

Supporting information

Contents

1. Study 1: Appearance Advantage in the 2012 California House Primaries	3
1.1: Sample Characteristics	3
Survey election results predict actual election outcomes	3
Sample demographics and comparison to population of interest	4
1.2: Study 1: State Senate Races.....	5
1.3: Candidate Appearance Ratings: Gathering and Editing Images and Attention Test	8
Study 1: Appearance Survey for the California Primary Experiment	9
1.4: Study 1: List of Races.....	11
1.5: Study 1 candidate-level results with alternative coding for appearance	12
Appearance measured as raw rating, rescaled 0-1, rather than <i>Appearance Advantage (within district)</i>	12
2. Study 2: Appearance Advantage in the 2012 General Elections.....	13
2.1: Descriptive Statistics for Study 2	13
2.1.1 List of Races Used in Study 2.....	13
2.1.2 Summary Statistics Study 2 Races.....	14
2.1.3 Summary Statistics Study 2 Individual Respondents	14
2.1.4 Study 2 Respondent Demographics	15
2.2: Replication of Table 3 with All Respondents	16
2.3: Study 2 Appearance Ratings for the General Election Study.....	17
What do raters see in these photographs?	18
2.4: General election study: races with fewer or more than two candidates	20
Single Candidate Retention Elections (Study 2).....	20
Multiple Candidate Races (study 2)	21
3. Alternative Interpretations	22
3.1: Changes in Estimates over Time	22
Replication of Table 2 by Study Days (first half versus second half).....	22
3.2: Replication of Table 3 by Study Days (first half versus second half)	23
4. Candidate Appearance as a Low-Information Heuristic	24
4.1: Alternative specifications for individual-level analysis of state-level general races	24
4.1.1 Replication of Table 3 (with probit): Voter Responsiveness to Candidate Appearance in Statewide General Elections	24
4.1.2 Replication of Table 3 (with standard errors clustered by race): Voter Responsiveness to Candidate Appearance in Statewide General Elections	24

4.1.3 Replication of Table 3 (with standard errors clustered by respondent): Voter Responsiveness to Candidate Appearance in Statewide General Elections	25
4.2: Political Knowledge Measures for Study 2	26
4.3: Replication of Table 4 for Down Ballot Races	27
4.4 Replication of Table 4 with Local Knowledge	28
4.5: Fixed Effects Analysis for Study 2.....	29
4.6: Individual-Level Estimates for the Primaries Experiment (Study 1)	32
4.7: Political Knowledge Measures for Study 1	33
5. Additional Analyses and Material	34
5.1: Voters Claim to Ignore Appearance in Voting Decisions	34
5.15: Multicandidate and Retention Races from 2012 Statewide General Election Survey	35
Single Candidate Retention Elections	35
Multiple Candidate Races	36
5.2 Out-Of-State Experiment	37
5.3: Decomposition of Race-Level and Candidate-Level Effects in Studies 1 and 2	40
5.4: Reanalysis of Atkinson et al.	41

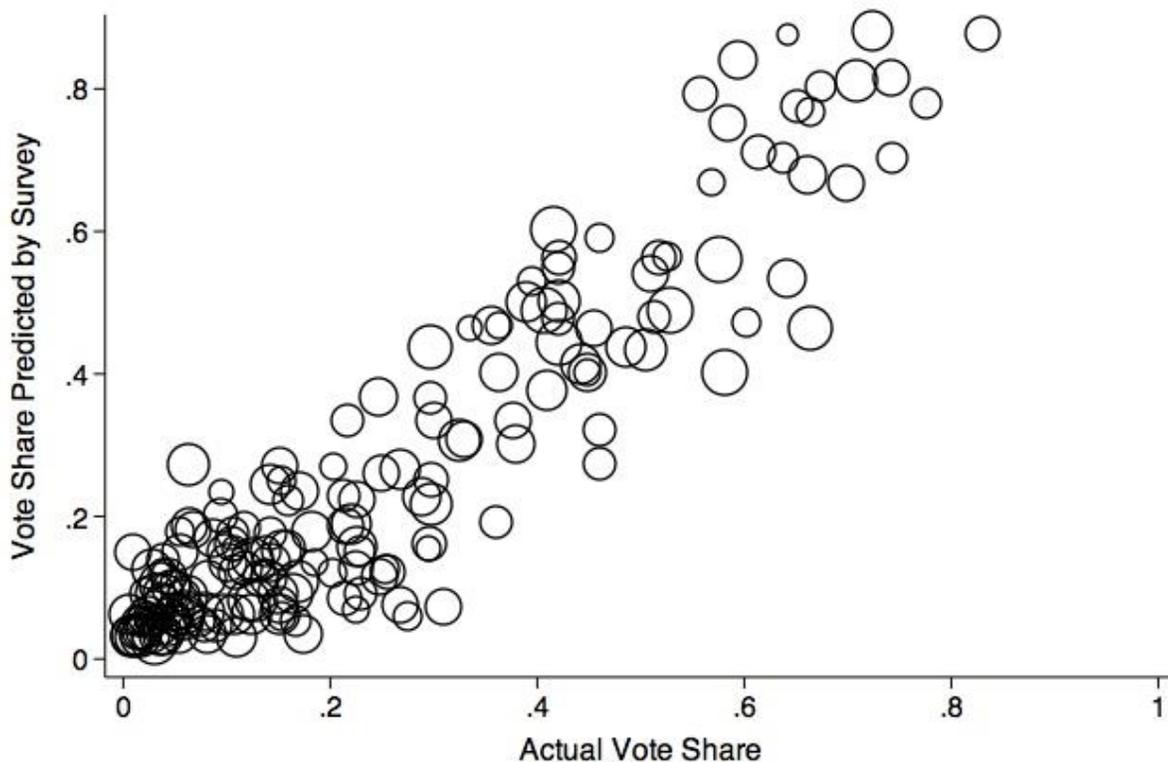
1. Study 1: Appearance Advantage in the 2012 California House Primaries

1.1: Sample Characteristics

Survey election results predict actual election outcomes

The data used are from a subset of races studied under the IGS California Top-Two Primary Survey. Responses to the IGS California Top-Two Primary Survey predicted actual outcomes from the June 5, 2012 election reasonably well.

Candidates' vote shares, as reported by the Secretary of State, were quite similar to the percentages of votes they received from survey respondents assigned to vote with open ballots without photographs. The scatterplot presented below shows this relationship.



Regressing candidates' predicted vote share on actual vote share confirms that the survey's results were in line with real-world outcomes. This regression produces a coefficient of 0.994 ($t=32.06$, $p<.001$), with a standard error of regression of .088.

Sample demographics and comparison to population of interest

Sampling frame	SSI Sample (May/June 2012) registered voters quota	IGS/Field Poll (September 2012) registered voters probability	CA Census (2010) population	CA Sec. of State registered voters population
Age				
18-19	1%	2%	4%	
20-24	1%	9%	10%	
25-34	17%	15%	19%	
35-49	21%	19%	28%	
50-64	42%	24%	24%	
65 & over	19%	30%	15%	
Gender				
Male	38%	49%	50%	
Female	62%	51%	50%	
Race				
White/Caucasian	81%	82%	75%	
African American	9%	8%	8%	
Asian	10%	11%	17%	
Ethnicity				
Hispanic or Latino	14%	21%	38%	
Not Hispanic or Latino	86%	79%	62%	
Education				
Some high school	2%	4%	12%	
High school/GED	12%	15%	24%	
Some college/2-year degree	50%	33%	34%	
4-year college degree	24%	26%	18%	
Graduate/professional degree	11%	21%	11%	
Party registration				
Democratic	45%	46%	43%	44%
Republican	31%	31%	30%	29%
No party preference	20%	18%	21%	21%
Other	4%	5%	5%	6%
Ideology				
Liberal (inc. leaners)	36%	35%		
Moderate	27%	22%		
Conservative (inc. leaners)	37%	42%		

We compare our sample to population benchmarks in the table above. Since our population of interest is California voters, we primarily compare our sample to a 2012 Field Poll probability sample of that population. The SSI sample matches the Field Poll sample reasonably well on several key covariates, including party registration and ideology, which are crucial for this study. (Note that the Field Poll survey used an 11-point measure of ideology rather than a 7-point measure, which likely accounts for the slightly lower percentage of self-reported moderates.)

1.2: Study 1: State Senate Races

In addition to running the experiment in US House primaries, we conducted the experiment in six primaries for California State Senate (Districts 7, 15, 17, 23, 25, 29). Across these races, we surveyed 741 registered voters (just 460 of whom, or 62%, actually voted in the study). The six races included 13 candidates.

We have two reasons to believe that the study was not as successfully administered in these state senate races as it was in the House races and the statewide races in Study 2. First, we were unable to obtain photographs of comparable quality across candidates in a majority of races. In four of the six contests, one candidate's photograph was noticeably lower in resolution. In each of these cases, the photo-advantaged candidate was either an incumbent or a member of the lower chamber of the legislature, while the photo-disadvantaged candidates were relatively unknown challengers. Indeed, when we regress our measure of relative appearance advantage on an incumbency indicator variable in these state senate races, we find that incumbency is a significant predictor (slope = 0.095, $p = 0.058$). By contrast, this is not the case in the House races (slope = 0.014, $p = 0.868$), implying that in addition to appearance, the photographs in the state senate contests may indeed signal effort (or lack thereof) that the actual campaigns (or lack thereof) are conveying in the real world.

