

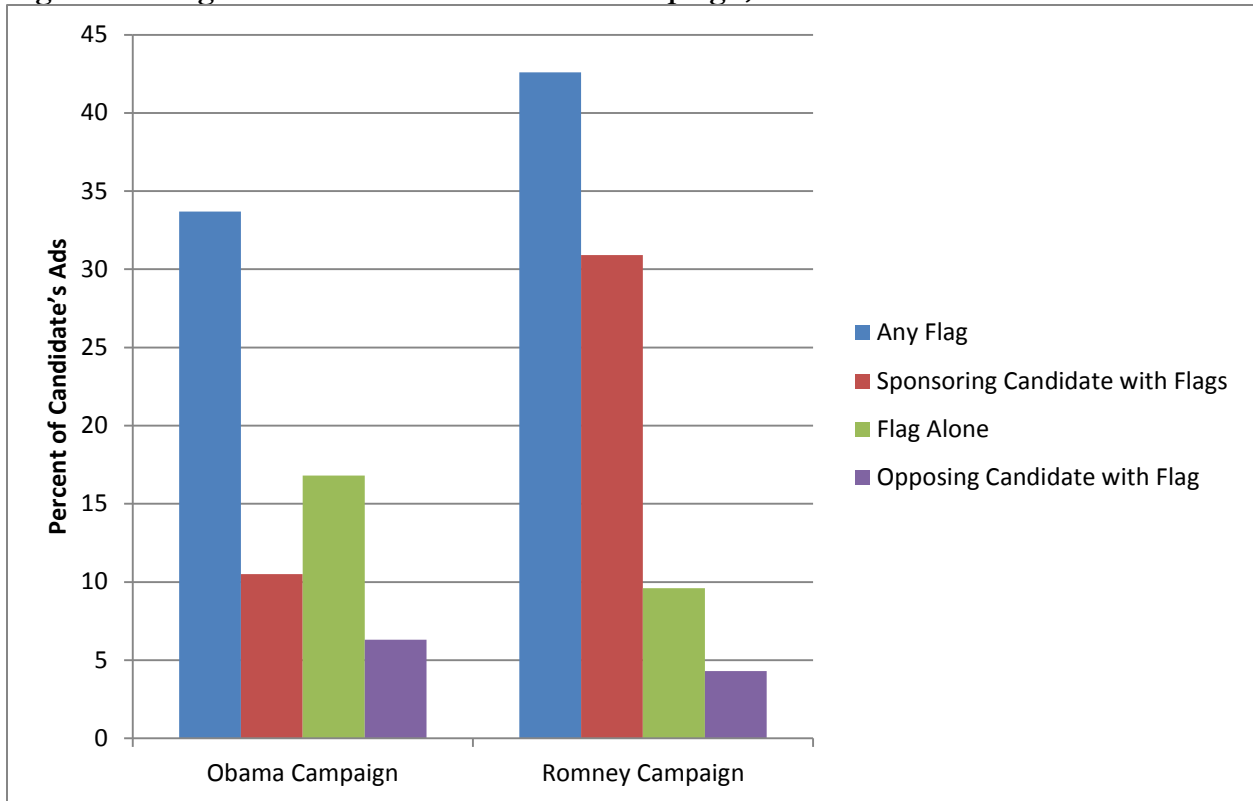
Online Supplementary Materials

Appendix 1: Flags in the 2012 Election Campaign

How do practitioners use flags in real world political campaigns? Are flags as ubiquitous as many assume? Does flag use vary substantially and meaningfully, or does it appear uniformly across campaigns? To answer these questions and to provide context for our experimental analysis, we collected exhaustive content analysis data on flag imagery in all 2012 presidential television ads and ad-like online videos produced by both major-party candidates and their “Super-PACs.”

Both campaigns and their Super-PACs posted all TV and web ads on YouTube. We coded ads from Crossroads GPS, Restore our Future, Americans for Prosperity, American Crossroads, and Priorities USA as well as the ads and online video content produced by the presidential campaigns themselves. A total of 256 television ads and 693 online ads were coded. We coded for the presence of any flag in an ad, for sponsoring candidates shown with the flag (subset of “any flag”), opponents shown with the flag (subset of “any flag”), flags without the candidates (subset of “any flag”) and separately for the presence of the candidates with a flag lapel pin. In addition, ads were coded for their topical focus, type, and whether they were negative, positive or contrast ads. A second coder independently coded a 10% sample of the ads to calculate intercoder reliability. Coders agreed on the presence of at least one scene with a flag in the ad 88% of the time (Scott’s $\pi = .78$); on the presence of the flag with the sponsoring candidate 98% of the time (Scott’s $\pi = .89$), and on the presence of a flag pin in the ad 92% of the time (Scott’s $\pi = .69$). It is worth noting that, like Geer (2005), we are looking at ads produced, unweighted by television airings or views. The data to weight by airings in 2012 is not yet available.

Figure A1: Flag Use in the 2012 Presidential Campaign, Television Ads



Note: This represents the percent of ads which have at least one scene with flag images across all television ads from the two presidential campaigns. The “any flag” category includes ads which have at least one scene where a flag appears with sponsoring candidate, a flag appears with the opposing candidate or the flag appears alone. Whether a candidate is shown with a flag lapel pin is treated as a separate category and not included as a flag image. Obama campaign ads (n=95), Romney campaign ads (n=94)

The Romney campaign was more likely to use flag imagery in their television advertisements – 43% of Romney ads versus 34% of Obama ads (Figure A1) – and they were three times more likely to show the sponsoring candidate with a flag in those ads.¹ However, the difference is mostly driven by the distinct ratios of positive, negative, and contrast ads between the campaigns as noted in the text.

¹ In fact, as Figure A1 shows, a greater percentage of Obama ads include scenes of a flag alone than scenes of Obama with the flag. Another source of flag imagery is flag pins, which are fairly common in ads, particularly on the sponsoring candidate. We did not include flag pins in our count of flags, but they do provide an additional flag cue. Obama’s TV ads featured Obama wearing a flag pin in 20% of ads, scenes of Romney wearing a flag pin in 11% of ads. Romney’s TV ads showed at least one scene with Romney wearing a flag pin in 30% of ads, Obama in a flag pin in 6% of ads.

