philadelphia County or entered what appeared to be an incomplete mailing address. Registrants who had already returned their mail ballot were also excluded. Appendix Table A.1 shows that selecting the experimental subjects in this way caused registrants included in the experimental universe to look different on several dimensions from the general population of registrants with an approved mail-ballot application, in that they were younger, more likely to be an ethnic/racial minority, and less likely to have extensive turnout history.

We used a block-randomized design to assign which 17,781 of the 38,155 registrants within the experimental universe were sent the mailer (Moore, 2012). There were 18 blocks based on each registrant's party of registration (i.e., Democrat, Republican, or other) and the date on which a mail ballot was sent to the registrant or when the mail ballot application was received (i.e., mail ballot sent on October 10, mail ballot sent on October 14, mail ballot sent on October 15, mail ballot application approved on October 13, mailed ballot application approved on October 14, or mail ballot application approved on October 15). Registrants within each block received the mailer with equal probability.

Despite our efforts to keep people who had already returned their mail ballot before receiving the mailer out of the experimental universe, this still happened frequently. Based on test mailings sent to people not included in the experimental universe, we believe that the earliest that anyone received the mailing was October 23. By October 22, at least 24.05 percent of treated registrants and 24.09 percent of untreated registrants had returned their mail ballots according to the data provided by the Pennsylvania Secretary of State's office. Given that, we present our experimental analyses both for all registrants in the experimental universe and all registrants in the experimental universe who had not returned their mail

ballot by October 22.

5 Data

We leverage four sources of data to analyze the characteristics of people who cast naked ballots in Philadelphia and the effects of receiving our mailer on voter behavior. These data include: 2021 Pennsylvania voter files, the 2019 ACS 5-year summary data on the demographics of census tracts, data from the Pennsylvania Secretary of State's office on which registrants had their mail ballot applications approved or had their approved mail ballots canceled in the 2020 general election, and data from the Philadelphia City Commissioners on the specific Philadelphia registrants who submitted a naked ballot in the 2020 general or 2021 primary and who cast mail ballots at an early in-person mail balloting site in the 2020 general election. We describe the variables that we obtained from each of these sources here and provide more details about each of these data sources in the Appendix.

The voter file provides information about a registrant's name, residential address, date of birth, and vote history by mode. We geocoded the residential addresses contained in the voter file, which allowed us to learn the census tract of the residential address and measure the demographics of that census tract using the 2019 ACS 5-year summary. We then imputed the probability of the registrant being from five racial and ethnic backgrounds by combining information about the racial composition of their census tract with their surname (Imai and Khanna, 2016).

We combine data from the Philadelphia City Commissioners and the Pennsylvania Secretary of State to construct a measure of who cast a naked ballot in the 2020 general election. Our primary measure focuses on whether an election administrator canceled the mail ballot because it was submitted outside of a secrecy envelope. However, the previous section highlights that some registrants who cast a naked ballot may have instead requested that their mail ballot be canceled or asked for a replacement mail ballot. Thus, when analyzing

the field experiment, we conduct a robustness analysis that also examines how receiving the mailer affects mail-ballot cancelation more broadly.

6   Results

Here, we report results of observational analyses of all Philadelphia registered voters whose request for a mail ballot was approved for the 2020 general election. We next analyze the predictors of casting naked ballots before turning to an analysis of the field experiment.

6.1   Who Returned Their Mail Ballot?

Table 1 looks at vote outcomes among Philadelphia registrants with an accepted mail-ballot application in the 2020 general election. The first column shows that there were a total of 434,012 registrants with an approved mail-ballot application. The next two rows show that 6,092 of these mail ballots were canceled, 4,188 of which were canceled by Philadelphia election officials because they were submitted outside of a secrecy envelope. The remaining three rows summarize the turnout records for those registrants with an accepted mail-ballot application in the voter file. 368,643 of these registrants (84.9%) had a record of casting a mail ballot, while an additional 33,681 registrants (7.8%) voted at the polls. This leaves only 31,683 registrants (7.3%) who had no record of voting in the voter file after having a mail-ballot application accepted.

Columns 2 through 4 of Table 1 show that past voting experience had substantial predictive power over who cast naked ballots and who returned their mail ballots. Column 2 shows that about 92.5% of registrants who had a record of voting by mail in at least one previous election returned their mail ballot once approved. This is a notably higher rate than the rate among registrants who only had a record of voting in person or registrants with no record of voting, which are 80.7% and 79.6% (columns 3 and 4), respectively. However, registrants with a history of in-person voting were still more likely to vote than registrants with no

Table 1: Vote Outcomes Among Philadelphians with an Approved Mail-Ballot Application in the 2020 General Election by Voter History and Party of Registration

Sample	(1) All Accepted Mail-Ballot Applications	(2) Previously Voted By Mail	(3) Previously Only Voted In-Person	(4) No Previous Vote History	(5) Democratic	(6) Republican	(7) Non-Major Party
		Vote Mode History			**Party of Registration**		
N	434,012	161,147	191,337	81,528	362,087	30,117	41,808
Cancelled Ballots:							
All Canceled Ballots	6,092	1,533	3,210	1,349	4,881	545	666
	(1.4)	(1.0)	(1.7)	(1.7)	(1.4)	(1.8)	(1.6)
Cast Naked Ballots	4,188	683	2,488	1,017	3,339	362	487
	(1.0)	(0.4)	(1.3)	(1.3)	(0.9)	(1.2)	(1.2)
Turnout Outcome:							
Voted by Mail	368,639	149,190	155,215	64,234	311,459	23,052	34,128
	(84.9)	(92.6)	(81.1)	(78.8)	(86.0)	(76.5)	(81.6)
Voted In-Person	33,682	7,216	19,932	6,534	26,176	4,105	3,401
	(7.8)	(4.5)	(10.4)	(8.0)	(7.2)	(13.6)	(8.1)
No Vote Record	31,691	4,741	16,190	10,760	24,452	2,960	4,279
	(7.3)	(2.9)	(8.5)	(13.2)	(6.8)	(9.8)	(10.2)

Note: Number in parentheses is the share of the mail-ballot applicants in the sample with that outcome.

vote history, because they were more likely to vote in person despite requesting a mail ballot (10.5% vs. 7.6%). Registrants who previously cast a mail ballot were also less likely to cast a naked ballot (0.4%) than registrants who had only voted in person or had no history of voting (both 1.3%). One implication of Table 1 is that one should not necessarily interpret data on unreturned ballots, like the EAVS reports, as a measure of abstention. This may be particularly true in the 2020 presidential election, as columns 5 through 7 highlight that there were sizable partisan differences in the rates of in-person voting after requesting a mail ballot. About 13.6% of Republican registrants who requested a mail ballot ultimately voted in person, as compared to 7.2% of Democrats and 8.1% of non-major party registrants. This means that about 58% of Republicans with an unreturned mail ballot voted in-person, as compared to about 52% and 44% of Democrats and non-major party registrants with unreturned mail ballots, respectively. Although we lack the data necessary to know for sure, we suspect that this was at least partially a consequence of Donald Trump's rhetoric against mail balloting. That rhetoric likely led some people, especially Republicans, to vote in person instead of returning their mail ballot. If Trump's rhetoric did cause more people than normal to vote in person instead of returning a requested mail ballot, this would make

unreturned ballots even less reflective of a lost mail vote in the 2020 general election than normal, particularly for Republicans.

6.2 Who Cast A Naked Ballot?

We next analyze what voter characteristics are predictive of returning a naked ballot in the 2020 general election. Our analysis focuses on the subset of Philadelphia registrants who we know returned a mail ballot in the 2020 general election, either because they are denoted in the voter file as casting a mail ballot or listed in the naked-ballot data provided by the Philadelphia City Commissioners. About 1.1% of voters in this sample are denoted as having cast a naked ballot. We model the likelihood that someone cast a naked ballot as a function of both voter characteristics and the characteristics of the voter's neighborhood. At the voter level, we examine how their age, race and ethnicity, party registration, and past turnout relate to their likelihood of casting a naked ballot. At the neighborhood level, we examine how census-tract measures of educational attainment, income, and the ability to speak English relate to the likelihood of casting a naked ballot.

The first column of Table 2 demonstrates that several voter-level variables are associated with the likelihood of casting a naked ballot. First, voters in general elections between 2016 and 2019 were less likely to cast naked ballots in 2020 than non-voters. Local election voters were almost 0.9 percentage points (p.p.) less likely to cast naked ballots, while people who only voted in federal elections were just over 0.2 p.p. less likely to cast a naked ballot. Second, voters who cast an in-person mail ballot at a satellite election office were 0.4 p.p. less likely to cast a naked ballot than voters who received their mail ballot in the mail. Third, voters with a higher probability of being a racial or ethnic minority were more likely to cast a naked ballot. Voters with a high probability of being Black, Hispanic, and Asian American are estimated to be about 0.9, 1.9, and 1.3 p.p. more likely to cast a naked ballot than voters with a high probability of being White, respectively.[8] Third, voters born in the 1970s or

---

[8] Appendix Table A.2 shows similar patterns, although estimated less precisely, if we instead measure race

17

earlier were more likely to cast a naked ballot than voters born in the 1980s or later. Finally, registered Republicans were about 0.5 p.p. more likely to cast a naked ballot than registered Democrats, with those lacking a major-party registration casting naked ballots at a rate in between registered Democrats and registered Republicans. Given that the base rate of casting a naked ballot is just over 1%, all of these differences represent large percentage changes in the likelihood of casting a naked ballot, in addition to being statistically significant.

Table 2's second column shows there are also neighborhood-level variables associated with the likelihood of casting a naked ballot. The share of households in a voter's census tract with a limited ability to speak English positively associates with their likelihood of casting a naked ballot, while the share of households in a voter's census tract with at least a bachelor's degree negatively associates with the likelihood of casting a naked ballot. The relationship between measures of census tract income and the likelihood a voter casts a naked ballot is less clear.

