Manuscript Draft:
09/14/2023

TWO TERMS OF ENDEARMENT:
INCUMBENT-PARTY PERFORMANCE IN U.S. PRESIDENTIAL ELECTIONS

John V. Kane
New York University
jvk221@nyu.edu

Word Count: 3998

Key words: incumbency; elections; voting; U.S. presidency

ABSTRACT

Presidential elections are arguably the most consequential recurring political event in the United States. Understanding the factors that determine their outcomes, therefore, is of substantial importance. One proposed factor pertains to candidates' incumbency status, yet its nature is complex and difficult to study with observational data. In particular, the individual-level mechanisms underlying incumbency effects remain surprisingly unclear. This Letter proposes many citizens generally believe that, ceteris paribus, presidents *should* be afforded two terms. Existence of such a norm implies that incumbency status possesses an *inherent* effect, operating independent of other mechanisms stemming from incumbency. A large, pre-registered survey experiment is employed to isolate incumbency status, finding evidence for a one-term advantage and clarifying the nature of the two-term disadvantage. The study thus uncovers a micro-foundational mechanism underlying incumbent-party performance in presidential elections. Finally, analyses of panel data explore which voters may be systematically inclined to vote based upon incumbency status.

Presidential elections are arguably the single largest, and most consequential, recurring political event in the United States. Understanding the factors that determine the outcomes of these elections is, therefore, of immense importance. For example, both mass media and scholarship tend to place great emphasis on the state of the national economy as a determinant of presidential approval and thus of election outcomes (Kane 2016; Lewis-Beck and Stegmaier 2000; Norpoth 2001), though the relative importance of the economy for presidential elections may be decreasing amid a more polarized public (Donovan et al. 2019).

Another factor cited by media and scholars is that of the presidential candidates' incumbency status. Scholarship on congressional elections provide clear evidence that incumbent-candidates possess a strong advantage (e.g., Abramowitz, Alexander, and Gunning 2006). However, the nature of incumbency advantage in U.S. presidential elections is, as Campbell (2014, 301) contends, "more complex than often thought."

To what extent does incumbency status matter for presidential elections? Scholarship on this question point to several empirical regularities. First, presidential candidates who have already served one term tend to win a second term at a much better-than-even chance (Mayhew 2008). However, while this "virtually assured" victory constitutes an advantage (Campbell 2014, 302), a complicating factor is that, per the 22nd Amendment to the Constitution, individual candidates are prohibited from serving more than two terms in office. Thus, incumbency can be studied not only at the level of incumbent-candidates, but also at the level of *incumbent-parties*. This analytical shift helps reveal a second empirical regularity in the forecasting literature: *incumbent-parties*—that is, the party occupying the presidency—tend to perform worse in national elections after two terms in office vis-à-vis after one term (Chatterjee and Eyigungor 2020; Sidman, Mak, and Lebo 2008).

Extant scholarship, then, suggests a one-term advantage (1TA) for incumbents over their opponents in the same election, but a two-term disadvantage (2TD) for a party (and, thus, the candidate it has nominated) after serving two (or more) terms.[1] Indeed, several presidential forecasting models in political science acknowledge this effect by accounting for whether or not a candidate is a one-term incumbent (e.g., Abramowitz 2016; Norpoth 2016).

Two important points remain unclear, however. The first is whether the 2TD represents an absolute disadvantage against the opponent in the same election (e.g., because voters believe it is "time for change" (Abramowitz 2016)), or if it merely reflects the *absence of* a 1TA. Because elections typically involve one party having either served one or two terms in office, it is difficult to investigate this question using observational data alone (e.g., aggregated election returns). That is, while we can compare a two-term incumbent-party's performance to four years prior, we cannot know how the same candidate would have fared in an election had their party *not* occupied the presidency for the previous eight years.

Second, and even more importantly, what, specifically, accounts for these predictable swings of the "electoral pendulum" (Norpoth 2016, 655)? Though scholarship has been able to point to empirical regularities, the micro-foundational mechanisms *underlying* the 1TA and 2TD remain unclear. This may be partly because potential mechanisms are indeed numerous. For example, incumbent presidents—much like incumbent members of Congress—may acquire various policymaking-related skills and campaigning advantages over potential opponents (Druckman, Kifer, and Parkin 2019; Levitt and Wolfram 1997; Mayhew 2008). Incumbent-party presidents are also able to selectively engage in credit-claiming for various national conditions (Campbell 2014). Another possibility is that of risk-aversion and/or status quo bias among members of the public

[1] For ease of exposition, any because of the rarity in recent decades of parties serving three consecutive terms, any reference to "two-term" will leave implied "or more".

(Mayhew 2008; Samuelson and Zeckhauser 1988). However, this latter mechanism would struggle to explain why, after *two* terms, incumbent-parties no longer appear to wield any advantage. Another potential mechanism might involve "thermostatic politics" theory (Wlezien 1995) wherein, after eight years of relatively liberal public policymaking, for example, a segment of the public shifts toward the party that advocates for more conservative public policymaking (see also Chatterjee and Eyigungor 2020). Governing elites' drift away from voters' policy preferences also helps explain the "cost of ruling" effect—i.e., the tendency for governing parties to lose political support over time (e.g., see Wlezien 2017). Additionally, candidates whose party has already served two terms may vary in their ability to garner credit (or blame) for their party's previous eight years in office (Sidman, Mak, and Lebo 2008).

In this Letter, I propose an additional mechanism—one that is *inherent* to incumbency itself. Specifically, I propose that beliefs about incumbency represent a social norm among a sizable portion of the U.S. public. Given the ratification of the 22nd Amendment in 1951, it is reasonable to suspect that many citizens simply believe that presidents *should* be able to serve two terms. In other words, *ceteris paribus*, a president serving only one term when s/he *is legally permitted to* serve two may seem inherently "wrong", especially if there also exists a perception that incumbents winning a second term is the *normal* course of events. This supposition coheres with broader findings regarding social norms. For example, extant evidence contends that prevailing social norms are substantially influenced by a nation's history (Acemoglu and Jackson 2015; Horne and Mollborn 2020), while experimental evidence demonstrates that conventions repeatedly observed in previous rounds can translate into an "intrinsic normative power" that shapes individuals' behavior going forward (Guala and Mittone 2010, 75).

By the same logic, there may also exist a social norm in which some citizens believe that, *ceteris paribus*, parties out-of-power deserve to "get a turn" after a member of the in-power party has served two terms (much akin to "time for change" (see Abramowitz 2016)). While such citizens may also be the same who privilege incumbents after one term, these two potential norms are theoretically distinct and imply different consequences for parties in control of the presidency.

To the extent these incumbency-related norms exist in the mass public, citizens wield a simple heuristic to assist them with voting in presidential elections—i.e., they can vote based upon how many terms a candidate, or candidate's party, has occupied the presidency. Again, and in contrast to the many proposed mechanisms identified above, emphasis on these potential norms implies that the 1TA and/or 2TD may be partly inherent to incumbency status itself, irrespective of downstream *consequences* of incumbency (e.g., campaign advantages, credit-claiming for economic conditions, or public demand for more or less liberal policymaking). This Letter therefore aims to test the following hypotheses:

> ***One-Term Advantage (1TA) Hypothesis (H1):*** *Candidates who have occupied the presidency for one term will attract a significantly greater vote share than candidates who have not served one term.*
>
> ***Two-Term Disadvantage (2TD) Hypothesis (H2):*** *Candidates whose party has occupied the presidency for two terms will attract a significantly smaller vote share than candidates whose party has not served two terms.*

OVERVIEW OF STUDIES

Studying the effects of incumbency status in presidential elections, as well as the potential mechanisms underlying it, is beset by multiple challenges. Innumerable factors vary from election to election, making direct comparisons difficult. For example, in addition to variation in national and international conditions every four years, electorates vary in demographic composition and

partisan balance (e.g., Bartels 2000). Changes also occur in the strategic behavior of out-party opponents (e.g., in deciding whether or not to run) from election to election (Mayhew 2008), rendering it difficult to determine the extent to which an election outcome is due to factors involving the incumbent vis-à-vis the incumbent's opponent. Similarly, presidents may be evaluated based upon the perceived competence of the *previous* administration (Green and Jennings 2017, 167–96). Beyond these challenges, elections typically feature a candidate whose party has held the presidency for one or two terms, which removes a baseline that would enable us to empirically separate a 1TA from a 2TD. As Wlezien (2001, 28) observes, "The two literally mirror each other—one "on," the other "off'—since 1952, with the exception of a single election, 1992."

The present study therefore relies primarily upon a large, pre-registered survey experiment (n=3,351) to test both **H1** and **H2** via isolating incumbency status from other contextual factors. This approach allows for a direct test of inherent incumbency effects while also permitting the empirical disentangling of a 1TA and 2TD. Finally, I analyze panel data to explore factors associated with voting based upon incumbency status of presidential candidates.

First, however, I present the outcomes of all presidential elections over the past 70 years, which help to succinctly illustrate the nature of the 1TA and 2TD. These descriptive analyses extend earlier analyses done by Abramowitz (2021) and Norpoth (2021), but focus only on incumbency status and include all presidential elections since the ratification of the 22nd Amendment.

INCUMBENCY EFFECTS IN MODERN PRESIDENTIAL ELECTIONS

Table 1 reports the outcomes of all presidential elections since 1952, both in terms of win rates and two-party vote shares. In short, any cell shaded in lighter gray is one that fits the pattern

TABLE 1. Incumbent-Party Advantage & 2-Term Incumbent-Party Disadvantage

	After 1 Term in Office		After 2 or More Terms in Office	
Year	*Outcome*	*2-Party Vote Share*	*Outcome*	*2-Party Vote Share*
1952			LOSS	44.71
1956	WIN	57.75%		
1960			LOSS	49.90
1964	WIN	61.35%		
1968			LOSS	49.60
1972	WIN	61.81%		
1976			LOSS	48.93%
1980	LOSS	44.71%		
1984	WIN	59.15%		
1988			WIN	53.94%
1992			LOSS	46.52%
1996	WIN	54.73%		
2000			WIN*	50.26%
2004	WIN	51.21%		
2008			LOSS	46.35%
2012	WIN	51.98%		
2016			WIN*	51.06%
2020	LOSS	47.73%		
TOTAL	77.78% W	54.49% Avg.	33.33% W	49.03% Avg.

Notes: Data from The American Presidency Project at University of California, Santa Barbara. Elections begin at 1952 as the 22nd Amendment (barring more than two terms) was ratified in 1951. * indicates the popular vote result conflicts with the Electoral College result.

expected by **H1** and **H2**, while darker cells are those that do not fit the pattern. As the table demonstrates, the win rate (in terms of popular vote[2]) is impressively high for one-term incumbent-party candidates (nearly 78%) and much lower (33.33%) for candidates whose party has occupied the presidency for two terms (or more, in the case of 1952). This constitutes a 44 percentage-point difference-in-proportions of popular-vote wins ($p<.05$).[3] In terms of respective vote shares, one-term incumbent-party candidates have averaged 54.5% of the two-party vote (i.e., a 9 percentage-point advantage over their competitors), while candidates whose party has served for two or more

[2] "Winning" here is defined as obtaining a majority of the two-party vote, regardless of the Electoral College outcome. Because the focus of this study is on behavior in the mass public, only the popular vote is considered.

[3] Given the directional hypotheses, one-tailed *p*-values are reported.

terms average only 49% of the two-party vote (i.e., a 2 percentage-point disadvantage). The difference between these two vote shares is significant (p=.01). The one-term incumbent-party vote share is also significantly greater than the vote share we would expect by chance (50%; p=.03), while the two-term vote share is not significantly less than 50% (p=.16). That said, in all nine elections with a two-term incumbent-party candidate, we observe a decrease in vote share from the previous election.

This descriptive section thus illustrates the presence and potential magnitude of the 1TA and 2TD. The results suggest, however, that the 2TD is noticeably weaker, perhaps not reflecting a disadvantage so much as the *absence of* a one-term advantage given that the average (two-party) vote share is only slightly below 50%. The next section will investigate whether such incumbency effects are to some degree *inherent*—i.e., that the incumbency status of candidates is capable, in and of itself, of influencing voting behavior in presidential elections.