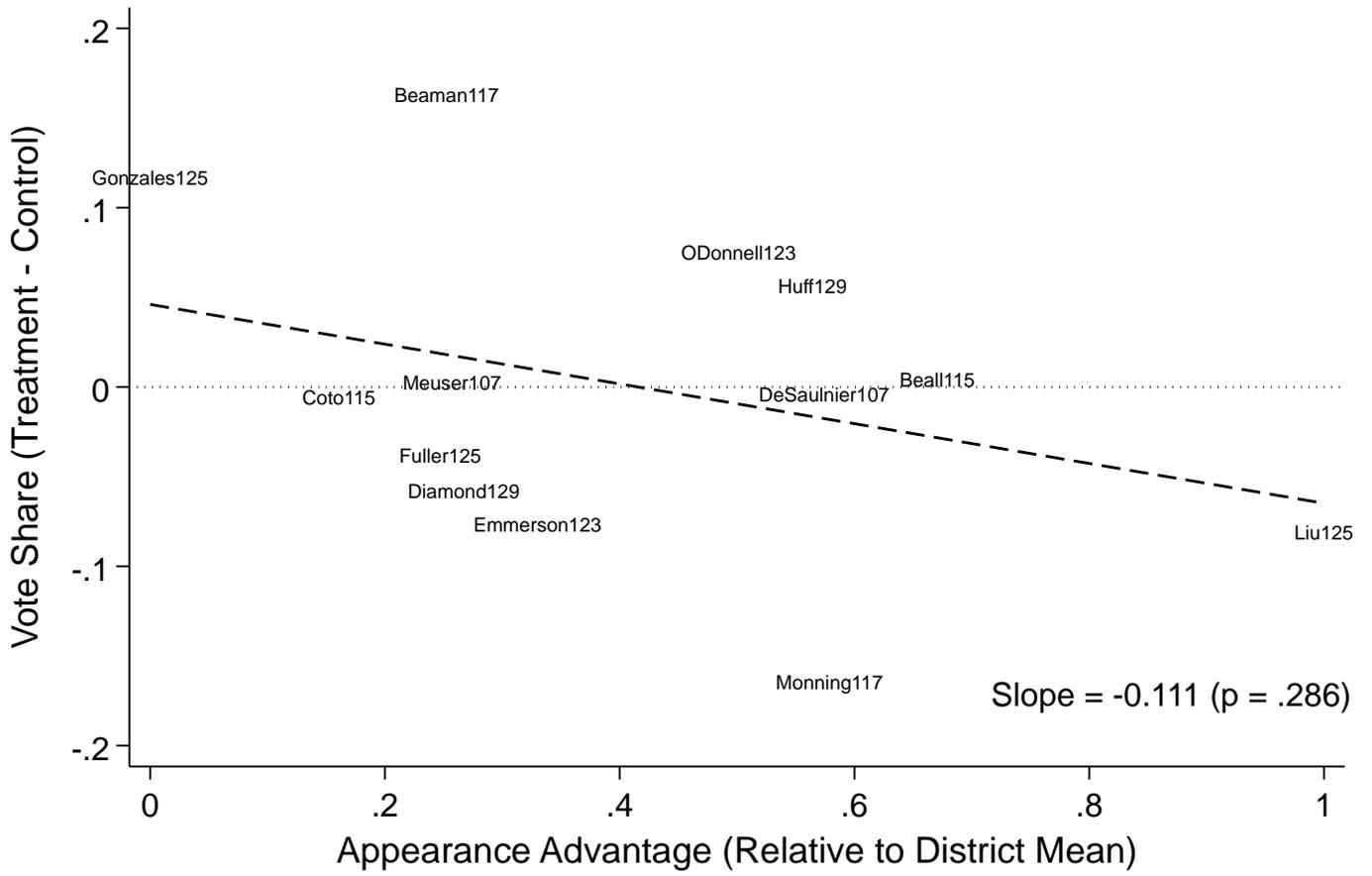
The experimental results are consistent with this reasoning. At the candidate level, the apparent effect of the treatment ballot is negative, implying that better-looking candidates do worse when we show photographs. (See the first figure below.) However, this difference is likely a fluke. First, the apparent negative effect is statistically indistinguishable from zero. Second, when we break out the apparent effect of candidate relative appearance advantage on candidates' fortunes in treatment and candidates' fortunes in control, we observe that there is a sharp upward slope in both trends, a phenomenon we do not observe in the House races. (See the second figure below.) This provides circumstantial evidence that the difference in photograph quality mirrors a noticeable difference in campaign quality in the real world (i.e., among control voters.)

The pattern we find at the candidate level and show in the first figure above holds up at the individual level: participants assigned to the treatment condition select candidates 3.3 percentage points more appearance-disadvantaged, on average, a difference that is not statistically significant ($p = 0.29$)

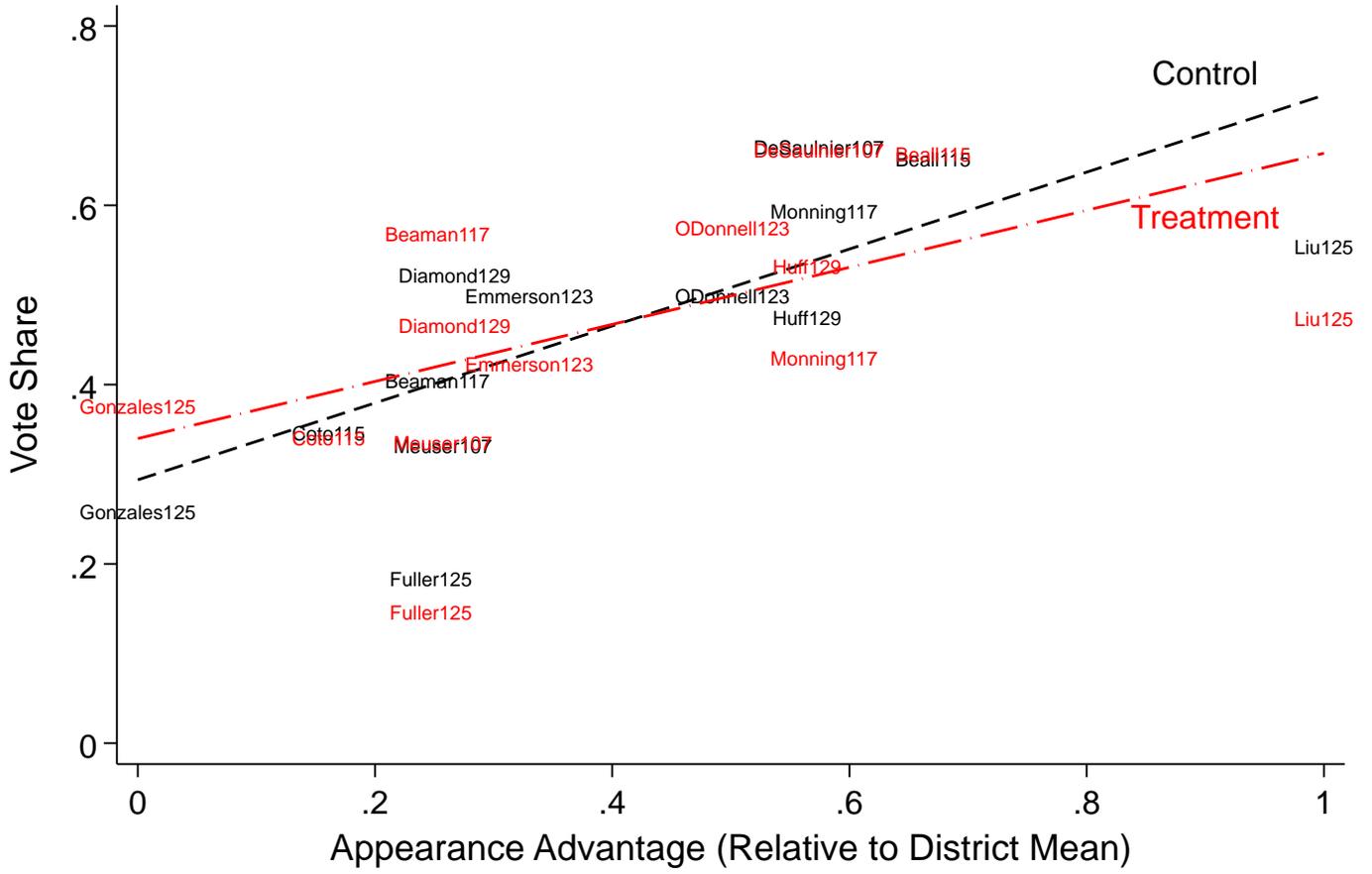
A second potential issue is that we administered the experiment concurrently with an experiment on the effect of the new top-two primary format on moderate candidates' fortunes and voters' decisions. This joint administration cannot affect House races because we always asked voters about their preferences in the House races first. But it has the potential to affect choices in state senate contests. State senate districts and US House districts are drawn without perfect overlap, so we assigned districts to the ballot study or the faces study for reasons other than respondents being assigned to the same experiment in both races (as doing so would have been impossible). Thus, participants who take part in both studies experience an odd combination of treatments. Among these participants in the state senate faces study, those assigned to control have previously seen a counterfactual partisan primary ballot before observing a top-two ballot without faces. Those assigned to treatment have previously seen a top-two ballot without faces (essentially, the control ballot in the faces study) before observing the top-two ballot with faces as the treatment in the senate races. It is not immediately clear what the actual net effect, if any, of this contamination is, but it cannot be as cleanly interpreted as the effect of faces on the ballot in the House races.

Most importantly, pooling the state senate races with the house races leaves the overall finding the same (i.e. it is still statistically significant).

Apparent effect of the appearance advantage on candidates' fortunes in state Senate races



Apparent effect of appearance advantage within conditions in state Senate races



1.3: Candidate Appearance Ratings: Gathering and Editing Images and Attention Test

Gathering Images

Two steps:

- (1) We had research assistance or students in a class for a research project look for an image that is as large as possible.
- (2) We asked them to find as many images as possible for each candidate and then choose a representative one.

Editing Images

Three steps:

- (1) We cropped images so that the edges of photograph reach to just past the edges of the candidate's face. This includes the ears on the right and left, hair on the top, and chin on the bottom. We used a pixel ratio of approximately 120 x 160.
- (2) We replaced backgrounds of candidate images so that they were all neutral. We removed background features in the photos using a layer of grey underneath the candidate image then simply erased the background around the head, neck and, shoulders.
- (3) To remove any influence of clothing or other differences in color, we also grey-scaled the images.