Column three of Table 2 shows that patterns are similar when voter-level and neighborhood-level variables are included in the same regression, although with a few exceptions. While there is still a positive association between the probabilities that voters are racial or ethnic minorities and the likelihood of casting a naked ballot, the magnitude of the relationship decreases when the contextual variables are also included in the regression. This suggests that language and comprehension barriers, potentially stemming from less education and/or a greater likelihood of being a non-native English speaker, are one of the mechanisms that leads racial or ethnic minority voters to be more likely to cast naked ballots than White voters.

The fourth column of Table 2 suggests that more experience using mail ballots is at least part of the reason why vote history negatively associates with the likelihood of casting a naked ballot. The regression reported in column 4 replicates the regression in column 3 except that it also includes an indicator for whether the registrant has previously cast a mail

and ethnicity using census-tract shares.

Table 2: Which Mail Ballot Voters Cast a Naked Ballot in the 2020 General Election in Philadelphia

Dependent Variable			Cast Naked Ballot		
	(1)	(2)	(3)	(4)	(5)
N	360,064	365,555	360,040	360,040	360,040
Registrant-level variables:					
Voted in 2017 or 2019 General Election	-0.009***		-0.009***	-0.005***	-0.004***
	(0.001)		(0.001)	(0.001)	(0.001)
Only Voted in 2016 or 2018 General Election	-0.002***		-0.003***	-0.002**	-0.002***
	(0.001)		(0.001)	(0.001)	(0.001)
In-Person Mail Ballot	-0.004***		-0.004***	-0.006***	-0.006***
	(0.001)		(0.001)	(0.001)	(0.001)
Probability Black	0.009***		0.004***	0.003***	0.003***
	(0.001)		(0.001)	(0.001)	(0.001)
Probability Hispanic	0.019***		0.012***	0.011***	0.011***
	(0.001)		(0.001)	(0.001)	(0.001)
Probability Asian	0.013***		0.009***	0.009***	0.008***
	(0.002)		(0.002)	(0.001)	(0.001)
Probability Other Race	0.000		-0.002	-0.002	-0.002
	(0.003)		(0.003)	(0.003)	(0.003)
Born in 1920s	0.009***		0.010***	0.010***	0.010***
	(0.002)		(0.002)	(0.002)	(0.002)
Born in 1930s	0.010***		0.011***	0.011***	0.011***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1940s	0.009***		0.010***	0.010***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1950s	0.009***		0.010***	0.010***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1960s	0.010***		0.011***	0.010***	0.011***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1970s	0.009***		0.011***	0.010***	0.011***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1980s	0.005***		0.007***	0.007***	0.007***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1990s	0.003***		0.005***	0.005***	0.004***
	(0.001)		(0.001)	(0.001)	(0.001)
Missing or Likely Erroneous Decade of Birth	0.008		0.009	0.009	0.009
	(0.011)		(0.012)	(0.011)	(0.011)
Republican Registrant	0.005***		0.004***	0.003***	0.002**
	(0.001)		(0.001)	(0.001)	(0.001)
Non-Major Party Registrant	0.002***		0.002***	0.001	0.000
	(0.001)		(0.001)	(0.001)	(0.001)
Had Cast Mail Ballot Before 2020 General Election				-0.009***	-0.007***
				(0.000)	(0.000)
Voted in 2021 Primary Election					-0.005***
					(0.000)
Census-tract-level variables:					
Share Limited Language		0.031***	0.013***	0.011**	0.011**
		(0.005)	(0.005)	(0.004)	(0.004)
Share BA Degree		-0.014***	-0.010***	-0.008***	-0.007***
		(0.002)	(0.002)	(0.002)	(0.002)
Share Below Poverty Line		0.001	0.000	0.000	0.000
		(0.006)	(0.005)	(0.005)	(0.005)
ln(Median Income)		-0.003*	-0.003*	-0.002*	-0.002*
		(0.002)	(0.001)	(0.001)	(0.001)
Constant	0.002**	0.042**	0.035**	0.035**	0.036**
	(0.001)	(0.017)	(0.015)	(0.015)	(0.014)
R-squared	0.004	0.003	0.005	0.006	0.007

Note: Excluded group is a registered Democrat born in 2000s, who did not vote in general election between 2016-2019, who didn't cast an in-person mail ballot, and who has not previously vote by mail. Robust standard errors clustered by census tract in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

ballot in any statewide election held between the 2012 primary and 2020 primary elections. People who previously cast a mail ballot are 0.8 p.p. less likely to have cast a naked ballot than people who had not.

We suspect that some of the variables that relate to the likelihood of casting a naked ballot in Table 2 are generally predictive of a voter's likelihood of casting a naked ballot, while others were a result of the specific 2020 general election context. In Section 3, we suggested that both Democrats and voters who use social media might be more likely to be exposed to campaigns about the importance of secrecy envelopes than Republicans and voters who do not use social media. Given the well-known link between age and social media usage, we contend that this is at least one reason why Democrats and younger voters were less likely to cast naked ballots than Republicans and older voters in the 2020 general election. To further test this argument, we explore whether these variables remain predictive of the likelihood of casting a naked ballot in the 2021 primary election when this communication was no longer present.

Table 3 shows that both party and age were less predictive of casting a naked ballot in the 2021 primary election than in the 2020 general election. Table 3 focuses on the approximately 60,000 mail-ballot voters in Philadelphia's 2021 primary election, about 0.8% of whom had a record of casting a naked ballot. This share of returned mail ballots that were outside of a secrecy envelope was down slightly from 1.1% in the 2020 general. Column 1 of Table 3 shows that Republican registrants were about 0.27 p.p. more likely than Democratic registrants to cast a naked ballot. As a point of comparison, the same column in Table 2 demonstrates that Republican registrants were about 0.5 p.p. more likely than Democratic registrants to cast a naked ballot in the 2020 general. So either in terms of percentage point or percentage change, party of registration was less predictive of naked ballots in the 2021 primary than in the 2020 general. A similar pattern is found with respect to decade of birth. While the coefficients on decade of birth in Table 3 generally show that older voters were slightly more likely than younger voters to cast naked ballots in the 2021 primary election,

the relationship between age and likelihood of casting a naked ballot is not nearly as large or as consistent as Table 2 demonstrated for the 2020 general election.

Differences in who casts mail ballots in a local primary election versus a presidential election also likely cause some of the differences in Tables 2 and 3. Whereas the share of households in a voter's census tract with a limited ability to speak English and with at least a bachelor's degree were most predictive during the 2020 general election, the median income of the census tract was the only consistent predictor in the 2021 primary. The fact that the share of households in a voter's census tract with a limited ability to speak English only predicted naked ballots in the 2020 general could be evidence of campaign effects, as nearly all of the campaign activity that we uncovered in the lead-up to the 2020 general election to inform people about naked ballots was done in English. However, we suspect that compositional change in the electorate mostly explains the noticeably different patterns in how the contextual variables relate to the likelihood of casting naked ballots in the 2020 general and 2021 primary elections. People who vote in a local primary election are likely to be more politically knowledgeable and experienced with voting, on average, than people who vote in a presidential election. As a consequence, we expect that local primary voters will be less likely to cast a naked ballot than presidential election voters. Consistent with this expectation, Column 4 of Table 2 shows that people who went on to vote in the 2021 primary election were almost 0.5 p.p. less likely to cast a naked ballot in the 2020 general election than people who didn't vote in the 2021 primary, even after conditioning on all of the other controls included in the regressions reported in Tables 2 and 3. Given that voting in a future election cannot cause people to act differently in the present, we interpret this as evidence that the types of people who vote in local primary elections are generally less likely to cast a naked ballot. Thus, while the share of returned mail ballots that were naked decreased from about 1.1% in the 2020 general election to 0.8% in the 2021 primary, this decrease is less than might have been expected given the previous behavior of 2021 primary voters. This reinforces the conclusion that the salience of naked ballots during the 2020

Table 3: Which Mail Ballot Voters Cast a Naked Ballot in the 2021 Primary Election in Philadelphia