AN EXPERIMENTAL TEST OF INHERENT INCUMBENCY EFFECTS

In contrast to the aggregated observational data above, a survey experiment enables us to *isolate* incumbency status from other potential factors that might explain the 1TA and 2TD at the individual level, such as the candidates' qualities, name recognition (Abramowitz 1975), the national context (e.g., the economy), and downstream advantages related to campaigning and credit-claiming. A survey experiment was therefore pre-registered and fielded in early 2023 via Lucid Theorem.[4] The total sample size was 3,351, with quotas in place for race/ethnicity, gender, age, and geographic region (see SA for demographic details).

[4] The pre-registration document can be found here: https://osf.io/3s2ty/?view_only=b7789d0da8a747918b71635d6cbbd2c7

TABLE 2: Incumbency Norms in the American Public

Survey Questions Regarding Incumbency Norms	
One-Term Norm	*The U.S. Constitution allows presidents to serve up to two (4-year) terms in office. Please rate your level of agreement with the following general statement, being sure to disregard any considerations relevant to the current U.S. president:* ***Regardless of their political party, as a general rule, I feel that presidents should be re-elected so that they can serve a second term in office.***
Two-Term Norm	*Now please rate your level of agreement with the following general statement, being sure to disregard any considerations relevant to the current U.S. president:* ***Regardless of their political party, as a general rule, I feel that once a political party has had control of the presidency for two full (4-year) terms, the other party should get a chance at being president.***

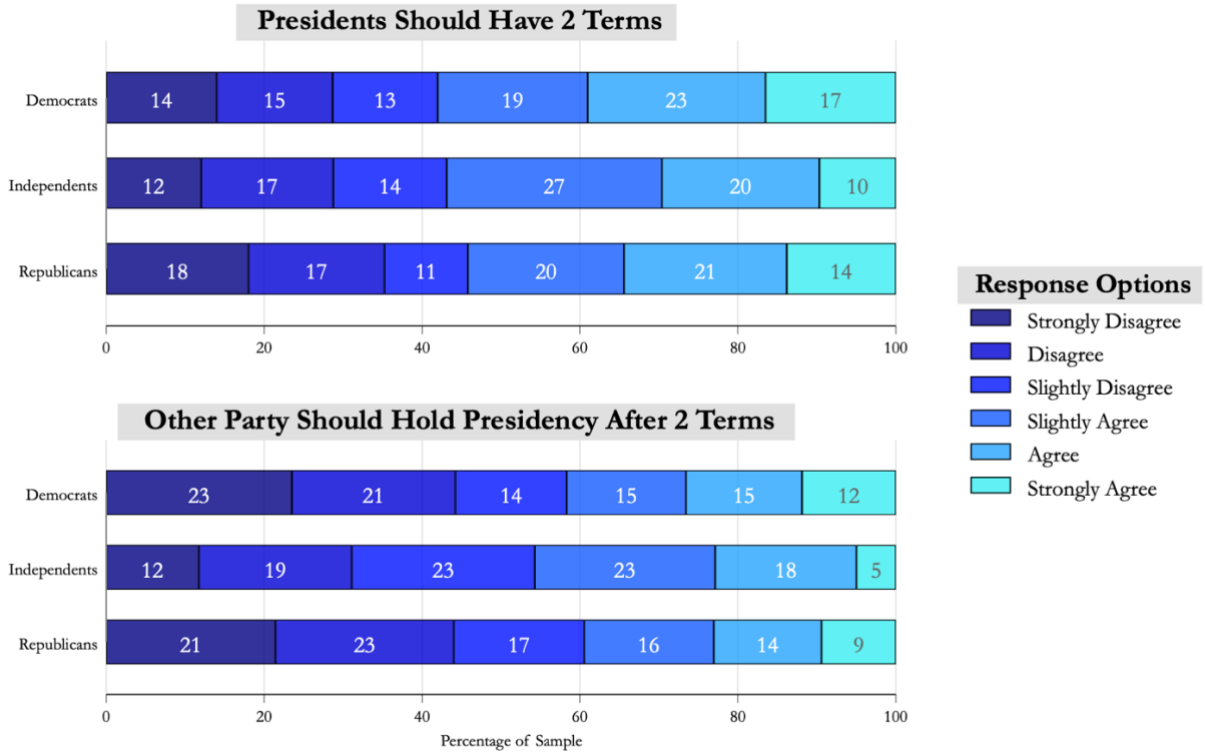
Notes: Questions asked early in the survey (prior to experiment, with an unrelated set of questions appearing after). Each question featured a six-point agreement scale ranging from Strongly Disagree (1) to Strongly Agree (6).

Earlier in the survey, a subset of respondents was asked two questions about incumbency norms.[5] These items are featured in Table 2. Both questions were followed by six-point agreement scales (ranging from "Strongly Disagree" to "Strongly Agree").

Figure 1 displays the results for these two items, disaggregated by partisan identity. Several features are worth noting. First, there is a remarkable level of agreement, particularly for the notion that presidents "should be re-elected". A significantly larger proportion of respondents expressed agreement with the first item compared to the second (56.4% versus 41.7%, $p<.001$)—a pattern that holds for each partisan group.[6] Second, across partisan groups, there is notable homogeneity in response patterns.[7] This suggests that a sizable share of the public subscribes to a norm in which presidents should serve a second terms and, to a lesser extent, that the opposing party should

[5] 100% of control group respondents were asked (as the control condition did not reference incumbency) while only 50% of treatment group respondents were asked.
[6] The overall correlation between the two items was .33 ($p<.001$).
[7] Chi-squared tests confirmed no significant relationship between party identification and each item at the .10 level. See SA for additional analyses examining correlates of these normative beliefs.

FIGURE 1. Incumbency Norms in the U.S. Public (by Partisanship)



Notes: Bars indicate percentages. Both questions asked at start of survey (pre-treatment) Total *n* for each panel = 1,672. Lucid data.

occupy the presidency after these two terms conclude. To the extent this is true, it follows that manipulating incumbency status itself should alter respondents' voting behavior.

The survey experiment therefore randomized 80% of the sample to one of two treatment conditions, and the remaining 20% to a control condition. These conditions describe a (fictional) presidential election from "many decades ago" that featured two candidates (Democratic and Republican), with the goal of learning how respondents would have voted. The key manipulation, per above, is the incumbency status of the candidates; however, the party of the incumbent is also randomized to account for the strong influence of partisanship on vote choice. Table 3 features the content of these conditions. To minimize the possibility of other potential considerations about

TABLE 3: Design of Survey Experiment

Content Featured in Survey Experiment Vignettes	
Non-Manipulated Content	We would like to know how you would have voted in a presidential election many decades ago in US history. The economy at the time was performing about average, with unemployment and inflation both at fairly typical levels. The US was not engaged in any major conflicts overseas. Two candidates were running for president at the time: one a Democrat, the other a Republican. The two candidates had roughly an equal number of years working in government.
Control Condition (Non-Incumbent Candidate)	[Respondents proceeded to outcome measure after exposure to content from above]
One-Term Incumbent Condition	The [Democratic / Republican] candidate was the incumbent president and was running for reelection to the presidency against the [Republican / Democratic] challenger.
Two-Term Candidate Condition	Neither of the two candidates had held the office of the presidency before the election. However, the president for the previous 8 years (two full presidential terms) was a member of the [Democratic / Republican] Party.
Outcome Measure	Had you been a voter during the time of this election, how do you think you would have voted? [Options: Democratic candidate, Republican candidate, voted but not for either the Democrat nor Republican, or abstained]

Notes: Table features content featured in the experimental vignettes and outcome measure. In the "One-Term Incumbent" and "Two-Term Candidate" conditions, the party of each candidate was randomized.

context being introduced, the vignette also featured brief, *non-manipulated* statements about economic and foreign policy conditions.

Importantly, within the treatment conditions, respondents were randomly assigned to observe either a Democratic or Republican incumbent-party candidate. Respondents could then choose to: 1) vote for the Democratic candidate, 2) vote for the Republican candidate, 3) vote for a different candidate, or 4) abstain (see Supplemental Appendix (SA) for exact wording).

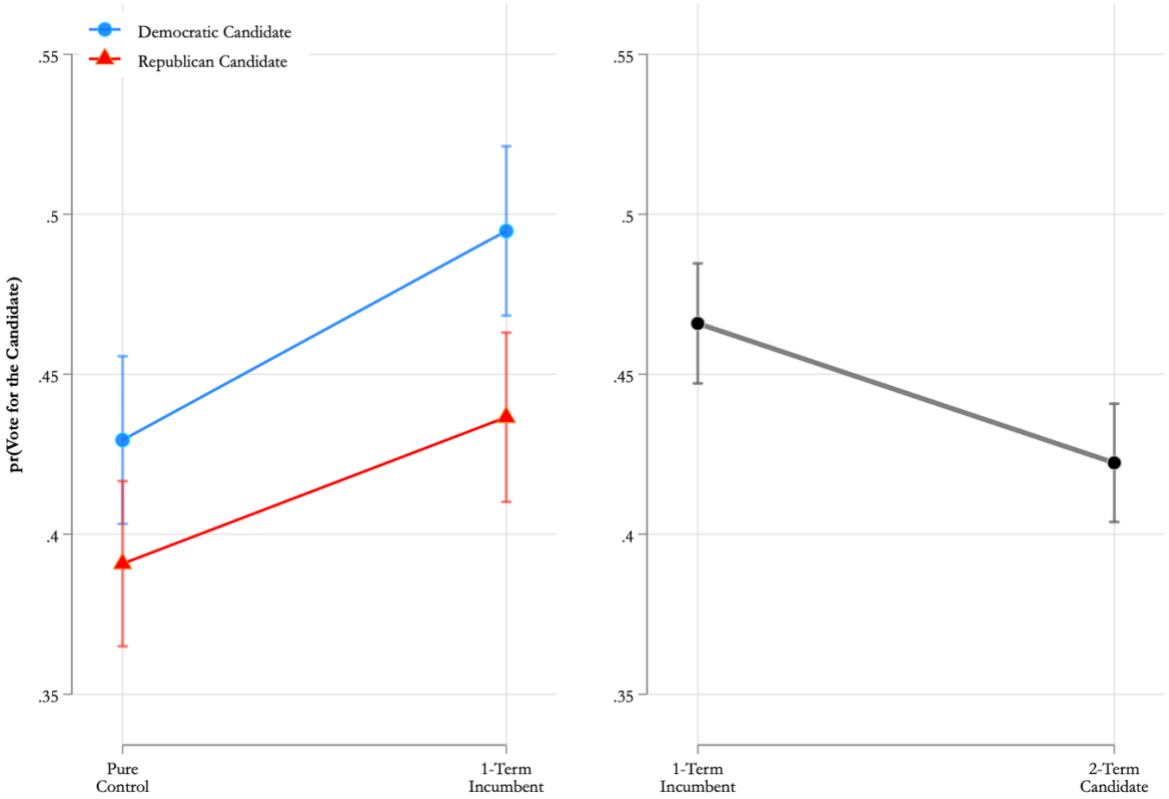
Experimental Results

Because the *Control* condition did not feature an incumbent, two dichotomous outcome measures were constructed from the four response options: 1) vote for the Democratic candidate (or not), 2) vote for the Republican candidate (or not). Examining these measures across the *Control* and *One-Term* conditions thus functions as a first test of **H1**. The left panel of Figure 2 displays these results, which demonstrate that manipulating candidates' incumbency status (from unspecified to one-term) yields an increase in the probability that respondents will vote for that candidate. The effect occurs regardless of the candidate's party: though average support for the Democratic candidate is higher in the sample than for the Republican candidate, the effect sizes are roughly similar: 6.5 percentage points for the Democratic candidate ($p<.01$) and 4.6 percentage points for the Republican candidate ($p<.05$).[8] The weighted average for these two effects is 5.6 percentage points, which amounts to a sizable advantage for one-term incumbent-party candidates (per **H1**).[9]

The right panel of Figure 2 compares the two treatment conditions as a means of examining how one-term candidates perform relative to candidates that are explicitly described as two-term incumbent-party candidates (rather than, as in the *Control*, incumbency status not being explicitly stated). As both conditions featured an incumbent, the outcome is whether the respondent voted for the incumbent candidate (=1) or not (=0). As shown in the figure, changing from a one-term to a two-term incumbent-party candidate yields a decrease in the likelihood that respondents vote for the candidate. The estimated effect is a 4.4 decrease in probability ($p=.01$), thus providing

[8] Because outcomes were dichotomous, logistic regression models were used to estimate effect sizes and obtain *p*-values.
[9] Note that the estimated probabilities can fall below 50% because third-party voting and abstention were alternative options.

