Email Mobilization Messages Suppress Turnout Among Black and Latino Voters: Experimental Evidence From the 2016 General Election: Supplemental Information

Michael U. Rivera, D. Alex Hughes, Micah Gell-Redman

October 1, 2020

Contents

1 Introduction **2**

2 Data **2**
2.1 Data Acquisition . . . 2
2.2 Voting Modes . . . 2
2.3 Early Voting . . . 3

3 Experiment Design **3**
3.1 Analytic Sample . . . 3
3.1.1 Table 1: Analytic Sample . . . 3
3.2 Randomization . . . 3
3.3 Randomization Check . . . 4
3.3.1 Table 2: Randomization Check . . . 4
3.4 Email Distribution Details . . . 5
3.5 Power Analysis . . . 5
3.5.1 Power: All Voters, Any Message vs. Control . . . 5
3.5.2 Figure 1: Left Plot: Any Message vs. Control . . . 5
3.5.3 Figure 1: Right Plot: Message vs. Message . . . 5
3.5.4 Power: Racial/Ethnic Groups, Any Message vs. Control . . . 5
3.5.5 Figure 2: Any Message vs. Control, by Racial/Ethnic group . . . 7
3.5.6 Power: Racial/Ethnic Groups, Message vs. Message . . . 7
3.5.7 Figure 3: Messages vs. Message, by Racial/Ethnic Group . . . 7
3.6 Contact Language Details . . . 8

4 Results **8**
4.1 Main Effects . . . 8
4.1.1 Difference in Means: Main Effects of Treatment . . . 8
4.1.2 Estimation Details For Main Model . . . 10
4.1.3 Table 4: Main Model of Effects . . . 12
4.2 Message Effects . . . 12
4.2.1 Estimation Details for Message Effects . . . 12
4.2.2 Results for Message Effects . . . 12
4.2.3 Table 5: Message Effects . . . 13
4.2.4 Figure 6: Message Effects . . . 13
4.3 Subgroups . . . 13
4.3.1 Estimation Details for Subgroups . . . 13
4.3.2 Results for Subgroups . . . 13

4.3.3 Table 6: Message Effect, by Racial/Ethnic Subgroup . . . 15
4.3.4 Figure 7: Subgroup Effects . . . 15
4.4 Subgroup Treatment Heterogeneity . . . 15
4.4.1 Table 7: Heterogeneous Effects of Treatmetn by Racial/Ethnic Group . . . 17
4.5 Message by Subgroup Effects . . . 17
4.5.1 Turnout and Observations Message-By-Group . . . 18
4.5.2 Table 8: Voting Rate and Observations, By Group . . . 18
4.5.3 Table 9: Voting Rate and Observations Among White Voters . . . 18
4.5.4 Table 10: Voting Rate and Observations Among Latino Voters . . . 18
4.5.5 Table 11: Voting Rate and Observations Among Black Voters . . . 18
4.5.6 Results for Message by Subgroup Effects . . . 19
4.5.7 Table 12: Message Effects, By Racial/Ethnic Subgroup . . . 19
4.5.8 Figure 8: Message by Subgroup Effects . . . 19
4.6 Results Are Robust to Exclusion of Early and All Vote-By-Mail Voters . . . 19
4.6.1 Table 13: Treatment Estimates are Little Changed Excluding Vote-By-Mail Voters . . . 21
4.6.2 Table 14: Subgroup Treatment Effects are Little Changed Excluding Vote-By-Mail Voters . . . 21
4.7 In Text Result: How many people were kept from the polls? . . . 21

References **21**

```
library(data.table)
library(knitr)
library(stargazer)
library(lfe)
library(bookdown)
```

```
conduct_power <- FALSE
```

1 Introduction

This document supports the the main results reported in "Does email suppress turnout on election day?". The main, published, results have been peer reviewed. The results reported in this supplementary information provide additional context and clarification.

2 Data

2.1 Data Acquisition

We use two voter extracts provided directly by the Florida Division of Elections. Before election day – to conduct randomization and assignment of treatment messages – we use data that is provided in the October 10, 2016 voter roll. This is the final extract that the Division of Elections made available prior to election day. It contained at that time the most up-to-date, official information about who is registered to vote.

After election day – to calculate treatment effects – we use the June 14, 2017 voter extract provided by the Florida Division of Elections. Notably, this data is extracted more than seven months after the election. We chose to use an extract this far after the election because earlier voter extracts change as final tallies come in. Voter extract data is comprehensive, so using this later version ensures that all counties had fully reported their data to the state Division of Elections.

2.2 Voting Modes

Among the people who voted in the 2016 election in Florida, how did they choose to vote? This result is reported in *Section 2, Experimental setting, approach and data.*

```
count_of_voters <- d[!(historyCode2016 == ''), .N]

d[!(historyCode2016 == ''), .(
  voting_rate = round(.N / count_of_voters, 3)),
  keyby = .(historyCode2016)][ , .(
    historyCode2016 = c(
      'Absentee', 'Ballot Not Counted', 'Early',
      'Provisional Ballot', 'In-person'),
    'Voting Rate' = voting_rate)]
```

```
##       historyCode2016 Voting Rate
## 1:           Absentee       0.294
## 2: Ballot Not Counted       0.003
## 3:              Early       0.398
## 4: Provisional Ballot       0.000
## 5:          In-person       0.305
```

2.3 Early Voting

Information about whether a voter has cast an early ballot is not published by the state until after the election. As a result it was not possible to exclude these voters from treatment assignment. However, the window for early voting closes the day before we send our treatment, and so, we are able to treat early voting as a pre-treatment covariate, and thus drop these voters from the analysis without risk of bias.

```
d[
  treat %in% 0:4 ,
  .('Early Voting Rate' = mean(historyCode2016 == 'E')),
  keyby = .('Assigned to Treatment' = any_message)
  ]
```

```
##    Assigned to Treatment Early Voting Rate
## 1:                 FALSE         0.3485366
## 2:                  TRUE         0.3486974
```

3 Experiment Design

3.1 Analytic Sample

The analytic sample for this manuscript is more restrictive than being just the set of registered voters at the time of the 2016 election. Specifically, this analysis excludes voters who voted early-in-person. These individuals are excluded because their behavior could not possibly have been affected by receiving an email message.

Approximately 4.86 percent of voters provide either a valid email, or an email that could be rationalized by simple regular expression rules.

How similar are voters who provided an email when compared to those who do not provide an email? As we report in the following table, those who provide an email are just as likely to be democratic, female, and non-white as those who do not provide an email. Voters that provide an email are, on average 4.5 years younger, and also less likely to have voted in the 2012 election than those who do not provide an email.