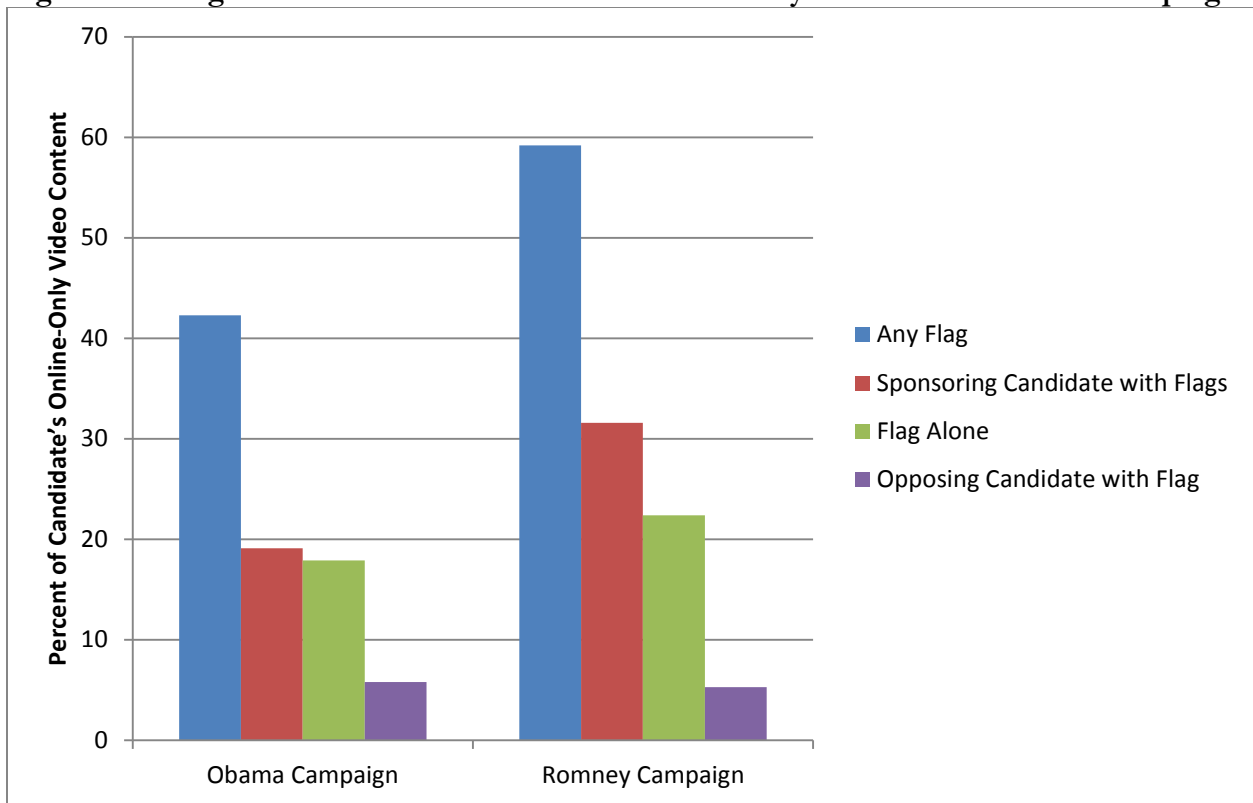
The Obama campaign ads were more likely to be negative (55%), and flags were far less common in negative advertising. Romney was more likely to produce contrast ads (51%), which also meant that he was more likely to appear in his ads. Both candidates had similar proportions of positive ads (20-21%). After accounting for tone, both candidates used the flag in similar proportions: the only difference that emerges is within positive ads where Romney was more likely to feature flags than Obama. Thus, the flag difference between the two campaigns was largely determined by the strategic choice between producing negative and contrast ads.

While candidates made substantial use of flag images, television ads by “Super-PACs” had less flag imagery. Only 15 ads (22% of Super-PAC ads) included any flag scene. Again, this is best explained by predominantly negative tones (90% of all Super-PAC ads were coded as negative). Turning to online-only video content, a relatively new phenomenon in 2012, we find a similar pattern to what we found in television ads. Romney’s campaign used flag imagery in more of their online content, and those ads were more likely to show him with the flag (Figure A2).²

In sum, flag use in the 2012 presidential campaigns appears to be driven mainly by strategic decisions about the tone of the television advertising campaign. While Obama does not avoid appearing with flags, Romney uses more flags and is more likely to appear with flags because Romney ran fewer negative ads.

² The two campaigns seem to have had different approaches to online video content. We limited our analysis of online content to videos less than five minutes in length for comparability. First, and most striking, the Obama campaign had many more videos. The Romney online videos were more issue oriented pieces, clips from campaign stops and testimonials such as the series from small business owners “We Built This”. The Obama online videos had issue and campaign clips, but also included video produced by particular field offices, celebrity endorsements, GOTV videos and informational videos explaining the campaign’s online volunteer tools. So while campaign TV ads are similar (in number, length, and purpose), the online video population may be less comparable.

Figure A2. Flag Use in Online Video Content Produced by the 2012 Presidential Campaigns



Note: This represents the percent of ads that have at least one scene with a flag image across all online ads from each of the two presidential campaigns. We restrict the analysis to online ads less than 5 minutes in length. Whether the candidate is shown with a flag lapel pin is treated as a separate category and not included as a flag image. “Any flag” includes ads where there are scenes where sponsoring candidate appears with a flag, opposing candidate appears with a flag or flag appears alone. Obama campaign ads (n=530), Romney campaign ads (n=76).

Appendix 2: Experimental Studies Sample Characteristics and Design

Table A1: Social & Demographic Attributes of the Samples & the Population

	2012 MTurk	2012 College	2013 MTurk	2010 National Pop.
Edu College Degree	47%	--	46%	30%
Some College	43%	--	44%	28%
H.S. Grad	13%	--	9%	31%
Age Median	18-29	--	18-29	45 to 49
% over 40	19%	--	18%	61%
Family Income Median	32-57k	--	32-57k	\$52,762
Female	45%	65%	40%	51%
White	79%	73%	75%	63%

Note: 2012 MTurk and 2012 College samples reflect wave 2 traits; 2013 MTurk is from wave 2 and included a Republican over-sample through selective invitations to complete wave 2.

Table A2: Political Attributes of the Samples & the 2004ANES

	2012 MTurk	2012 College	2013 MTurk	2004 ANES
Flag Feel (0 to 1)	.48 (.29)	.55 (.27)	.49 (.30)	.76 (.28)
Ashamed (0 to 1)	.26 (.23)	.21 (.19)	.22 (.22)	.39 (.46)
Patriotic (0 to 1)	.45 (.27)	.53 (.27)	.48 (.26)	--
Patriot Index (0 to 1)	.39 (.22)	.43 (.20)	.47 (.17)	.58 (.30)
White-Black (-1 to 1)	.08 (.20)	.06 (.20)	.10 (.24)	.01 (.19)
PID (-1 to 1)	-.25 (.50)	-.34 (.53)	-.13 (.55)	-.04 (.69)

Note: Cells show mean, standard deviation in parentheses. 2012 MTurk and 2012 College samples reflect wave 2 traits; 2013 MTurk is from wave 2 and included a Republican over-sample through selective invitations to complete wave 2. 2004 American National Election Study.