FIGURE 2. The Inherent (Dis)Advantages of Incumbency Status in Presidential Elections



Notes: Point estimates indicate predicted probability of voting for a (hypothetical) candidate given assignment to a particular condition. "Pure Control" condition features no mention of incumbency; "1-Term Incumbent" specifies one candidate as having held the presidency for previous four years; "2-Term Candidate" specifies that one candidate is from a party that has held the presidency for previous eight years. Model = logistic regression with 83% confidence intervals shown to allow for comparisons with other point estimates. Total n=3,351. Lucid data.

additional empirical evidence for **H1** while also demonstrating that two-term incumbent-party candidates—even with identical descriptions and in identical contexts—do not enjoy the same inherent advantage as their one-term counterparts.

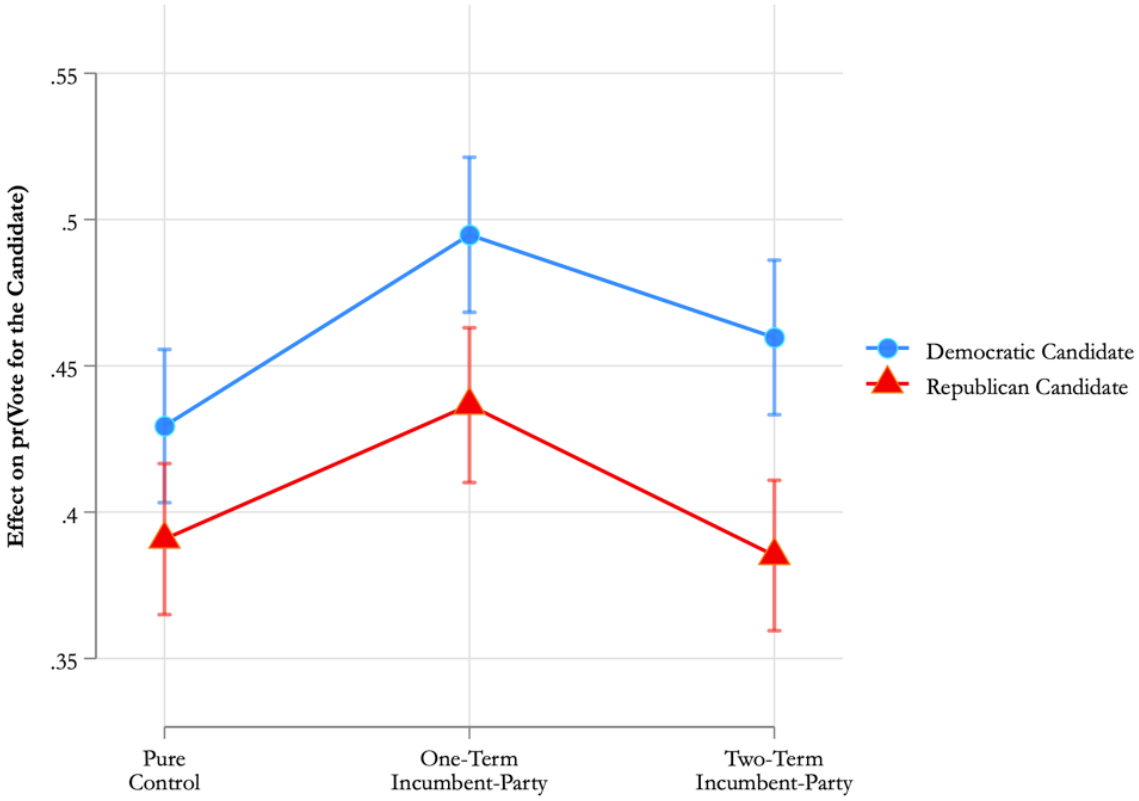
The SA examines whether the effects displayed in Figure 2 were moderated by respondents' party identification as well as by their normative beliefs regarding incumbency (see SA for details). Further, a series of additional tests confirm several important validity-related features of the study. First, given that (roughly half of) respondents were asked questions about incumbency prior to the experiment, the possibility of a priming effect was investigated. As reported in the SA, whether

or not a respondent was asked these questions does not substantially alter the effect of being a two-term versus one-term incumbent-party candidate on vote choice.

Second, to confirm sufficient respondent attentiveness to the experiment, factual manipulation checks were asked after the outcome measure (see Kane and Barabas 2019). A chi-squared test confirms that respondents' answers significantly covaried with treatment assignment ($p<.001$). Nevertheless, the overall percent answering correctly was 71%, which implies that the estimated treatment effects may be somewhat attenuated. To investigate this, responses to static, pre-treatment factual questions ("mock vignette checks"; see Kane, Velez and Barabas (2023)) were employed to estimate treatment effects across attentiveness levels. Among the most attentive respondents, the treatment effects observed in Figure 2 are substantially larger (see SA for details). Together, these results provide even stronger evidence for **H1**, as those who were most likely to have observed the candidates' incumbency status are those who display the strongest treatment effects.

Finally, because the experiment featured a *Control* condition, we can directly test **H2** and, thus, more closely investigate the nature of the 2TD. In short, is it a disadvantage in the absolute sense, or should it instead be viewed as the absence of a 1TA (i.e., a disadvantage relative only to the party's performance in the previous election)? Figure 3 displays the results of this analysis (again, separated by party of the candidate since no incumbent was specified in the *Control* condition).

FIGURE 3. Two-Term Disadvantage as the Absence of a One-Term Advantage



Notes: Point estimates indicate predicted probability of voting for a (hypothetical) candidate given assignment to a particular condition. Model = logistic regression with 83% confidence intervals shown to allow for comparisons with other point estimates. Total n=3,351. Lucid data.

The key pattern to observe is that a 2TD exists, but *only relative to the one-term candidate*, not to the *Control* condition wherein incumbency status was not mentioned. The implication, therefore, is that there exists no *inherent* disadvantage—relative to the opponent—arising from a party having already served two terms. Rather than constituting an inherent "punishment" in any given election, therefore, this supposed disadvantage for two-term incumbent-party candidates appears to simply represent the *absence* of the electoral advantage enjoyed by one-term incumbent-party candidates.[10]

[10] As in the test of **H1**, the SA provides analyses of how these effects vary by party and attentiveness.

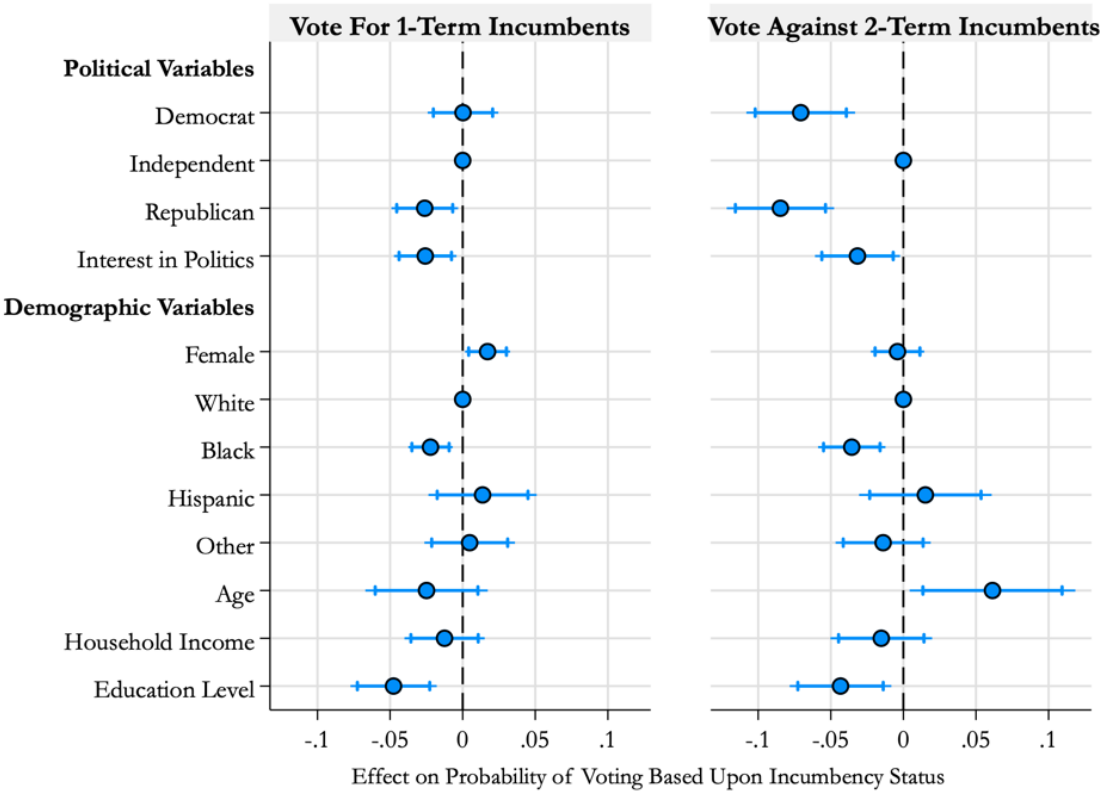
EXPLORING PREDICTORS OF INCUMBENCY-BASED VOTING IN U.S. ELECTIONS

Given the experimental findings, this final section leverages panel data to explore the types of voters who might, in real-world information environments, be especially inclined to re-elect a one-term incumbent-party presidential candidate, or vote against a two-term incumbent-party candidate, based upon these candidates' incumbency status. The Voter Study Group (VSG) panel data set features thousands of respondents across multiple waves ranging from 2011 to 2020 (the 2011 wave asked about vote choice in 2008). This permits the identification of respondents who voted for the one-term incumbent-party candidate on two separate occasions (Obama in 2012 and Trump in 2020), or *against* the two-term incumbent-party candidate on two separate occasions (Obama in 2008 and Trump in 2016).[11] Thus, a key feature of the present data is that voting based upon incumbency status requires voting for different parties.

For this exploratory analysis, standard political and socio-demographic variables, all measured in the baseline year (2011), are included as predictor variables and are recoded to range from 0 to 1 (see SA for additional details). Logistic regression is used to predict voting for one-term incumbents, as well as against two-term incumbents (0=no, 1=yes).[12] Only respondents who voted in all four elections are included in the analysis (n=2,527).[13]

[11] Hereafter the study will refer to "one-term incumbents" and "two-term incumbents" for ease of exposition.
[12] Approximately 3.2% and 5.6% of respondents fall into these two categories, respectively.

FIGURE 4. Predictors of Voting for the Two Incumbent Types (2008 – 2020)



Notes: "Vote For 1-Term Incumbents" = voting for Obama in 2012 and Trump in 2020; "Vote Against 2-Term Incumbents" = voting for Obama in 2008 and for Trump in 2016. Point estimates indicate change in probability of voting for the 1-term incumbent (left), and against the two-term incumbent (right), with covariates at observed values. 90% (vertical caps) and 95% confidence intervals shown. Base categories appear on the vertical 0 lines. Models = logistic regression. Only included respondents who voted in all four elections from 2008-2020. Total n =2,527. Voter Study Group data (2020).

First, the correlation between one-term incumbent voting and two-term incumbent voting is positive but modest (.45, p<.001), indicating that some—though certainly not all—voters may be susceptible to both types of incumbency effects in real-world elections.

The main results of this analysis are displayed in Figure 4, with the estimated effect of each predictor (on the probability of voting for both incumbent-types) represented on the x-axis (reference categories are represented by the points on the vertical zero line).

Regarding political variables, Republicans are approximately two percentage-points less likely than Independents to vote for one-term incumbents. In the right panel, we observe that both

Republicans and Democrats are far less likely than independents (over seven percentage points) to vote against two-term incumbents. Greater political interest is associated with a several percentage-point decrease in likelihood of voting for one-term incumbents and against two-term incumbents. We also observe that greater education predicts a nearly 5 percentage-point lower likelihood of voting for one-term incumbents and against two-term incumbents. These findings, while exploratory in nature, are quite consistent with existing literature (e.g., see Lewis-Beck et al. 2008): those who do not affiliate with a party, and/or might lack the interest or cognitive resources to deeply engage with election information, should indeed be most inclined to rely upon a simple heuristic—i.e., the incumbency status of a candidate—to make a decision on Election Day. Lastly, the consistently negative effects for Black persons (relative to Whites), as well as the strong (over 6 percentage-point) effect of age for two-term incumbents, are also noteworthy, further suggesting that the tendency to vote based upon incumbency status across U.S. elections has a systematic, predictable component.