Attention test

We included attention tests in both surveys. In our instructions we told participants to ignore the question itself and choose designated responses. The question asks which number is the largest out of numbers 1, 2, and 3. Our instructions prompt them to ignore the question and instead choose numbers 1 and 2. We excluded respondents who failed the attention test.

Study 1: Appearance Survey for the California Primary Experiment

On the next page, we show the instructions and question wording for the appearance-rating survey.

We recruited 154 respondents on Mechanical Turk and excluded 11 because they were from California, three because they failed the attention test, and one because he or she rated all candidates as average. Each participant rated a random set of 25 candidates. Very few respondents said they recognized candidates, with respondents recognizing only 0.25 on average (of the 25 they rated). After excluding respondents for these reasons, we have 52.1 ratings per candidate on average.

At the end of the survey, we asked the attention test described above, state of residence, and gender.

Instructions for the CA primary experiment appearance-rating survey (screenshot)

Instructions

**For the candidates on the following 25 pages,
please tell us who you would be most inclined to vote for.**



Example of screen and question wording for the CA primary experiment appearance-rating survey (screenshot)



How good of a Congressperson do think this person would be?

- Much better than average
- Somewhat better than average
- Average
- Somewhat worse than average
- Much worse than average
- (I recognize this candidate)



1.4: Study 1: List of Races

<i>Congressional Districts</i>	<i>Number of candidates</i>	<i>Number of respondents</i>
1	7	75
5	3	71
6	3	72
9	3	67
11	4	63
17	3	39
18	4	44
22	2	51
28	4	54
30	5	88
33	8	49
36	2	92
44	2	39
49	3	47
Total	53	851

1.5: Study 1 candidate-level results with alternative coding for appearance

Appearance measured as raw rating, rescaled 0-1, rather than *Appearance Advantage (within district)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent: Variable: Picture condition minus no-picture condition vote share								
	All	Challengers	Inc.	Non-viable	Viable	Dem.	Rep.	All with controls
Appearance rating	0.58*** (0.21)	0.36 (0.23)	1.23*** (0.37)	0.66** (0.26)	0.41 (0.29)	0.49 (0.34)	0.80*** (0.26)	0.48** (0.22)
Incumbent								-0.099*** (0.029)
White								0.063** (0.030)
Male								-0.11*** (0.036)
Constant	-0.28*** (0.10)	-0.16 (0.11)	-0.67*** (0.19)	-0.29** (0.13)	-0.23 (0.14)	-0.29 (0.17)	-0.35** (0.12)	-0.17 (0.13)
Observations	53	39	14	24	29	23	23	53
R-squared	0.129	0.061	0.476	0.225	0.066	0.090	0.312	0.384

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2. Study 2: Appearance Advantage in the 2012 General Elections

2.1: Descriptive Statistics for Study 2

2.1.1 List of Races Used in Study 2

The table shows the number of respondents (non-unique) for each race for each state.

	<i>AG</i>	<i>CoAgr</i>	<i>Colns</i>	<i>CoLab</i>	<i>CoPubL</i>	<i>Gov</i>	<i>LtGov</i>	<i>JuCCA</i>	<i>RailCo</i>	<i>StAud</i>	<i>SecSt</i>	<i>Sen</i>	<i>SupIntEd</i>	<i>SecTres</i>	<i>UBR</i>	<i>Total</i>	<i>Number of races</i>
AZ	0	0	0	0	0	0	0	0	0	0	0	61	0	0	0	61	1
CA	0	0	0	0	0	0	0	0	0	0	0	243	0	0	0	243	1
CO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48	48	1
FL	0	0	0	0	0	0	0	0	0	0	0	162	0	0	0	162	1
IN	48	0	0	0	0	0	0	0	0	0	0	0	48	0	0	96	2
MA	0	0	0	0	0	0	0	0	0	0	0	75	0	0	0	75	1
MI	0	0	0	0	0	0	0	0	0	0	0	96	0	0	0	96	1
MO	57	0	0	0	0	57	57	0	0	0	57	57	0	57	0	342	6
NC	0	126	126	126	0	126	126	0	0	126	126	0	126	126	0	1,134	9
NJ	0	0	0	0	0	0	0	0	0	0	0	57	0	0	0	57	1
NY	0	0	0	0	0	0	0	0	0	0	0	184	0	0	0	184	1
OH	0	0	0	0	0	0	0	0	0	0	0	138	0	0	0	138	1
PA	148	0	0	0	0	0	0	0	0	0	0	148	0	148	0	444	3
TN	0	0	0	0	0	0	0	0	0	0	0	55	0	0	0	55	1
TX	0	0	0	0	0	0	0	135	135	0	0	135	0	0	0	405	3
VA	0	0	0	0	0	0	0	0	0	0	0	101	0	0	0	101	1
WA	122	0	122	0	122	122	122	0	0	122	122	122	0	122	0	1,098	9
WI	0	0	0	0	0	0	0	0	0	0	0	77	0	0	0	77	1
Total	375	126	248	126	122	305	305	135	135	248	305	1,711	174	453	48	4,816	44

Six states have more than one race.

In addition to the races reported above, we also included in the data collection phase of the experiment several single candidate and multi-candidate (greater than two) races that we excluded from the main analysis, which only included two-candidate races for ease of analysis. The races excluded include three separate single-candidate Court of appeals retention elections, three separate single-candidate Florida Supreme Court races, multi-candidate races for Michigan State Board of Education, Michigan State University Board of Trustees, Michigan University Board of Regents, and Washington Superintendent of Public Instruction. Unless otherwise noted, all results reported for Study 2 exclude these races which have more than or less than two candidates. See section 2.4 of this document for the results in these races.

2.1.2 Summary Statistics Study 2 Races

Variable	Obs	Weight	Mean	SD	Min	Max
Republican Vote Share (Treatment - Control)	44	4816	0.014	0.097	-0.185	0.240
Appearance Advantage (for Republican)	44	4816	0.523	0.189	0	1
Senate	44	4816	0.355	0.484	0	1
Governor	44	4816	0.063	0.246	0	1
Down ballot	44	4816	0.581	0.499	0	1
Republican female	44	4816	0.248	0.437	0	1
Democratic female	44	4816	0.427	0.500	0	1
White Republican	44	4816	0.975	0.159	0	1
White Democrat	44	4816	0.974	0.161	0	1
Matched on Race & Gender	44	4816	0.606	0.494	0	1

Results presented at the race-level, weighted by respondents.

2.1.3 Summary Statistics Study 2 Individual Respondents

Variable	Obs	Mean	SD	Min	Max
Vote Republican	4816	0.268	0.443	0	1
Treatment	4816	0.491	0.500	0	1
Appearance Advantage (for Republican)	4816	0.523	0.187	0	1
Treatment X Appearance Advantage	4816	0.256	0.291	0	1
Party ID (1 Strong Democrat to 7 Strong Republican)	4452	5.086	1.886	1	7
Knowledge	4816	2.376	1.332	0	4
Local Knowledge	4816	0.465	0.257	0	1

Results presented at the individual-response level.

2.1.4 Study 2 Respondent Demographics

Age	Freq.	Percent	Cum.
18-25	1,884	39.12	39.12
26-34	1,464	30.4	69.52
35-54	1,215	25.23	94.75
55-64	180	3.74	98.48
65 or over	73	1.52	100
Total	4,816	100	

Education	Freq.	Percent	Cum.
Less than High School	70	1.45	1.45
High School / GED	376	7.81	9.26
Some College	1,729	35.9	45.16
2-year College Degree	513	10.65	55.81
4-year College Degree	1,601	33.24	89.06
Masters Degree	409	8.49	97.55
Doctoral Degree	35	0.73	98.28
Professional Degree (JD, MD)	83	1.72	100
Total	4,816	100	

2.2: Replication of Table 3 with All Respondents

In the paper, we drop Study 2 respondents (Mechanical Turk workers) who failed the attention test and who said they will not vote. Below are the results when we don't drop those respondents. Not all the estimates are statistically significant, but they're close. Interestingly, if we estimate the effects only dropping those who failed the attention test, the effect is statistically significant in every column.

Appearance advantage in the General Election

Dependent variable: Picture condition minus no-picture condition vote share

	(1)	(2)	(3)	(4)	(5)
	All	Matched on race and gender	All	Senate & Governor	Other
Appearance Advantage (for Republican)	0.17** (0.08)	0.24* (0.13)	0.11 (0.09)	0.23 (0.14)	0.15* (0.08)
Incumbent			0.00 (0.02)		
Female Republican			0.07* (0.04)		
Female Democrat			-0.00 (0.03)		
White Republican			-0.00 (0.10)		
White Democrat			-0.01 (0.10)		
Constant	-0.07 (0.04)	-0.11 (0.07)	-0.04 (0.13)	-0.13 (0.08)	-0.03 (0.04)
Observations	44	26	44	18	26
R-squared	0.11	0.12	0.19	0.14	0.12

This table shows candidate-level regressions. Dependent variable: picture condition minus no-picture condition vote share (coded so that higher values indicate greater Republican vote share). *** p<0.01, ** p<0.05, * p<0.1

2.3: Study 2 Appearance Ratings for the General Election Study

On the next page, we show the instructions and question wording for the appearance-rating survey.

We recruited 253 respondents on Mechanical Turk and excluded 11 because they failed the attention test. Each participant rated a random set of 25 candidates. Very few respondents said they recognized candidates, with respondents recognizing only 0.3 on average (of the 25 they rated). After excluding respondents because they failed the attention test and after excluding "recognize" responses, we have 53.8 ratings per candidate on average.

At the end of the survey, we asked the attention test described above, state of residence, and gender.

Instructions for the general election experiment appearance-rating survey (screenshot)

Instructions

**For the candidates on the following 25 pages,
please tell us how good of an elected official (such as Senator, Governor,
House Member, etc.) you think each would be.**

>>

Example of screen and question wording for the general election study appearance-rating survey (screenshot)



How good of an elected official (e.g., Sen. or Gov.) do think this person would be?

- Much better than average
- Somewhat better than average
- Average
- Somewhat worse than average
- Much worse than average
- (I recognize this candidate)

>>

What do raters see in these photographs?