Figure A3: Experimental Treatment Conditions and Images

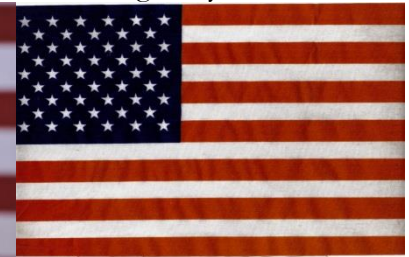
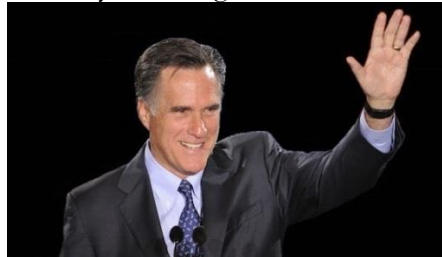
Experimental Treatment Conditions – Studies 1 & 2

Romney: No Flag Obama: No Flag	Romney: Flag Obama: No Flag	Romney: No Flag Obama: Flag	Romney: Flag Obama: Flag	Flag Only	No Images
16.7%	16.7%	16.7%	16.7%	16.7%	16.7%

Romney: No Flag

Romney: Flag

Flag Only



Obama: No Flag

Obama: Flag



Experimental Treatment Conditions – Study 3

Ryan: No Flag Biden: No Flag	Ryan: Flag Biden: No Flag	Ryan: No Flag Biden: Flag	Ryan: Flag Biden: Flag	Flag Only	No Images
16.7%	16.7%	16.7%	16.7%	16.7%	16.7%

Ryan: No Flag

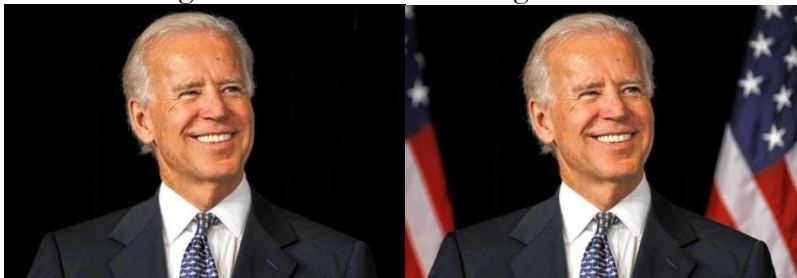
Ryan: Flag

Flag Only



Biden: No Flag

Biden: Flag



Note. Subjects were assigned to one of six experimental treatment conditions. The four main treatment conditions vary the presence or absence of the flag with an image of the candidate. Two additional conditions were a flag image only condition and a no image condition. Images are displayed individually on sequential survey pages accompanying a question asking respondents whether they “have a favorable or unfavorable impression” of evaluation of the candidate (on a 5 point favorability scale. Thus the treatment is exposure to two survey pages taken together. We randomize the order of presentation of the candidates. The no images conditions include only the evaluation questions. The flag only conditions have the flag image on both candidate evaluation pages. Our key dependent variable is asked after respondents have been exposed to all images in a treatment condition. Each photo was displayed at ~175 x 110 pixels.

Although the flag images differ across candidates, they are paired with the identical image without the flag, perfectly balancing the other in the pairing. This would be a problem if we asserted different effects for each candidate. However, our claim is the opposite – that all flag effects are effectively equivalent – so the variability in flag images actually strengthens the external validity of our results and does not weaken the internal validity of our comparisons.

Appendix 3: Measurement Characteristics

Patriotism:

We measured *patriotism* with an index of three common survey items from the American National Election Studies and other sources: self-described patriotism, feelings when seeing the flag flying, and whether some things make the respondent feel ashamed about America, each on a 5-point response scale. These items created fairly reliable indices of symbolic patriotism (College $\alpha = .69$; MT¹² $\alpha = .77$; MT¹³ $\alpha = .69$). Each ranges from 0 to 1 with 1 representing the most patriotism (College: $M = .43$, $SD = .20$; MT¹²: $M = .39$, $SD = .22$; MT¹³: $M = .47$, $SD = .17$). These measurement properties and those that follow are for subjects participating in the second wave of the studies.

Prejudice:

We measured *prejudice* as the tendency of respondents to derogate African Americans relative to whites on feeling thermometers ranging from 0 to 100. We construct a difference score of prejudice by subtracting respondents' rating for blacks from their rating for whites, rescaled ranging from -1 to +1 (College: $M = .06$, $SD = .19$; MT¹²: $M = .08$, $SD = .20$; MT¹³: $M = .10$, $SD = .24$). The correlation between patriotism and prejudice is moderately positive in each study (College $r = .23$, MT¹² $r = .12$, MT¹³ $r = .24$).

Partisanship:

The *partisanship* item combines the traditional 7-point branching measure into a single question, coded -1 (Strong Democrat) to +1 (Strong Republican) (College: $M = -.34$, $SD = .53$; MT¹²: $M = -.25$, $SD = .50$; MT¹³: $M = -.13$, $SD = .55$). The correlation between partisanship and patriotism is moderate (College $r = .38$, MT¹² $r = .33$, MT¹³ $r = .31$), as is the correlation between partisanship and prejudice (College $r = .24$, MT¹² $r = .23$, MT¹³ $r = .28$). These relationships are consistent in size and direction with similar comparisons in the 2004 American National Election Studies (PID-patriot = .34, PID-prejudice = .15, patriot-prejudice = .19), providing more evidence of external validity.

Correlations between independent variables:

The correlation between patriotism and prejudice is moderately positive in each study (College $r = .23$, MT¹² $r = .12$, MT¹³ $r = .24$). The correlation between partisanship and patriotism is moderate (College $r = .38$, MT¹² $r = .33$, MT¹³ $r = .31$), as is the correlation between partisanship and prejudice (College $r = .24$, MT¹² $r = .23$, MT¹³ $r = .28$). These relationships are consistent in size and direction with similar

comparisons in the 2004 American National Election Studies (PID-patriot=.34, PID-prejudice=.15, patriot-prejudice=.19), providing more evidence of external validity.