3.1.1 Table 1: Analytic Sample

Table 1: Analytic Sample

	All Data	Valid Email	Assigned Treatment	Analytic Sample
Two_Party_Dem_VS	0.513	0.515	0.521	0.513
SE	0.00000	0.00000	0.00000	0.00000
Age_in_2016	51.235	45.726	45.502	45.259
SE.1	0.005	0.023	0.026	0.033
Voted_in_2012	0.712	0.526	0.517	0.536
SE.2	0.0001	0.0004	0.0005	0.001
Proportion_Female	0.528	0.523	0.522	0.521
SE.3	0.0001	0.001	0.001	0.001
Proportion_Non.White	0.288	0.314	0.322	0.304
SE.4	0.0001	0.001	0.001	0.001
Number_of_Observations	12,969,440	629,738	503,859	328,181

3.2 Randomization

We block random assign voters who provide a valid email within levels of the following blocking variables: congressional district and self-reported race.

We reproduce the randomization code here. We do not execute it, to maintain a single randomization. Randomization is conducted within blocks defined by Congressional District and racial identification categories.

```
## for individuals with a unique houseID; this is a simple statement
## make a draw from 0 (pure control) to 4 (treamtent 3).
##   - note that this is blocked on _individual_ characteristics
##     which we're able to do because we don't have any super-individual
##     problems with the data.
d[houseID <= 999999, treat := sample(0:4, size = .N, replace = TRUE),
  by = .(congressionalDistrict, raceCategory)
  ]
```

If there is more than one person living at a house, then we cluster randomize each person at that household into the same condition, using their Congressional District as a blocking factor.

```
## for individuals with a shared houseID we take a single treatment
## assignment draw and give it to everyone at the same house.
##   - note that this is blocked on congressionalDistrict, but that we
##     can't do any better that this beacuse we now have super-individual
##     information that is encoded in the clusters.
d[houseID > 999999, treat := rep(sample(0:4, size = 1), each = .N),
  by = .(congressionalDistrict)
  ]
```

3.3 Randomization Check

Distributions of covariates across treatment conditions do not raise any concerns that randomization did not proceed as planned. The following table reports a covariate balance table.

3.3.1 Table 2: Randomization Check

Table 2: Randomization Check

	Control	Any Email
Two_Party_Dem_VS	0.51	0.51
SE	0.0000	0.0000
Age_in_2016	45.26	45.26
SE.1	0.07	0.04
Voted_in_2012	0.54	0.54
SE.2	0.002	0.001
Proportion_Female	0.52	0.52
SE.3	0.002	0.001
Proportion_Non.White	0.30	0.31
SE.4	0.002	0.001
Number_of_Observations	66,132	262,049

Green and Gerber (2012) suggest examining whether assignment to treatment can be predicted by any on-hand covariates. Here too, there is no evidence to support a hypothesis that the randomization did not proceed as planned. We estimate two models:

$$any_message = \beta_0 + \epsilon \tag{1}$$

$$any_message = \beta_0 + \beta_1 voted_2012 + \beta_2 gender + \beta_3 race + \beta_4 party + \epsilon \tag{2}$$

The F-test p-value for a comparison of these model is 0.26. That is, the probability of observing these data under the null hypothesis that randomization was not problematic, is 0.26. This is insufficient to reject the null hypothesis.

3.4 Email Distribution Details

All treatment emails were distributed from a Qualtrics email management service, through University of Texas at Austin servers. The nature of our delivery does not permit us to observe whether emails were actually received by voters, nor whether they were opened. However, the randomization of treatment ensures that, in expectation, rates of opening are balanced across conditions. The technical staff at the University were able to confirm that the batch of emails was handled by their servers, and that the University email servers did not have their white-list status changed as a consequence of this mailing.

Although the limitations of the delivery infrastructure lead our estimates to be less precise than what may have been possible under a placebo-controlled design (Gerber and Green (2012)), they remain unbiased estimates of the change in voting that a campaign or other actor might expect to cause through a single, direct-email campaign.

3.5 Power Analysis

3.5.1 Power: All Voters, Any Message vs. Control

In this section, we report *ex post* power analysis information drawn from the analysis reported in the main text of the paper. In our preferred model specification, our experimental messaging caused a control vs. any treatment difference in turnout rates of 0.5 percentage points. We use this treatment effect estimate together with the number of registered voters in each condition to compute a simulated value for the achieved power in this test.[1] As we report in Figure 1, we find that, when comparing the 400,000 registered voters who received any form of message against the 100,000 registered voters who did not receive any contact this test is adequately powered.

3.5.2 Figure 1: Left Plot: Any Message vs. Control

3.5.3 Figure 1: Right Plot: Message vs. Message

In the next plot we report the power achieved when comparing any one condition against any other condition. More specifically, each comparison is of any treatment group (100,000 voters per message) against any other treatment or control group (also, 100,000 voters per message).

3.5.4 Power: Racial/Ethnic Groups, Any Message vs. Control

In this section we estimate the power that is achieved in this design for groups of voters that share a self-reported ethnic identity. Specifically, we report the power comparing the number of voters assigned to receive any message against similar voters assigned to receive control.

- Among white voters, there are approximately 40,000 voters that are assigned to control and 160,000 assigned to treatment.
- Among Latino voters, there are approximately 12,500 voters that are assigned to control and 50,000 assigned to treatment.
- Among black voters, there are approximately 7,300 voters that are assigned to control and 30,000 assigned to treatment.