Rather than asking about particular traits, we asked raters the broader question of how good of an elected official they thought the person depicted would be. We did this to sidestep the debate about what trait in candidates' faces voters respond to. Researchers have found results consistent with competence (Todorov et al. 2005), attractiveness (Banducci et al. 2008), and dominance (Rule et al. 2010). After conducting the initial study, we collected additional ratings of 40 candidate pictures from Study 2 (all senatorial and gubernatorial candidates, plus six down-ticket state executive races to add variation in race, gender, and salience). To investigate what raters see in these photographs, we randomly asked raters in the second survey to assess the competence, dominance, or attractiveness of the person in the photograph with a 5-point scale.

We conducted the study through Survey Sampling International (SSI) in May 2012. We surveyed 515 individuals and yielded 2072 candidate rating-respondent pairs. We collapse these ratings by candidate-characteristic and examine the correlations between our original ratings and the ratings on these three characteristics to assess what people see in candidates' faces.

As the tables below show, perceptions of whether a candidate would be a good elected official, based on a photograph alone, correlate most strongly with perceptions of competence and attractiveness. When we introduce all three characteristic ratings—competence, dominance, and attractiveness—into a regression analysis (OLS) with our original measure as the dependent variable, we observe that perceptions of competence have the largest coefficient.

Summary of Rating Measures

Rating	<i>n</i> <i>Candidates</i>	<i>Mean</i>	<i>Standard</i> <i>Deviation</i>	<i>Min</i>	<i>Max</i>
Overall (Orig. Ratings)	40	0.55	0.21	0.12	1.00
Attractiveness	40	3.28	0.46	2.45	4.16
Competence	40	3.45	0.27	2.86	3.84
Dominance	40	3.53	0.35	2.67	4.38

Correlations of Rating Measures

	<i>Overall (Original Ratings)</i>	<i>Perceived Attractiveness</i>	<i>Perceived Competence</i>	<i>Perceived Dominance</i>
Overall	1			
Perceived Attractiveness	0.60	1		
Perceived Competence	0.61	0.57	1	
Perceived Dominance	0.19	0.29	0.29	1

Regression (OLS) Analysis: Overall (original) rating on specific ratings

(1)	
DV: Overall Face Rating	
Attractiveness	0.17** (0.07)
Competence	0.32** (0.12)
Dominance	-0.02 (0.08)
Constant	-1.03*** (0.38)
Candidates	40
R-squared	0.46

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

2.4: General election study: races with fewer or more than two candidates

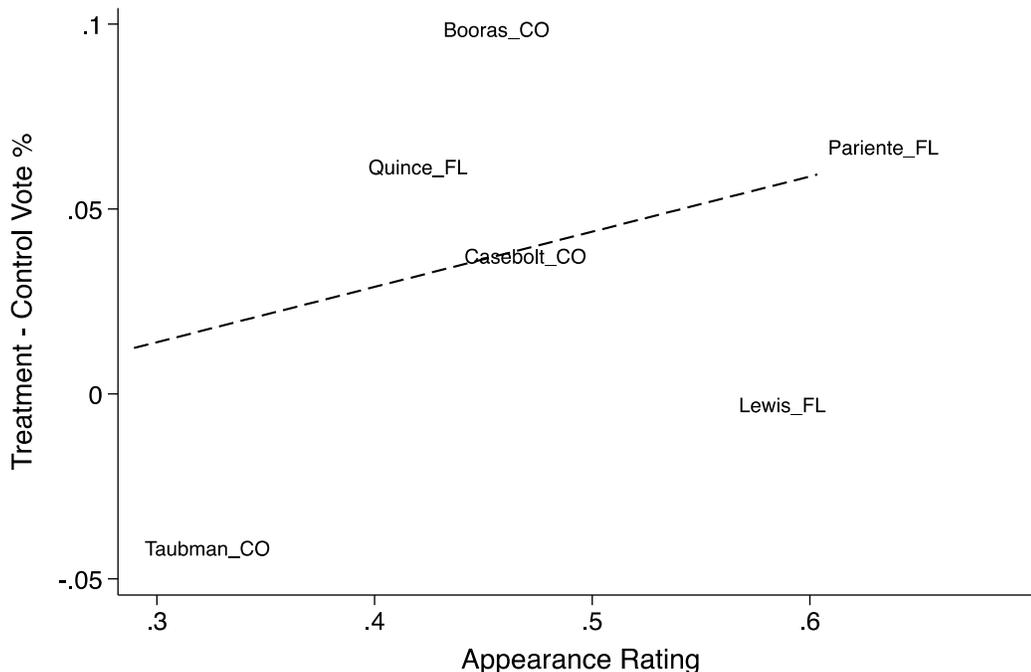
In Study 2, the general election survey, we asked respondents in a handful of states about their preferences in retention races (which feature only one candidate) and in races with more than two candidates. We do not present the results of the experiment in these races in the paper because the analysis is necessarily different from the two-candidate races. In particular, we cannot use “vote for the Republican” as the dependent variable. As such, race-level analysis is more complicated. Candidate-level analysis (as used in the primary study) is more easily interpreted, so we present those results here.

Single Candidate Retention Elections (Study 2)

The general election survey asked voters in two states about three retention races each. Voters in Colorado reported whether they wanted to retain three judges in the Colorado Court of Appeals, while voters in Florida reported whether they wanted to retain three justices from the Florida Supreme Court.

In these races, our candidate-level dependent variable is the difference in vote share a candidate received under the two conditions, $\text{Vote\%}_T - \text{Vote\%}_C$. Since candidates in retention elections lack opponents, we use appearance rather than appearance advantage as the independent variable. (This is the average score the candidate received in the MTurk rating survey. Recall that we asked respondents, “How good of an elected official do you think this person would be?”) We rescale both variables 0-1.

As shown in the scatterplot below, we observe the expected positive trend between $\text{Vote\%}_T - \text{Vote\%}_C$ and appearance. This apparent upward slope is imprecisely estimated because we only have six candidates in these races, so it fails to reach statistical significance. However, it is substantively large. A bivariate regression of $\text{Vote\%}_T - \text{Vote\%}_C$ on appearance demonstrates that this slope is 0.15 ($p = 0.52$). When we only examine within-state variation (state fixed effects), the estimated coefficient is 0.20 ($p = 0.57$).

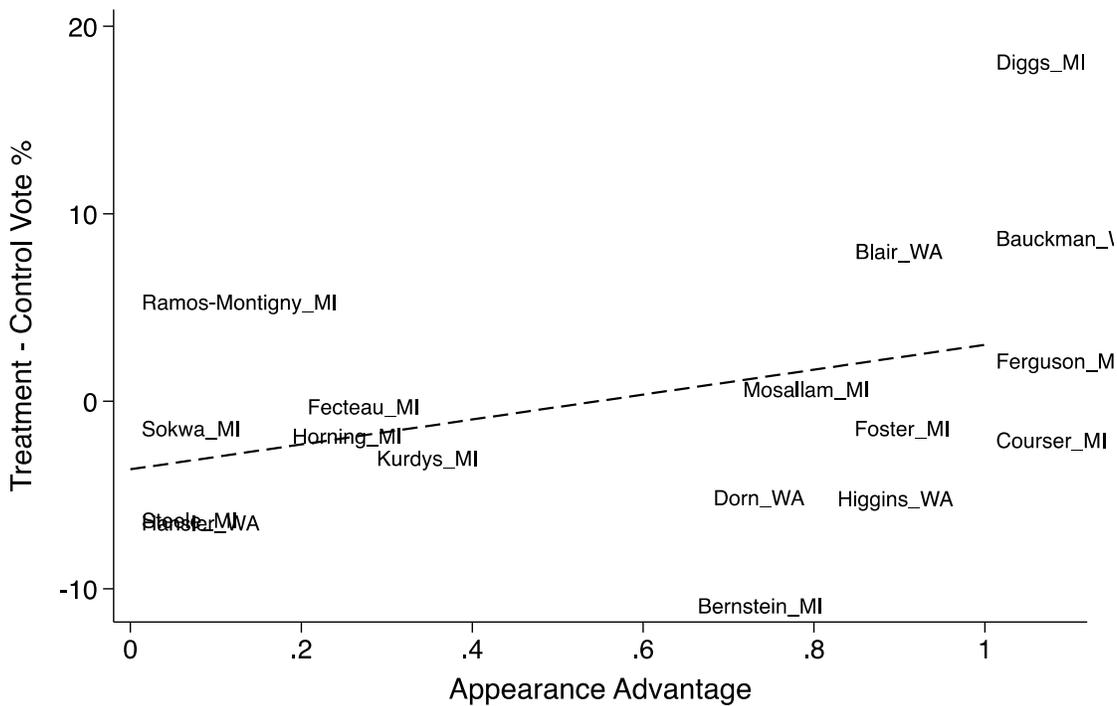


Multiple Candidate Races (study 2)

The general election survey asked respondents in Michigan about three races featuring four candidates (State Board of Education, University Board of Regents, and Michigan State University Board of Trustees) and respondents in Washington about the race for Superintendent of Public Instruction, which featured five candidates.

The dependent variable in these races is the difference in vote share a candidate received under the two conditions, $\text{Vote\%}_T - \text{Vote\%}_C$. The independent variable is the candidate's within-race appearance advantage, $(\text{face rating} - \text{race minimum face rating}) / (\text{race maximum face rating} - \text{race minimum face rating})$. Both variables are scaled 0-1.

As shown in the scatterplot below, we observe the expected positive trend between $\text{Vote\%}_T - \text{Vote\%}_C$ and appearance. This apparent upward slope is imprecisely estimated because we only have seventeen candidates in these races, but we come closer to conventional levels of statistical significance here than in the single-candidate races. The estimated slope from the regression of $\text{Vote\%}_T - \text{Vote\%}_C$ on appearance advantage is 0.07 ($p = 0.12$), implying that the most appearance advantaged candidate would be expected to see a net seven-point benefit on the ballot with photographs over the most appearance disadvantaged candidate. The estimated apparent effect is consistent when we only examine within-race variation (race fixed effects) ($b = 0.07$, $p = 0.15$).