Presidential Vote Choice:

Our primary outcome is vote choice in the upcoming election. We asked subjects about their likelihood of voting and their intended presidential vote choice. The display logic for vote choice was limited to subjects who said they were “Definitely,” “Probably,” or “Maybe” going to vote. Subjects were asked, “If you vote, who do you think you will vote for in the election for President?” Subjects could choose “Barack Obama,” “Mitt Romney,” “Other candidate,” or “Don’t know.” The 2013 study substituted “Joe Biden” and “Paul Ryan” and the hypothetical, “If Paul Ryan and Joe Biden are the Republican and Democratic nominees for President in 2016...”). Those who selected a candidate were asked a follow-up to indicate preference strength on a 3-point scale: “Very strong,” “Somewhat strong,” “Not very strong.” Subjects who indicated “Don’t know” for candidate choice were asked, “If you had to choose today, which candidate would you vote for?” These questions were assembled to form a 9-point vote choice scale ranging from very strong Republican preference (+1) to very strong Democratic preference (0), with the few third party choices excluded (College: $M=.20$, $SD=.34$; MT¹²: $M=.25$, $SD=.34$; MT¹³: $M=.41$, $SD=.38$). In 2012, Romney won 47% of the popular vote. Our subjects are more Democratic than the general electorate, which might be expected to provide a hard test of any Republican advantage hypothesis. When asked their vote preference, in Study 1, 17% chose Romney, and 21% chose Romney in Study 2. In Study 3, which included the extra effort to recruit Republicans, 40% said they would vote for Ryan.

Appendix 4: Additional Tests Isolating Democrat-Only Flag Conditions

Net treatment effects: Bivariate OLS models comparing the Democrat-only flag condition to the two non-flag treatments combined reveal no evidence of a Democratic benefit (College: $b=.07, s.e.=.07, p=.29$; MT'12: $b=.03, s.e.=.04, p=.49$; MT'13: $b=.04, s.e.=.04, p=.28$; Pooled: $b=.04, s.e.=.03, p=.15$). In fact, if anything, Democratic flag use consistently benefits the Republican candidate, with a pooled point estimate *identical* to the marginally-significant result for all flag exposure.

Table A3: Democratic Candidate Flag Effects on Presidential Vote by Patriotism

Study	Presidential Vote (Republican)			
	College	MT'12	MT'13	Pooled
Patriotism	.32 (.23)	.56* (.10)	.62* (.11)	.63* (.07)
Dem Flag Only	-.20 (.16)	-.09 (.08)	-.19^ (.12)	-.13* (.07)
Dem Flag Only* Patriotism	.57 (.35)	.30 (.18)	.48* (.23)	.37* (.13)
Constant	.02 (.10)	.00 (.05)	.08 (.06)	.02 (.04)
R ²	.15	.18	.14	.17
N	85	287	389	761

Note. Unstandardized OLS coefficients with standard errors in parentheses. Excluded category: any treatment without flags. * $p < .05$ ^ $p < .10$, two-sided.

Table A4: Democratic Candidate Flag Effects on Presidential Vote by Prejudice

Study	Presidential Vote (Republican)			
	College	MT'12	MT'13	Pooled
Prejudice	.27 (.22)	.10 (.12)	.36* (.09)	.30* (.07)
Dem Flag Only	.03 (.07)	-.01 (.05)	.03 (.04)	.01 (.03)
Dem Flag Only* Prejudice	.43 (.37)	.43* (.20)	.11 (.17)	.23^ (.12)
Constant	.14* (.04)	.23* (.02)	.35* (.02)	.28* (.02)
R ²	.09	.04	.06	.06
N	85	287	389	761

Note. Unstandardized OLS coefficients with standard errors in parentheses. Excluded category: any treatment without flags. * $p < .05$ ^ $p < .10$, two-sided.

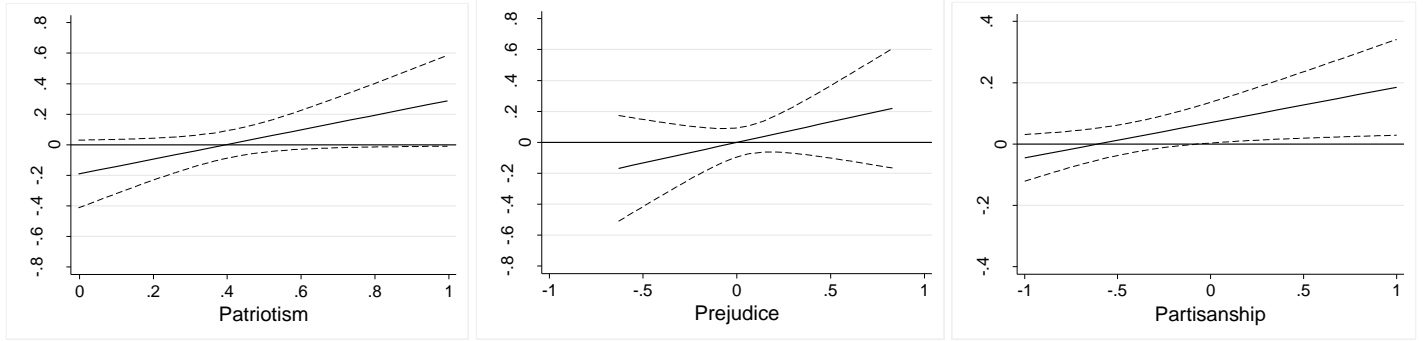
Table A5: Democratic Candidate Flag Effects on Presidential Vote by Partisanship

Study	Presidential Vote (Republican)			
	College	MT'12	MT'13	Pooled
Partisanship	.49* (.05)	.45* (.03)	.48* (.03)	.48* (.02)
Dem Flag Only	.06 (.05)	.01 (.03)	-.04 (.03)	.03 (.02)
Dem Flag Only* Partisanship	.13^ (.07)	.06 (.06)	.04 (.05)	.06^ (.03)
Constant	.37* (.03)	.39* (.02)	.45* (.02)	.42* (.02)
R ²	.74	.52	.54	.57
N	85	286	389	760

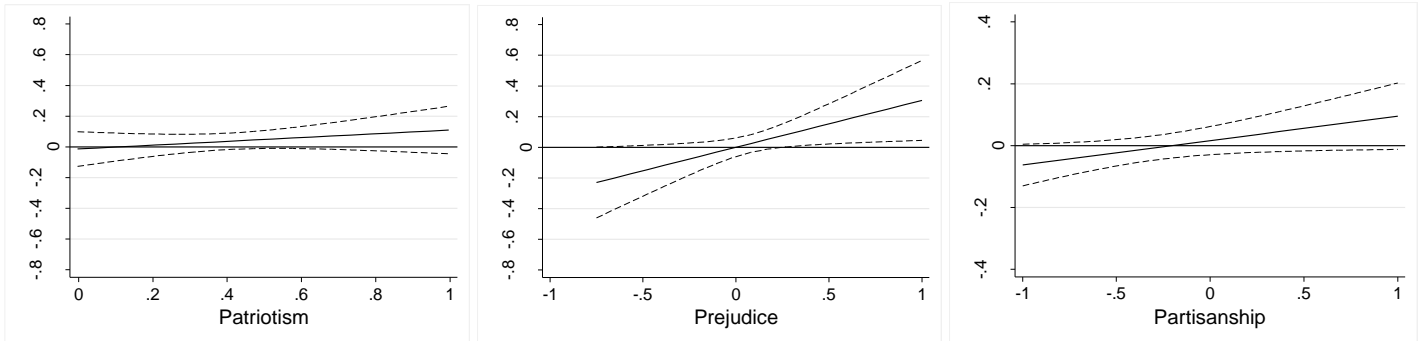
Note. Unstandardized OLS coefficients with standard errors in parentheses. Excluded category: any treatment without flags. * $p < .05$ ^ $p < .10$, two-sided.