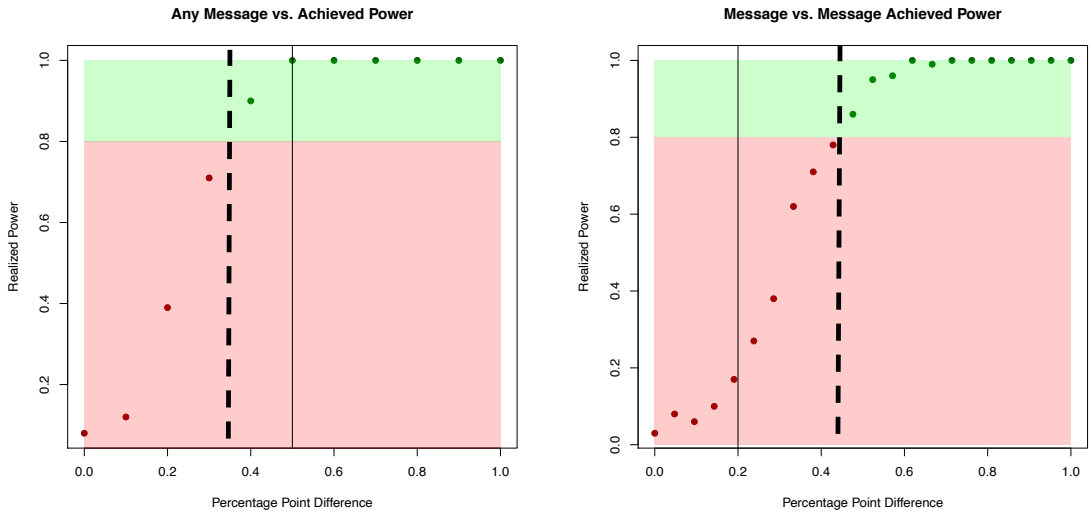


Figure 1: **Left Plot: Any Message vs. Control**. Plotted is the power curve for a scenario comparing the 400,000 voters in any treatment group (100,000 per message) against the 100,000 voters in the control group. The vertical line is drawn at the achieved treatment effect reported in the main body of the text. The dashed vertical line is drawn at the minimum detectable effect given this sample and design. The solid vertical line is drawn at the achieved treatment effect reported in the main body of the text. **Right Plot: Message vs. Message**. Plotted is the power curve for a scenario comparing 100,000 voters in any one message group against 100,000 from another message group. The vertical line is drawn at the largest difference in message effects observed in this analysis. Even the largest effect is under-powered.

3.5.5 Figure 2: Any Message vs. Control, by Racial/Ethnic group

3.5.6 Power: Racial/Ethnic Groups, Message vs. Message

In this section we estimate the power that is achieved in this design for groups of voters that share a self-reported ethnic identity. Distinct from the last section, here, we report the power achieved in a comparison between voters assigned to receive one message against another message. Whereas in the last section there were 4-times the number of people in the any message condition than control, in this section they are balanced.

- Among white voters, there are approximately 40,000 voters assigned to each condition
- Among Latino voters, there are approximately 12,500 voters that are assigned to each condition
- Among black voters, there are approximately 7,300 voters that are assigned to control and 7,300 assigned to treatment.

As is clear from this power analysis, there are very few people in each message-by-racial-group comparison cell. As a result, for any comparison that is this detailed to be adequately powered, the message-vs-message difference in voting would need to be **very** large. Plainly stated, this design is poorly powered to detect differences in response to messages, within racial/ethnic subgroups. Later in this SI, we report the estimates for these differences. We do so because we pre-registered these effects, but we would encourage readers to interpret any differences in behavior with a clear view for the limitations.

3.5.7 Figure 3: Messages vs. Message, by Racial/Ethnic Group

```
## This chunk produces the left most plot in Figure 3 titled
## "Messages vs. Message, by racial/ethnic group" in the SI

if(conduct_power == TRUE) {
  plot_power(
      power = power_by_group_white_by_message,
      effect_size = effect_size,
      treat_size = 0.000,
      main = 'White Voters: Message vs. Message Achieved Power',
      file = 'power_plot_by_group_white_by_condition.pdf'
  )
}
```

```
## This chunk produces the center plot in Figure 3 titled
## "Messages vs. Message, by racial/ethnic group" in the SI

if(conduct_power == TRUE) {
  plot_power(
      power = power_by_group_latino_by_message,
      effect_size = effect_size,
      treat_size = 0.000,
      main = 'Latino Voters: Message vs. Message Achieved Power',
      file = 'power_plot_by_group_latino_by_condition.pdf'
  )
}
```

```
## This chunk produces the right most plot in Figure 3 titled
## "Messages vs. Message, by racial/ethnic group" in the SI

if(conduct_power == TRUE) {
  plot_power(
```

[1]Code to support this is included in the repository for this project.

```
    power = power_by_group_black_by_message,
    effect_size = effect_size,
    treat_size = 0.000,
    main = 'Black Voters: Message vs. Message Achieved Power',
    file = 'power_plot_by_group_black_by_condition.pdf'
  )
}
```

3.6 Contact Language Details

The experimental treatment consisted of the text at the end of the first paragraph of the email (see Table 3). For clarity, 20% of the analytic sample did not receive any email; 20% of the analytic sample received the *Baseline* message, and 20% of the sample received each of the three descriptive social norm messages.

Figure 4 shows the general form of the email, while Figure 5 reproduces an example of one of the treatments as it would have appeared to subjects. All emails were unformatted text, contained no images or hyperlinks, and contained the same emailed header and signature information.

Table 3: Cue texts

Cue name	Text
Baseline	Democracy depends on citizens like you – so please vote!
General Descriptive Social Norm	We know that many people will be turning out to vote. Democracy depends on citizens like you – so please vote.
Ethnic Descriptive Social Norm 1	We know that Blacks and Latinos will be voting in large numbers. Democracy depends on citizens like you – so please vote.
Ethnic Descriptive Social Norm 2	We know that many people with bad opinions of Latinos will be voting in large numbers. Democracy depends on citizens like you – so please vote.

4 Results

4.1 Main Effects

4.1.1 Difference in Means: Main Effects of Treatment

Here, we first report the difference in means calculated between the groups that were assigned to receive any email message against those who were assigned to the control group. This comparison, otherwise unconditioned by additional covariates provides a straightforward, unbiased estimate of the causal effect of receiving an email message.

Throughout all estimated models, note that estimates for effects and standard errors have all been scaled by a factor of 100, and can be directly interpreted as percentage points.

```
effect_table <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E'),
  .(mean_vote = mean(outcome),
    sem_vote  = sem(outcome),
    variance = var(outcome),
```

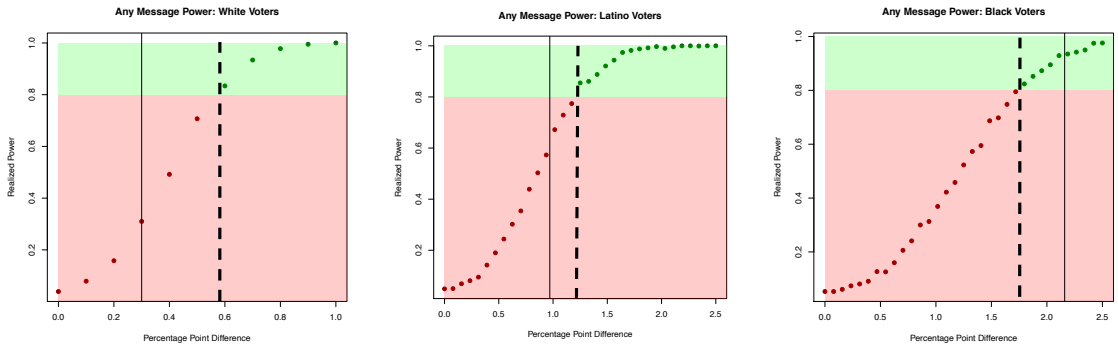


Figure 2: **Any message vs. Control, by racial/ethnic group**. Plotted is the power achieved comparing voters within a racial/ethnic group who received any message against voters from the same group who were assigned to the control group. The vertical line is drawn at the largest difference in message effects observed in this analysis. The dashed vertical line is drawn at the minimum detectable effect given this sample and design. The solid vertical line is drawn at the achieved treatment effect reported in the main body of the text.