Dependent Variable			Cast Naked Ballot		
	(1)	(2)	(3)	(4)	(5)
N	59,986	60,698	59,980	59,980	59,980
Registrant-level variables:					
Voted in 2017 or 2019 General Election	-0.007***		-0.007***	-0.003	-0.002
	(0.002)		(0.002)	(0.002)	(0.002)
Only Voted in 2016 or 2018 General Election	-0.003*		-0.003*	-0.001	-0.001
	(0.002)		(0.002)	(0.002)	(0.002)
Probability Black	0.008***		0.004***	0.004***	0.004***
	(0.001)		(0.001)	(0.001)	(0.001)
Probability Hispanic	0.012***		0.010***	0.009***	0.009***
	(0.003)		(0.003)	(0.003)	(0.003)
Probability Asian	0.006**		0.006*	0.005	0.005
	(0.003)		(0.003)	(0.003)	(0.003)
Probability Other Race	0.014		0.013	0.013	0.014
	(0.009)		(0.009)	(0.009)	(0.009)
Born in 1920s	0.002		0.002	0.003	0.002
	(0.006)		(0.006)	(0.006)	(0.006)
Born in 1930s	0.005		0.005	0.005	0.005
	(0.006)		(0.006)	(0.006)	(0.006)
Born in 1940s	0.002		0.002	0.003	0.002
	(0.005)		(0.005)	(0.005)	(0.005)
Born in 1950s	0.000		0.001	0.001	0.001
	(0.005)		(0.005)	(0.005)	(0.005)
Born in 1960s	0.003		0.004	0.004	0.004
	(0.005)		(0.005)	(0.005)	(0.005)
Born in 1970s	0.002		0.002	0.003	0.003
	(0.005)		(0.005)	(0.005)	(0.005)
Born in 1980s	-0.002		-0.001	-0.001	-0.001
	(0.005)		(0.005)	(0.005)	(0.005)
Born in 1990s	-0.002		-0.002	-0.002	-0.002
	(0.005)		(0.005)	(0.005)	(0.005)
Missing or Likely Erroneous Decade of Birth	-0.008		-0.008	-0.006	-0.006
	(0.005)		(0.005)	(0.005)	(0.005)
Republican Registrant	0.003		0.002	0.001	0.000
	(0.002)		(0.002)	(0.002)	(0.002)
Non-Major Party Registrant	-0.000		-0.000	-0.003*	-0.003*
	(0.002)		(0.002)	(0.002)	(0.002)
Had Cast Mail Ballot Before 2020 General Election				-0.011***	-0.019***
				(0.001)	(0.006)
Cast Mail Ballot in 2020 General Election					-0.011**
					(0.005)
Had Cast Mail Ballot Before 2020 General Election X Cast Mail Ballot in 2020 General Election					0.009
					(0.006)
Voted in 2020 General Election					0.004
					(0.007)
Cast Naked Ballot in 2020 General Election					0.064***
					(0.016)
Census-tract-level variables:					
Share Limited Language		-0.004	-0.007	-0.008	-0.008
		(0.007)	(0.007)	(0.007)	(0.007)
Share BA Degree		-0.002	0.000	0.002	0.002
		(0.002)	(0.002)	(0.002)	(0.002)
Share Below Poverty Line		0.000	0.000	-0.000	-0.000
		(0.006)	(0.006)	(0.006)	(0.006)
ln(Median Income)		-0.006***	-0.005**	-0.004**	-0.004**
		(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.008*	0.075***	0.061***	0.063***	0.068***
	(0.005)	(0.021)	(0.023)	(0.023)	(0.025)
R-squared	0.003	0.001	0.003	0.006	0.008

Note: Excluded group is a registered Democrat born in 2000s, who did not vote in general election between 2016-2019, and who has not previously vote by mail. Robust standard errors clustered by census tract in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

general election reduced their incidence.

Column 5 of Table 3 examines how a voter's experience in the 2020 general election associates with their likelihood of casting a naked ballot in the 2021 primary. It does this by adding the following as independent variables: an indicator for whether they voted by mail in the 2020 general election, the interaction between this indicator and the indicator for whether they voted by mail prior to the 2020 general, an indicator for whether they voted using any mode in the 2020 general, and an indicator for whether they cast a naked ballot in the 2020 general. Most notably, this regression shows that voters who cast a naked ballot in the 2020 general election were about 6 p.p. more likely to cast a naked ballot in the 2021 primary than voters who submitted a mail ballot in a secrecy envelope in the 2020 general. This speaks to the importance of providing information to people whose mail ballot is rejected to prevent them from having their mail ballot rejected for the same reason in future elections.

6.3 How Postcards Affected Recipients' Voting Behavior

We next analyze how being sent the instructional postcard displayed in Figure 1 affected the recipients' voter behavior. Our first analysis uses data from the Pennsylvania Secretary of State to consider how being sent the mailer affected the speed with which people returned mail ballots. Next, we combine data on canceled ballots from the Pennsylvania Secretary of State and naked ballots from the Philadelphia City Commissioners to examine whether being sent the postcard reduced the likelihood of a mail ballot being canceled. Finally, we use data from the voter file to examine how being sent the mailer affected the likelihood of a registrant voting by mail, voting in person, or voting overall.

Figure 2 shows that people who were sent the informational postcard returned their mail ballot more quickly than people who were not. Figure 2a shows the share of mail ballots that had been returned by a given date separately for the groups that were and were not

mailed a postcard.[9] It shows that just under 80 percent of the people in our experimental universe ultimately returned their mail ballot. Figure 2b shows the difference in the share of mail ballots returned by date in the group randomly mailed a postcard relative to the group that was not. Starting on October 24, we observe a higher percentage of the ballots being returned in the group that was mailed a postcard than in the group that was not. This is consistent with our evidence that the first day anyone received the postcard was October 23. We observe the peak *difference* between the share returned in the mailed and control groups on October 30, when those sent the mailer were about 0.89 p.p. ($p = 0.035$, one-tailed) more likely to have returned their mail ballot than those who were not. The magnitude of this difference declines until it ultimately settles at 0.59 p.p. ($p = 0.080$, one-tailed) in the final update.

In Section 4, we conjectured that our postcard may provide more novel information for people who have more limited previous voting experience. Consequentially, we speculate that receipt of our mailing might affect behavior more if the registrant has not previously voted by mail or voted in any way. To explore this hypothesis, Appendix Figures A.1 and A.2 show how being sent the mailer affected the speed of mail ballot return among people who had no previous record of voting by mail or voting in any form, respectively. While the experiment is not sufficiently powered to statistically test this hypothesis, these figures show larger differences in the speed of return between the treatment and control groups among people with less voting experience. This suggests that future work should more thoroughly investigate whether there are particular benefits of sending a follow-up instructional mailer to less experienced voters.

Having observed that people returned their mail ballots more quickly when they were sent a mailer, we next turn to looking at whether the postcard reduced clerical errors on ballots. For the regressions reported in columns one through three of Table 4, the dependent

---

[9]Anecdotally, there often was a lag between the date a mail ballot was returned and the date a mail ballot is noted as being returned in the mail-ballot application data. Thus, this graph shows a lower bound on the share ballots returned by the date.

Figure 2: Mailer Caused Recipients to Return Mail Ballots in Full Experimental Universe (N = 38,155)
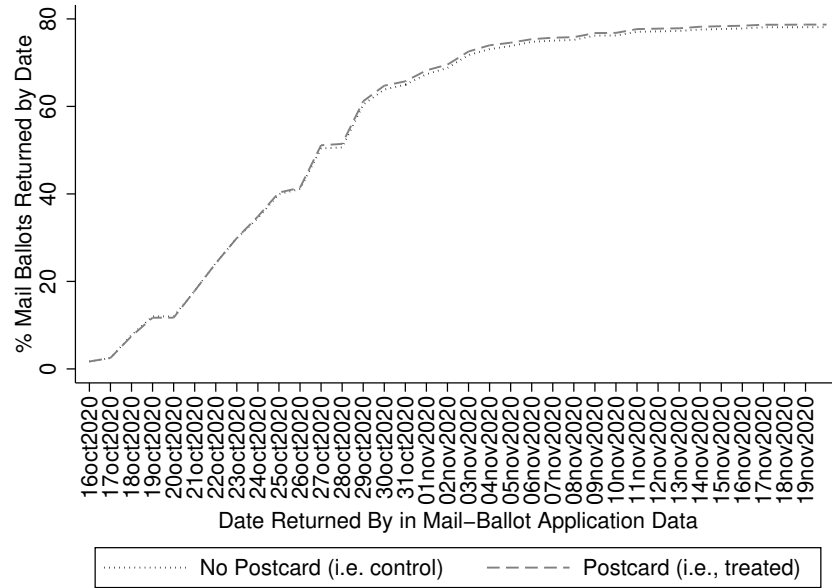
(a) Share of Mail Ballots Returned by Date
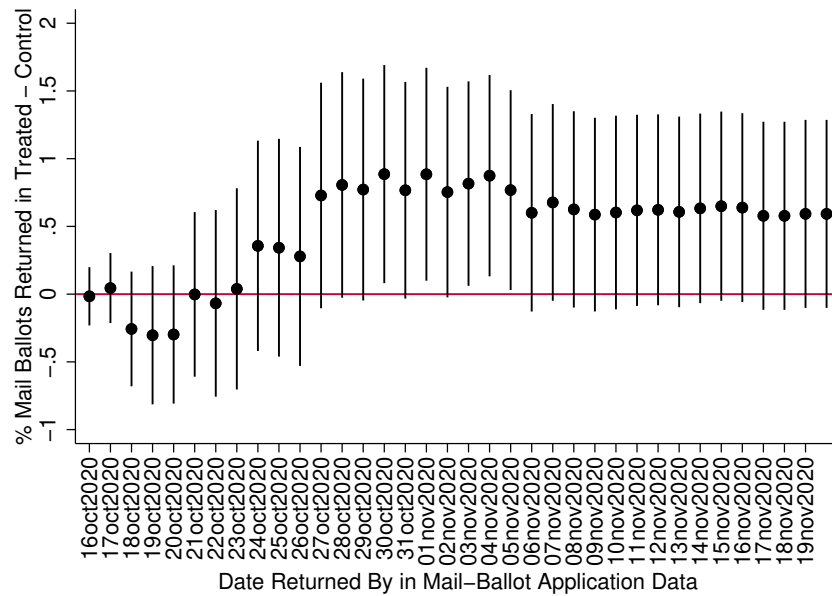
(b) Difference in Share of Mail Ballots Returned by Date in Mailed and Control

Note: Bottom figure represents a 90 percent symmetric confidence interval on the difference.