These interactive models focusing just on flag use by the Democratic candidate show the same results as the full models of flag effects. Thus, Democrats using the flag appear to aid Republican candidates just as much as Republican candidates with the flag do and the flag alone (without candidates) does.

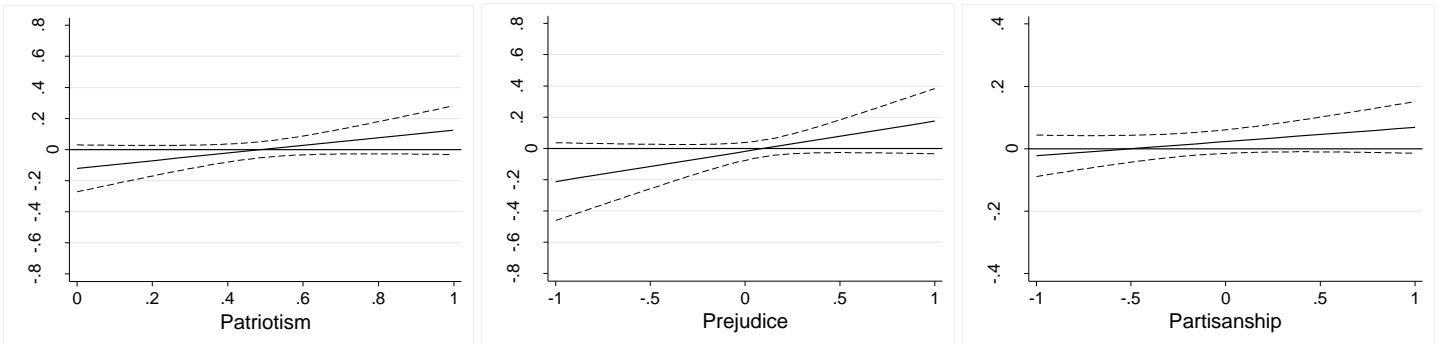
Figure A4: Effects of a Democrat Appearing with Flag on Vote Choice
College 2012



Mechanical Turk 2012



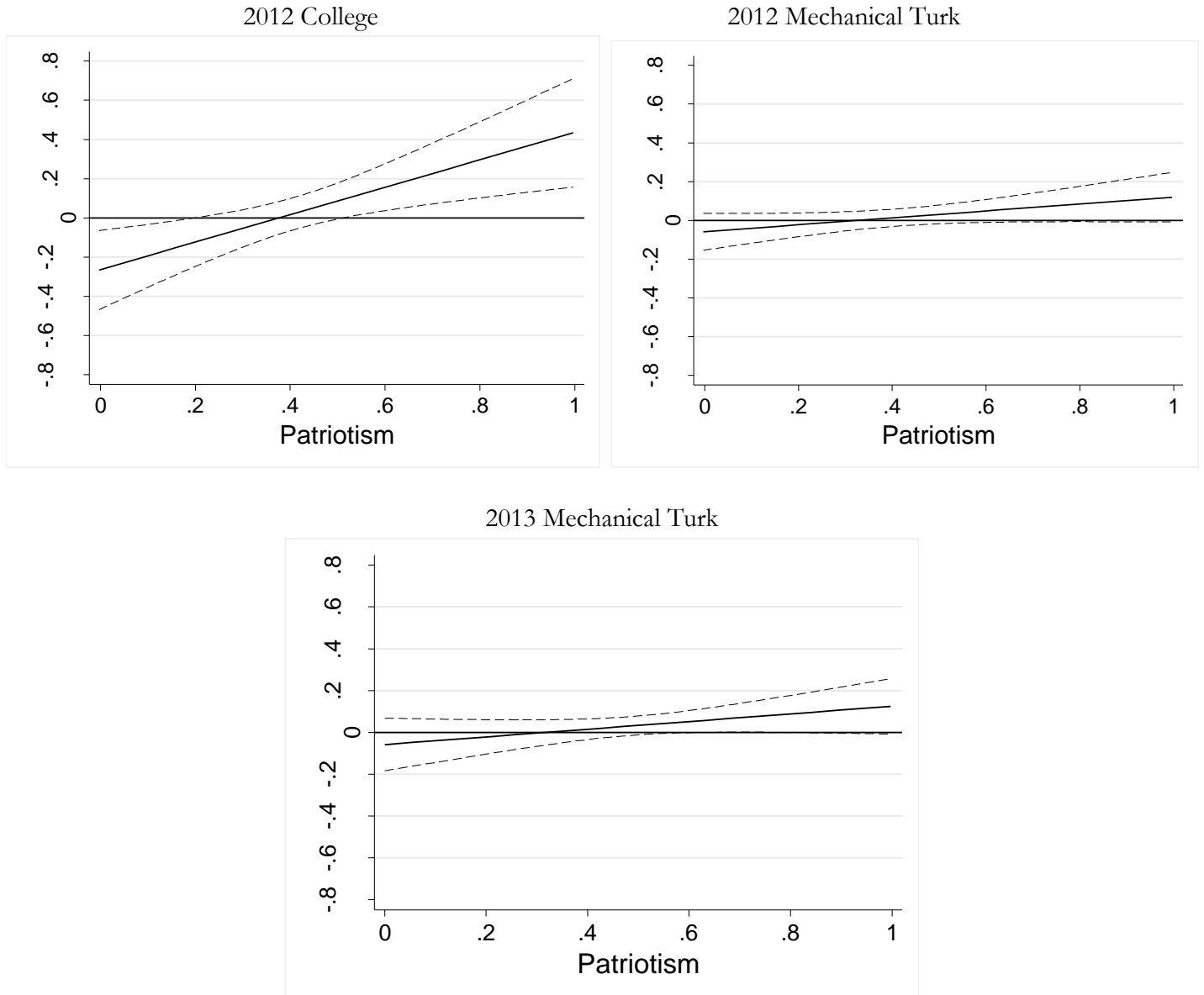
Mechanical Turk 2013



Note: Dashed lines indicate 90% confidence intervals, two-sided. This represents the marginal treatment effect of the Democrat-only flag, relative to seeing a treatment with no flags.