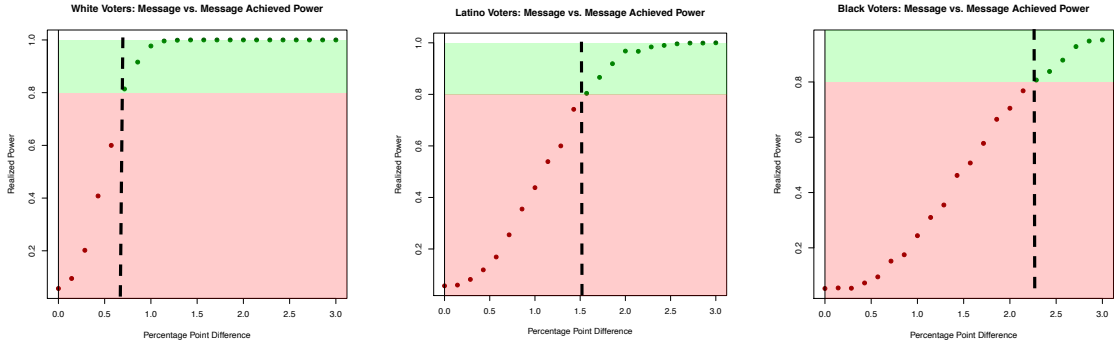


Figure 3: **Messages vs. Message, by racial/ethnic group**. Plotted is the power achieved comparing voters within a racial/ethnic group who were assigned to receive a particular message against voters in that same group, who were assigned to receive a different specific message. The dashed vertical line is drawn at the minimum detectable effect given this sample and design. The solid vertical line is drawn at the achieved treatment effect reported in the main body of the text.

Figure 4: Email Text

```
Dear Registered Voter,

You are currently a registered voter in the State of Florida. This is
a friendly reminder that tomorrow is Election Day. <Democracy depends
on citizens like you -- so please vote!>

If you have any questions about the voting process, please call the
Florida Department of State voter assistance hotline toll free at
1-866-308-6739.

Sincerely,

Michael Rivera
Department of Government
The University of Texas at Austin
```

Note: The text in <> is one of four randomly assigned messages.

```
    observations = .N),
  keyby = .(any_message)]
```

```
effect_table
```

```
##    any_message mean_vote    sem_vote  variance observations
## 1:       FALSE 0.7865179 0.001593428 0.1679100        66132
## 2:        TRUE 0.7811974 0.000807637 0.1709287       262049
```

This simple difference in means produces an estimate of the treatment effect of -0.532 and an estimate of the standard error of this difference of 0.17864.

4.1.2 Estimation Details For Main Model

We use regression with robust standard errors to estimate the causal effects of email contact Wooldridge (2010), Gerber and Green (2012). To measure the effects of any form of contact, we estimate the following regression:

$$Y_i = \alpha + \bar{\tau} Treat_i + \gamma X_i + \phi District_i + \epsilon_i, \tag{3}$$

where $Treat_i$ is an indicator that the subject received any form of treatment, and $\bar{\tau}$ is the average difference in the propensity to turnout for a registered voter assigned to any of the four treatment groups compared to those in the control group. *This quantity, $\bar{\tau}$ is the primary causal quantity of interest in this analysis.* X is a vector of controls to improve the efficiency of the estimates, *District* is a vector of congressional districts and ϵ captures idiosyncratic errors at the individual level.

```
model_1 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
               lm(outcome ~ any_message)]
model_2 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
             felm(outcome ~ any_message | congressionalDistrict)]
model_3 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
             felm(outcome ~ any_message + factor(major_party) + race3 +
                    age + I(age^2) | congressionalDistrict)]
```

Table 4: Main Model of Effects

	Dependent variable:		
	outcome		
	OLS	*felm*	
	(1)	(2)	(3)
Assigned Message	-0.005^{***}	-0.005^{***}	-0.005^{***}
	(0.002)	(0.002)	(0.002)
Republican			-0.0003
			(0.002)
Independent			-0.109^{***}
			(0.002)
Black			-0.126^{***}
			(0.002)
Latino			-0.050^{***}
			(0.002)
Other			-0.069^{***}
			(0.003)
Age			0.017^{***}
			(0.0002)
Age^2			-0.0001^{***}
			(0.00000)
Intercept	0.787^{***}		
	(0.002)		
District FE	No	Yes	Yes
Observations	328,181	328,181	327,952
R^2	0.00003	0.017	0.125
Adjusted R^2	0.00002	0.017	0.125
F Statistic	8.776^{***} (df = 1; 328179)		

Note: *p<0.1; **p<0.05; ***p<0.01
Democrats and Whites are baseline values.

Please Remember to Vote Tomorrow

Michael Rivera <noreply@utexas.edu> Mon, Nov 7, 2016 at 7:00 AM
Reply-To: Michael Rivera <m.rivera@utexas.edu>
To: [redacted]@gmail.com

Dear Registered Voter,

You are currently a registered voter in the State of Florida. This is a friendly reminder that tomorrow is Election Day. Democracy depends on citizens like you—so please vote!

If you have any questions about the voting process, please call the Florida Department of State voter assistance hotline toll free at 1-866-308-6739.

Sincerely,

Michael Rivera
Department of Government
The University of Texas at Austin

Figure 5: In other treatment conditions, the phrase *Democracy depends on citizens like you–so please vote!* is replaced with the text in Table 3.

4.1.3 Table 4: Main Model of Effects

4.2 Message Effects

4.2.1 Estimation Details for Message Effects

In our pre-analysis plan, we proposed to test for whether different messages would have distinct effects on voter turnout. We do not report estimates of these effects in the main body because such estimates are under-powered and we urge considerable care in interpreting these results. For completeness, we include these results here. These models are similar to those in the first estimated model, but in the following equation, we estimate a separate indicator for each message type sent.

$$\begin{aligned} Y_i = \alpha &+ \tau_1 Baseline_i \\ &+ \tau_2 DescriptiveSocialNorm_i \\ &+ \tau_3 EthnicDescriptiveSocialNorm1_i \\ &+ \tau_4 EthnicDescriptiveSocialNorm2_i \\ &+ \gamma X_i + \phi District_i + \epsilon_i \end{aligned}$$

where τ_1, τ_2, τ_3, and τ_4 are the average difference in the propensity to turnout between a registered voter in control condition (reference category) compared to the *Baseline*, *Descriptive Social Norm*, *Ethnic Descriptive Social Norm 1*, and *Ethnic Descriptive Social Norm 2*, respectively.