CONCLUSION

The key contribution of this Letter is experimental evidence for the notion that a one-term advantage (1TA) for presidential incumbents is, at least in part, *inherent*. That is, that the 1TA may arise from the mere fact of incumbency itself, irrespective of other factors that can stem from incumbency—e.g., name recognition, perceived responsibility for national economic and/or policy conditions, and campaign advantages. Corroborating this finding, the Letter offers suggestive evidence that there exists widespread, bipartisan endorsement of a norm that one-term presidential incumbents *should* be re-elected to a second term.

A second contribution of this Letter's experimental findings is that the two-term disadvantage (2TD) can perhaps be better understood as the *absence of* a one-term advantage rather than as citizens possessing a principled desire to change party control of the presidency. This potentially offers insight as to why two-term incumbent-party candidates (i.e., candidates whose party has controlled the presidency for the past two terms) can apparently "fail to get credit" for their party's previous eight years in office (Sidman, Mak, and Lebo 2008). Put differently, the experimental findings suggest that two-term incumbent-party candidates do *not* get credit for their party's previous eight years in office, which obviates the need for any explanation (e.g., a poorly run campaign) as to how such credit might have been squandered (e.g., see Bartels and Zaller 2001). This may help explain why, despite their party having a popular president at the time of the election, two-term incumbent-party candidates—such as Gore in 2000 and Clinton in 2016—experienced remarkably close elections.[14]

That said, while there may not be an *inherent* disadvantage for a party that has served two terms, it is of course possible for this disadvantage to manifest via *downstream* effects of the party's eight years in the presidency (e.g., stronger challengers or a relatively more conservative/liberal mass public). This may help explain why, in contrast to the experiment results, we nevertheless do observe some evidence for a (albeit slight) 2TD in election results since 1952 (per Table 1).

Of course, there exists a multitude of candidate-oriented and contextual factors involved in any presidential election. Candidates' incumbency status represents only one such factor. However, in an age of high partisan polarization, elections are likely to remain close contests. Thus, identifying

[14] Per Gallup data, in early November of the respective election year, Bill Clinton's approval was 63% while Obama's was 57%. Available at: https://news.gallup.com/interactives/507569/presidential-job-approval-center.aspx

factors that can alter outcomes by even a couple of percentage points, and *why*, assists us in better determining which candidates will go on to occupy the U.S. presidency.

REFERENCES

Abramowitz, Alan I. 1975. "Name Familiarity, Reputation, and the Incumbency Effect in a Congressional Election." *Western Political Quarterly* 28 (4): 668–84.

———. 2016. "Will Time for Change Mean Time for Trump?" *PS: Political Science & Politics* 49 (4): 659–60.

———. 2021. "It's the Pandemic, Stupid! A Simplified Model for Forecasting the 2020 Presidential Election." *PS: Political Science & Politics* 54 (1): 52–54.

Abramowitz, Alan I., Brad Alexander, and Matthew Gunning. 2006. "Incumbency, Redistricting, and the Decline of Competition in U.S. House Elections." *The Journal of Politics* 68 (1): 75–88.

Acemoglu, Daron, and Matthew O. Jackson. 2015. "History, Expectations, and Leadership in the Evolution of Social Norms." *The Review of Economic Studies* 82 (2): 423–56.

Bartels, Larry M. 2000. "Partisanship and Voting Behavior, 1952-1996." *American Journal of Political Science* 44 (1): 35.

Bartels, Larry M., and John Zaller. 2001. "Presidential Vote Models: A Recount." *PS: Political Science & Politics* (01): 9–20. https://doi.org/10.1017/S1049096501000026.

Campbell, James E. 2014. "Issues in Presidential Election Forecasting: Election Margins, Incumbency, and Model Credibility." *PS: Political Science & Politics* 47 (2): 301–3.

Chatterjee, Satyajit, and Burcu Eyigungor. 2020. "Policy Inertia, Election Uncertainty, and Incumbency Disadvantage of Political Parties." *The Review of Economic Studies* 87 (6): 2600–2638.

Donovan, Kathleen, Paul M. Kellstedt, Ellen M. Key, and Matthew J. Lebo. 2019. "Motivated Reasoning, Public Opinion, and Presidential Approval." *Political Behavior*, February.

Druckman, James, Martin Kifer, and Michael Parkin. 2019. "Campaign Rhetoric and the Incumbency Advantage." *American Politics Research* 48 (1): 22–43.

Green, Jane, and Will Jennings. 2017. *The Politics of Competence: Parties, Public Opinion and Voters*. Cambridge, United Kingdom ; New York, NY, USA: Cambridge University Press.

Guala, Francesco, and Luigi Mittone. 2010. "How History and Convention Create Norms: An Experimental Study." *Journal of Economic Psychology* 31 (4): 749–56.

Horne, Christine, and Stefanie Mollborn. 2020. "Norms: An Integrated Framework." *Annual Review of Sociology* 46 (1): 467–87.

Kane, John V. 2016. "Control, Accountability, and Constraints: Rethinking Perceptions of Presidential Responsibility for the Economy." *Presidential Studies Quarterly* 46 (2): 335–64.

Kane, John V., and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science* 63 (1): 234–49.

Kane, John V., Yamil R. Velez, and Jason Barabas. 2023. "Analyze the Attentive and Bypass Bias: Mock Vignette Checks in Survey Experiments." *Political Science Research and Methods* 11 (2): 293–310.

Levitt, Steven D., and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U. S. House." *Legislative Studies Quarterly* 22 (1): 45–60.

Lewis-Beck, Michael S., William G. Jacoby, Helmut Norpoth, and Herbert F. Weisberg. 2008. *The American Voter Revisited*. University of Michigan Press.

Lewis-Beck, Michael S., and Mary Stegmaier. 2000. "Economic Determinants of Electoral Outcomes." *Annual Review of Political Science* 3 (1): 183–219.

Mayhew, David R. 2008. "Incumbency Advantage in U.S. Presidential Elections: The Historical Record." *Political Science Quarterly* 123 (2): 201–28.

Norpoth, Helmut. 2001. "Divided Government and Economic Voting." *Journal of Politics* 63 (2): 414–35.

———. 2016. "Primary Model Predicts Trump Victory." *PS: Political Science & Politics* 49 (4): 655–58.

———. 2021. "Primary Model Predicts Trump Reelection." *PS: Political Science & Politics* 54 (1): 63–66.

Samuelson, William, and Richard Zeckhauser. 1988. "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty* 1 (1): 7–59.

Sidman, Andrew H., Maxwell Mak, and Matthew J. Lebo. 2008. "Forecasting Non-Incumbent Presidential Elections: Lessons Learned from the 2000 Election." *International Journal of Forecasting*, US Presidential Election Forecasting, 24 (2): 237–58.

Wlezien, Christopher. 1995. "The Public as Thermostat: Dynamics of Preferences for Spending." *American Journal of Political Science* 39 (4): 981–1000.

———. 2001. "On Forecasting the Presidential Vote." *PS: Political Science and Politics* 34 (1): 25–31.

———. 2017. “Policy (Mis) Representation and the Cost of Ruling: U.S. Presidential Elections in Comparative Perspective.” *Comparative Political Studies* 50 (6): 711–38.

SUPPLEMENTAL APPENDIX

TWO TERMS OF ENDEARMENT:
INCUMBENT-PARTY PERFORMANCE IN U.S. PRESIDENTIAL ELECTIONS

SUPPLEMENTAL APPENDIX A:

ETHICAL STATEMENT & DESCRIPTIVE STATISTICS FOR LUCID SAMPLE

The present study obtained Institutional Review Board (IRB) approval from the researcher's University prior to fielding the study. The survey lasted approximately 10 minutes and concluded by thanking and debriefing respondents about the general purpose of the study. Respondents participated voluntarily and were compensated for their participation in an ethical manner, and in a way consistent with existing research practice (e.g., see Berinsky, Huber and Lenz 2012). Researchers paid $1.50 per respondent for participation in the study (approximately 125% of the U.S. federal minimum wage (equal to $7.25 at the time of the study)).

Table A1 provides descriptive statistics for respondents in the Lucid sample. As a reference, the table also includes national benchmarks from recent US Census data.

TABLE A1. Descriptive Statistics, Lucid Sample

	Lucid Study (N=3,351)	National Benchmarks
Median HH Income	$45k-49k	$67k
Median Age	45	38.1
Female	51.66%	51%
White	71.47%	76.3%
Black	9.73%	13.4%
Hispanic	11.97%	18.5%
Democrat	43.48%	42%
Independent	21.19%	11%
Republican	35.33%	47%
Liberal	27.01%	--
Moderate	43.48%	--
Conservative	29.51%	--
Northeast	20.14%	20%
South	37.15%	34%
Midwest	18.59%	20%
West	24.11%	26%

Notes: The Lucid sample was selected to mirror U.S. Census data on Age (18-24; 25-34; 35-44; 45-54; 55-64; 65+), Race (White; Black; Hispanic; Asian; Other), Gender, and Geographic Region (West; Midwest; Northeast; South). Household Income, Age, Gender, and Race/Ethnicity national benchmarks are from most recent US Census data available. Party identification benchmarks are from Gallup (2021 data). Regional benchmarks are from Lucid's targets based upon Census data.

SUPPLEMENTAL APPENDIX B:

SURVEY EXPERIMENT OUTCOME MEASURES & CORRELATES OF NORM ENDORSEMENT

Survey Measures

Party Identification

Generally speaking, do you consider yourself to be a(n):

- ○ Strong Democrat (1)
- ○ Democrat (2)
- ○ Independent, but Leaning Democratic (3)
- ○ Independent (4)
- ○ Independent, but Leaning Republican (5)
- ○ Republican (6)
- ○ Strong Republican (7)

Incumbency Norms

The U.S. Constitution allows presidents to serve up to two (4-year) terms in office. Please rate your level of agreement with the following general statement, being sure to disregard any considerations relevant to the current U.S. president:

Regardless of their political party, as a general rule, I feel that presidents should be re-elected so that they can serve a second term in office.

- ○ Strongly Disagree (1)
- ○ Disagree (2)
- ○ Slightly Disagree (3)
- ○ Slightly Agree (4)
- ○ Agree (5)
- ○ Strongly Agree (6)

Now please rate your level of agreement with the following general statement, being sure to disregard any considerations relevant to the current U.S. president:

Regardless of their political party, as a general rule, I feel that once a political party has had control of the presidency for two full (4-year) terms, the other party should get a chance at being president.

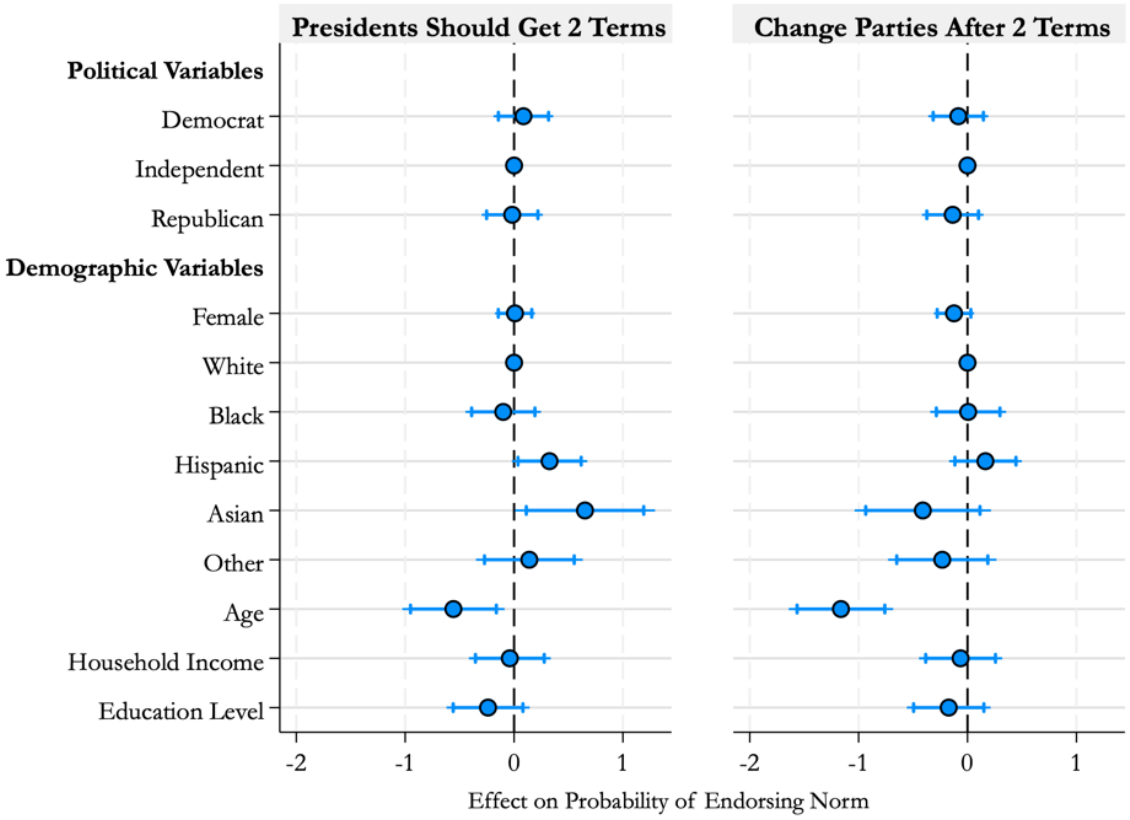
- ○ Strongly Disagree (1)
- ○ Disagree (2)
- ○ Slightly Disagree (3)
- ○ Slightly Agree (4)
- ○ Agree (5)
- ○ Strongly Agree (6)

Outcome Measure

Had you been a voter during the time of this election, how do you think you would have voted?