Appendix 5: Additional Analyses, Marginal Effects by Study, Flag Effects on Vote Choice by Predisposition, Non-Linear Estimates

Figure A5: Marginal Flag Effects on Presidential Vote Choice by Patriotism

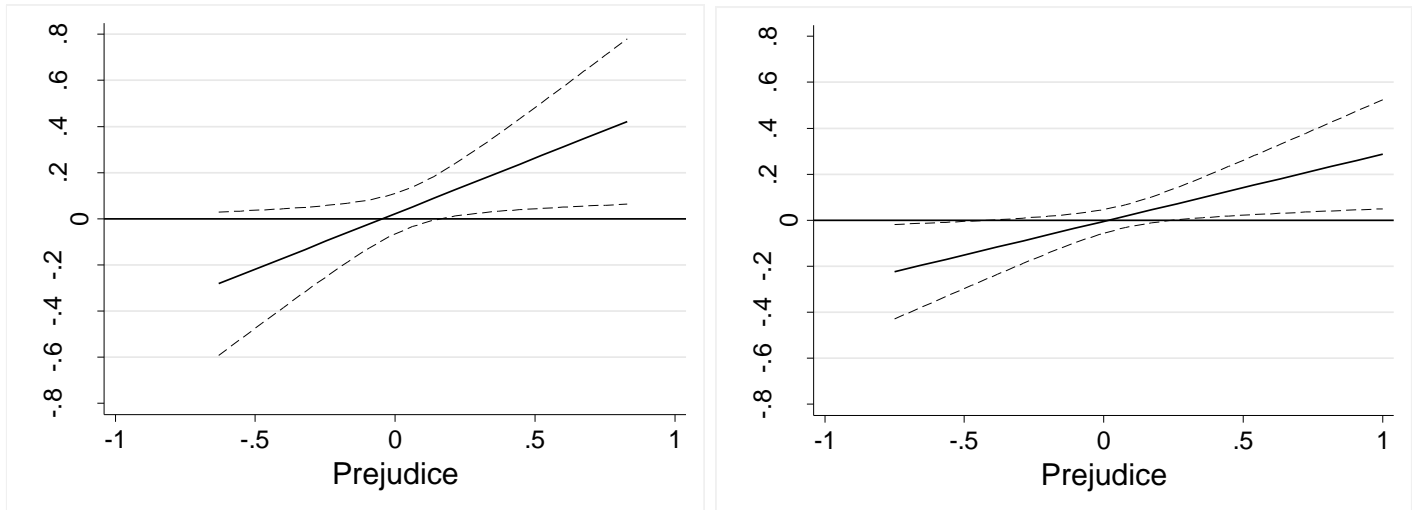


Note: Dashed lines indicate 90% confidence intervals, two-sided. Marginal effect of any flag relative to treatments with no flags.

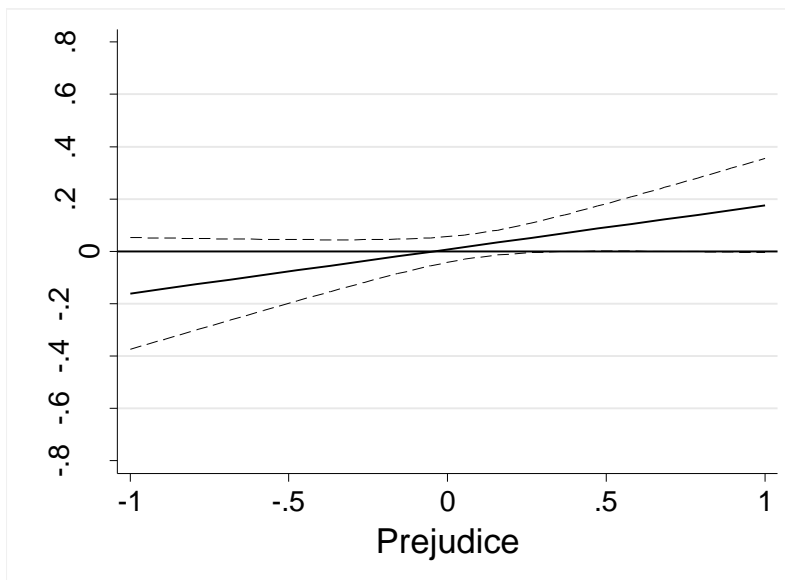
Figure A6: Marginal Flag Effects on Presidential Vote Choice by Prejudice

2012 College

2012 Mechanical Turk



2013 Mechanical Turk

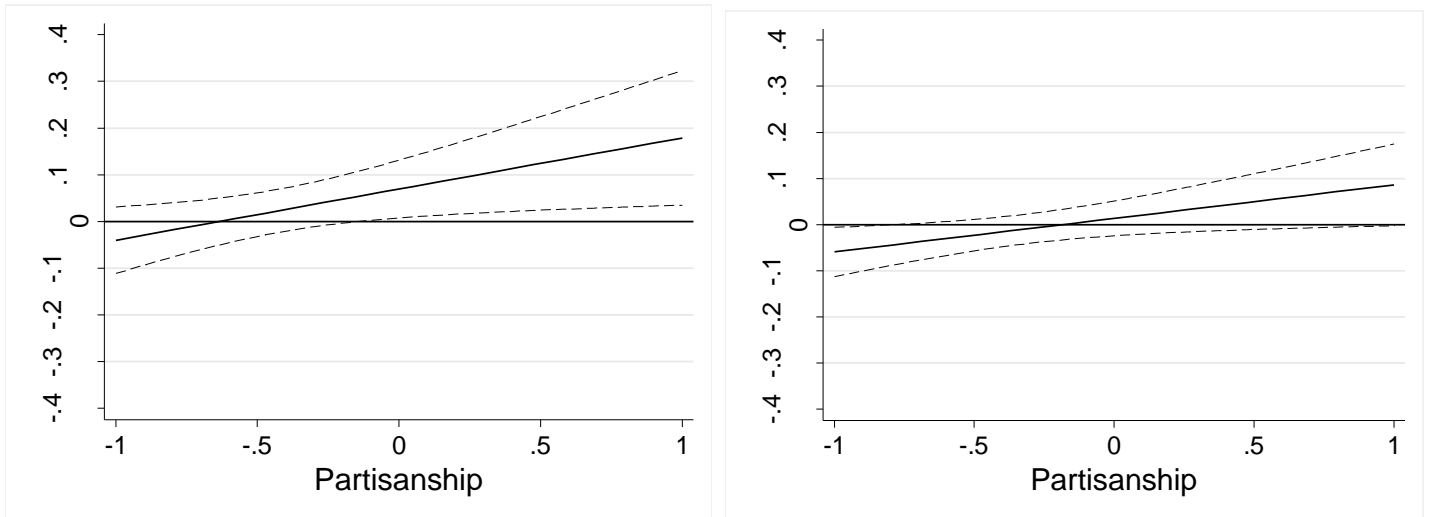


Note: Dashed lines indicate 90% confidence intervals, two-sided. Marginal effect of any flag relative to treatments with no flags.