As reported in Table 5 and shown in Figure 6, there is little evidence to support a hypothesis that message content causes differential turnout in voters.

4.2.2 Results for Message Effects

```
model_4 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
             lm(outcome ~ any_message)]
```

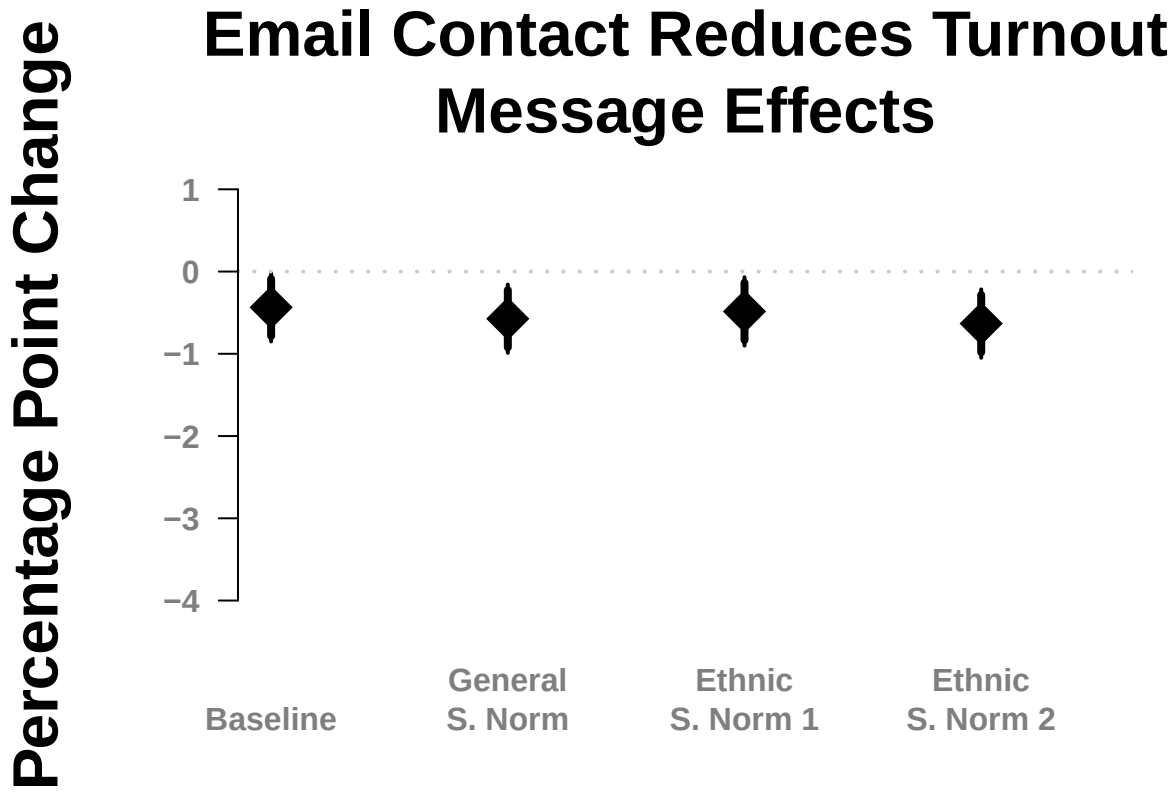
```
model_5 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
               lm(outcome ~ factor(treat))]
model_6 <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
             felm(outcome ~ factor(treat) | congressionalDistrict)]
model_6a <- d[treat %in% 0:4 & !(historyCode2016 == 'E'),
             felm(outcome ~ factor(treat) + factor(major_party) +
                    race3 + age + I(age^2) + registrationYear| congressionalDistrict)]

message_anova <- anova(model_4, model_5, test = 'F')
```

Indeed, an F-test for whether individual coefficients for each message improve the fit of the model above a single $\bar{\tau}$ returns a p-value of 0.78.

4.2.3 Table 5: Message Effects

4.2.4 Figure 6: Message Effects



4.3 Subgroups

4.3.1 Estimation Details for Subgroups

In Table 12 we report results from models that estimate the effect of each stimulus among voters who are able to vote on election day, broken out by whether that voter has chosen to identify as White, Black, or Latino in their voter registration. These results are also plotted in Figure 8.

4.3.2 Results for Subgroups

```
model_7 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "White",
```

Table 5: Message Effects

	Dependent variable:			
	outcome			
	OLS		*felm*	
	(1)	(2)	(3)	(4)
Any Message	-0.005^{***}			
	(0.002)			
Baseline		-0.005^{**}	-0.005^{**}	-0.004^{**}
		(0.002)	(0.002)	(0.002)
Descriptive Social Norm		-0.006^{***}	-0.006^{***}	-0.006^{***}
		(0.002)	(0.002)	(0.002)
Ethnic Descriptive Social Norm 1		-0.004^{*}	-0.004^{*}	-0.005^{**}
		(0.002)	(0.002)	(0.002)
Ethnic Descriptive Social Norm 2		-0.006^{***}	-0.006^{***}	-0.006^{***}
		(0.002)	(0.002)	(0.002)
Republican				-0.002
				(0.002)
Independent				-0.102^{***}
				(0.002)
Black				-0.125^{***}
				(0.002)
Latino				-0.044^{***}
				(0.002)
Other				-0.060^{***}
				(0.003)
Age				0.016^{***}
				(0.0002)
Age^2				-0.0001^{***}
				(0.00000)
Intercept				-0.003^{***}
				(0.0001)
Constant	0.787^{***}	0.787^{***}		
	(0.002)	(0.002)		
District FE	No	No	Yes	Yes
Observations	328,181	328,181	328,181	327,952
R^2	0.00003	0.00003	0.017	0.130
Adjusted R^2	0.00002	0.00002	0.017	0.130

Note: *p<0.1; **p<0.05; ***p<0.01

15 Democrats and Whites are baseline values.

```
  felm(outcome ~ any_message + factor(major_party)
         + age + I(age^2)  | congressionalDistrict)]
model_8 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "Latino",
  felm(outcome ~ any_message + factor(major_party)
       + age + I(age^2) | congressionalDistrict)]
model_9 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "Black",
  felm(outcome ~ any_message + factor(major_party)
       + age + I(age^2) | congressionalDistrict)]
```