Table 4: Effect of Mailer on Naked Ballots and Mail Ballot Cancelation in Full Experimental Universe

Dependent Variable	Cast Naked Ballot			Canceled Mail Ballot		
	(1)	(2)	(3)	(4)	(5)	(6)
N	38,155	36,789	36,789	38,155	36,789	36,789
Sent Mailer	-0.0010	-0.0012	-0.0020	-0.0033**	-0.0037***	-0.0061*
	(0.0011)	(0.0011)	(0.0022)	(0.0014)	(0.0014)	(0.0035)
Likelihood of Casting a Naked Ballot		0.6129***	0.5882***		0.0862	0.0142
		(0.0790)	(0.1076)		(0.1071)	(0.1477)
Sent Mailer X Likelihood of Casting a Naked Ballot			0.0535			0.1560
			(0.1563)			(0.2131)
Constant	0.0112***	0.0021*	0.0024	0.0209***	0.0196***	0.0207***
	(0.0007)	(0.0012)	(0.0015)	(0.0010)	(0.0019)	(0.0025)
R-squared	0.0004	0.0023	0.0023	0.0014	0.0014	0.0014

Note: All regressions also include block fixed effects. Likelihood of casting a naked ballot is measured using fitted value from regression reported in Column 3 of Table 2. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

variable equals one if someone in our experimental universe returned a naked ballot and zero otherwise. Column one of Table 4 shows that about 1.1 percent of the control group in the experiment cast a naked ballot as compared to about 1.0 percent of the treatment group. This 0.1 p.p. reduction in the share of naked ballots in the treatment group relative to the control group is not statistically significant at conventional levels ($p = 0.171$, one-tailed).

We continue to find no significant effect of being sent a mailer on the likelihood of casting a naked ballot when we account for a voter's prior likelihood of casting a naked ballot based on their observable characteristics. In Column 2, we include the fitted likelihood that a voter casts a naked ballot from the regression reported in Column 3 of Table 2 as an explanatory variable. In essence, this is a one-dimensional score derived from individual- and tract-level variables which indicates registered voters whose demographics and place of residence make them especially likely to cast a naked ballot. This variable is strongly predictive of a voter's likelihood of casting a naked ballot, indicating that the same variables that predict a voter's likelihood of casting a naked ballot in the full sample of mail ballot voters also predict a voter's likelihood of casting a naked ballot within the experimental universe. But given that the mailer was randomly assigned to registrants within the experimental universe, it is unsurprising that including this as a control has minimal effects on the estimated impact of being sent a mailer on the likelihood of casting a naked ballot.

Information about how to cast a ballot has greatest potential to reduce errors among

voters who are most likely to commit them. Thus, the third column of Table 2 explores whether there is any evidence that the mailer helped to reduce naked ballots among the subset of voters in our experimental universe who our observational analysis suggests would be particularly likely to cast naked ballots. It does this by including an interaction term between being assigned the treatment and the fitted value from the regression reported in Column 3 of Table 2 as an explanatory variable. While the experiment is underpowered to reach any definitive conclusion, the fact that our estimated coefficient on this interaction term is positive is inconsistent with the mailer being more effective on voters who were particularly likely to cast a naked ballot.

While we cannot reject the null hypothesis that the mailer had no effect on the incidence of naked ballots, the fourth and fifth columns of Table 2 show that the mailer appears to have reduced the likelihood that a mail ballot was canceled for various reasons. The dependent variable in these regressions equals one if someone in our experimental universe had their mail ballot canceled for any reason besides it being undeliverable to their requested mailing address. Things that could trigger a mail ballot being canceled, in addition to the mail ballot being returned outside of a secrecy envelope, include the voter forgetting to sign the mail ballot declaration or the voter requesting a replacement mail ballot. Column 4 of Table 4 shows that about 2.1 percent of the control group had their mail ballot canceled, as compared to about 1.8 percent of the treatment group. The 0.33 p.p. reduction in the share of cancelled ballots in the treatment group relative to the control group is statistically significant at conventional levels ($p = 0.010$, one-tailed). The regression reported in Column 5 of Table 4 reaches a similar conclusion when the regression also controls for the fitted likelihood that a voter casts a naked ballot from the regression reported in Column 3 of Table 2 as an explanatory variable. Interestingly, this variable is not very related to the likelihood of having a cancelled mail ballot, suggesting that the voter characteristics that positively predict naked ballots negatively predict other reasons for mail ballot cancelation and vice versa. The sixth column shows no meaningful interaction between receiving the mailer and

Table 5: Effect of Mailer on Vote Mode and Turnout in Full Experimental Universe (N = 38,155)

Dependent Variable	Voted	Voted By Mail	Voted In-Person	Voted	Voted By Mail	Voted In-Person
	(1)	(2)	(3)	(4)	(5)	(6)
Sent Mailer	0.0033	0.0059	-0.0026	0.0036	0.0066	-0.0029
	(0.0034)	(0.0042)	(0.0029)	(0.0034)	(0.0042)	(0.0029)
Voted in 2017 or 2019 General Election				0.0873***	0.0645***	0.0228***
				(0.0045)	(0.0056)	(0.0040)
Only voted in 2016 or 2018 General Election				0.0472***	0.0210***	0.0263***
				(0.0043)	(0.0052)	(0.0035)
Had Cast Mail Ballot Before 2020 General Election				0.0595***	0.1326***	-0.0730***
				(0.0039)	(0.0049)	(0.0033)
Constant	0.8694***	0.7795***	0.0899***	0.8195***	0.7351***	0.0844***
	(0.0024)	(0.0029)	(0.0020)	(0.0037)	(0.0043)	(0.0027)
R-squared	0.0026	0.0021	0.0009	0.0203	0.0225	0.0092

Note: All regressions also include block fixed effect. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

this fitted likelihood that a voter casts a naked ballot, although again the experiment is underpowered to reach any definitive conclusion.

Table 5 examines how postcard receipt affects the likelihood that someone in the experimental universe voted according to the statewide voter file. Column 1 shows that people sent a postcard were about 0.33 p.p. ($p = 0.170$, one-tailed) more likely to have been recorded as voting in the 2020 general election than people not sent a postcard. Thus, we cannot reject the null hypothesis of no difference in turnout between people who were and were not sent a postcard. Column 2 shows that people sent a postcard were 0.59 p.p. ($p = 0.083$, one-tailed) more likely to vote by mail than people who were not. But this did not translate into higher turnout because Column 3 shows that people sent a postcard were 0.26 p.p. ($p = 0.188$, one-tailed) less likely to vote in person than people who were not sent a postcard. This suggests that receiving a postcard may have caused some people to return mail ballots who would have otherwise voted in person.

Because Table 1 shows that past voting experience relates to both the likelihood of voting and vote mode, the fourth through sixth columns in Table 5 replicate the regression reported in the first three columns with additional controls for past turnout and past mail ballot

28

usage.[10]   These regressions show substantively similar findings to the regressions without these controls.

A substantial number of people in the experimental universe could not have been affected by the postcard because they returned their mail ballot before anyone began receiving our mailing on October 23. Consequentially, we replicate the analysis we reported in Tables 4 and 5 on the subset of respondents in our experimental universe who did not return a mail ballot by October 22. The findings, which are reported in Appendix Tables A.3 and A.4, respectively, are substantively similar to the findings reported in the main text.[11]

7   Conclusion

The unique circumstances surrounding the 2020 elections raised the salience of lost votes by mail. Using data from Philadelphia County, Pennsylvania in the 2020 general election, we combined observational and experimental analyses to advance our understanding of this issue in numerous ways. First, we examined how often people who requested a mail ballot ultimately substituted to casting an in-person ballot, showing that this accounted for almost 8 percent of the people who requested a mail ballot in the 2020 general election. Moreover, nearly 10 percent of the people in our experimental universe ultimately cast an in-person ballot, despite nearly everyone in the universe having requested a mail ballot within one month of Election Day. An implication is that an unreturned mail ballot should not necessarily be interpreted as a lost vote, as substitution to in-person balloting occurs with some frequency. This was likely a more common phenomenon than usual in 2020 because of Donald Trump's rhetoric against mail balloting, which probably contributed to our finding that Republicans were more likely than Democrats or non-major party registrants to vote in person after requesting a mail ballot.

---

[10]Because blocks were constructed based on party of registration, controls for party of registration are already incorporated through the inclusion of block fixed effects.

[11]Figure A.3 shows that this conclusion is similar if any date between October 22 through October 26 is used to remove registrants from the experimental universe.

This paper also adds to the growing literature that establishes the concerning pattern that racial and ethnic minority voters are more likely to have lost mail votes than White voters (Baringer, Herron and Smith, 2020; Cottrell, Herron and Smith, 2021; Shino, Suttmann-Lea and Smith, 2022). We demonstrate that racial or ethnic minority voters are more likely to submit a naked ballot than White voters. Contextual variables suggest that one reason for this is that racial and ethnic minority voters are more likely to face language barriers than White voters. Spanish was the only non-English language in which Philadelphia County was required to provide election materials during 2020 based on Section 203 of the Voting Rights Act. While Chinese was subsequently also added, it is estimated that in Philadelphia tens of thousands of voters still cannot access election materials in their primary language (Nassef and To, 2021). Thus, lawmakers and election administrators must continue to work to structure the mail-balloting process to minimize the importance of English-language skills to successfully cast a mail ballot (see also Jones-Correa and Waismel-Manor, 2006; Hopkins, 2011)

Our findings also demonstrate that outreach by campaigns, interest groups, and election officials reduce the number of lost mail votes. In contrast to recent work, we find that younger voters were less likely to submit a naked ballot in the 2020 general election than older voters. This inversion of the expected pattern may result from the widespread communication efforts that arose, often on social media, to inform people about naked ballots. We advance this interpretation in part because we observe smaller differences in the incidence of naked ballots among younger and older voters in a local primary election that occurred months after the presidential election. Likewise, we observe that Democrats were less likely to cast naked ballots than Republicans only during the 2020 general election, which we again suspect resulted from Democrats being more exposed to communication efforts than Republicans. The experimental analysis reinforces the conclusion that outreach can reduce the number of lost mail votes by showing that people who were randomly assigned to be mailed a postcard with instructions about how to successfully cast a mail ballot returned their mail ballot more

quickly and were less likely to have their mail ballot canceled.