- ○ I would have voted for the Democratic candidate (1)
- ○ I would have voted for the Republican candidate (2)
- ○ I would have voted, but not for either the Democratic candidate nor Republican candidate (3)
- ○ I would have abstained (not voted at all) (4)

FIGURE B1. Correlates Of Norm Endorsement



Notes: Outcomes are coded as agreement (1) vs disagreement (0) with each norm (see survey questions above). Point estimates indicate change in probability of endorsing the norm of re-electing presidents who have only served one term (left), and changing parties after a party has held office for two terms (right) with covariates at observed values. 90% (vertical caps) and 95% confidence intervals shown. Base categories appear on the vertical 0 lines. Continuous variables are recoded to range from 0 to 1. Models = logistic regression with 1,622 respondents each. Overall the models suggest that age and race are associated with endorsement of each norm, while other variables exhibit smaller (non-significant) estimated effects.

SUPPLEMENTAL APPENDIX C:

SURVEY EXPERIMENT REGRESSION MODEL OUTPUT & INVESTIGATING THE POSSIBILITY OF PRIMING

Table C1 provides the regression output underlying Figures 2 and 3 in the manuscript.

TABLE C1. Regression Output (Lucid Experiment)

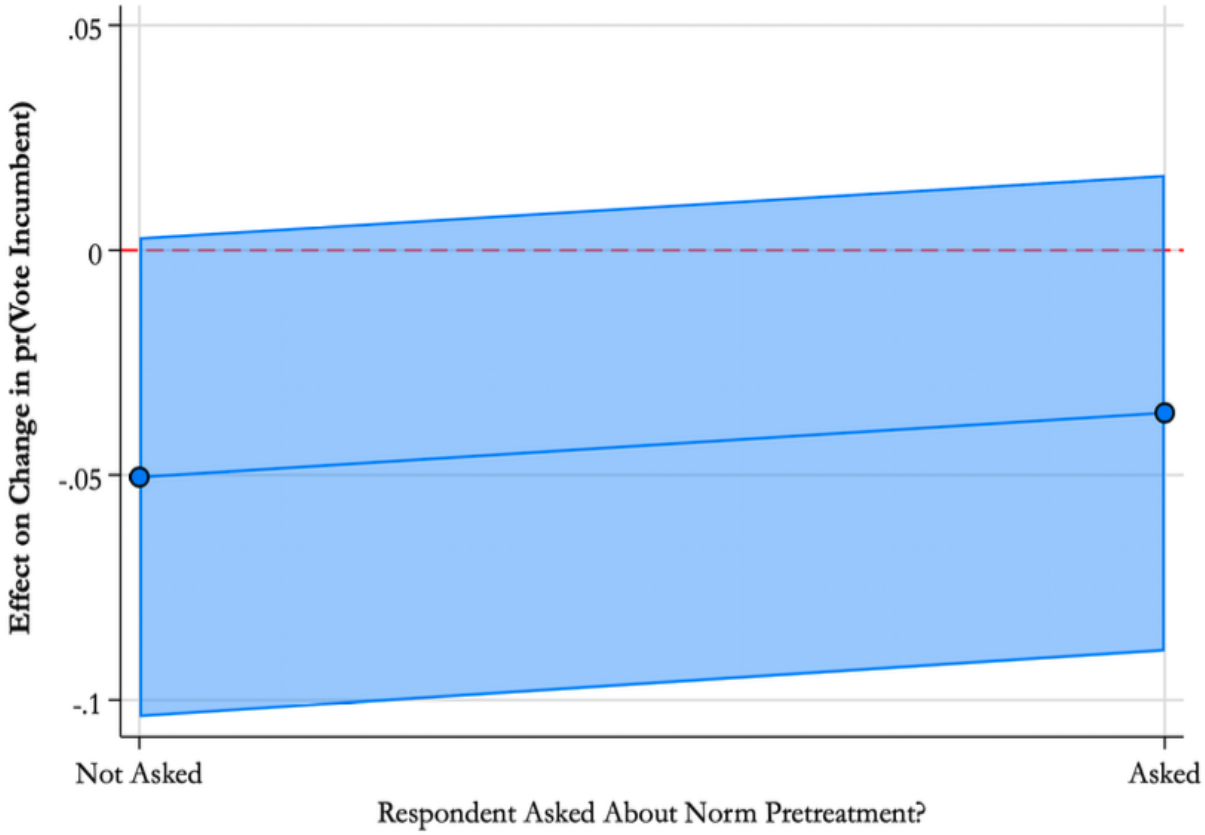
	Democratic Candidate	*Republican Candidate*	*2T vs. 1T Candidate*
One-Term Incumbent	0.26**	0.19*	--
	(0.11)	(0.11)	--
Two-Term Incumbent	0.12	-0.02	-0.18*
	(0.11)	(0.11)	(0.08)
Constant	-0.28***	-0.44***	-0.14*
	(0.08)	(0.08)	(0.05)
Observations	2,014	2,010	2,678

*Notes: These are the results underlying manuscript Figures 2 and 3. The control condition is the excluded category for the first two models; one-term incumbency is the excluded category for the third model. Logistic regression model coefficients shown with standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, ^ p<0.10 (one-tailed hypothesis tests).*

As noted in the manuscript, a subset of respondents was asked (pre-treatment) about their normative beliefs regarding incumbency in presidential elections. Specifically, 100% of control group respondents were asked these items, while only 50% of respondents in the treatment groups (i.e., a one-term or two-term incumbent-party candidate) were randomly selected to answer these items. Importantly, a series of demographic and unrelated questions were asked after the measures involving norms and before the experiment. This approach, along with asking only a subset of respondents about their normative beliefs, were deliberately taken to minimize the risk of priming respondents in a way that might augment effect sizes.

Nevertheless, the possibility of a priming effect is something that can be investigated empirically. As such, Figure C1 shows the effect of changing from a one-term party-incumbent candidate to a

FIGURE C1. Investigating the Possibility of Priming for 2T vs. 1T Contrast



Notes: Models are logistic regression with an interaction specified between treatment assignment and a whether pre-treatment norm questions were asked. 95% CIs shown. N=2,678.

two-term party incumbent candidate (see *y*-axis).[1] The *x*-axis separates this effect by whether a respondent was or was not asked the normative belief questions pre-treatment. As reported in the manuscript, this effect is negative. However, for our purposes here, what Figure C1 reveals is that the estimated negative effect is very similar regardless of whether a respondent was asked the pre-treatment normative belief question or not (the effect among those not asked (asked) rounds to 5 (4) percentage points; the interaction term between treatment and being asked was also non-significant (p=.72)). Thus, there is no evidence here that a priming effect might explain the results reported in the paper (indeed, the slightly smaller effect among those "Asked" is opposite what we would expect had priming occurred.

[1] Because all control group respondents were asked about their normative beliefs, this analysis cannot be performed using a treatment vs. control contrast. However, the 2T vs. 1T contrast is a strong effect that should still exhibit a priming effect if indeed there was one.

SUPPLEMENTAL APPENDIX D:

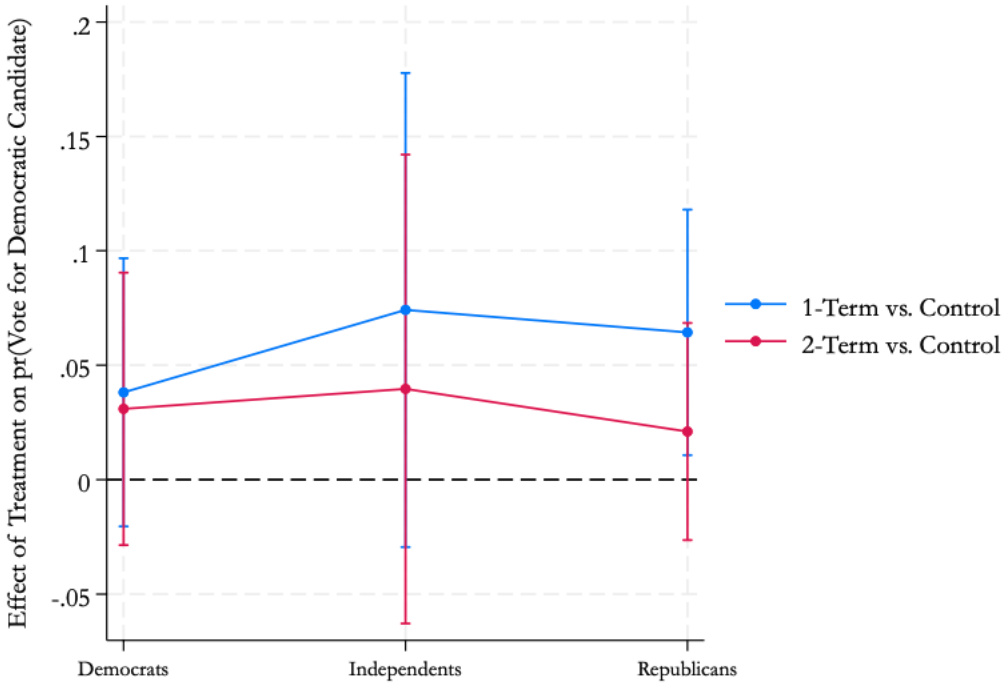
TREATMENT EFFECTS BY RESPONDENTS' PARTY IDENTIFICATION & NORMATIVE BELIEFS

Does Party Moderate Treatment Effects?

As the Lucid experiment features partisan candidates, it is reasonable that we might observe noticeably stronger or weaker treatment effects depending upon the party identification of the respondent. As such, the results featured in manuscript Figures 2 and 3 were re-analyzed via allowing the treatment to interact with respondents' party identification (Democrat, Independent, or Republican (leaners were coded as partisans)).

The results for the Democratic candidate are featured in Figure D1, which plots the estimated treatment effect of 1) the 1-term (1T) vs. Control contrast, and 2) the 2-term (2T) vs. Control contrast, for each partisan group. First, echoing the results shown in the manuscript, it is clear that the 2T vs. Control contrast (see red point estimates) exhibits noticeably weaker effects compared to the 1T vs. Control contrast (see blue point estimates). Second, treatment estimates tend to be smallest for Democratic respondents, likely because these respondents were inclined to vote for the Democratic candidate regardless of incumbency status. Third, it is notable that while the largest estimate is for Independents (i.e., those who might be most inclined to vote based upon incumbency status), we nevertheless observe a sizable 1T effect among Republican respondents.

