Candidate Faces and Election Outcomes[1]

Prepared for Publication on Andrew Gelman's Blog

Matthew Atkinson Ryan D. Enos Seth J. Hill[2]

[1]The authors thank Mike Franks, Andrew Gelman, R. Brian Law, Jeff Lewis, Elisabeth Michaels, David Sears, Alexander Todorov, Lynn Vavreck, John Zaller, the UCLA Workshop on Political Methodology, and our survey participants from February, March, and May 2007.

[2]All authors: Department of Political Science, University of California, Los Angeles; matthewa@ucla.edu, renos@ucla.edu, sjhill@ucla.edu.

Abstract

Recent research finds that inferences from candidate faces predict aggregate vote margins. Many have concluded this to mean that voters choose the candidate with the better face. We implement a survey with participant evaluations of over 167,000 candidate face pairings. Through regression analysis using individual- and district-level vote data we find that the face-vote correlation is explained by a relationship between candidate faces, incumbency, and district partisanship. We argue that the face-vote correlation is not just the product of simple voter reactions to faces, but also of party and candidate behavior that affects which candidates compete in which contests.

Recent work in psychology demonstrates that the naive, rapid evaluations by survey participants of the facial competence of candidates linearly predict United States Congressional candidate vote share (Todorov, Mandisodza, Goren, & Hall 2005). Todorov et al. (2005) showed college students the faces of candidates contesting United States House and Senate elections in 2000, 2002, and 2004, and asked them to unreflectively choose the more competent looking candidate in each contest. They find that the candidate more frequently selected as appearing competent won the actual election in 66.8 percent of House and 71.6 percent of Senate contests. They also find that the proportion of paired evaluations in which one candidate's face is judged more competent than the opponent's correlates to the difference in vote share between the two candidates in both the House ($r = .40$) and in the Senate ($r = .44$).

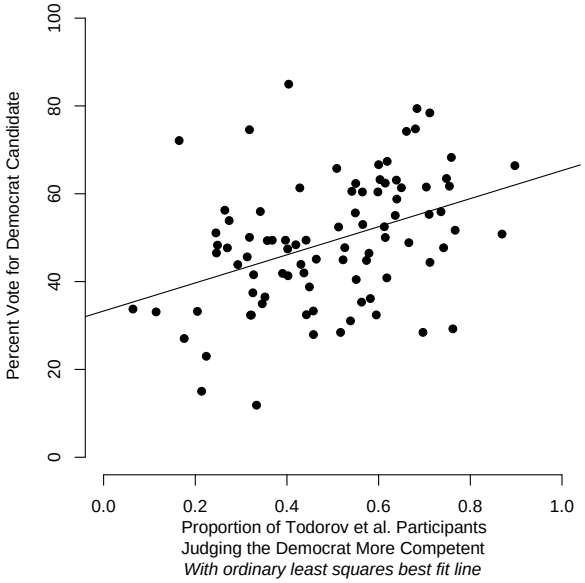
In Figure 1 we reproduce Todorov et al.'s (2005) main finding using participant evaluations to predict Democratic vote share.[1] Similar findings have subsequently shown that candidate faces predict gubernatorial elections (Benjamin & Shapiro 2006, Ballew II & Todorov 2007) and that executive faces predict corporate profits (Rule & Ambady 2008).

Although Todorov et al. (2005) were careful to discuss their results as predictive rather than causal, there has been a tendency among popular commentators and some scholars to interpret the results as an indictment of the ability of American voters to make reasoned decisions. An item in *The New York Times Magazine*, writing about Little, Burriss, Jones, & Roberts (2007), declared that "Faces Decide Elections" (Skloot 2007), and a *National Public Radio* segment suggested we "[f]orget political polls ... voters prefer candidates who look competent, even if they are not" (Hamilton 2005). In fact, some have implied the need for electoral reform so that voters will not be duped by candidates' faces (Zebrowitz & Montepare 2005).

Todorov et al. (2005) measure relative candidate facial competence in a laboratory environment. This innovative approach to measuring a candidate characteristic may have led

[1]We thank Alex Todorov for generously sharing with us his data.