3. Alternative Interpretations

3.1: Changes in Estimates over Time

The general election study (Study 2) was conducted over 17 days. This table replicates Table 2 for participants interviewed in the first half of the study (first nine days) and then in the second half. The estimates show that treatment effect declines considerably. On the next page, we test whether this change is statistically significant using individual-level data (see the interaction: Second half x Treatment x Appearance Advantage), which shows that it is in some but not all specifications.

Replication of Table 2 by Study Days (first half versus second half)

First half					
	(1)	(2)	(3)	(4)	(5)
	All	Matched on race and gender	All	Senate	Other
Appearance Advantage (for Republican)	0.323** (0.140)	0.434* (0.245)	0.239 (0.166)	0.423 (0.264)	0.343** (0.144)
Incumbent			0.001 (0.039)		
Female Republican			0.082 (0.073)		
Female Democrat			-0.024 (0.060)		
White Republican			-0.144 (0.196)		
White Democrat			-0.039 (0.184)		
Constant	-0.100 (0.079)	-0.171 (0.141)	0.111 (0.259)	-0.240 (0.157)	-0.055 (0.079)
Observations	43	25	43	15	28
R-squared	0.115	0.120	0.177	0.165	0.180
Second half					
Appearance Advantage (for Republican)	0.117 (0.087)	0.078 (0.147)	0.089 (0.102)	0.215 (0.176)	0.073 (0.101)
Incumbent			0.011 (0.025)		
Female Republican			0.023 (0.048)		
Female Democrat			0.013 (0.038)		
White Republican			0.020 (0.117)		
White Democrat			0.043 (0.109)		
Constant	-0.074 (0.048)	-0.043 (0.082)	-0.127 (0.154)	-0.129 (0.102)	-0.050 (0.054)
Observations	44	26	44	15	29
R-squared	0.042	0.012	0.061	0.103	0.019

Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

3.2: Replication of Table 3 by Study Days (first half versus second half)

	(1)	(2)	(3)	(4)	(5)	(6)
	All	Matched on	Low	High	Weak/	Strong
VARIABLES	respondents	Race and	Knowledge	Knowledge	Indep.	Partisan
	vote_R	Gender	vote_R	vote_R	vote_R	vote_R
Treatment	-0.12 (0.07)	-0.21* (0.12)	-0.23*** (0.09)	-0.00 (0.11)	-0.23*** (0.08)	0.09 (0.14)
Appearance Advantage (for Republican)	-0.11** (0.06)	-0.08 (0.13)	-0.20*** (0.06)	-0.05 (0.09)	-0.27*** (0.08)	0.07 (0.14)
Treatment x Appearance Advantage (for Republican)	0.35*** (0.09)	0.50*** (0.18)	0.56*** (0.09)	0.12 (0.16)	0.44*** (0.11)	0.10 (0.20)
Second half	0.00 (0.06)	-0.04 (0.10)	-0.05 (0.08)	0.04 (0.10)	-0.11 (0.08)	0.19 (0.13)
Second half x Treatment	0.05 (0.10)	0.17 (0.17)	0.06 (0.12)	0.02 (0.15)	0.14 (0.12)	-0.19 (0.18)
Second half x Appearance Advantage (for Republican)	0.13* (0.07)	0.19 (0.15)	0.20* (0.11)	0.07 (0.13)	0.35*** (0.09)	-0.19 (0.21)
Second half x Treatment x Appearance Advantage (for Republican)	-0.23 (0.14)	-0.42 (0.30)	-0.27* (0.16)	-0.17 (0.22)	-0.35** (0.17)	0.09 (0.26)
Constant	0.28*** (0.05)	0.25*** (0.09)	0.33*** (0.06)	0.23*** (0.07)	0.41*** (0.07)	0.06 (0.09)
Observations	4,816	2,918	2,324	2,492	2,826	1,626
R-squared	0.01	0.01	0.01	0.00	0.01	0.01

Standard errors clustered by respondent and race in parentheses. No fixed effects *** p<0.01, ** p<0.05, * p<0.1

As can be seen in the replication of Table 2 on the previous page, appearance advantage has a substantively large and statistically significant effect on vote choice during the first half of our study period, while in the second half the effect is considerably smaller and statistically insignificant. This result is consistent with one-way non-compliance, i.e. in the second half (closer to the election) more of the respondents in the control condition have in fact been treated with candidate appearance in the real world. The replication of Table 4 (just above) presents similar results at the individual level where the dependent variable is individual vote choice. If campaigns are increasingly exposing voters to candidate appearance as we approach Election Day, then we would expect that (1) respondents in the control condition will be more likely to vote for the appearance advantaged candidate in the period closer to the election and that (2) being exposed to candidate photos in the treatment condition will have less of an effect in the second half. Indeed, respondents in the control condition are more likely to vote for the appearance advantaged candidate in the second half of the study, consistent with the idea that they are being increasingly exposed to candidate appearance in the real world. In contrast, the effect of exposing respondents to candidate photographs decreases in the second half of the study (Second half x Treatment x Appearance Advantage), consistent with the idea that increasing number of respondents assigned to treatment have already been treated with candidate appearance in the real world. This difference is not always statistically significant at conventional levels, but usually close to significance.

We also examined whether partisanship and incumbency became more or less important between the first and second halves, but they did not.

4. Candidate Appearance as a Low-Information Heuristic

4.1: Alternative specifications for individual-level analysis of state-level general races

4.1.1 Replication of Table 3 (with probit): Voter Responsiveness to Candidate Appearance in Statewide General Elections

DV: Vote Republican indicator variable	(1) All respondents	(2) Matched on race and gender	(3) Low knowledge	(4) High knowledge	(5) Weak/indep.	(6) Strong partisan
Treatment	-0.28* (0.15)	-0.32 (0.22)	-0.61*** (0.21)	0.03 (0.19)	-0.41** (0.17)	-0.18 (0.26)
Appearance advantage (for Republican)	-0.10 (0.19)	0.10 (0.43)	-0.19 (0.25)	-0.03 (0.29)	-0.14 (0.25)	-0.27 (0.31)
Treatment x Appearance advantage (for Republican)	0.59*** (0.22)	0.69** (0.34)	1.18*** (0.33)	0.04 (0.25)	0.64** (0.25)	0.66* (0.37)
Observations	4,816	2,918	2,324	2,492	2,826	1,626

Note: This table shows individual-level probit regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the individual and race-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.1.2 Replication of Table 3 (with standard errors clustered by race): Voter Responsiveness to Candidate Appearance in Statewide General Elections

DV: Vote Republican indicator variable	(1) All respondents	(2) Matched on race and gender	(3) Low knowledge	(4) High knowledge	(5) Weak/indep.	(6) Strong partisan
Treatment	-0.09** (0.04)	-0.11* (0.06)	-0.20*** (0.05)	0.01 (0.04)	-0.14*** (0.04)	-0.05 (0.05)
Appearance advantage (for Republican)	-0.03 (0.06)	0.03 (0.13)	-0.06 (0.08)	-0.01 (0.10)	-0.05 (0.09)	-0.07 (0.08)
Treatment x Appearance advantage (for Republican)	0.20*** (0.07)	0.24** (0.10)	0.39*** (0.11)	0.01 (0.08)	0.22*** (0.08)	0.18* (0.10)
Observations	4,816	2,918	2,324	2,492	2,826	1,626
R-squared	0.00	0.00	0.01	0.00	0.00	0.00

Note: This table shows individual-level OLS regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the race-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.1.3 Replication of Table 3 (with standard errors clustered by respondent): Voter Responsiveness to Candidate Appearance in Statewide General Elections

DV: Vote Republican indicator variable	(1) All respondents	(2) Matched on race and gender	(3) Low knowledge	(4) High knowledge	(5) Weak/indep.	(6) Strong partisan
Treatment	-0.09* (0.05)	-0.11 (0.07)	-0.20*** (0.07)	0.01 (0.07)	-0.14** (0.06)	-0.05 (0.08)
Appearance advantage (for Republican)	-0.03 (0.05)	0.03 (0.08)	-0.06 (0.07)	-0.01 (0.07)	-0.05 (0.07)	-0.07 (0.07)
Treatment x Appearance advantage (for Republican)	0.20*** (0.07)	0.24** (0.12)	0.39*** (0.11)	0.01 (0.10)	0.22** (0.10)	0.18* (0.11)
Observations	4,816	2,918	2,324	2,492	2,826	1,626
R-squared	0.00	0.00	0.01	0.00	0.00	0.00

Note: This table shows individual-level OLS regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the respondent-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.2: Political Knowledge Measures for Study 2

We have a few questions about the federal government. Many people don't know the answers to these questions. If you don't know, just say so.

Do you happen to remember what job John Boehner holds?

- Speaker of the US House
- Governor of Texas
- Chief Justice of the US Supreme Court
- Vice President of the United States
- Prime Minister of Canada
- I'm not sure

Do you happen to know which party has the most members in the US House of Representatives right now?

- Republicans
- Democrats
- Neither
- I'm not sure

Do you happen to remember what political party President Franklin Roosevelt was a member of?

- Democratic Party
- Republican Party
- Whig Party
- Some other party
- I'm not sure

Do you happen to remember what industry the Dodd-Frank Act regulates?

- Finance
- Healthcare
- Insurance
- Oil and Natural gas
- I'm not sure

Of the 1,933 respondents, 66% correctly answered the Boehner question, 68% did so for the U.S. House of Rep. question, only 59% new FDR's party, and only 45% said that Dodd-Frank regulates finance. 51% answer three or more questions correctly and only 28% answered all four correctly.