These results shed light on why more mail votes weren't lost in the 2020 general election. Prior to the election, many expressed concern that the substantial increase in demand for mail balloting resulting from the COVID-19 pandemic would generate more lost votes than normal both because voters wouldn't be able to access mail ballots and because mail ballot counting rules would generate many rejected ballots. Yet, people who lacked access to mail voting largely substituted to in-person voting (Yoder et al., 2021) and the share of returned mail ballots that were rejected declined between 2016 and 2020 (Rakich, 2021; Jose Altamirano AND Tova Wang, 2022). Our results suggest that the informational campaigns that emerged out of these concerns likely played a role in limiting the number of lost votes. An implication is that we should not assume that the issue of lost mail votes has been solved permanently. Voters will not necessarily be aware of naked ballots in future elections just because they learned about them in 2020. Future elections may differ dramatically with respect to lost votes, especially as rules about counting mail ballots become stricter and similarly extensive informational campaigns are not conducted. The impacts of election administration policies and procedures are not necessarily static—they hinge on how election administrators and other political actors respond. Thus, it is crucial that researchers, election administrators, and lawmakers continue to think about how to set up rules, design mail ballots, and structure communication efforts to minimize the chances of this occurring.

The experimental results suggest that sending registrants instructional postcards after they are sent a mail ballot may reduce burdens on election administrators and increase the number of votes that count. From an election administration standpoint, there is value in getting people to return mail ballots more quickly even if there is no ultimate effect on total ballots cast. Mail ballots that are not received in a timely way are at a greater risk of being rejected for being received too late. Relatedly, people may decide to vote in-person when they don't return their mail ballot quickly enough because they are concerned that their mail ballot will not be received in time to count. When they cast provisional ballots, this

uses up both poll worker and election administrator bandwidth. Finally, the more swiftly a mail ballot is returned, the more time there is for the ballot to be cured if a clerical error is discovered that would otherwise cause it to be rejected.

Moreover, the evidence is consistent with—although certainly not conclusive of—the claim that follow-up postcards can increase the number of ballots that ultimately count. In the experimental universe, about one in eight registrants failed to vote even though they had all expressed an interest in voting by requesting a mail ballot within a month of the election. So there is a set of people who request mail ballots that may require additional mobilization to ultimately cast that ballot. Moreover, registrants who are able to receive a mail ballot also might be more easily mobilized through mail than the average registrant. While we cannot reject the null hypothesis that receiving a postcard has no effect on turnout, the point estimate of a 0.33 p.p. increase in turnout is consistent with our postcard increasing the number of votes cast by roughly the same amount that Green and Gerber's (2019) meta analysis finds that non-partisan mailers increase turnout.

We hope that our experimental analysis leads to future studies of how sending mailers to people who just received their mail ballot affects the number of lost votes by mail and counted ballots. In that vein, we want to highlight several factors that future work might consider. First, are there specific profiles of registrants who are most likely to benefit from receiving a follow-up postcard? We hypothesize that people with less experience with mail balloting—and voting more broadly—might especially benefit from receiving a follow-up postcard. While our study was underpowered to statistically test this hypothesis, the patterns we observe are consistent with this being the case. Second, are such mailings more effective in certain electoral contexts? The postcards studied here were a blip in the broader campaign environment given the importance of Pennsylvania in the 2020 presidential election, where many partisan and non-partisan actors were disseminating information similar to what was included in the postcard that we analyzed. Thus, we think it is possible that follow-up postcards could be more impactful if distributed in less salient elections (Arceneaux

and Nickerson, 2009), when people are less likely to encounter messaging about how to cast a mail ballot elsewhere. Finally, how much time should there be between the distribution of the mail ballots and the follow-up postcard? Logistical constraints meant that people likely received the postcard about a week after receiving their mail ballot. We suspect that postcards may have been more effective if there was less time between when the mail ballot and the postcard were received.

A final question is who should be responsible for sending mailers and engaging in the other forms of voter education necessary to reduce lost votes by mail. We think there are several reasons why it makes sense for this to be done by election officials. First, election officials are incentivized to educate all voters before every election, unlike campaigns who are primarily incentivized to educate voters who they think will support their preferred candidates in elections that they believe could be competitive. Second, most election officials have experience engaging in many forms of voter outreach, including outreach by direct mail and through their broader efforts to educate voters about the voting process (Mann and Bryant, 2020; Bryant et al., 2022; Merivaki and Suttmann-Lea, 2022). Third, elections officials possess informational advantages about the timing of mail ballot distribution that may make it easier for them to educate voters before they return a mail ballot. However, educational efforts to reduce lost votes by mail require money that is unlikely to exist in the current budgets of elections officials. So, if we do expect elections officials to inform voters about how to avoid lost mail votes, we also must provide them the resources necessary to do so.

References

Arceneaux, Kevin and David W Nickerson. 2009. "Who is mobilized to vote? A re-analysis of 11 field experiments." *American Journal of Political Science* 53(1):1–16.

Baringer, Anna, Michael C. Herron and Daniel A. Smith. 2020. "Voting by Mail and Ballot Rejection: Lessons from Florida for Elections in the Age of the Coronavirus." *Election Law Journal: Rules, Politics, and Policy* 19(3):289–320.

Bryant, Lisa A, Michael J Hanmer, Alauna C Safarpour and Jared McDonald. 2022. "The Power of the State: How Postcards from the State Increased Registration and Turnout in Pennsylvania." *Political Behavior* 44:535–549.

Caltech/MIT Voting Technology Project. 2001. "Voting: What Is/What Could Be.".
URL: *https://vote.caltech.edu/reports/1*

Commonwealth of Pennsylvania Governor's Office. 2020. "Executive Order 2020-02." `https://www.oa.pa.gov/Policies/eo/Documents/2020-02.pdf`.

Cottrell, David, Michael C. Herron and Daniel A. Smith. 2021. "Voting-by-mail Ballot Rejection and Experience with Mail-in Voting." *American Politics Research* 49(6):577–590.

Couloumbis, Angela and Jamie Martine. 2020. "Republicans Seek to Sideline Pa. Mail Ballots that Voters Were Allowed to Fix." *Philadelphia Inquirer* .
URL: *https://www.inquirer.com/politics/election/spl/pennsylvania-mail-ballots-legal-challenge-naked-ballots-fixed-cured-20201103.html*

Deeley, Lisa M. 2020. "Letter to Rep. Brian Cutler and Sen. Joseph B. Scarnati, September 21, 2020.".
URL: *https://files7.philadelphiavotes.com/announcements/921-2020_Deeley_Letter_on_Secrecy_Envelope.pdf*

Gerber, Alan S. and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–663.

Green, Donald P. and Alan S. Gerber. 2019. *Get Out the Vote: How to Increase Voter Turnout.* 4 ed. Brookings Institution Press.

Hatmaker, Julia. 2020. "Bradley Cooper and Phillies' Star Andrew McCutchen Want You to Vote." *pennlive.com* .
URL: *https://www.pennlive.com/news/2020/10/bradley-cooper-wants-you-to-vote.html*

Hopkins, Daniel J. 2011. "Translating Into Votes: The Electoral Impacts of Spanish-Language Ballots." *American Journal of Political Science* 55(4):814–830.

Hopkins, Daniel J., Marc Meredith, Anjali Chainani, Nathaniel Olin and Tiffany Tse. 2021. "Results from a 2020 Field Experiment Encouraging Voting by Mail." *Proceedings of the National Academy of Sciences* 118(4).

Hopkins, Daniel J., Marc Meredith and Kira Wang. 2021. How Many Naked Ballots Were Cast in Pennsylvania's 2020 General Election? Technical report MIT Election Data and Science Lab.
URL: *https://medium.com/mit-election-lab/how-many-naked-ballots-were-cast-in-pennsylvanias-2020-general-election-91a277a1c19d*

Hopkins, Daniel J, Marc Meredith, Michael Morse, Sarah Smith and Jesse Yoder. 2017. "Voting but for the law: Evidence from Virginia on photo identification requirements." *Journal of Empirical Legal Studies* 14(1):79–128.

Hopkins, Daniel J, Susanne Schwarz and Anjali Chainani. Forthcoming. "Officially Mobilizing: Repeated Reminders and Feedback from Local Officials Increase Turnout." *Journal of Politics* .

Imai, Kosuke and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records." *Political Analysis* 24(2):263–272.

Jones-Correa, Michael and Israel Waismel-Manor. 2006. "Verifying Implementation of Language Provisions in the Voting Rights Act." *Voting Rights Act Reauthorization of* pp. 161–182.

Jones, Layla A. 2020. "What to Do If You Got a Notice Your PA Mail Ballot Was Canceled." *Billy Penn* .
URL: *https://billypenn.com/2020/10/30/pennsylvania-mail-ballot-cancelled-request-replacement-deadline/*

Jose Altamirano AND Tova Wang. 2022. Ensuring All Votes Count: Reducing Rejected Ballots. Technical report Harvard Kennedy School Ash Center for Democratic Governance and Innovation Policy Brief.
URL: *https: // ash. harvard. edu/ publications/ ensuring-all-all-votes-count-reducing-r*

Kiner, Deb. 2020. "Celebrities Appear in the Nude in Ads Explaining the 'Naked Ballot' Issue in Pa. and Other States." *pennlive.com* .
URL: *https://www.pennlive.com/elections/2020/10/celebrities-appear-in-the-nude-in-ads-explaining-the-naked-ballot-issue-in-pa-and-other-states.html*

Kurtz, Judy. 2020. "Celebrities Go Nude for PSA on Naked Ballots Celebrities." *The Hill* .
URL: *https://thehill.com/blogs/in-the-know/in-the-know/520036-celebrities-go-nude-for-psa-on-naked-ballots/*

Lai, Jonathan. 2020*a*. "Courts Extend Pa. Mail Ballot Deadlines in Bucks and Delaware Counties." *Philadelphia Inquirer* .
URL: *https://www.inquirer.com/politics/election/bucks-delaware-county-mail-ballot-deadlines-extended-20200602.html*

Lai, Jonathan. 2020*b*. "Tens of Thousands of Pennsylvania Mail Ballots Were Turned in After the Deadline. November Could Be Worse." *Philadelphia Inquirer* .
URL: *https://www.inquirer.com/politics/election/pa-mail-ballots-deadline-2020-primary-election-20200610.html*

Lai, Jonathan and Ellie Rushing. 2020. "USPS Says Pennsylvania Mail Ballots May Not Be Delivered On Time, and State Warns of 'Overwhelming' Risk to Voters." *Philadelphia Inquirer* .
URL: *https://www.inquirer.com/politics/election/pennsylvania-mail-voting-deadlines-post-office-lawsuit-20200813.html*

Mann, Christopher B. and Lisa A. Bryant. 2020. "If You Ask, They Will Come (to Register and Vote): Field Experiments with State Election Agencies on Encouraging Voter Registration." *Electoral Studies* 63:102021.