4.3.3 Table 6: Message Effect, by Racial/Ethnic Subgroup

Table 6: Message Effect, by Racial/Ethnic Subgroup

	Dependent variable:		
	outcome		
	White RV	Latino RV	Black RV
	(1)	(2)	(3)
Any Message	−0.003	−0.010**	−0.022***
	(0.002)	(0.004)	(0.006)
Republican	0.005***	−0.005	−0.072***
	(0.002)	(0.004)	(0.011)
Independent	−0.105***	−0.102***	−0.122***
	(0.002)	(0.004)	(0.006)
Age	0.014***	0.021***	0.023***
	(0.0002)	(0.001)	(0.001)
Age^2	−0.0001***	−0.0002***	−0.0002***
	(0.00000)	(0.00001)	(0.00001)
Observations	203,917	63,369	36,482
R^2	0.090	0.102	0.125
Adjusted R^2	0.090	0.101	0.124
Residual Std. Error	0.352 (df = 203885)	0.421 (df = 63337)	0.453 (df = 36450)
Note:			*p<0.1; **p<0.05; ***p<0.01

4.3.4 Figure 7: Subgroup Effects

4.4 Subgroup Treatment Heterogeneity

Throughout the manuscript, we principally report the subgroup treatment effects. Here, for completeness, we report a model that estimates whether voters in minority racial/ethnic groups are affected by treatment differently than white voters.

```
model_hte <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E'),
  felm(outcome ~ any_message * race3 + factor(major_party)
```

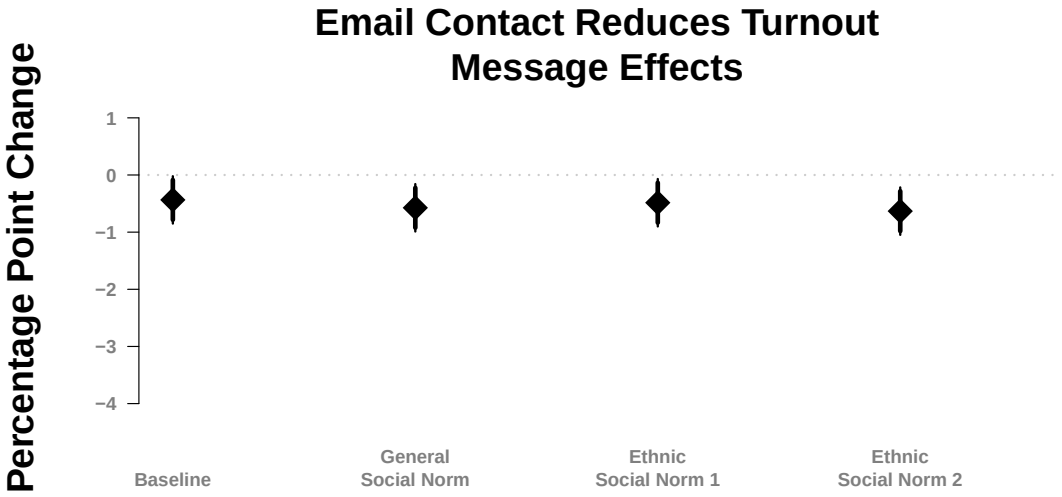


Figure 6: Message Effects

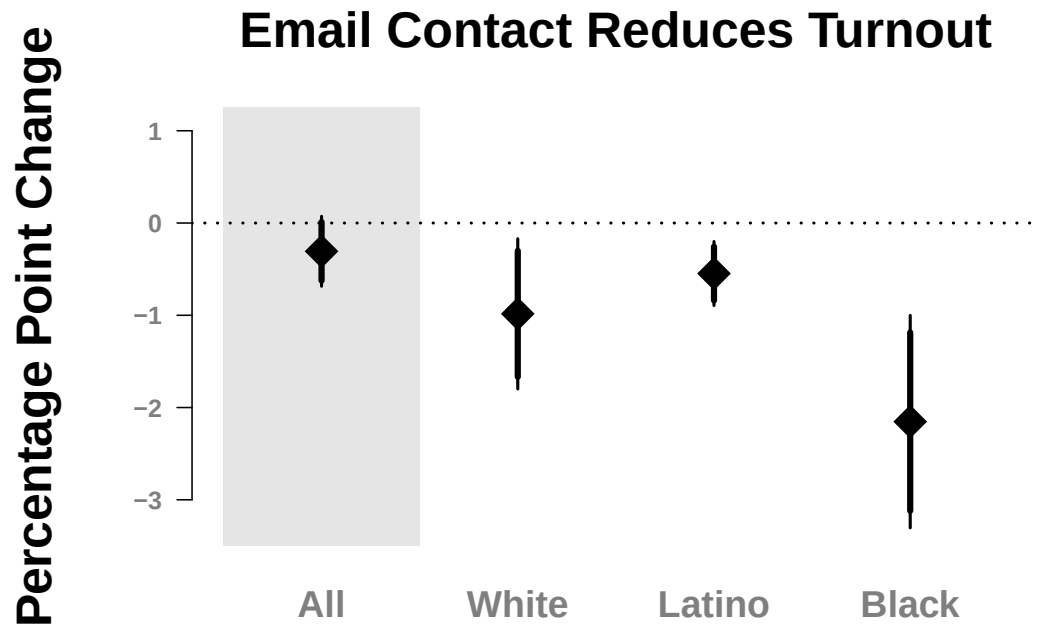


Figure 7: Subgroup Effects. All messages are combined and models are estimated on subgroups of voters. These points correspond to $\bar{\tau}_w, , \bar{\tau}_b$, and $\bar{\tau}_l$ reported in the main body.

```
        + age + I(age^2)  | congressionalDistrict)]
```

4.4.1 Table 7: Heterogeneous Effects of Treatmetn by Racial/Ethnic Group

Table 7: Heterogeneous Effects of Treatment by Racial/Ethnic Group

	Dependent variable:
	outcome
Any Message	−0.003
	(0.002)
Black Voter	−0.111***
	(0.005)
Latino Voter	−0.044***
	(0.004)
Other Voter	−0.080***
	(0.006)
Republican	−0.0003
	(0.002)
Independent	−0.109***
	(0.002)
Age	0.017***
	(0.0002)
Age^2	−0.0001***
	(0.00000)
Any Message * Black Voter	−0.018***
	(0.005)
Any Message * Latino Voter	−0.007*
	(0.004)
Any Message * Other Voter	0.014**
	(0.007)
Observations	327,952
R^2	0.125
Adjusted R^2	0.125
Residual Std. Error	0.386 (df = 327914)
Note:	*p<0.1; **p<0.05; ***p<0.01

4.5 Message by Subgroup Effects

There is rich scholarship that examines how white (Alvarez and Bedolla (2004), Hajnal and Rivera (2014), Jardina (2019)), Latino (García-Bedolla (2005), García-Bedolla and Michelson (2012)), and black (McGowen

(2010), Philpot, Shaw, and McGowen (2009)) identities shape that way voters engage in politics.