Figure 1: The Relationship of Vote Share to Inferred Competence, United States Senate Elections 2000-2004.



some secondary authors to perceive that the correspondence between facial competence and vote share was based on experimental evidence. In fact, the predictive results presented in Todorov et al. (2005) are observational. An experiment to identify the effect of face on vote would randomly allocate faces to electoral contests. In the absence of this randomization, we investigate the effect of candidate facial competence using statistical control variables.

We find evidence that candidate facial competence has a small but significant effect on individual-level vote choice. But despite the variable's predictive power, its estimated influence is small compared to other political variables such as incumbency and partisanship. Furthermore, we find evidence that candidate faces are not distributed randomly across districts because political circumstances influence the allocation of facial competence to electoral contests. We argue that the most important role of faces in elections operates at the elite-level, not at the level of individual voters responding to candidate faces.

Incumbency, Partisanship, and the Congressional Vote

Political science research on elections generally emphasizes a standard set of causes for congressional election outcomes and vote choice such as incumbency, partisanship, and the economy. Perhaps most importantly, the incumbency status of a member of the House or Senate is a consistent electoral advantage. Nearly 90 percent of House incumbents who stand are reelected, and researchers have estimated that incumbency is worth 4 to 6 percentage points of vote share in House elections (Erikson 1971, Mayhew 1974, Gelman & King 1990, Jacobson 2004). For individual voters and at the district level, party identification consistently predicts the presidential and congressional vote choice (Campbell, Converse, Miller, & Stokes 1960, Jacobson 2004).

In spite of the focus on these political factors, American political behavior research has a long tradition of considering candidate-specific factors. Most prominently, the authors of *The American Voter* provide a framework for understanding the vote choice that integrates political and candidate-specific factors in what they term the "funnel of causality" (Campbell et al. 1960, ch. 2). They describe voting behavior as an output from a "multitude of prior factors" that creates a "converging sequence of causal chains." An analyst can work backwards from the actual vote event to causes of greater and greater temporal distance — from election day activities to long-term predispositions such as partisanship. The authors believed that each vote could theoretically be dissected into its constituent causal parts.

The "funnel of causality" framework implies that a factor influencing voting behavior should be studied in terms of its immediate direct effects and in terms of the causally prior factors. Because they fall near the end of the funnel of causality, candidate characteristics such as facial competence may be influenced by prior causal factors. Jacobson & Kernell (1983), for example, suggest that the types of candidates who choose to run in a given election are caused by external political and economic conditions.

To identify the effect of candidate faces in the presence of other factors that may cause variation in the allocation of faces, we require a measure of facial competence that enables

comparison across contests and candidates. The measure of Todorov et al. (2005) does not provide this because it only compares the candidate faces in each specific contest. We implement a new survey to measure the individual competence of a set of candidate faces, pairing each face against faces from many other contests. Our measure is the result of a "round robin" tournament where two randomly drawn faces are paired against one another, rather than a single set of a head-to-head contests. This enables us to place each face in our pool of candidates on a common scale of facial competence and use these measures to assess the interplay of face and politics.

Our contribution to the understanding of elections is threefold. First, taking account of candidate face and political variables together, we find that district partisanship and incumbency remain strong and consistent predictors of the congressional vote and that the effect of facial evaluations is much smaller than some interpretations have suggested. Second, using our original measure of facial competence, we present evidence that candidate facial competence, incumbency, and partisanship are correlated in a way that suggests strategic behavior by political elites. Third, we use surveys to find the direct effect of candidate face on the individual vote choice. We show that the direct effect of face is small, varies by the partisan identification of the voter, and that the remaining effect is driven by the face of the challenging candidate.

Candidate Faces in Congressional Elections

In this section, we estimate the effect of candidate facial competence on vote share. We begin our evaluation of the effect by comparing the bivariate relationship of facial competence with vote from Todorov et al. (2005) against models predicting vote share based on incumbency status and district partisanship. Following Todorov et al. (2005), we analyze data from the 2000, 2002 and 2004 Senate elections and the 2004 House elections.[2] We operationalize

[2]Because we use presidential vote share to measure district partisanship, we omit the 2002 House elections due to redistricting. For the House analysis in this paper, we include contests in which both major party

difference in facial competence using the proportion of evaluations in which the Todorov et al. (2005) participants chose the Democratic candidate as more competent than the Republican candidate.