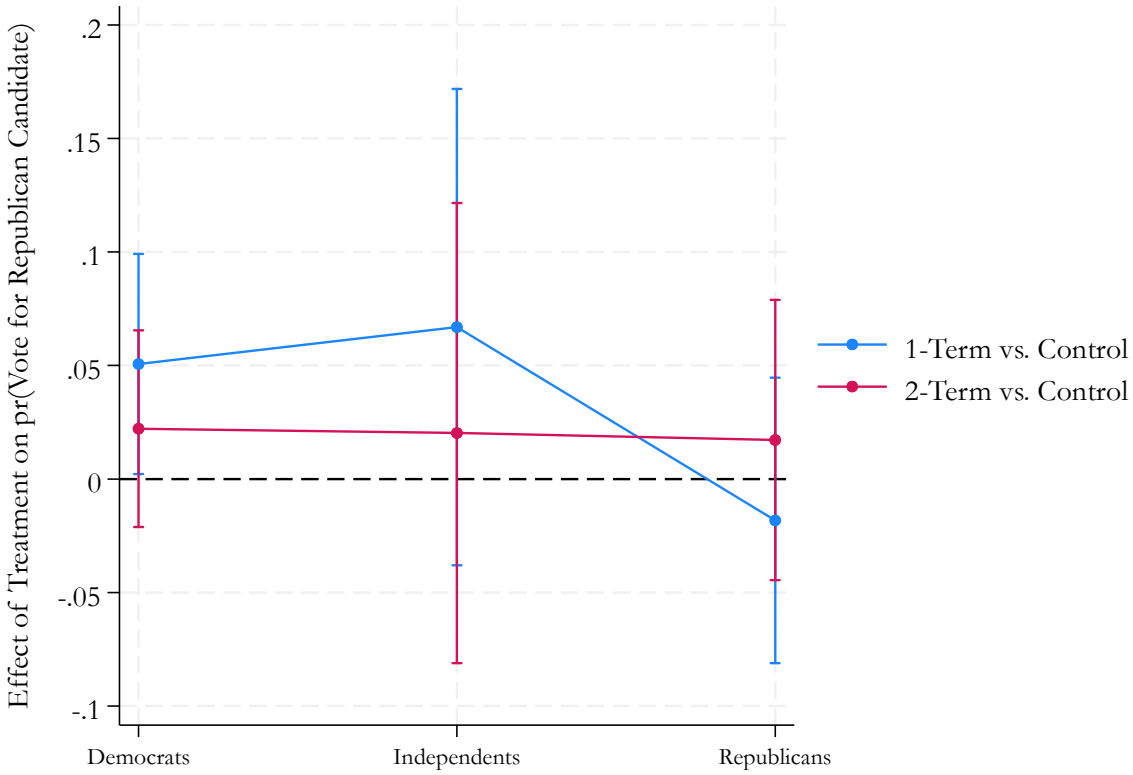
FIGURE D1. Treatment Effects on pr(Voting for Democratic Candidate) by Party



Notes: Models are logistic regression with an interaction specified between treatment assignment and a trichotomous measure of party identification (leaners coded as partisans). 95% CIs shown.

The results for the Republican candidate are featured in Figure D2. First, we again observe that the 2T vs. Control contrast (see red point estimates) tends to exhibit weaker effects compared to the 1T vs. Control contrast (see blue point estimates). Second, mirroring the previous graph, the weakest results tend to be for Republican respondents (again, likely because these respondents were inclined to vote for the Republican candidate irrespective of the treatment). Third, we again see that the largest estimate is for Independents and that a sizable 1T effect exists among Democratic respondents.

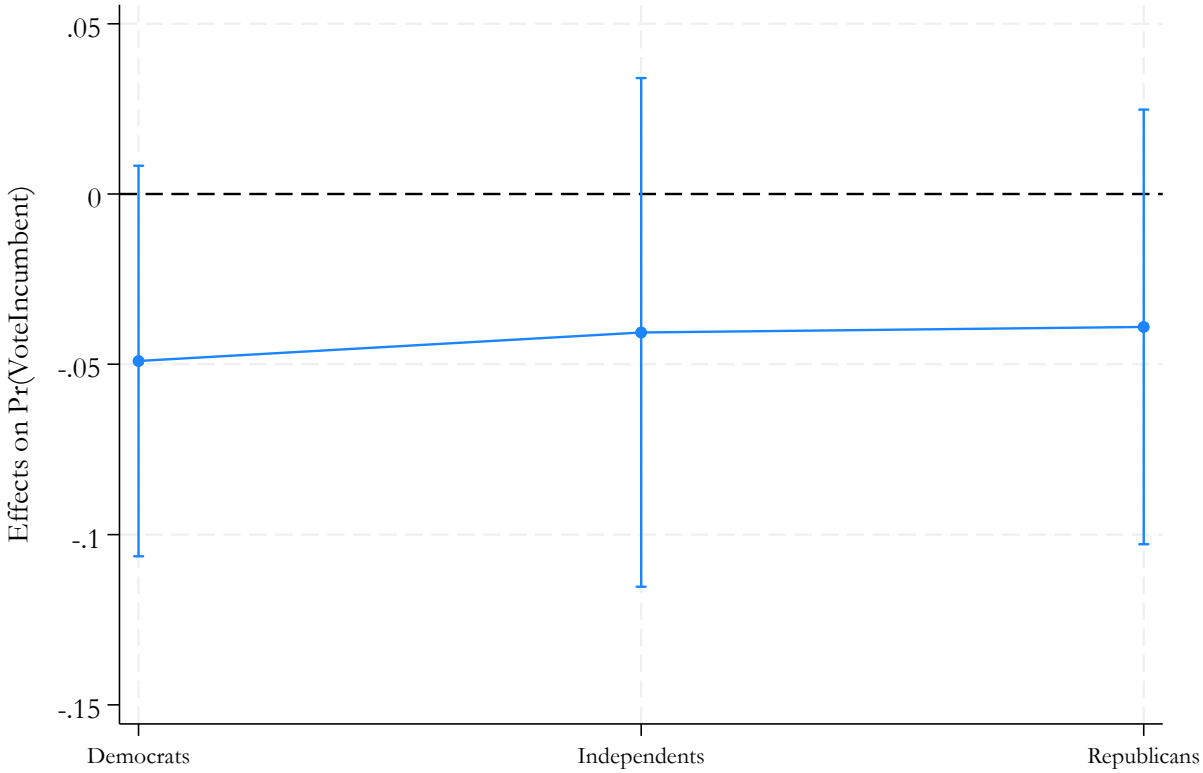
FIGURE D2. Treatment Effects on pr(Voting for Republican Candidate) by Party



Notes: Models are logistic regression with an interaction specified between treatment assignment and a trichotomous measure of party identification (leaners coded as partisans). 95% CIs shown.

Lastly, Figure D3 reports, for each partisan group, the effect of the 2T vs. 1T contrast on the probability of voting for the incumbent. Here we observe remarkably similar results across partisan categories. Compared to one-term incumbents, candidates whose party has occupied the presidency for the two previous terms receive between 4 and 5 percentage-points less in terms of vote share, regardless of the party of the respondent.

FIGURE D3. Effect of 2T vs. 1T on pr(Voting for Incumbent-Party Candidate) by Party

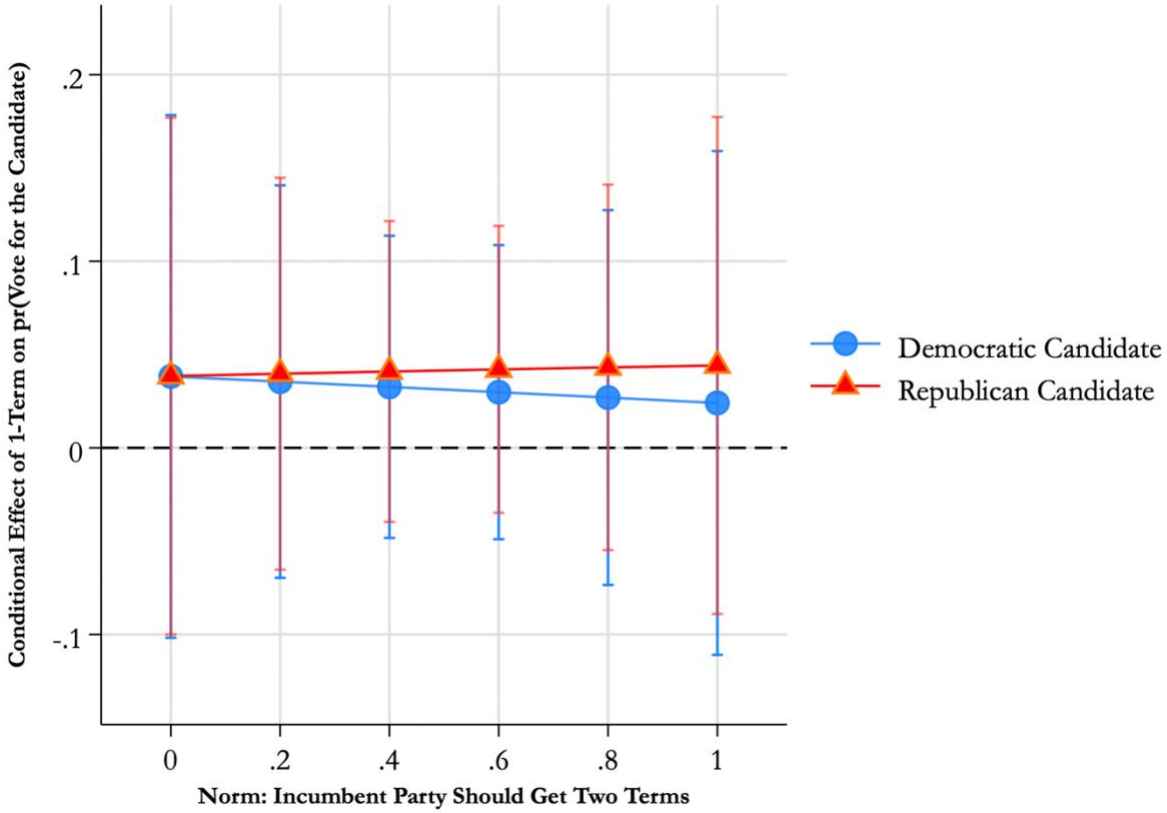


Notes: Models are logistic regression with an interaction specified between treatment assignment and a trichotomous measure of party identification (leaners coded as partisans). 95% CIs shown.

How Much do Treatment Effects Vary by Pre-Treatment Normative Beliefs?

Given the pre-registered analyses, as well as the pre-treatment measure of respondents' normative beliefs regarding incumbency, we can explore the extent to which the effects observed in the experiment vary based respondents' normative beliefs. Logistic models were run—one for the Democratic candidate and one for the Republican candidate—which specified an interaction between treatment assignment (1-term party incumbent versus Control) and the 1T normative belief measure. In Figure D4, we observe that the effect of changing from the Control condition to a 1-term party incumbent (see y-axis) does not vary a great deal depending on respondents' pre-treatment normative beliefs about 1-term incumbents. Regardless of whether the one-term party incumbent was a Democrat or Republican, the slopes are quite flat. Both interactions terms are non-significant (p=.90 for the Democrat and p=.98 for the Republican) and the total estimated

FIGURE D4. Moderating Effects of 1-Term Advantage Norm on 1T vs. Control Contrast