4.5.1 Turnout and Observations Message-By-Group

Despite the relatively large sample size available in this experiment, we are extremely limited in what we are able to conclude due to limitations of statistical power. Even with an analytic sample containing more than 260,000 individuals, to answer the question about how four messages might differently affect three identity groups leaves approximately only 20,000 voters in each comparison. Given experimental cells of this size, using standard criteria (e.g. $\alpha = 0.05$, $\beta = 0.8$) it would only be possible to reliably measure effects that are larger than 1.2 percentage points. This effect size is more than twice the well-established direct-mail effect Green and Gerber (2019).

4.5.2 Table 8: Voting Rate and Observations, By Group

Table 8: Voting Rate and Observations, By Group

	Race	Voters	Turnout
1	White	204,054	0.837
2	Black	36,518	0.627
3	Latino	63,404	0.730

4.5.3 Table 9: Voting Rate and Observations Among White Voters

Table 9: Voting Rate and Observations Among White Voters

	Message Label	White Voters	White Turnout
1	Baseline	40,806	0.837
2	Control	41,223	0.840
3	Ethnic Descriptive Norm 1	40,599	0.837
4	Ethnic Descriptive Norm 2	40,764	0.836
5	General Descriptive Norm	40,662	0.837

4.5.4 Table 10: Voting Rate and Observations Among Latino Voters

Table 10: Voting Rate and Observations Among Latino Voters

	Message Label	Latino Voters	Latino Turnout
1	Baseline	12,777	0.728
2	Control	12,637	0.739
3	Ethnic Descriptive Norm 1	12,648	0.733
4	Ethnic Descriptive Norm 2	12,669	0.722
5	General Descriptive Norm	12,673	0.728

4.5.5 Table 11: Voting Rate and Observations Among Black Voters

Nevertheless, for completeness, we report estimates of these models here in the appendix, but we urge due caution in interpreting these effects, as their under-powered nature makes quite possible that any measured *message-by-group* difference make be a false-positive result.

Table 11: Voting Rate and Observations Among Black Voters

	Message Label	Black Voters	Black Turnout
1	Baseline	7,470	0.622
2	Control	7,359	0.640
3	Ethnic Descriptive Norm 1	7,171	0.625
4	Ethnic Descriptive Norm 2	7,218	0.628
5	General Descriptive Norm	7,300	0.620

4.5.6 Results for Message by Subgroup Effects

```
model_10 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "White",
  felm(outcome ~ factor(treat) + factor(major_party)
         + age + I(age^2)  | congressionalDistrict)]
model_11 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "Latino",
  felm(outcome ~ factor(treat) + factor(major_party)
       + age + I(age^2)| congressionalDistrict)]
model_12 <- d[
  treat %in% 0:4 & !(historyCode2016 == 'E') & race3 == "Black",
  felm(outcome ~ factor(treat) + factor(major_party)
       + age + I(age^2) | congressionalDistrict)]
```

4.5.7 Table 12: Message Effects, By Racial/Ethnic Subgroup

4.5.8 Figure 8: Message by Subgroup Effects

4.6 Results Are Robust to Exclusion of Early and All Vote-By-Mail Voters

In the main analysis that we report, we exclude all voters who voted early. It is possible – however very unlikely – that voters could have voted by mail *after* having received our email messages. In our investigation, we found that there are some postal areas where a ballot mailed by a registered voter *could have* been placed into a USPS blue box for pickup on the afternoon that we sent our email message, and been received by the officials on time. We believe that perhaps 10% of registered voters reside in areas where this is possible, but we must acknowledge that this is an imprecise estimate.

Here, we examine whether estimates change as a result of modifying the set of voters who are *included* or *excluded*. In particular, in this set of models we newly exclude all individuals who voted-by-mail. This is because it is not possible for us to ascertain the time that a ballot was mailed.

The analysis that is reported in the main body of the paper is *extremely causally conservative*. We include the 29% of voters who were very unlikely to have their behavior altered by our messages (because they are very likely to have already mailed in their ballot) in our sample. As a result, what we report in the main body is a causally sound, but lower-bound on the treatment effect that is possible for these messages.

In this section, we replicate the same results as reported in the main tables above, but in this instance we exclude those individuals who voted by mail. We encourage caution in interpreting these results, because it is possible (though probably unlikely) that these estimates contain post-treatment bias.

We summarise the results of this robustness check as follows:

- In all cases the direction of the effect is the same
- The magnitude of estimated effects are slightly larger in these models that do not include vote-by-mail voters

Table 12: Message Effects, By Racial/Ethnic Subgroup

	Dependent variable:		
	outcome		
	White	Latino	Black
	(1)	(2)	(3)
Baseline	−0.003	−0.009*	−0.024***
	(0.002)	(0.005)	(0.007)
Descriptive Norm	−0.003	−0.010*	−0.023***
	(0.002)	(0.005)	(0.007)
Ethnic Descriptive Social Norm 1	−0.003	−0.006	−0.023***
	(0.002)	(0.005)	(0.008)
Ethnic Descriptive Social Norm 2	−0.004	−0.014***	−0.016**
	(0.002)	(0.005)	(0.008)
Age	0.005***	−0.005	−0.072***
	(0.002)	(0.004)	(0.011)
Age^2	−0.105***	−0.102***	−0.122***
	(0.002)	(0.004)	(0.006)
age	0.014***	0.021***	0.023***
	(0.0002)	(0.001)	(0.001)
I(agê2)	−0.0001***	−0.0002***	−0.0002***
	(0.00000)	(0.00001)	(0.00001)
Observations	203,917	63,369	36,482
R^2	0.090	0.102	0.125
Adjusted R^2	0.090	0.101	0.124
Note:			*p<0.1; **p<0.05; ***p<0.01

- There is less precision in these estimates, due to the reduction of sample size

Overall, we belive a fair interpretation is that we learn similar things under both schemes: messages reduce turnout among black and Latino, but not white voters.

```
model_no_absentee_1 <- d[treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')),
               lm(outcome ~ any_message)]
model_no_absentee_2 <- d[treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')),
             felm(outcome ~ any_message | congressionalDistrict)]
model_no_absentee_3 <- d[treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')),
             felm(outcome ~ any_message + factor(major_party) + race3 +
                    age + I(age^2) | congressionalDistrict)]
model_no_absentee_subgroup_1 <- d[
  treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')) & race3 == "White",
  felm(outcome ~ any_message + factor(major_party)
         + age + I(age^2)  | congressionalDistrict)]
model_no_absentee_subgroup_2 <- d[
  treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')) & race3 == "Latino",
  felm(outcome ~ any_message + factor(major_party)
       + age + I(age^2) | congressionalDistrict)]
model_no_absentee_subgroup_3 <- d[
  treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')) & race3 == "Black",
  felm(outcome ~ any_message + factor(major_party)
       + age + I(age^2) | congressionalDistrict)]
model_no_absentee_subgroup_hte <- d[
  treat %in% 0:4 & !(historyCode2016 %in% c('E', 'A')),
  felm(outcome ~ any_message * race3 + factor(major_party)
         + age + I(age^2)  | congressionalDistrict)]
```

4.6.1 Table 13: Treatment Estimates are Little Changed Excluding Vote-By-Mail Voters

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, Oct 01, 2020 - 08:55:15 PM

4.6.2 Table 14: Subgroup Treatment Effects are Little Changed Excluding Vote-By-Mail Voters

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, Oct 01, 2020 - 08:55:15 PM

4.7 In Text Result: How many people were kept from the polls?

How many registered voters did these email messages keep from voting? Including all individuals who were assigned to receive an email – including those who voted early – produces the largest estimate of this demobilizing effect.