Because the method of Todorov et al. (2005) does not have experimental control, the strong predictive power of the facial competence variable may not accurately reflect the true causal influence on vote choice. To be precise, the estimated effect of facial competence will be biased if it is correlated with some variable omitted from the model, for example district partisanship or incumbency, that causes election outcomes.

In fact, participant ratings comparing the Democrat and Republican candidate faces are positively correlated with both incumbency and partisanship. For both House and Senate candidates, the correlation between the proportion of Todorov et al. (2005) participants who choose the Democrat more competent and Democratic presidential vote is about $r = .2$. The correlation with a dichotomous indicator variable for a Democratic incumbent is about $r = .5$ and with a Republican incumbent indicator is about $r = -.4$.[3]

In order to identify the relative contributions of face, incumbency, and partisanship, we present regression models predicting Democratic vote share in Table 1. For each chamber, the first column is a model with facial competence difference only. Our regression model in column one replicates the impressive correlations between face and vote noted in Todorov et al. (2005). Moving from none of the participants choosing the Democratic face more competent to all of the participants choosing the Democratic face more competent is associated with an estimated 29 point increase in Democratic vote share in the House, and an estimated 32 point increase in the Senate. These are surprisingly large effects.

The second and third columns for each chamber, which use incumbency and district partisanship, present models that also predict large differences in candidate vote share.

candidates were white males. This is due to power concerns for our survey measurement tool to be described later in the paper, and as a simple attempt to control for potential race or gender effects that might influence both vote share and facial competence inference.

[3]Correlations with incumbency are polyserial as the incumbency variable is dichotomous and the competence variable is numeric (Fox 2007), while the correlation to district partisanship is Pearson's r.

Moving from a Democratic incumbent to a Republican incumbent is associated with a change in Democratic vote of almost 30 points in the House and of 25 points in the Senate. In the third column, the estimated effect of district partisanship is also large, with each additional point of presidential vote translating into about a point of House and Senate vote share. In both House and Senate, the incumbency-only model and the partisanship-only model explain a much greater amount of variance in vote shares than the faces-only model.

In the fourth column for each chamber we present a multiple regression model to jointly-estimate the effect of faces, incumbency, and partisanship. When controlling incumbency and district partisanship, the estimated effect of facial competence on vote share decreases more than 5-fold in the House and in the Senate, and is no longer statistically significant. In both the House and Senate the estimated effect of incumbency and partisanship are both decreased. Moving from a Democratic incumbent to a Republican incumbent is now worth about 23 points in the House and 21 points in the Senate, and presidential vote now translates about two-to-one into congressional vote.

The results of Table 1 suggest that the bivariate comparison of face and vote suffers from omitted variable bias and that incumbency and partisanship remain important causes of the congressional vote. They do not necessarily indicate that faces do not matter, as in both cases the coefficient remains positive and of meaningfully substantive size. The greater uncertainty in the estimate and the small effect compared to the effect of other variables leads us, however, to question the claim that "faces decide elections."

The correlation of candidate facial competence to incumbency, partisanship, and the vote outcome indicates that faces must matter somewhere along the causal chain. We next discuss research that indicates ways in which faces might matter. The research suggests that to understand the relationship of face to vote, we must move beyond the relative comparison of candidate faces and consider the relationship of individual candidate faces to political context. When we control for the incumbent candidate's facial competence, we find evidence that challenger facial competence has an effect on individual-level vote choice, but only when

Table 1: Using Facial Competence, Incumbency, and Partisanship to Predict Democratic Vote Share.