4.3: Replication of Table 4 for Down Ballot Races

While we find some evidence in the paper that partisanship inoculates people against appearance effects, the results in Table 4 fail to clearly support this finding. When we restrict the analysis to down-ballot races (i.e., excluding senate or gubernatorial races) we get a similar result: candidate appearance still has a substantively large (though not statistically significant) effect on the vote choice of low knowledge, strong partisans. When we perform the same analysis dividing the sample by respondents' level of knowledge about local candidates (see section 4.4), we also find that strong partisanship alone fails to inoculate voters against candidate appearance. However, local knowledge absent strong partisanship also fails to shield voters' from candidates' looks. Across all specifications in the paper, and in the following tables (sections 4.3-4.5), we consistently find that candidate appearance influences low-knowledge, non-strong partisans. In contrast, we never find an effect among voters with both high knowledge and strong partisanship. In total, these findings imply that if voters had more information, whether in the form of political knowledge or strong partisanship, they would not rely on candidate appearance.

Is Partisanship Protective in Down-Ballot Races? Study 2: Replication of Table 4 for Down Ballot Races

	(1)	(2)	(3)	(4)
DV: Vote Republican indicator variable	Low knowledge & Non-strong partisan	Low knowledge & Strong partisan	High knowledge & Non-strong partisan	High knowledge & Strong partisan
Treatment	-0.17** (0.08)	-0.12 (0.13)	-0.10 (0.11)	0.12 (0.11)
Appearance advantage (for Republican)	0.01 (0.12)	0.04 (0.13)	-0.07 (.)	-0.07 (0.10)
Treatment x Appearance advantage (for Republican)	0.35*** (0.13)	0.25 (0.22)	0.13 (0.13)	-0.03 (0.07)
Constant	0.31*** (0.08)	0.15 (0.10)	0.36*** (0.06)	0.21** (0.09)
Observations	821	406	778	558
R-squared	0.01	0.01	0.00	0.02

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Standard errors clustered at the individual and race-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.4 Replication of Table 4 with Local Knowledge

Respondents were coded as high in local knowledge if they answered more than 50% of knowledge questions about their local candidates correctly.

	(1)	(2)	(3)	(4)
DV: Vote Republican indicator variable	Low knowledge & Non-strong partisan	Low knowledge & Strong partisan	High knowledge & Non-strong partisan	High knowledge & Strong partisan
Treatment	-0.09 (0.07)	-0.07 (0.07)	-0.26** (0.11)	0.01 (0.15)
Appearance advantage (for Republican)	-0.07 (0.11)	-0.09 (0.08)	0.01 (0.10)	-0.02 (0.09)
Treatment x Appearance advantage (for Republican)	0.16 (0.11)	0.24*** (0.08)	0.36** (0.17)	0.01 (0.20)
Constant	0.33*** (0.07)	0.21*** (0.06)	0.35*** (0.08)	0.18** (0.08)
Observations	1,951	1,122	875	504
R-squared	0.00	0.01	0.01	0.00

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Standard errors clustered at the individual and race-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.5: Fixed Effects Analysis for Study 2

Since six of 18 states had multiple state-wide elections, some respondents in Study 2 voted for multiple candidates (636 of 1,885). This fact allows us to conduct another individual-level test of appearance voting: do we find a similar effect when we only examine within-subject variation in appearance voting? More precisely, do respondents in the photo condition vote more based on faces than those in the no-photo condition when we only examine variation in vote choice within respondent, not across respondents. We conduct this test by including indicator variables for each respondent (respondent fixed effects). We also include race indicator variables (race fixed effects) so that we only examine within-respondent variation within races, not across races. Since we include these fixed effects, we cannot estimate the treatment indicator (it is collinear with respondent-fixed effects) or appearance advantage (it is collinear with race-fixed effects). We can of course continue to estimate the interaction between the treatment indicator and appearance advantage, which is the coefficient of interest. The first table on the next page presents this fixed-effect estimate. It reassuringly yields an almost identical coefficient, 0.19, but the estimate is now even more precisely estimated, presumably because we have excluded much irrelevant variation across respondents in vote choice. Furthermore, when we only examine races matched on race and gender, Column 3 shows the estimate rises to a large 0.34. Since we only examine variation within respondent, these results provide a robustness check.

The next four columns of the first table on next page test the low-information heuristic predictions. As expected, candidate appearance affects low information voters (Column 3) far more than high knowledge voters, who exhibit no significant effect of appearance advantage on vote choice (Column 4). Similarly, candidate appearance affects weak partisans and independents (Column 5), but not strong partisans (Column 6).

When we examine the impact of candidate appearance on vote choice among subsets of voters based on political knowledge and political partisanship with fixed effects, the results change somewhat. The second table on the next page shows those results, replicating Table 4 but with respondent and race fixed effects. The main change is that low knowledge, strong partisans (Column 2) no longer rely on appearance. We don't know why but, to speculate, voters may be willing to try to cast an informed vote in Senate and gubernatorial races but end up relying on appearance. In down-ballot races, however, they may not be willing to try and so just vote with their party. The six states where we have multiple races—and so can estimate within-respondents—tend to have mostly down-ballot races.

Replication of Table 3 with Fixed Effects

DV:	(1) Vote R	(2) Vote R	(3) Vote R	(4) Vote R	(5) Vote R	(6) Vote R
	All respondents	Matched on Race and Gender	Low Knowledge	High Knowlege	Weak/ Indep.	Strong Partisan
Treatment x Appearance Advantage (for Republican)	0.19*** (0.06)	0.34** (0.15)	0.31*** (0.09)	0.09 (0.10)	0.31*** (0.09)	0.01 (0.07)
Observations	4,816	2,918	2,324	2,492	2,826	1,626
Respondents (effective in f.e.)	636	588	309	327	367	223
R-squared	0.81	0.85	0.78	0.84	0.78	0.89
Race fixed effects	X	X	X	X	X	X
Individual fixed effects	X	X	X	X	X	X

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the individual and race-level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Replication of Table 4 with Fixed Effects

	(1) Vote R	(2) Vote R	(3) Vote R	(4) Vote R
	Low Knowledge & Non- Strong Partisan	Low Knowledge & Strong Partisan	High Knowledge and Non- Strong Partisan	High Knowledge & Strong Partisan
Treatment x Appearance Advantage (for Republican)	0.417*** (0.152)	0.058 (0.123)	0.201 (0.138)	-0.054 (0.096)
Observations	1,475	694	1,351	932
Respondents (effective in f. e.)	198	91	169	132
R-squared	0.765	0.841	0.796	0.937
Race fixed effects	X	X	X	X
Individual fixed effects	X	X	X	X

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the individual and race-level in parentheses *** p<0.01, ** p<0.05, * p<0.1

Replication of Table 4 with Fixed Effects for Down Ballot Races

(Excludes Senate races and gubernatorial races)

	(1) Vote R	(2) Vote R	(3) Vote R	(4) Vote R
	Low Knowledge & Non- Strong Partisan	Low Knowledge & Strong Partisan	High Knowledge and Non- Strong Partisan	High Knowledge & Strong Partisan
Treatment x Appearance Advantage (for Republican)	0.426*** (0.147)	0.073 (0.125)	0.226* (0.128)	-0.059 (0.097)
Observations	821	406	778	558
Respondents (effective in f. e.)	198	91	169	132
R-squared	0.692	0.779	0.716	0.915
Race fixed effects	X	X	X	X
Individual fixed effects	X	X	X	X

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Constant not shown. Standard errors clustered at the individual and race-level in parentheses *** p<0.01, ** p<0.05, * p<0.1

We only have two states with Senate and gubernatorial races (see SI section 2.1), so we don't replicate the table above for these up ballot races.

Replication of Table 4 with Fixed Effects With Local Knowledge

	(1)	(2)	(3)	(4)
DV: Vote Republican indicator variable	Low knowledge & Non-strong partisan	Low knowledge & Strong partisan	High knowledge & Non-strong partisan	High knowledge & Strong partisan
Treatment x Appearance advantage (for Republican)	0.321*** (0.121)	-0.047 (0.093)	0.235 (0.150)	0.043 (0.169)
Constant	-0.306*** (0.108)	0.031 (0.061)	0.000 (0.000)	0.957*** (0.169)
Observations	1,951	1,122	875	504
Respondents (effective in f.e.)	198	91	169	132
R-squared	0.791	0.877	0.767	0.939
Race fixed effects	X	X	X	X
Individual fixed effects	X	X	X	X

Note: This table shows individual-level regressions. The dependent variable is coded Republican vote 1 and Democratic vote 0. Standard errors clustered at the individual and race-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.6: Individual-Level Estimates for the Primaries Experiment (Study 1)

Unlike in general elections, we would not necessarily expect partisan and high-knowledge voters to eschew candidate appearance in House primary contests. While party labels may serve as informative cues in general election races, they are significantly less valuable in primary contests, which usually feature multiple candidates from the same party. (Ten of the 14 races included in Study 1 did so.) Furthermore, voters know much less about congressional candidates than gubernatorial and senatorial candidates, like those included in the general election study (Krasno 1997). In fact, knowledge is so low in congressional primaries that even politically knowledgeable voters appear largely ignorant of candidates' positions, except for what they can glean from candidate partisanship (Ahler, Citrin, and Lenz 2014). Consequently, unlike general elections, all voters may fall back on appearance in their voting decisions, even politically knowledgeable strong partisans, because they know so little else about the candidates and cannot rely on party.

Since the primary featured multiple candidates from the same party in most races, we cannot use the individual-level dependent variable from the general elections—Republican vote choice—in analyzing the experimental results from the House primaries. Instead, we assign respondents values on the dependent variable based on the appearance of the candidate they voted for. We assign a vote for the candidate with the best-rated appearance in each district to 1 and a vote for the candidate with the worst to 0, and assign votes for candidates in between values according to a simple formula. We assign all other candidates a value as follows: $(\text{candidate rating} - \text{district minimum rating}) / (\text{district maximum rating} - \text{district minimum rating})$. The dependent variable in this analysis, then, is the average relative (within-district) appearance of vote choices. We include district fixed effects and cluster the standard errors at the district level. We measure general political knowledge with a 4-item scale and classify respondents as highly knowledgeable if they answered three or more correctly (see SI section 4.7 for wording).