In a model that does not condition on whether someone voted absentee – necessary because it is *in principle* possible for a voter to receive our message and then decide the modality to use to cast their vote, a post-treatment decision – we estimate that this message experiment reduced turnout by at most -2249, and by at least -1390.

References

Alvarez, R Michael, and Lisa Garcı'a Bedolla. 2004. "The Revolution Against Affirmative Action in California: Racism, Economics, and Proposition 209." *State Politics & Policy Quarterly* 4 (1). SAGE Publications Sage CA: Los Angeles, CA: 1–17.

Table 13: Treatment Estimates Are Little Changed Excluding Vote-By-Mail Voters

	Dependent variable:			
	outcome			
	OLS		*felm*	
	(1)	(2)	(3)	(4)
Any Message	-0.008^{***}	-0.007^{**}	-0.007^{***}	-0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Republican			0.051^{***}	0.051^{***}
			(0.003)	(0.003)
Independent			-0.099^{***}	-0.099^{***}
			(0.003)	(0.003)
Black			-0.128^{***}	-0.115^{***}
			(0.004)	(0.008)
Latino			-0.055^{***}	-0.049^{***}
			(0.003)	(0.007)
Other			-0.094^{***}	-0.101^{***}
			(0.004)	(0.009)
Age			0.018^{***}	0.018^{***}
			(0.0003)	(0.0003)
Age2			-0.0002^{***}	-0.0002^{***}
			(0.00000)	(0.00000)
Any Message * Black				-0.016^{**}
				(0.008)
Any Message * Latino				-0.008
				(0.007)
Any Message * Other				0.010
				(0.010)
Constant	0.602^{***}			
	(0.003)			
Observations	176,719	176,719	176,639	176,639
R^2	0.00004	0.022	0.095	0.095
Adjusted R^2	0.00003	0.021	0.094	0.094
Note:				*p<0.1; **p<0.05; ***p<0.01

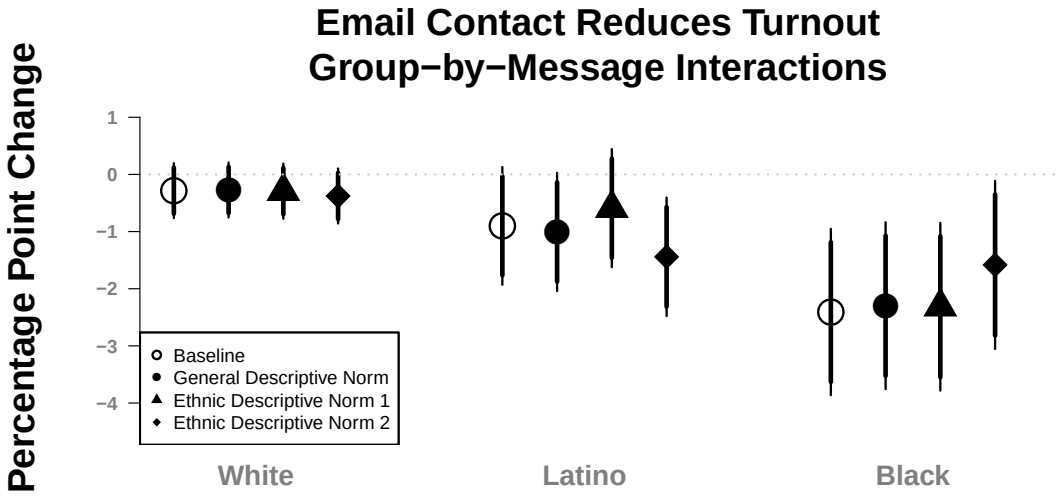


Figure 8: Message by Subgroup effects

.

Table 14: Subgroup Treatment Estimates Are Little Changed Excluding Vote-By-Mail Voters

	Dependent variable:		
	outcome		
	White RV	Black RV	Latino RV
	(1)	(2)	(3)
Any Message	−0.004	−0.012*	−0.019**
	(0.004)	(0.006)	(0.008)
Republican	0.061***	0.028***	−0.053***
	(0.004)	(0.007)	(0.014)
Independent	−0.101***	−0.081***	−0.113***
	(0.004)	(0.006)	(0.007)
Age	0.017***	0.024***	0.022***
	(0.0004)	(0.001)	(0.001)
Age2	−0.0001***	−0.0002***	−0.0002***
	(0.00000)	(0.00001)	(0.00001)
Observations	100,731	36,051	24,920
R^2	0.073	0.058	0.065
Adjusted R^2	0.072	0.057	0.064
Note:			*p<0.1; **p<0.05; ***p<0.01

García-Bedolla, Lisa. 2005. *Fluid Borders: Latino Power, Identity, and Politics in Los Angeles.* Los Angeles: University of California Press.

García-Bedolla, Lisa, and Melissa R. Michelson. 2012. *Mobilizing Inclusion: Transforming the Electorate Through Get-Out-the-Vote Campaigns.* Yale University Press.

Gerber, Alan S, and Donald P Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* WW Norton.

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

Hajnal, Zoltan, and Michael U Rivera. 2014. "Immigration, Latinos, and White Partisan Politics: The New Democratic Defection." *American Journal of Political Science* 58 (4). Wiley Online Library: 773–89.

Jardina, Ashley. 2019. *White Identity Politics.* Cambridge University Press.

McGowen, Ernest B. 2010. "Depressing the Vote: The Environmental Determinants of Turnout in African-American Precincts in the 2000 Presidential Election." In *African-American Political Psychology*, edited by White I.K. Philpot T.S. Palgrave Macmillan.

Philpot, Tasha S, Daron R Shaw, and Ernest B McGowen. 2009. "Winning the Race: Black Voter Turnout in the 2008 Presidential Election." *Public Opinion Quarterly* 73 (5). Oxford University Press: 995–1022.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* MIT press.