Merivaki, Thessalia and Daniel A. Smith. 2020. "A Failsafe for Voters? Cast and Rejected Provisional Ballots in North Carolina." *Political Research Quarterly* 73(1):65–78.

Merivaki, Thessalia and Mara Suttmann-Lea. 2022. "Can Electoral Management Bodies Expand the Pool of Registered Voters? Examining the Effects of Face-to-face, Remote, Traditional, and Social Media Outreach." *Policy Studies* pp. 1–31.

Meyer, Katie. 2020. "'Naked Ballot' Rule Could Lead To Thousands Of Pa. Votes Getting Rejected." *NPR* .
URL: *https://www.npr.org/2020/10/01/918368319/naked-ballot-rule-could-lead-to-thousands-of-pa-votes-getting-rejected*

Moore, Ryan T. 2012. "Multivariate Continuous Blocking to Improve Political Science Experiments." *Political Analysis* 20(4):460–479.

Murphy, Jan. 2020. "Glitches in Pa. Voter Services Website Not the Result of 'Any Malicious Activity,' State's Top Election Official Says." *pennlive.com* .
URL: *https://www.pennlive.com/elections/2020/09/glitches-in-pa-voter-services-website-not-the-result-of-any-malicious-activity-states-top-election-official-says.html*

Nassef, Al-Sharif and Andy To. 2021. "Philadelphia's Voters Need Materials in Their Own Languages." *Philadelphia Inquirer* .
URL: *https://www.inquirer.com/opinion/commentary/voting-materials-other-languages-suppresses-vote-20211213.html*

National Conference of State Legislatures. 2020. Voting Outside the Polling Place: Absentee, All-Mail and other Voting at Home Options. Technical report National Conference of State Legislatures.
URL: `https: // www. ncsl. org/ research/ elections-and-campaigns/ absentee-and-early-voting. aspx`

Otterbein, Holly. 2020. "Pennsylvania Election Officials Warn of 2000 Florida Redux." *Politico* .
URL: *https://www.politico.com/news/2020/09/22/pennsylvania-officials-warn-2000-florida-redux-420275*

Pennsylvania Department of State. 2020. Guidance Concerning Civilian Absentee and Mail-in Ballot Procedures. Technical report Pennsylvania Department of State.

URL: *https://www.dos.pa.gov/VotingElections/OtherServicesEvents/Documents/DOS%20Guidance%2*
In%20Ballot%20Procedures.pdf

Pennsylvania Department of State. 2021. Report on the 2020 General Election. Technical
report Pennsylvania Department of State.
URL: `https: // www. dos. pa. gov/ VotingElections/ Documents/`
`2020-General-Election-Report. pdf`

Rakich, Nathaniel. 2021. "Why So Few Absentee Ballots Were Rejected in 2020." *FiveThir-*
tyEight .
URL: *https://fivethirtyeight.com/features/why-so-few-absentee-ballots-were-rejected-in-*
2020/

Shino, Enrijeta, Mara Suttmann-Lea and Daniel A. Smith. 2022. "Determinants of Rejected
Mail Ballots in Georgia's 2018 General Election." *Political Research Quarterly* 75(1):231–
243.

Spencer, Douglas M. and Zachary S. Markovits. 2010. "Long Lines at Polling Stations?
Observations from an Election Day Field Study." *Election Law Journal: Rules, Politics,*
and Policy 9(1):3–17.

Stein, Robert M., Christopher Mann, Charles Stewart III, Zachary Birenbaum, Anson Fung,
Jed Greenberg, Farhan Kawsar, Gayle Alberda, R. Michael Alvarez, Lonna Atkeson, Emily
Beaulieu, Nathaniel A. Birkhead, Frederick J. Boehmke, Joshua Boston, Barry C. Burden,
Francisco Cantu, Rachael Cobb, David Darmofal, Thomas C. Ellington, Terri Susan Fine,
Charles J. Finocchiaro, Michael D. Gilbert, Victor Haynes, Brian Janssen, David Kimball,
Charles Kromkowski, Elena Llaudet, Kenneth R. Mayer, Matthew R. Miles, David Miller,
Lindsay Nielson, Yu Ouyang, Costas Panagopoulos, Andrew Reeves, Min Hee Seo, Haley
Simmons, Corwin Smidt, Farrah M. Stone, Rachel VanSickle-Ward, Jennifer Nicoll Victor,
Abby Wood and Julie Wronski. 2020. "Waiting to Vote in the 2016 Presidential Election:
Evidence from a Multi-county Study." *Political Research Quarterly* 73(2):439–453.

Stewart, Charles, III. 2010. "Losing Votes by Mail." *Journal of Legislation and Public Policy*
3(13):573–601.

Stewart, Charles, III. 2020. "Reconsidering Lost Votes by Mail." *Harvard Data Science*
Review 2(4).

U.S. Election Assistance Commission. 2017. The Election Administration and Voting Sur-
vey: 2016 Comprehensive Report. Technical report U.S. Election Assistance Commission.
URL: `https: // www. eac. gov/ sites/ default/ files/ eac_ assets/ 1/ 6/ 2016_`
`EAVS_ Comprehensive_ Report. pdf`

Wachter, Debbie. 2020. "Lawyer Withdraws Secrecy Envelope Objection." *New Castle*
News .
URL: *https://www.ncnewsonline.com/news/local_news/lawyer-withdraws-secrecy-*
envelope-objection/article_da779e3f-d05c-5a60-91af-bcb23179c88c.html

Walters, Joanna. 2020. "Pennsylvania Politicians Go Topless to Warn Voters: Don't Mail in 'Naked Ballots'." *The Guardian* .
URL: *https://www.theguardian.com/us-news/2020/sep/29/pennsylvania-politicians-naked-ballots-election-secrecy-envelope*

Yoder, Jesse, Cassandra Handan-Nader, Andrew Myers, Tobias Nowacki, Daniel M. Thompson, Jennifer A. Wu, Chenoa Yorgason and Andrew B. Hall. 2021. "How Did Absentee Voting Affect the 2020 U.S. Election?" *Science Advances* 7(52).

8   Appendix

8.1   Data

We use information contained in two versions of the Pennsylvania voter file for Philadelphia County to measure turnout by mode in the 2020 general election. The base voter file is from January 18, 2021 and was supplemented with information on 2020 general election turnout by mode from the March 1, 2021 voter file. This is done because Philadelphia County had not yet included 2020 general election turnout by mode as of January 18, 2021. The voter file includes turnout records for 368,683 registrants who are denoted as casting mail ballots and 367,801 registrants who are denoted as casting in-person ballots. This represents about 97 percent of 380,895 mail ballots and about 99 percent of the total ballots that the Pennsylvania Secretary of State's office reported were cast in Philadelphia County in this election (Pennsylvania Department of State, 2021). Based on our analysis, we conclude that some (but not all) of the people who submitted naked ballots in Philadelphia County in the 2020 general election were recorded as having voted in the Pennsylvania voter file. We also merged in turnout by mode for the 2021 primary election from the July 19, 2021 voter file for those registration numbers in the base voter file.

We merged into the voter file demographic information on the Census tract that registrants reside in from the 2019 ACS 5-year summary data from Philadelphia County. Because the voter file does not contain information on what Census tract a registrant resides in, we geocoded the residential addresses contained in the voter file through Geocodio.io. This returns the longitude and latitude of a registrant's address, the Census block that the residence is contained within, and measures of the accuracy geocode. We exclude from our analysis a small number of registrants with an accuracy score less than 0.9, with an accuracy type of "place", or who were coded into a Census tract that is not contained in the ACS data. The likely race and ethnicity of individual registrants is imputed by applying the approach of Imai and Khanna (2016). Imai and Khanna's method combines information on the demographics of the Census block that a registrant resides in with the demographics of people with the registrant's surname to estimate the probabilities that the registrant is White, Black, Hispanic, Asian American, and all other races and ethnicities.

The date on which a mail ballot was received by county elections officials was taken from data on accepted mail ballot applications reported by the Pennsylvania Secretary of State. On many days in September, October, and November, the Pennsylvania Secretary of State distributed a file listing all the mail ballot applications that had been approved in Pennsylvania as of that date. Among other things, these data contained information on which registrants had submitted an approved application for a mail ballot, the date that the request was approved, the address the registrant requested the ballot be mailed to, the date the ballot was sent to the registrant, and the date that the ballot was received back by the county elections officials.[12] We used data contained in the file distributed on October 16 to draw the universe for our postcard experiment and the data contained in the file distributed

---

[12]Based on our analysis of the data, we believe that some mail ballots were returned earlier than they are noted as being received in these data.

on November 20—the final file distributed—to measure the date that the ballot was received back by election officials.