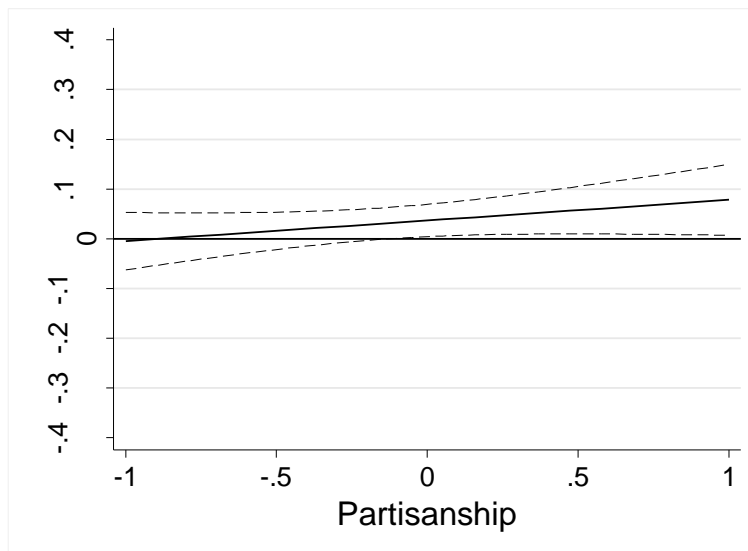
Figure A7: Marginal Flag Effects on Presidential Vote Choice by Partisanship

2012 College

2012 Mechanical Turk

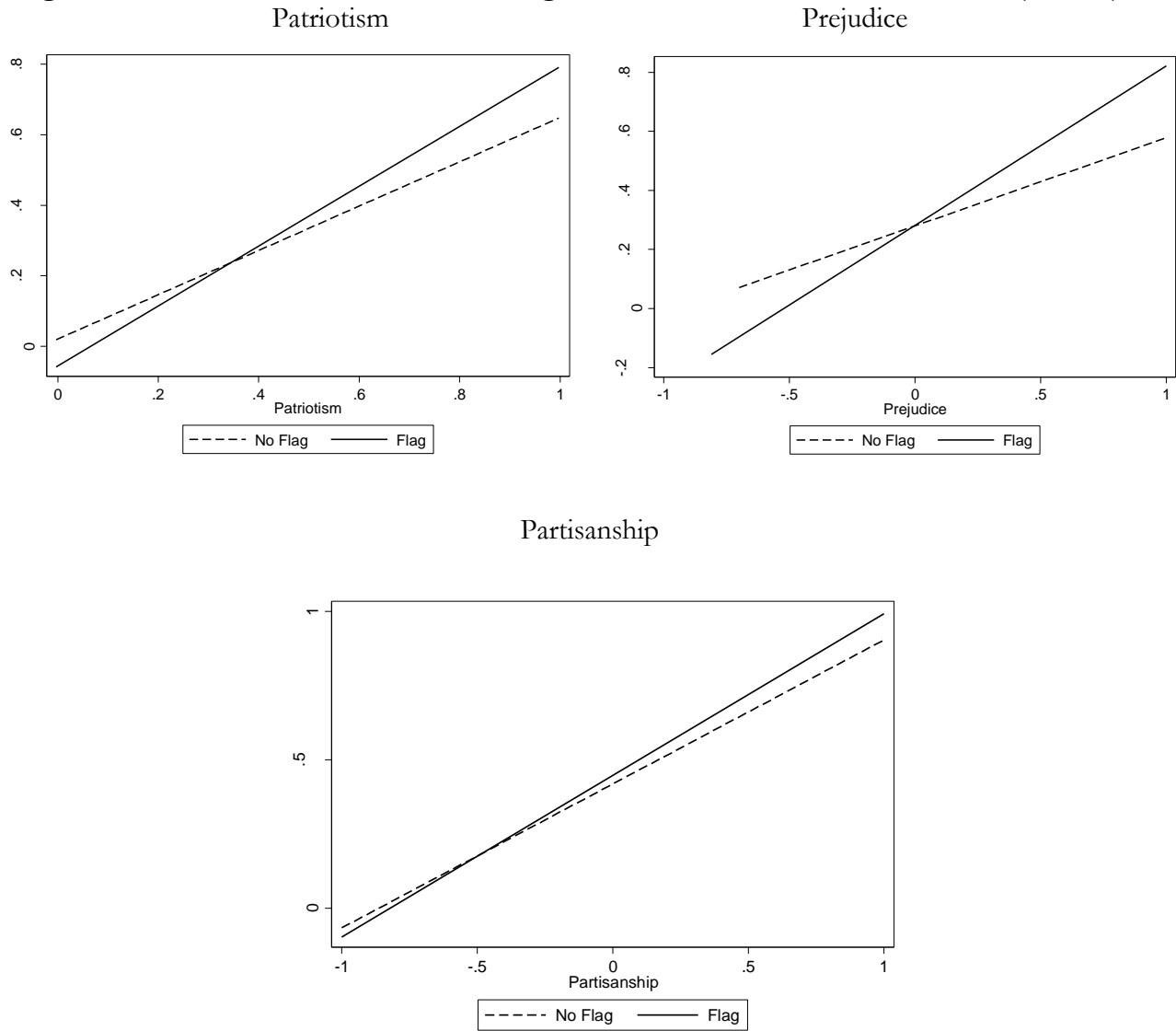


2013 Mechanical Turk



Note: Dashed lines indicate 90% confidence intervals, two-sided. Marginal effect of any flag relative to treatments with no flags.

Figure A8: Predicted Linear Values for Flag Effects on Presidential Vote Choice (Pooled)



Note: The difference between solid and dashed lines here is the marginal treatment effect displayed in Figure 1 of the main text.