We present this individual-level analysis in the table on the next page. The first column shows the treatment effect among all participants. The coefficient of 0.05 indicates that, on average, participants randomly assigned to the ballot with photographs voted for candidates who were 0.05 points more appearance-advantaged on a one-point scale than did participants assigned to the standard ballot (95% CI 0.02 to 0.08).

Consistent with our expectations in the primary context, voters appear roughly equally susceptible to candidate appearance regardless of partisanship and knowledge. As the table reports, high-knowledge and low-knowledge voters who saw the photo ballot voted for appearance-advantaged candidates more frequently. The difference in treatment effects between these two groups is not statistically significant. Similarly, we find that strong partisans (those who self-place at 1 or 7 on the 7-point party ID scale) vote for appearance-advantaged candidates at a similarly elevated rate to that of independents and weak partisans when they see photos of the candidates.

Voter Responsiveness to the Appearance Experiment in House Primaries

	(1)	(2)	(3)	(4)	(5)
DV: Vote appearance	All respondents	Political knowledge		Partisanship	
		Low	High	Weak/indep.	Strong
Treatment	0.05*** (0.02)	0.07** (0.03)	0.05** (0.02)	0.05** (0.02)	0.05*** (0.02)
Constant	0.37*** (0.03)	0.36*** (0.03)	0.38*** (0.03)	0.39*** (0.03)	0.35*** (0.03)
Observations	851	267	538	432	386
R-squared	0.29	0.26	0.33	0.35	0.27
District f.e.	X	X	X	X	X

Note: This table shows individual-level regressions. The dependent variable is the average relative (within-district) appearance of vote choices. Robust standard errors clustered at candidate level in parentheses (see note on Table 1 for more detail). *** p<0.01, ** p<0.05, * p<0.1.

4.7: Political Knowledge Measures for Study 1

We constructed our index measuring political knowledge from four items. Three of these items were multiple choice: identification of Senate Majority Leader as Harry Reid's political office, identification of Treasury Secretary as Timothy Geithner's political office, and identification of the Supreme Court as the institution with the responsibility to decide if a law is constitutional or not. We constructed our fourth item from seven-point party placement questions. Respondents who correctly placed the Democratic Party to the left of the Republican Party scored 1 on this item, while those who incorrectly placed the parties scored 0. Cronbach's alpha for these four items is 0.73. The distribution of political knowledge based on this index is described as follows: 5.7% of respondents scored 0 out of 4, 14.5% scored 1 out of 4, 21.7% scored 2 out of 4, 25.7% scored 3 out of 4, and 32.4% answered all questions correctly. We classified respondents as “high knowledge” if they scored at least 3 out of 4.

5. Additional Analyses and Material

5.1: Voters Claim to Ignore Appearance in Voting Decisions

Respondents in Study 2 (General Election) were asked how much candidate appearance, personality, the economy, candidates' party, and candidates' issue positions influenced their vote choices with options ranging from not at all to a great deal (coded 1-4).

Voters choose candidates for many different reasons. Please tell us how much each of the following factors influenced your choices on this survey:

	A great deal	Somewhat	A little	Not at all
Candidates' appearance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candidates' personality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The economy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candidates' party	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candidates' issue positions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Of the five reasons, respondents placed the least weight on candidate appearance:

Variable	Obs	Mean	Std. Dev.	Min	Max
Appearance	1933	1.70	0.83	1	4
Personality	1933	2.58	0.92	1	4
Economy	1933	3.26	0.86	1	4
Party	1933	2.87	0.98	1	4
Position	1933	3.61	0.74	1	4

Half of respondents say appearance should not matter at all, with 82% saying a little or not at all. This result is not an artifact of respondents in the treatment condition understanding the subject of the study and reporting that appearance is unimportant. Respondents in the treatment condition actually placed greater weight on candidate appearance than respondents in the control condition (mean difference = 0.14, 95% CI [.07, 0.22]).

5.15: Multicandidate and Retention Races from 2012 Statewide General Election Survey

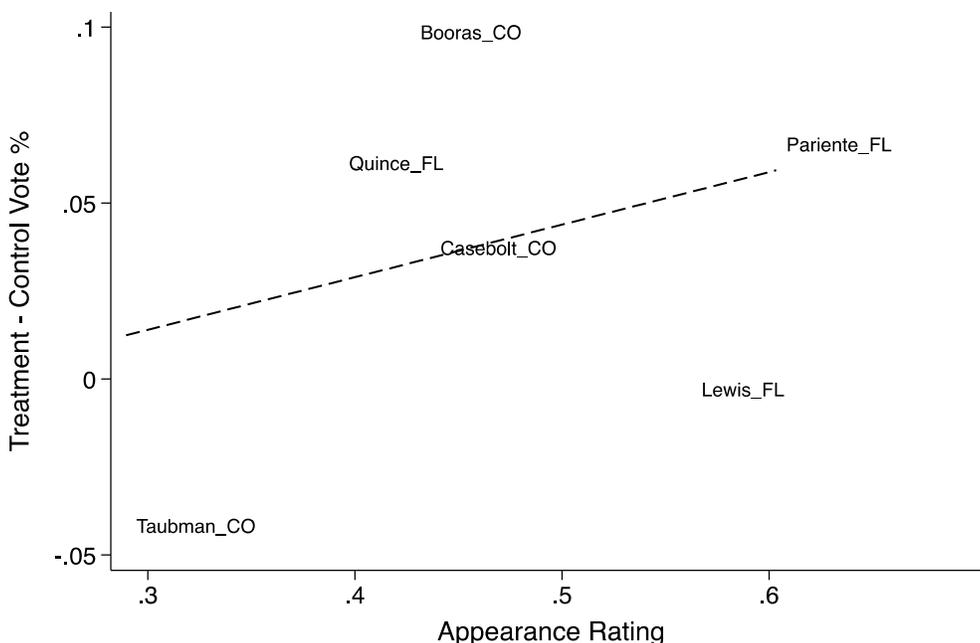
In Study 2, the general election survey, we asked respondents in a handful of states about their preferences in retention races (which feature only one candidate) and in races with more than two candidates. We do not present the results of the experiment in these races in the paper because the analysis is necessarily different from the two-candidate races. We cannot use “vote for the Republican” as the dependent variable. As such, race-level analysis is more complicated. Candidate-level analysis (as used in the primary study) is more easily interpreted, so we present those results here. Given that we only have a handful of such races, the estimates are necessarily imprecise. They are, however, consistent with the overall results in the paper, that is, appearance advantaged candidates tend to win more votes in retention and multicandidate races (in the photo condition compared to the control condition).

Single Candidate Retention Elections

The general election survey asked voters in two states about three retention races each. Voters in Colorado reported whether they wanted to retain three judges in the Colorado Court of Appeals, while voters in Florida reported whether they wanted to retain three justices from the Florida Supreme Court.

In these races, our candidate-level dependent variable is the difference in vote share a candidate received under the two conditions, $\text{Vote\%}_T - \text{Vote\%}_C$. Since candidates in retention elections lack opponents, we use appearance rather than appearance advantage as the independent variable. (This is the average score the candidate received in the MTurk rating survey. Recall that we asked respondents, “How good of an elected official do you think this person would be?”) We rescale both variables 0-1.

As shown in the scatterplot below, we observe the expected positive trend between $\text{Vote\%}_T - \text{Vote\%}_C$ and appearance. This apparent upward slope is imprecisely estimated because we only have six candidates in these races, so it fails to reach statistical significance. However, it is substantively large. A bivariate regression of $\text{Vote\%}_T - \text{Vote\%}_C$ on appearance demonstrates that this slope is 0.15 ($p = 0.52$). When we only examine within-state variation (state fixed effects), the estimated coefficient is 0.20 ($p = 0.57$).

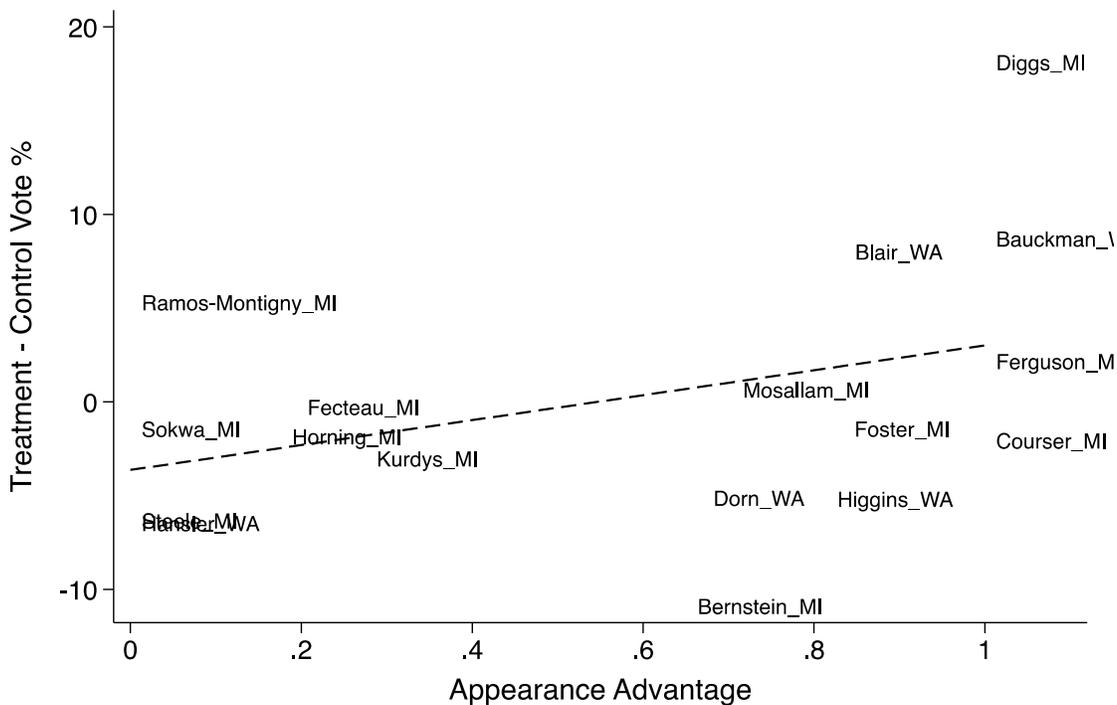


Multiple Candidate Races

The general election survey asked respondents in Michigan about three races featuring four candidates (State Board of Education, University Board of Regents, and Michigan State University Board of Trustees) and respondents in Washington about the race for Superintendent of Public Instruction, which featured five candidates.