Information on rejected mail ballots comes from two sources. The Pennsylvania Secretary of State reported information on approved mail ballot applications that had been canceled on many days in September, October, and November. We use data from a file distributed on November 20, which was the final file distributed, to identify 8,258 canceled mail ballots from Philadelphia County. Among other things, these data include information on the reason the registrant's mail ballot was canceled. The five reasons that were denoted more than 10 times were: 1) undeliverable (n = 3,370), 2) vote canceled (n = 2,880), 3) replaced (n = 1,168), 4) label canceled (n = 619), and 5) no signature (n = 211). "Undeliverable" indicated that the ballot was returned to election officials because it was not successfully delivered to the registrant who requested it. "Vote canceled" was how Philadelphia County denoted that a registrant had their mail ballot canceled because it was returned outside of a security envelope. "Replaced" and "label canceled" were both ways that Philadelphia County denoted that a registrant had their mail ballot canceled because they were issued a replacement mail ballot. Finally, no signature indicated that the mail ballot was canceled because the declaration on the mail ballot envelope was not signed.

We also received information from the Philadelphia City Commissioners' Office on the voter registration numbers for voters who cast naked ballots in Philadelphia County in the 2020 general and 2021 primary elections. 4,191 registration numbers were provided for the 2020 general election, which represented about 1.1% of mail ballots cast in this election. This rate of mail ballots is similar to the statewide average estimated by Hopkins, Meredith and Wang (2021). 491 registration numbers were provided for the 2021 primary election, which represented about 0.8 percent of the mail ballots cast in this election. We were able to match 4,189 of the 4,191 and 485 of the 491 listed voter registration numbers into our voter files for the 2020 general and 2021 primary elections, respectively. Almost all registrants denoted as having a mail ballot canceled for submitting a naked ballot in the 2020 general election in data provided by the Pennsylvania Secretary of State are also listed in the data provided by the Philadelphia City Commissioners' Office, as well as some registrants that were not listed in the Pennsylvania Secretary of State data. Additionally, the Philadelphia City Commissioners' Office provided us the voter registration numbers of the voters who cast an in-person mail ballot at one of the satellite election offices in the 2020 general election.

We construct two variables using the data contained in the two sources on rejected mail ballots. The first variable in an indicator for whether the voter cast a naked ballot. This variable is set equal to one if either the registration number was listed in the Secretary of State data as "voted canceled" or contained in the Philadelphia City Commissioners' Office data, and zero otherwise. The second variable is an indicator for whether the voter canceled a mail ballot that was successfully delivered to them. It is set equal to one if either the registration number was listed in the Secretary of State data as "voted canceled", "replaced", "label canceled" or "no signature" or contained in the Philadelphia City Commissioners' Office data, and zero otherwise.

Table A.1: Comparing Experimental Universe to Population of Approved Mail-Ballot Applicants

	All Approved Mail Ballot Applicants	Experimental Universe
N	434012	37924
Voted in 2017 or 2019 General Election	0.452	0.269
Only Voted in 2016 or 2018 General Election	0.321	0.381
Voted in No General Elections from 2016-2019	0.228	0.350
Had Cast Mail Ballot Before 2020 General Election	0.371	0.143
Probability White	0.370	0.295
Probability Black	0.414	0.443
Probability Hispanic	0.111	0.150
Probability Asian	0.073	0.082
Probability Other Race	0.031	0.029
Born in 1910s	< 0.001	< 0.001
Born in 1920s	0.009	0.011
Born in 1930s	0.044	0.035
Born in 1940s	0.107	0.074
Born in 1950s	0.169	0.134
Born in 1960s	0.150	0.150
Born in 1970s	0.124	0.138
Born in 1980s	0.187	0.192
Born in 1990s	0.184	0.229
Missing or Likely Erroneous Decade of Birth	< 0.001	< 0.001
Democratic Registrant	0.834	0.780
Republican Registrant	0.069	0.092
Non-Major Party Registrant	0.096	0.128

Note: Excludes 231 registrants from the experimental universe for whom we cannot observe this information.

8.2   Additional Tables and Figures

Table A.2: Replicating Table 2 With Census Tract Racial Share Instead of Imputed Race

Dependent Variables	Cast Naked Ballot in 2020 General Election				
	(1)	(2)	(3)	(4)	(5)
N	365,573	365,555	365,555	365,555	365,555
Registrant-level variables:					
Voted in 2017 or 2019 General Election	-0.009***		-0.009***	-0.005***	-0.004***
	(0.001)		(0.001)	(0.001)	(0.001)
Only Voted in 2016 or 2018 General Election	-0.002***		-0.003***	-0.002**	-0.002***
	(0.001)		(0.001)	(0.001)	(0.001)
In-Person Mail Ballot	-0.004***		-0.004***	-0.006***	-0.006***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1920s	0.009***		0.009***	0.010***	0.009***
	(0.002)		(0.002)	(0.002)	(0.002)
Born in 1930s	0.010***		0.010***	0.010***	0.011***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1940s	0.008***		0.009***	0.009***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1950s	0.009***		0.009***	0.009***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1960s	0.010***		0.010***	0.010***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1970s	0.009***		0.010***	0.010***	0.010***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1980s	0.005***		0.006***	0.006***	0.006***
	(0.001)		(0.001)	(0.001)	(0.001)
Born in 1990s	0.003***		0.004***	0.004***	0.004***
	(0.001)		(0.001)	(0.001)	(0.001)
Missing or Likely Erroneous Decade of Birth	0.007		0.008	0.009	0.008
	(0.011)		(0.011)	(0.011)	(0.011)
Republican Registrant	0.005***		0.004***	0.003***	0.002**
	(0.001)		(0.001)	(0.001)	(0.001)
Non-Major Party Registrant	0.003***		0.003***	0.001	0.000
	(0.001)		(0.001)	(0.001)	(0.001)
Had Cast Mail Ballot Before 2020 General Election				-0.009***	-0.008***
				(0.000)	(0.000)
Voted in 2021 Primary Election					-0.005***
					(0.000)
Census-tract-level variables:					
Share Black	0.010***	0.001	0.002**	0.002*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Share Asian	0.023***	0.012*	0.013*	0.012*	0.012*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Share Hispanic	0.034***	0.022***	0.021***	0.020***	0.020***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Share Other Race	-0.006	0.002	0.003	0.004	0.003
	(0.017)	(0.016)	(0.015)	(0.014)	(0.014)
Share Limited Language		0.004	0.001	-0.000	-0.001
		(0.007)	(0.007)	(0.006)	(0.006)
Share BA Degree		-0.012***	-0.009***	-0.008***	-0.007***
		(0.002)	(0.002)	(0.002)	(0.002)
Share Below Poverty Line		-0.008	-0.004	-0.004	-0.004
		(0.006)	(0.005)	(0.005)	(0.005)
ln(Median Income)		-0.004***	-0.004***	-0.004***	-0.004***
		(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.001	0.059***	0.052***	0.050***	0.050***
	(0.001)	(0.016)	(0.015)	(0.015)	(0.015)
R-squared	0.004	0.003	0.005	0.006	0.006

Note: Excluded group is a registered Democrat born in 2000s, who did not vote in general election between 2016-2019, who didn't cast an in-person mail ballot, and who has not previously vote by mail. Robust standard errors clustered by census tract in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Effect of Mailer on Naked Ballots and Mail Ballot Cancelation Among Most Plausibly Affected Registrants in Experimental Universe

| Dependent Variable	(1) Cast Naked Ballot	(2)	(3)	(4) Canceled Mail Ballot	(5)	(6)
N	28,934	27,831	27,831	28,934	27,831	27,831
Sent Mailer	-0.0009	-0.0011	-0.0029	-0.0031*	-0.0036**	-0.0057
	(0.0014)	(0.0014)	(0.0029)	(0.0017)	(0.0018)	(0.0044)
Likelihood of Casting a Naked Ballot		0.7505***	0.6972***		0.2128	0.1495
		(0.1017)	(0.1375)		(0.1335)	(0.1818)
Sent Mailer X			0.1153			0.1369
Likelihood of Casting a Naked Ballot			(0.2016)			(0.2665)
Constant	0.0139***	0.0025	0.0033*	0.0243***	0.0212***	0.0222***
	(0.0009)	(0.0016)	(0.0020)	(0.0012)	(0.0024)	(0.0030)
R-squared	0.0009	0.0033	0.0033	0.0031	0.0032	0.0032

Note: Excludes registrants from experimental universe that had their mail ballot canceled for being undeliverable or returned their mail ballot by October 22. All regressions also include block fixed effect. Likelihood of casting a naked ballot is measured using fitted value from regression reported in Column 3 of Table 2. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Effect of Mailer on Vote Mode and Turnout Among Most Plausibly Affected Registrants in Experimental Universe (N = 28,934)

Dependent Variable	Voted (1)	Voted By Mail (2)	Voted In-Person (3)	Voted (4)	Voted By Mail (5)	Voted In-Person (6)
Sent Mailer	0.0033	0.0068	-0.0035	0.0038	0.0075	-0.0037
	(0.0044)	(0.0053)	(0.0038)	(0.0044)	(0.0053)	(0.0038)
Voted in 2017 or 2019 General Election				0.0952***	0.0521***	0.0432***
				(0.0058)	(0.0072)	(0.0053)
Only Voted in 2016 or 2018 General Election				0.0541***	0.0188***	0.0353***
				(0.0053)	(0.0062)	(0.0044)
Had Cast Mail Ballot Before 2020 General Election				0.0671***	0.1552***	-0.0881***
				(0.0056)	(0.0070)	(0.0048)
Constant	0.8312***	0.7125***	0.1186***	0.7793***	0.6742***	0.1051***
	(0.0030)	(0.0036)	(0.0026)	(0.0045)	(0.0052)	(0.0034)
R-squared	0.0036	0.0061	0.0035	0.0196	0.0229	0.0122