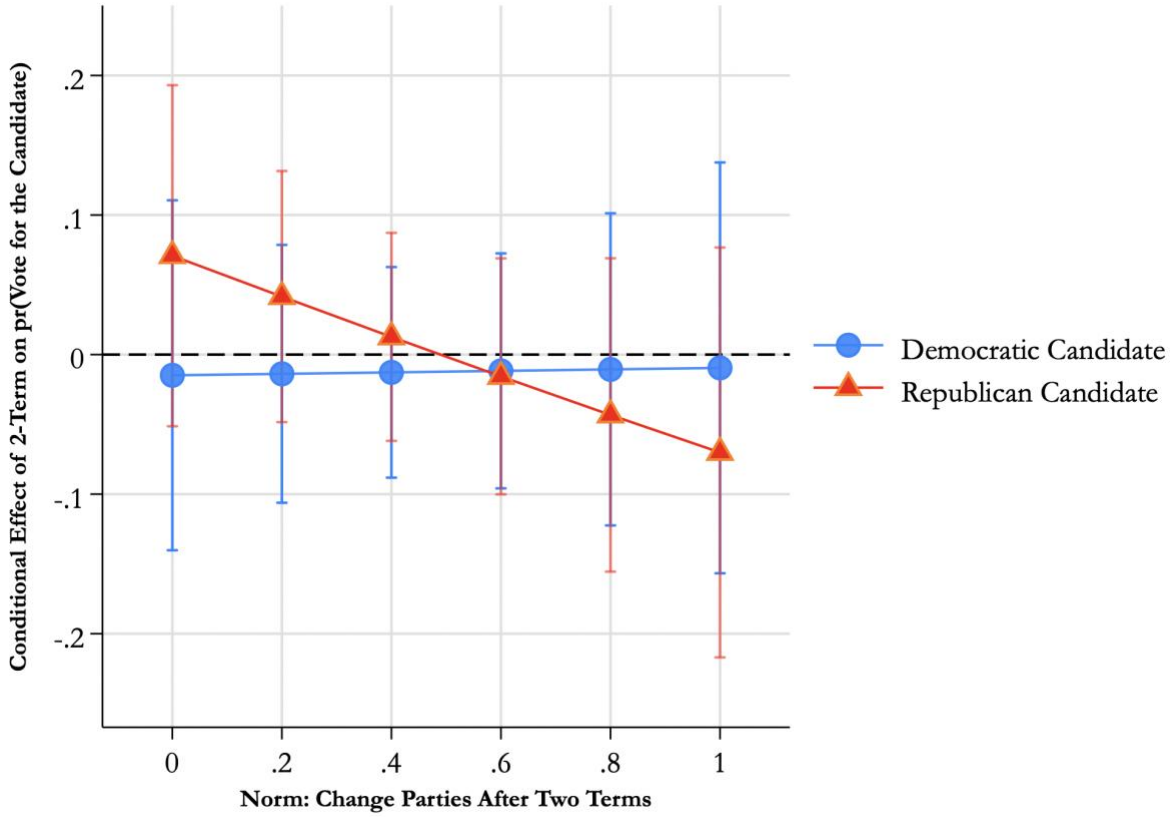


Notes: Models are logistic regression with an interaction specified between treatment assignment and the norm involving whether one-term incumbents should be re-elected. The latter is coded to range from 0 to 1 for interpretive ease, with higher values indicating greater agreement that one-term party candidates should be given a second term. 95% CIs shown. N=657 and 658 for the Democratic and Republican candidate models, respectively.

change in probability of voting for the candidate is small (about -1.4 percentage-points for the Democratic candidate and 0.6 percentage points for the Republican candidate). As such, this indicates that effects were reasonably similar across levels of normative beliefs.

Next, logistic models were run—again, one for the Democratic candidate and one for the Republican candidate—which specified an interaction between treatment assignment (2-term party incumbent versus Control) and the 2T normative belief measure. The results are presented in Figure D5. First, for the Democratic candidate, the slope is again quite flat, exhibiting a negligible estimated increase in effect size (about a half of a percentage point) from low agreement

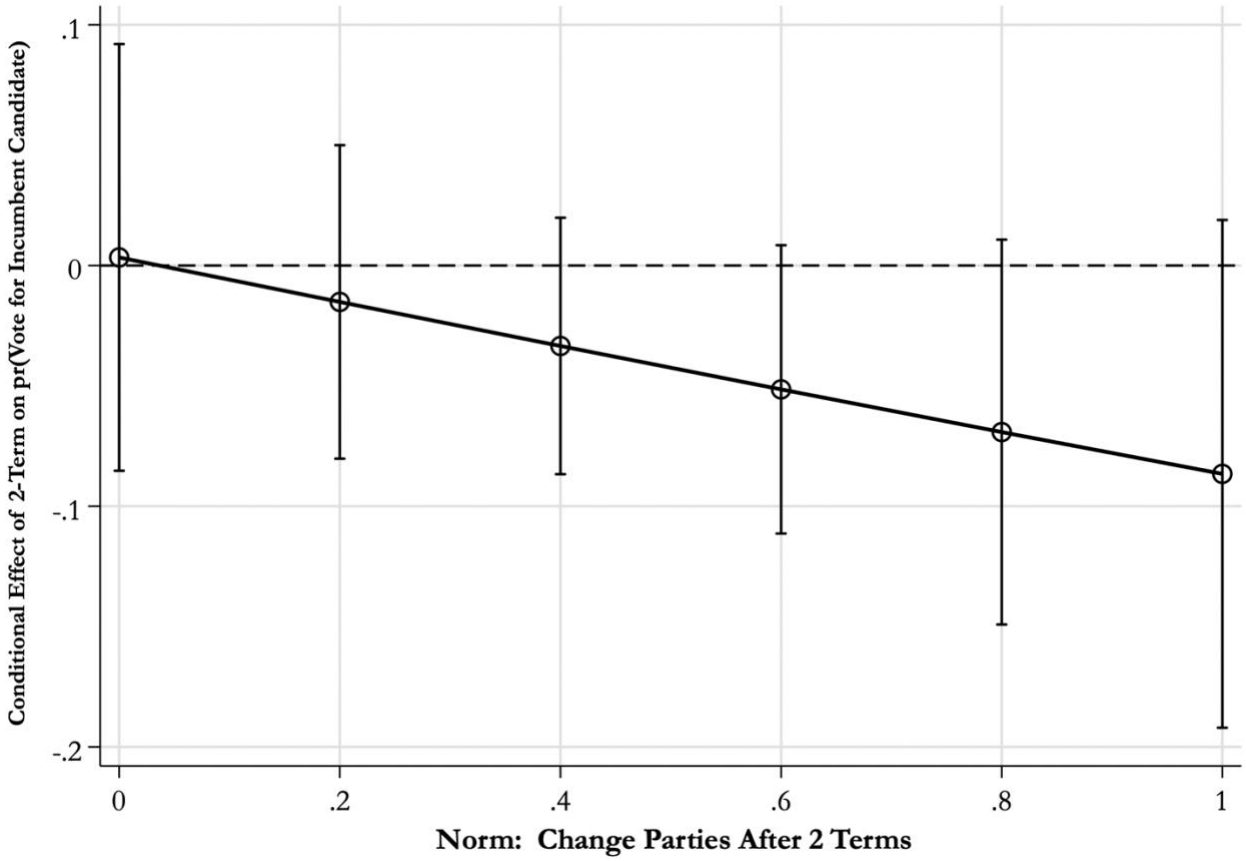
FIGURE D5. Moderating Effects of 2-Term Disadvantage Norm on 2T vs. Control Contrast



Notes: Models are logistic regression with an interaction specified between treatment assignment and the norm involving whether other party should be elected after one party has held office for two terms. The latter is coded to range from 0 to 1 for interpretive ease, with higher values indicating greater agreement that other party should be elected after incumbent party candidate has served two terms. N=676 and 665 for the Democratic and Republican candidate models, respectively.

with the 2T norm to high agreement (the interaction term p-value = .97). For the Republican candidate, however, the estimated change in effect size across the *x*-axis is far more substantial (approximately 14 percentage points). This is consistent with the notion of a two-term disadvantage: after a party's candidate has served two terms in office (versus the number of terms not being specified), their party performs poorest among those who believe that the presidency should change party hands once one party has served two terms. That said, the confidence interval for each point estimate overlaps with zero (the interaction term itself is also non-significant at conventional levels (p=.22)).

FIGURE D6. Moderating Effects of 2-Term Disadvantage Norm on 2T vs. 1T Contrast



Notes: Models are logistic regression with an interaction specified between treatment assignment and the norm involving whether other party should be elected after one party has held office for two terms. The latter is coded to range from 0 to 1 for interpretive ease, with higher values indicating greater agreement that other party should be elected after incumbent party candidate has served two terms. 95% CIs shown. N=1,344.

Finally, we can examine whether the effect of going from a one-term to a two-term incumbent party candidate on vote choice varies depending on the norm regarding two-term candidates (i.e., that the other party should have a chance to take office after one party has already served two terms). Logistic regression model results are shown in Figure D6 (these results average over the party of the incumbent). We observe that a normative belief that the presidency should change party hands after two terms tends to result in a more *negative* effect of going from a one-term party candidate to a two-term party candidate. In other words, respondents who strongly endorse this belief were even less likely to vote for the candidate whose party had already served two terms. At maximum belief in this norm, the effect is approximately an 8.6 percentage-point decrease in

the probability of voting for the incumbent party candidate. That said, the 95% confidence intervals again always overlap with 0, and the interaction term itself is non-significant (p=.27).

The overall results, therefore, suggest that the treatments operated fairly similarly regardless of pre-treatment normative attitudes toward incumbency. Although some substantively meaningful differences are observable in Figures D5 and D6, other effects seem quite homogeneous. In addition, statistically, there is too much uncertainty in the models to confidently reject the null hypothesis of homogeneous effects.

Why are the effects not more heterogeneous depending upon normative beliefs? Two points are worth highlighting. First, because the moderator (i.e., the normative belief) was not randomized, it may be correlated with other omitted factors that are also associated with the outcome. To the extent this is the case, some degree of statistical bias could be obscuring true heterogeneous effects. Second, and more importantly, it is worth stressing that since the manipulation (i.e., number of presidential terms) was isolated, little else could be driving the main effects reported in the manuscript beyond normative beliefs about incumbency in presidential elections. In other words, the narrowness of the manipulation allows for greater confidence that normative beliefs are the key mechanism. Thus, while the lack of a strong interaction here suggests that the effect was similar regardless of ostensible, general beliefs about incumbency, in the real world, wherein numerous additional types of information are present (party cues, economic indicators, candidate quality, etc.), such normative beliefs may play a much larger, more easily detectable role in determining presidential vote choice.

SUPPLEMENTAL APPENDIX E:

FACTUAL MANIPULATION CHECK RESULTS

To measure respondent inattentiveness, and confirm attentiveness to the manipulation itself, a factual manipulation check (FMC) was asked after the outcome measure (see Kane and Barabas 2019). Overall, 71.29% of the sample answered the FMC (shown below) correctly. While this constitutes a clear majority, it is nevertheless indicative of a substantial amount of inattentiveness to the experiment.

To confirm that the manipulation was attended to, a chi-squared test was performed. This test examined whether responses to the FMC significantly covaried with treatment assignment. As shown in Table E1, this is clearly the case: within each experimental condition, between 65% and 80% of respondents answered the FMC in a way that indicates they were significantly attentive to the key manipulated details of the experiment ($p<.001$).

Factual Manipulation Check

Which of the following statements accurately reflects the scenario you just read about?

- ○ There was no mention of whether either candidate had been president before (1)
- ○ The Democratic candidate was the incumbent president (2)
- ○ The Republican candidate was the incumbent president (3)
- ○ The Democratic Party had held the presidency for the previous 8 years (two terms) (4)
- ○ The Republican Party had held the presidency for the previous 8 years (two terms) (5)

TABLE E1. Responses to FMC Significantly Covary with Treatment Assignment

Response to FMC	**Experimental Condition**					
	Control	*1TD*	*1TR*	*2TD*	*2TR*	*Total*
Control	**80.53**	13.41	12.39	16.57	13.78	27.39
1Term Democrat (1TD)	5.35	**70.64**	7.10	7.61	4.30	19.01
1Term Republican (1TR)	4.16	7.00	**72.81**	3.88	7.11	18.83
2 Term Democrat (2TD)	5.35	6.26	4.08	**67.31**	9.63	18.53
2 Term Republican (2TR)	4.61	2.68	3.63	4.63	**65.19**	16.23
Total	100.00	100.00	100.00	100.00	100.00	100.00

Column percentages are shown. The share answering correctly in each condition appears in bold text. Pearson Chi2 = 5659.59. $p<.001$

SUPPLEMENTAL APPENDIX F:

ESTIMATING TREATMENT EFFECTS ACROSS LEVELS OF ATTENTIVENESS

Given the possibility of substantial respondent inattentiveness, respondents were shown—at the start of the survey and prior to the experiment—a mock vignette (shown below) as well as several mock vignette checks (MVCs) afterward (see Kane, Velez and Barabas 2023). The first MVC was used to exclude respondents from continuing the survey. The following two MVCs (shown below) were used to construct an additive scale, with higher scores indicating higher respondent attentiveness.

Per the findings of Kane et al. (2023), the attentiveness scale was interacted with the treatment condition to examine whether treatment effects were stronger among those who (likely) paid greater attention to the experimental vignette and outcome measures. Figure F1, which reexamines the results of Figure 2 in the manuscript, demonstrates this to be the case: those who were attentive show noticeably larger effects, with those who were least attentive (i.e., answered the two MVCs above incorrectly) showing near-zero effects. Specifically, the results from the manuscript's left panel of Figure 2 are re-estimated to be 12.7 and 9.1 percentage points for the Democratic and Republican candidates, respectively ($p<.05$ in both cases). The treatment effect in the right panel of Figure 2 is re-estimated to be 7.1 percentage points ($p<.05$).