The dependent variable in these races is the difference in vote share a candidate received under the two conditions, $\text{Vote\%}_T - \text{Vote\%}_C$. The independent variable is the candidate's within-race appearance advantage, $(\text{face rating} - \text{race minimum face rating}) / (\text{race maximum face rating} - \text{race minimum face rating})$. Both variables are scaled 0-1.

As shown in the scatterplot below, we observe the expected positive trend between $\text{Vote\%}_T - \text{Vote\%}_C$ and appearance. This apparent upward slope is imprecisely estimated because we only have seventeen candidates in these races, but we come closer to conventional levels of statistical significance here than in the single-candidate races. The estimated slope from the regression of $\text{Vote\%}_T - \text{Vote\%}_C$ on appearance advantage is 0.07 ($p = 0.12$), implying that the most appearance advantaged candidate would be expected to see a net seven-point benefit on the ballot with photographs over the most appearance disadvantaged candidate. The estimated apparent effect is consistent when we only examine within-race variation (race fixed effects) ($b = 0.07$, $p = 0.15$).



5.2 Out-Of-State Experiment

As we explain in the paper, we only ran the faces experiment in states with larger populations, large enough to get a reasonable sample size on Mechanical Turk. Since Mechanical Turk does not allow requesters to specify the state of respondent, we actually interviewed respondents in all states. For respondents in smaller states, we presented them with the following language:

We don't have races in your state. So we are going to ask you about races in another state.

Please just answer the questions as best as you can. **No need to look up candidates or answers!**

The behavior of these respondents in the experiment is potentially interesting. Since they presumably know even less (and usually nothing) about the candidates, they might fall back on appearance to a greater extent. On the other hand, they may just vote for their preferred party (the ballot showed candidate party).

We interviewed 1688 respondents in these smaller states without races in the experiment. Of these, we lose 6% because they failed the attention test and 10% because they said they would not vote (we drop these respondents for comparability, they could still vote in congressional or local races), leaving 1410 respondents.

For respondents in the experiment we present in the paper (in-state respondents), we have 44 races with two candidates. For the out-of-state experiment, we have only 42 races because we chose to not send these respondents to two of the states that had only one race: the New York Senate (Gillibrand versus Long) race and the Ohio Senate race (Mandell versus Brown).

The results for out-of-state respondents look generally similar to the results for in-state respondents. On the next page, we replicate Figure 3 for the out-of-state respondents, which shows a scatterplot of the relationship between appearance advantage and the difference in vote share between the treatment and control groups. The plot generally resembles the in-state plot (that is, Figure 3 in the paper). On the following page, we replicate the regressions from Table 3 for the out-of-state respondents. The overall effect size is somewhat smaller, 0.15 out-of-state versus 0.20 in-state. Given the smaller sample sizes, the estimates are generally noisier and not always statistically significant in the other columns.

Overall, these results suggest that people did not rely more on faces when they knew less about the candidates.

Replication of Table 3 for out-of-state respondents: Appearance advantage in the General Election

Dependent variable: Picture condition minus no-picture condition vote share	(1)	(2)	(3)	(4)	(5)
	All	Matched on race and gender	All	Senate & Governor	Other
Appearance Advantage (for Republican)	0.152** (0.058)	0.123 (0.075)	0.120* (0.062)	0.033 (0.091)	0.207** (0.075)
Incumbent			0.027* (0.015)		
Female Republican			0.012 (0.036)		
Female Democrat			0.007 (0.023)		
White Republican			0.144 (0.085)		
White Democrat			0.001 (0.078)		
Constant	-0.046 (0.032)	-0.032 (0.040)	-0.168 (0.115)	0.006 (0.051)	-0.069 (0.041)
Observations	42	24	42	16	26
R-squared	0.145	0.107	0.304	0.009	0.242

Note: This table shows candidate-level regressions. The number of respondents used in this analysis is 1,410. Dependent variable: picture condition minus no-picture condition vote share (coded so that higher values indicate greater Republican vote share). Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

5.3: Decomposition of Race-Level and Candidate-Level Effects in Studies 1 and 2

The dependent variable in Studies 1 and 2 is a difference in vote share between two conditions: one in which candidate photographs are on the ballot and one in which they are not. Thus, the apparent treatment effect that we estimate is the result of candidate appearance being differentially important to voters in the two conditions.

We can decompose this treatment effect by estimating the effect of candidate appearance advantage (in the primaries) or Republican appearance advantage (in the general elections) separately in each of the two conditions. When we do so in the primary, we find a positive (albeit statistically insignificant) relationship between appearance advantage and vote share in treatment (slope = 0.03, 95% CI -0.22 to 0.27). Interestingly, we find a negative relationship between appearance advantage and vote share in control (slope = -0.18, 95% CI -0.47 to 0.11).

When we do this in the general election study, we find a positive relationship between Republican appearance advantage and Republican vote share in treatment (slope = 0.08, 95% CI -0.13 to 0.28) and essentially no relationship in control (slope = -0.02, 95% CI -0.26 to 0.22). Interestingly, the increasing importance of appearance advantage in the control group that we document in the overestimating/underestimating section is a result control voters shifting from an apparent preference for appearance *disadvantaged* candidates to a lack of preference either way.

There are many potential explanations for these findings of a negative-to-zero relationship between appearance advantage and vote share in the control, which do run somewhat counter to previous observational findings of an appearance effect. Previous studies generally investigate high-salience, high-information races in which incumbents face quality opponents (e.g., senate races). By contrast, House primaries and down-ballot state races are low-salience, low-information contests in which incumbents may be of much higher quality than their challengers. It's also possible that appearance-disadvantaged incumbents are better politicians, as they may have to rely on things other than appearance to win re-election. While somewhat far-fetched, this is one potential explanation for the apparent pattern in the primaries and in the over-time pattern of results in the general election. Of course, it's also possible that these results are due to chance.

5.4: Reanalysis of Atkinson et al.

To further analyze whether candidate appearance could be influencing election outcomes, we reanalyzed the Atkinson et al. data on Senate elections from 1992 to 2006. To make the analysis more consistent with our experiment, we make somewhat different decisions about data and variables:

- Most importantly, we use a different DV. Atkinson et al. use individual-level exit poll data. We use actual Senate election results, so the analysis below is at the race-level, not the individual level. Specifically, we use the Democratic Senate candidate's vote margin over their opponent.
- Competence advantage is Democratic candidate facial competence minus Republican candidate facial competence, as measured by Atkinson et al. They estimate coefficients separately for incumbents and challenger appearance. We prefer our specification only because it more closely matches the general election analysis in our paper – the results are similar either way.
- Following Atkinson et al., Cook rating of competitiveness takes eight values, from solid Republican to solid Democrat.
- Other controls: Following Atkinson et al., we include year fixed effects. We tried controlling for incumbency, but the Cook report variable makes the effect of incumbency vanish, or sometimes even reverse sign. Including incumbency does not change the effect of appearance when controlling for Cook report. We also tried controlling for the partisan tendency of each state, using the average Democratic vote for president in the previous four elections, but the results were essentially unchanged.
- Atkinson et al. examined only Senate races with an incumbent, not open seats. We examine both, but the results are very similar if we exclude open races (and similar if we only examine open races).

The regression table on the next page shows three specifications, all of which find a positive coefficient on competence advantage, which is always statistically distinguishable from zero. The most relevant model is column 3, which finds that a shift from the most appearance disadvantaged to the most advanced corresponds with a 26 percentage point increase in the Democrats margin over the Republican.

Atkinson et al. find that the appearance effect, as measured by challenger appearance, does not change any Senate election results. With the estimates presented on the next page, do any change?

To calculate the number of race outcomes that would change when we use both candidates' faces and examined all 190 races, we calculated the effect of changing the Democrat's appearance advantage from 0.5 to the actual advantage observed (on a 0-1 variable), so we're examining the effect of changing from the midpoint to the observed value. We do so for the model in column 3 (on the next page). Instead of changing no outcomes, it changes 13 of 190 races or 6.8%. If we look at races that Cook rated as tossups, it's six of 34 (17%). If we look at races Cook rated as tossups or leaning, it's 10 of 75 (13%).

Since they focus on races with an incumbent, we reestimated the model in column 3 for those races. Based on these estimates, the appearance effect would change nine of 146 races (6%). If we limit to those cook classified as tossups, it is four of 15 (26.6%).

So, in competitive races (rated at least one year before Election Day), appearance is affecting 17-27% of the outcomes, which seems nontrivial.

After presenting the regression results on the next page, we then show a residual plot on the following page.

Reanalysis of Atkinson et al. for Senate elections, 1992-2006

	(1)	(2)	(3)
DV: Democratic Senate candidate vote margin			
Competence advantage (for Democrat)	0.63*** (0.091)	0.65*** (0.092)	0.26*** (0.058)
Cook rating of competitiveness (for Democrat)			0.55*** (0.030)
1994.year		-0.095 (0.075)	-0.087** (0.044)
1996.year		-0.13* (0.074)	-0.011 (0.044)
1998.year		-0.091 (0.078)	-0.025 (0.046)
2000.year		-0.093 (0.073)	0.0059 (0.043)
2002.year		-0.14** (0.070)	-0.037 (0.042)
2004.year		-0.047 (0.072)	0.012 (0.042)
2006.year		0.040 (0.074)	0.081* (0.043)
Constant	-0.35*** (0.053)	-0.29*** (0.073)	-0.42*** (0.044)
Observations	190	190	190
R-squared	0.204	0.253	0.743
SER	0.24	0.24	0.14

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Residual plot for all races based on column 3

This plot controls for race competitiveness, as measured by Cook ratings.