Figure A9: Predicted Fractional Polynomial Values for Flag Effects on Presidential Vote Choice (Pooled)

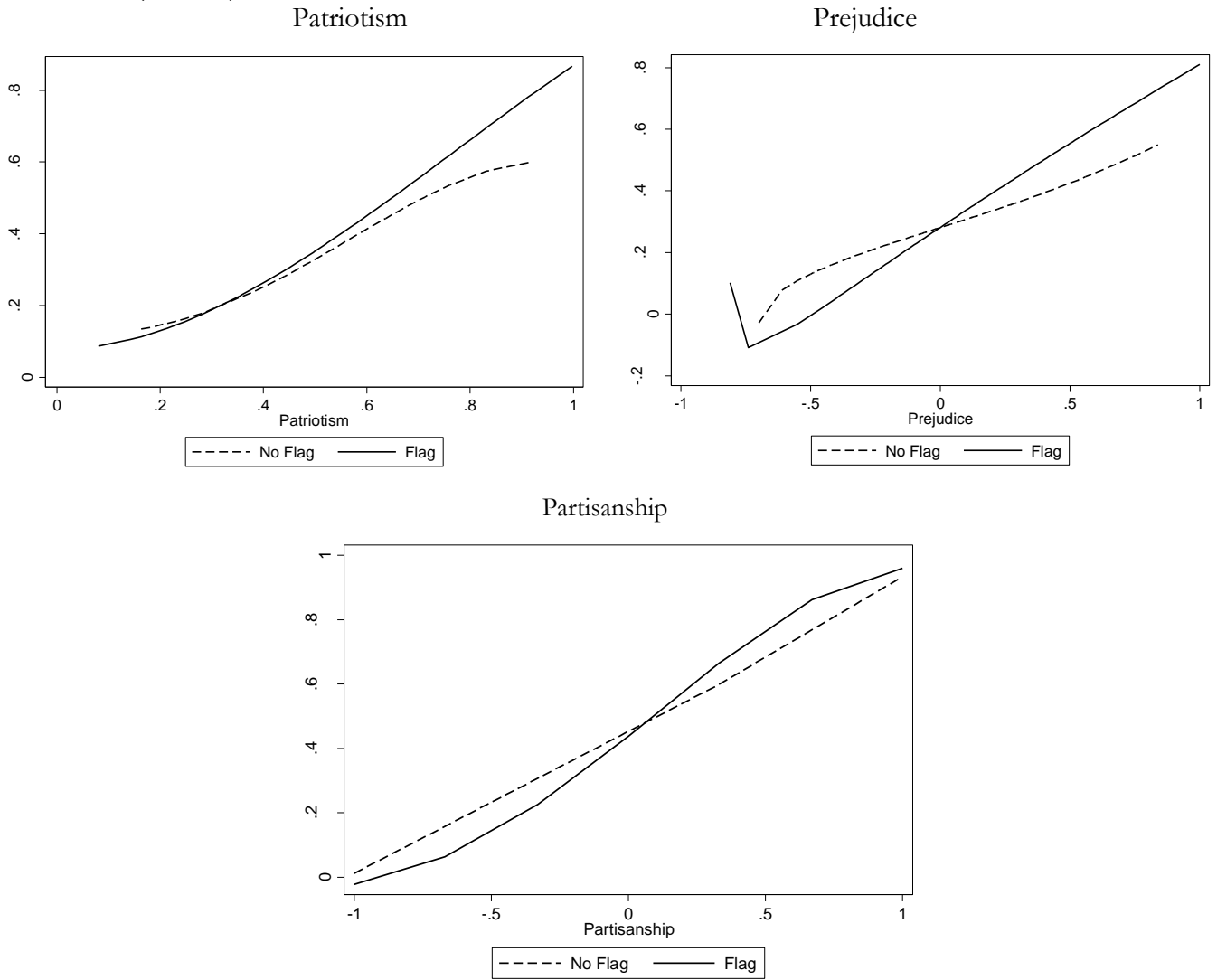


Table A6: Means of Republican Vote Choice by PID & Flag Condition (Pooled)

	No Flag	Flag	Difference	<i>p</i>	<i>N</i>
Strong Democrat	.02	.01	-.01	.25	187
Weak Democrat	.10	.10	.00	.83	404
Leans Democrat	.17	.15	-.02	.30	368
Independent	.39	.41	+.02	.76	236
Leans Republican	.65	.78	+.13	.004	171
Weak Republican	.79	.84	+.05	.19	172
Strong Republican	.91	.93	+.02	.85	55

Note: Means with standard deviations in parentheses. P-values calculated from two-sided t-tests. Differences may appear to diverge from means due to rounding.

Means of Republican Vote Choice by PID & Flag Condition (Pooled)

	No Flag	Flag	Difference	<i>p</i>	<i>N</i>
Democrat	.11	.10	-.01	.35	959
Independent	.39	.41	+.02	.76	236
Republican	.74	.83	+.08	.004	398

Note: Means with standard deviations in parentheses. P-values calculated from two-sided t-tests. Differences may appear to diverge from means due to rounding.

Appendix 6: Additional Discussion on Experimental Treatments & Analysis

It is important to note that there are subtle differences in the pictures across candidates for each experimental condition. Those differences could conceivably make a substantive difference in estimates of flag effects. However, these differences do not present confounds for our experimental analysis.

When “both candidates without flags” is the reference category, the conditions where one candidate appears with the flag and the other does not are within-candidate comparisons because the only change is the presence of a flag for one candidate. The photo of the other candidate is identical in both conditions. This tight comparison applies to the left columns in each table showing each individual experimental condition. The “both candidates with flags” condition changes two things relative to the baseline, in that both candidates now appear with flags, and the flag-only condition changes multiple dimensions by not even showing a candidate. If we didn’t have the two “one candidate with flag only” conditions in each study, these other flag conditions would pose an interpretive challenge relative to the baseline. However, we generally see the same results across all flag conditions (regardless of one or both candidates with the flag, even just the flag with no candidate). We would arrive at the same conclusions about flag effects if we dropped the two-candidates-with-flags and the flag-only conditions from analysis (accepting the argument that they vary more than one thing compared to the baseline). However, they both show similar flag effects observed for the conditions where only one candidate has a flag, and so we think including “both flag” and “flag only” broadens the finding. It is flag exposure that matters, not details of which candidates appear with the flag, how the flag is presented, or whether any candidate appears at all.

We would have more inferential trouble if we were arguing that flag effects depended on other factors in the image (e.g. candidate identity). In retrospect, if we had found such differences, it would have been better to have photos superimposing candidates on the same flag/non-flag backgrounds rather than using real pictures of the candidates with the flag edited out of the background. But since we argue for similar results with any photo featuring a flag compared to those that don’t we do not view these differences as confounds for our interpretation of flag effects. Those slight differences extend the generalizability of the

results in our view, in that they suggest the flag doesn't need to be positioned just so with a particular set of candidate gestures and facial framing to find the flag effect.

Finally, it is quite likely that differences in the flag background would matter if they made it difficult for participants to tell whether it was an American flag or not, but we see these images as clear in their flag display on their face despite differences, and the similar results support that conclusion empirically.

We thank our perceptive reviewers and editor for encouraging a more detailed discussion of these issues.