Figure F2 includes the analyses featured in Figure 3 of the manuscript, with comparisons between 2T candidates and the Control condition. Though there is a tendency for the more attentive to exhibit slightly more positive effects (see dotted lines), averaging across the Republican and Democratic candidates, these changes in estimated effects are relatively flat compared to the changes observed in Figure F1 and not statistically distinguishable from zero.

Mock Vignette

A Passage from a Recent Magazine Article:
More than 125 scientific societies and journal publishers are urgently warning lawmakers not to move forward with a rumored policy that would make all research supported by federal funding immediately free to the public. In three separate letters, they argue such a move would be costly, could bankrupt many scientific societies that rely on income from journal subscriptions, and would harm the scientific enterprise. Lawmakers won't comment on whether they are considering a

policy that would change publishing rules, and society officials say they have learned no details. But if the rumor is accurate, the order would represent a major change from current U.S. policy, which allows publishers to withhold federally-funded research from the general public for up to 1 year.

Mock Vignette Checks

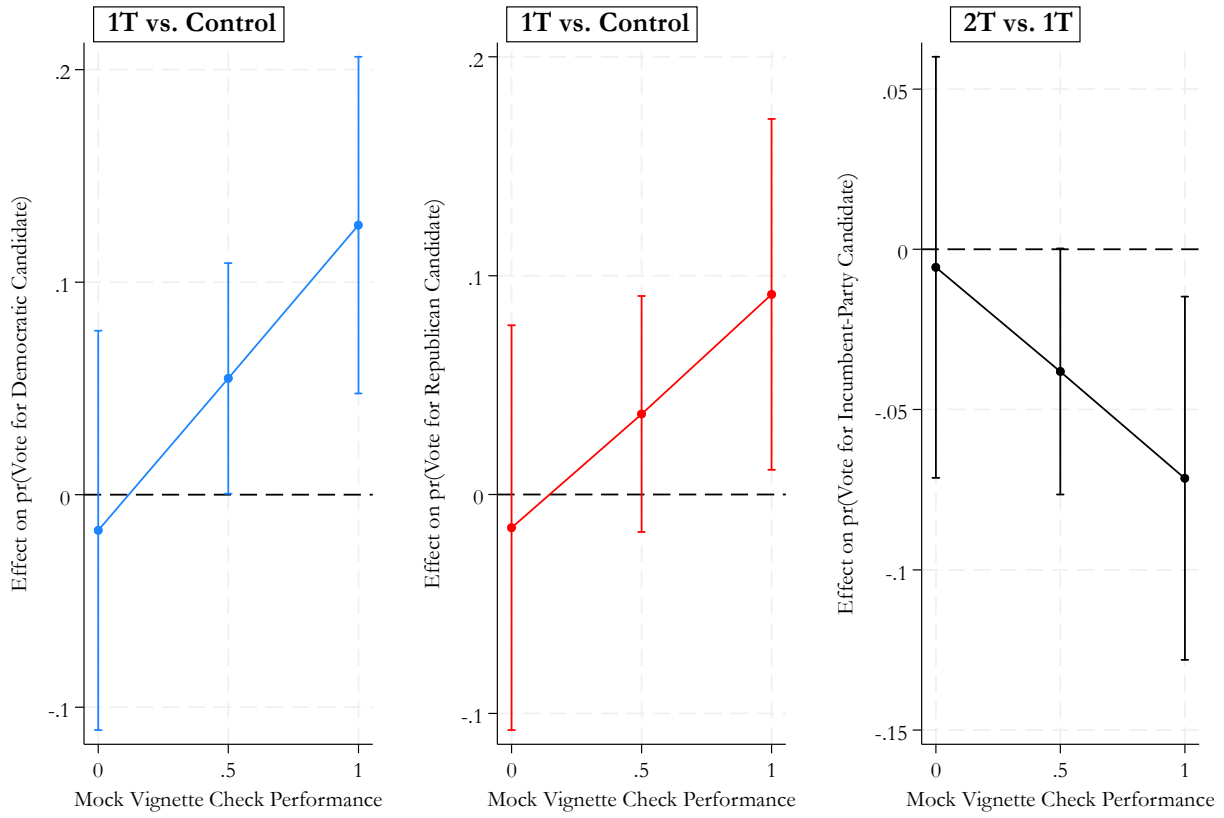
Regarding the rumored change in policy that was discussed, the magazine excerpt indicated that:

- ○ Lawmakers won't comment on whether they are considering it (1)
- ○ Legal scholars stated the change in policy would be challenged in courts (2)
- ○ Journal publishers have already begun preparing for the change in policy (3)
- ○ Scientific researchers are divided in terms of their support for the policy (4)
- ○ All of the above (5)
- ○ None of the above (6)

According to the magazine excerpt you just read, current policy allows federally-funded research to be withheld from the general public for up to:

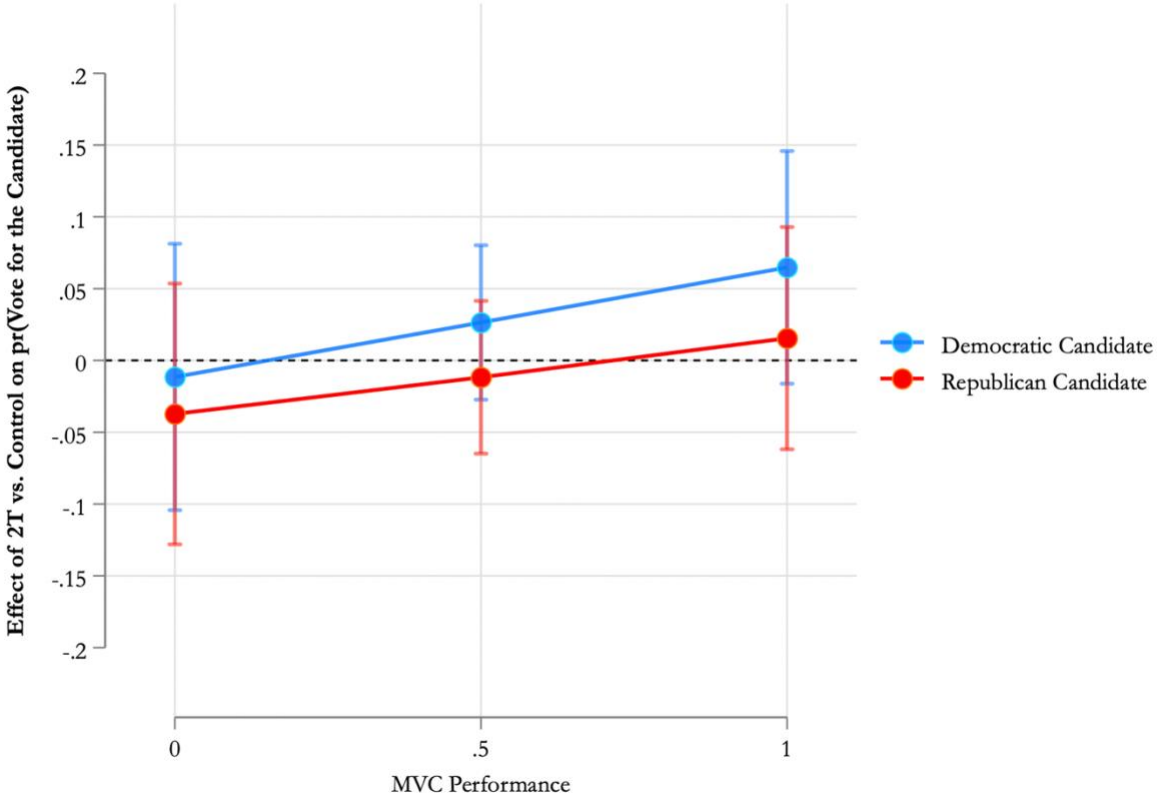
- ○ 1 month (1)
- ○ 6 months (2)
- ○ 1 year (3)
- ○ 3 years (4)
- ○ 5 years (5)
- ○ None of the above (6)

FIGURE F1. Larger Effects Among the More Attentive (see MS Figure 2)



Notes: Models are logistic regression with an interaction specified between treatment assignment and the additive attentiveness scale. Vertical axis displays effect of treatment on predicted probability of voting for a particular candidate. 95% CIs shown.

FIGURE F2. No Evidence of a Two-Term Disadvantage Across Levels of Attentiveness



Notes: Models are logistic regression with an interaction specified between treatment assignment and the additive attentiveness scale. Vertical axis displays effect of two-term treatment (vs. Control) on predicted probability of voting for a particular candidate. 95% CIs shown.

SUPPLEMENTAL APPENDIX G:

SURVEY ITEMS & REGRESSION MODEL OUTPUT (VSG DATA)

The manuscript features analyses using Voter Study Group panel data. The dataset is the "September-November 2020 Voter Full Data Set", which features survey data on the same set of respondents between 2011 (the baseline year) and 2020 (respondents in 2011 were asked about their vote choice in 2008). The data are available at: https://www.voterstudygroup.org/data. Table G1 describes each variable used in the analysis, as well as its corresponding survey item in the VSG data set and details regarding any recoding of the original variable. All continuous variables were recoded to range between 0 and 1 for the analyses featured in the manuscript. Table G2 features the regression output underlying Figure 4 in the manuscript.

TABLE G1. VSG Survey Items and Coding Details

Description	Survey Item	Coding Details
Vote choice in 2008	*presvote08_2011*	Coded as voting for Obama (1) or not (0). Non-voters excluded.
Vote choice in 2012	*presvote_2012*	Coded as voting for Obama (1) or not (0). Non-voters excluded.
Vote choice in 2016	*presvote_2016*	Coded as voting for Trump (1) or not (0). Non-voters excluded.
Vote choice in 2020	*presvote_2020Nov*	Coded as voting for Trump (1) or not (0). Non-voters excluded.
Respondent's party (baseline year)	*pid7_2011*	Leaners coded as partisans. "Not sure" coded as independents. "Skipped" and "Not asked" excluded.
Political interest (baseline year)	*polinterest_2011*	Reverse coded (so that higher values indicate greater interest). "Not sure" "Skipped" and "Not asked" excluded.
Gender (baseline year)	*gender_2011*	Males coded as 0, females coded as 1. "Skipped" and "Not asked" excluded.
Racial identification (baseline year)	*race_2011*	Because they constituted less than a combined 5% of the total sample, Asians, Native Americans, "Mixed", "Other" and Middle Eastern were coded as "Other". "Skipped" and "Not Asked" excluded.
Age (baseline year)	*birthyr_2011*	Subtracted birth year from 2011. "Skipped" and "Not asked" excluded.
Family income (baseline year)	*faminc_2011*	The value of 31 ("150,000 or more") was changed to 12 to create a continuous variable. "Prefer not to say", "Skipped" and "Not asked" excluded.
Highest level of education completed	*educ_2011*	"Skipped" and "Not asked" excluded.

TABLE G2. Regression Output Underlying Manuscript Figure 4

	Vote for 1-Term Incumbents	Vote Against 2-Term Incumbents
Democrat	0.01	-0.97***
	(0.32)	(0.22)
Republican	-1.02**	-1.32***
	(0.38)	(0.24)
Political Interest	-0.84*	-0.63*
	(0.35)	(0.30)
Female	0.56*	-0.08
	(0.25)	(0.19)
Black	-1.11^	-1.03*
	(0.60)	(0.52)
Hispanic	0.37	0.26
	(0.44)	(0.36)
Other Race	0.14	-0.30
	(0.45)	(0.41)
Age	-0.81	1.23*
	(0.70)	(0.58)
Family Income	-0.41	-0.31
	(0.46)	(0.36)
Educational Attainment	-1.56**	-0.87*
	(0.48)	(0.36)
Constant	-1.43**	-1.36**
	(0.55)	(0.44)
Observations	2,527	2,527

*Notes: These are the results underlying manuscript Figure 4. Logistic regression model coefficients shown with standard errors in parentheses. *** p<0.001, ** p<0.01, * p<0.05, ^ p<0.10 (two-tailed hypothesis tests).*