Note: All regressions also include block fixed effect. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure A.1: Effect of Postcards on Ballot Return Speed Among Registrants in Experimental Universe Without Previous Mail-Ballot Usage
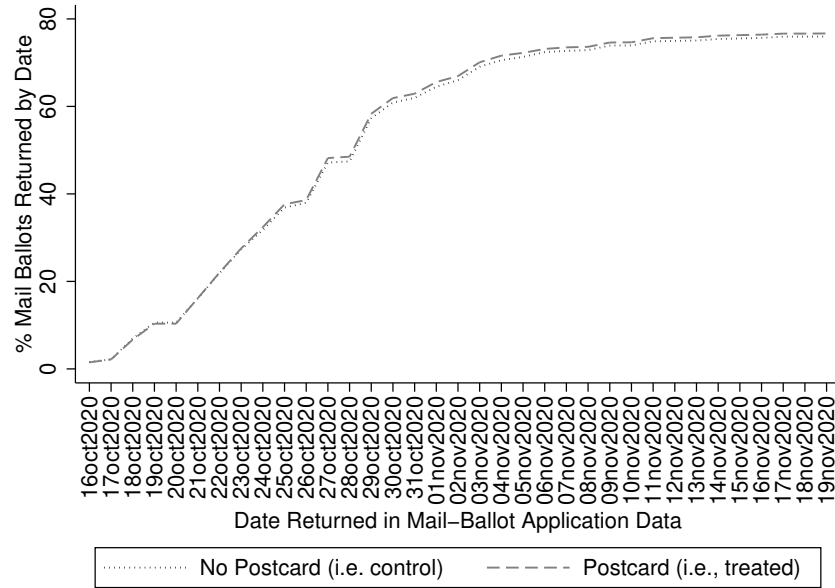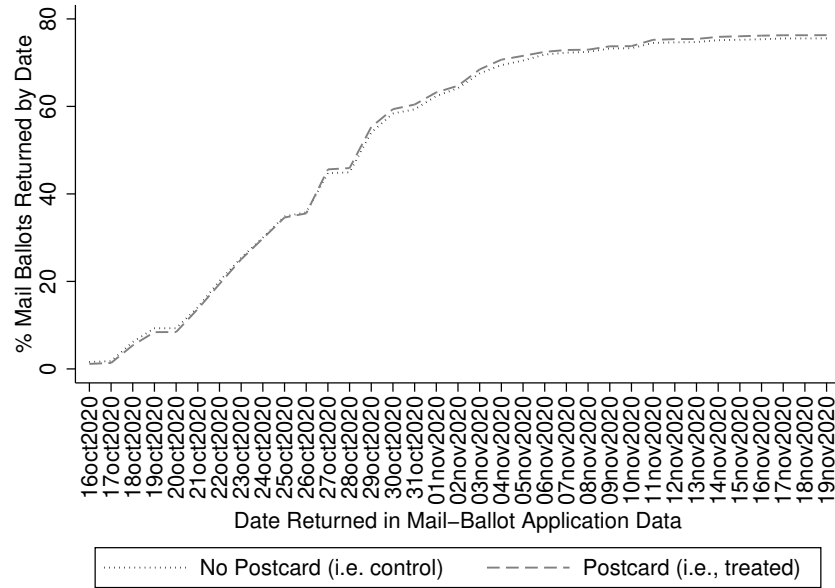

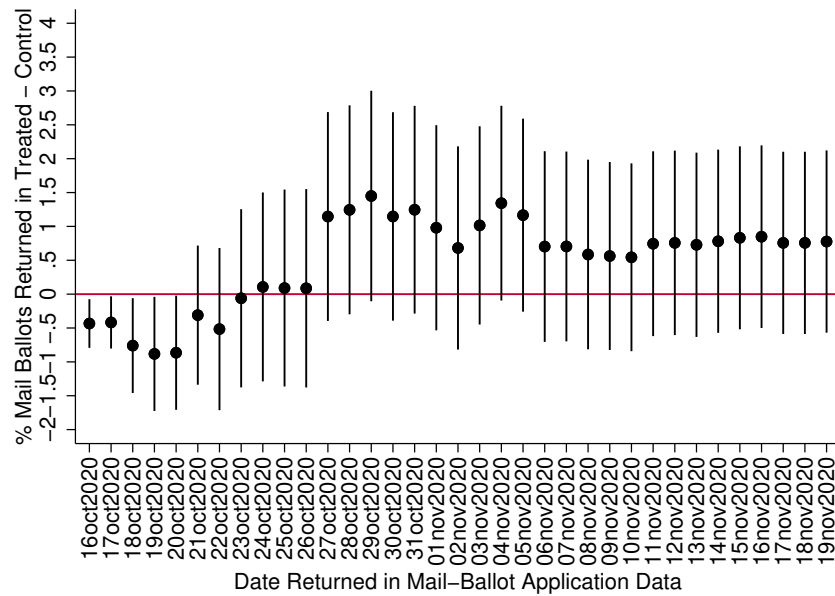(a) Share of Mail Ballots Returned by Date
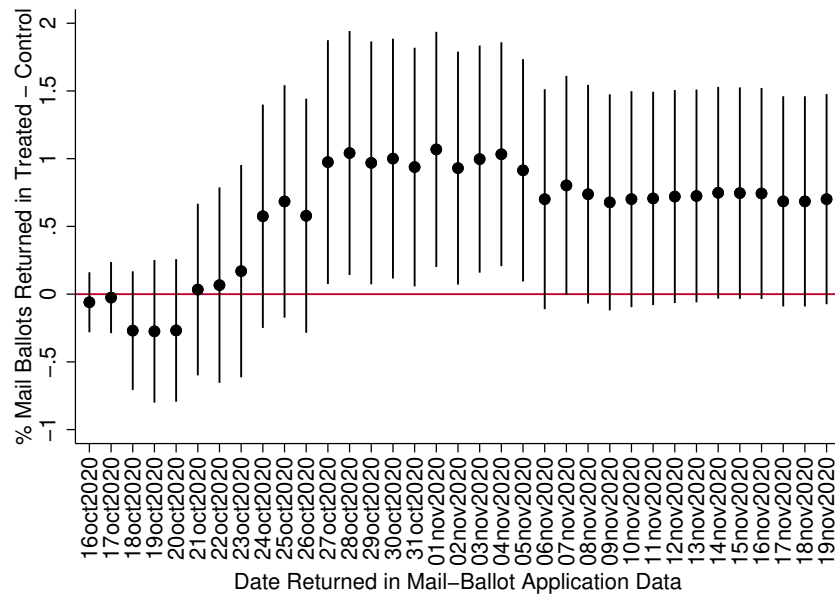
(b) Difference in Share of Mail Ballots Returned by Date in Mailed and Control

Note: Bottom figure represents a 90 percent symmetric confidence interval on the difference.

Figure A.2: Effect of Postcards on Ballot Return Speed Among Registrants in Experimental Universe Without Previous Vote History

(a) Share of Mail Ballots Returned by Date

(b) Difference in Share of Mail Ballots Returned by Date in Mailed and Control

Note: Bottom figure represents a 90 percent symmetric confidence interval on the difference.
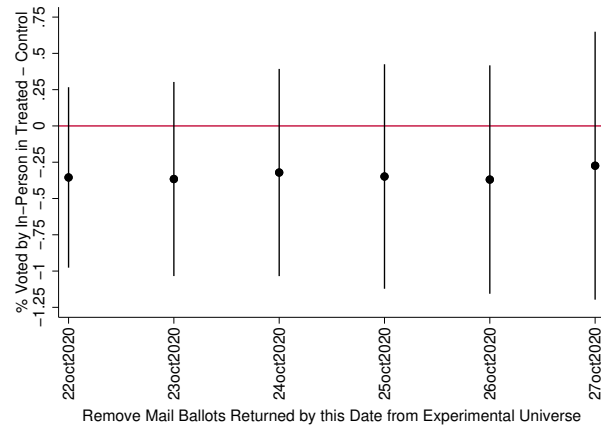
Figure A.3: Robustness of Estimates in Table A.4 to Alternate Sample Exclusions
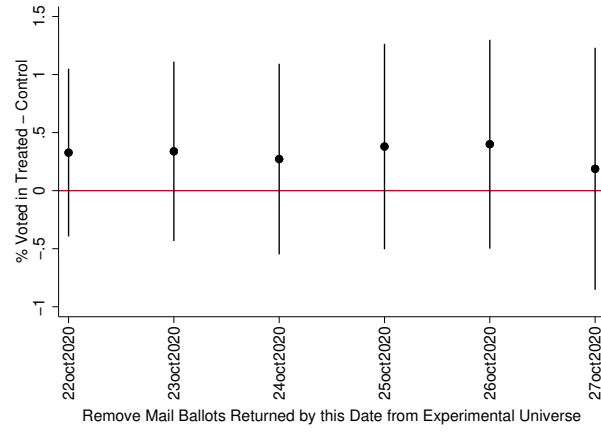
(a) Dependent variable is voted
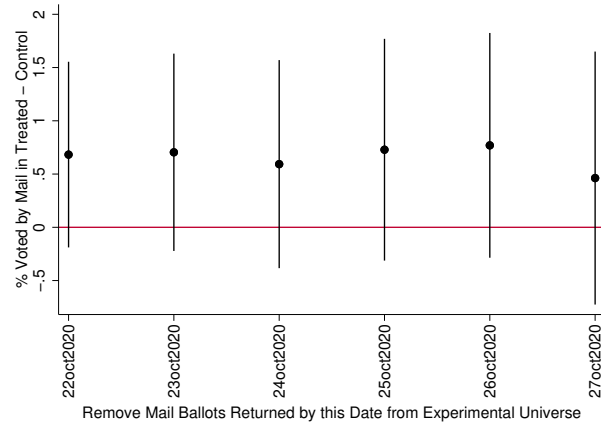
(b) Dependent variable is voted by mail

(c) Dependent variable is voted in-person

Note: Bars represent a 90 percent symmetric confidence interval on the difference.