	House				**Senate**			
	Faces Only	Incumbency Only	Partisanship Only	All	Faces Only	Incumbency Only	Partisanship Only	All
Intercept	33.44	41.52	1.27	20.80	33.30	48.03	5.93	19.36
	(2.83)	(1.49)	(4.17)	(2.47)	(4.13)	(2.23)	(7.05)	(5.50)
Todorov Democrat Competence	28.86			4.43	32.00			6.65
	(5.56)			(2.28)	(7.86)			(5.25)
Democrat Incumbent		22.80		15.42		14.31		10.76
		(1.76)		(1.58)		(2.76)		(2.53)
Republican Incumbent		-6.71		-8.09		-10.70		-9.69
		(1.71)		(1.32)		(2.72)		(2.36)
District Partisanship			1.02	0.50			0.93	0.57
			(0.09)	(0.05)			(0.15)	(0.11)
N	142	142	142	142	89	89	89	89
R^2	0.16	0.80	0.47	0.89	0.16	0.59	0.31	0.70
Adjusted R^2	0.16	0.80	0.47	0.88	0.15	0.58	0.30	0.69
Std. Error of Regression	14.09	6.84	11.17	5.27	13.49	9.48	12.24	8.17

Ordinary least squares with standard errors in parentheses. Dependent variable is Democratic Congressional vote share for 2004 House races or 2000, 2002 and 2004 Senate Races. District partisanship measured by contemporaneous (2000, 2004) or lag (2002) Democratic two-party presidential vote share.

each candidate's face is considered as a separate factor.

The Association Of Political Factors and Facial Competence

The causes of the association between incumbency and district partisanship and facial competence are suggested by findings in psychology about individuals' ability to infer traits from faces and in economics on physical appearance and vocational success. Social science research has consistently shown that facial characteristics are associated not just with higher salaries, but with greater human capital as well (Biddle & Hamermesh 1998, Mobius & Rosenblat 2006, Hamermesh & Biddle 1994, Rule & Ambady 2008). Among the more innovative recent studies demonstrating the connection between human capital and facial appearance, Zebrowitz, Hall, Murphy, & Rhodes (2002) asked research participants to evaluate the intelligence of individuals in photographs of whom the participants had no prior knowledge. Zebrowitz et al. (2002) found that attractiveness not only correlated to perceived intelligence but correlated to actual I.Q. test scores. Hamermesh & Biddle (1994, p.1191) report in their article on physical appearance and vocational success: "The effects of an individual's looks on his or her earnings are very robust." In short, facial characteristics are associated with a variety of measures of human capital and success.

These insights from psychology and economics motivate several possible explanations for the association between facial appearance and political success. One explanation for the correlation of facial competence to incumbency is that incumbents have higher human capital. If it is true that most challengers are low-quality candidates (Jacobson 2004), then it should also be true that incumbents will have greater average levels of human capital. If incumbents have higher levels of human capital than challengers, then the economics and psychology findings would indicate that, on average, incumbents will have better facial characteristics than challengers.

The correlation of face to district partisanship may also have a human capital-based explanation. If individuals with more human capital, on average, have better facial charac-

teristics, then the candidate advantaged by the district's party identification should tend to have better facial characteristics. We might see this for two, non-mutually exclusive reasons: strategic candidate behavior or strategic party behavior.

The strategic candidate behavior hypothesis focuses on the costs and benefits available to prospective congressional candidates. Prospective candidates with more human capital — and by implication better average facial qualities — encounter higher opportunity costs by participating in a congressional contests that they do not win. Thus strategic candidates with better faces may be more likely to enter a contest when the expected probability of winning is higher and less likely to enter a contest when the expected probability of winning is lower (Jacobson & Kernell 1983). Therefore, in districts with lopsided partisan proclivities or popular incumbents, the disadvantaged challenger would be less likely to possess high facial competence.

The strategic party behavior hypothesis focuses on the costs and benefits parties face in trying to recruit quality congressional candidates. Strategic political parties may choose to recruit with more effort in congressional contests that they expect will be the most competitive (Herrnson 1986, Jacobson 1996). This effort may lead to higher-quality faces in the competitive districts for at least two reasons. Parties may anticipate that a candidate with better facial qualities will appeal to voters and therefore select explicitly on candidate appearance. Or, in the process of recruiting candidates with high levels of human capital, parties may unintentionally recruit candidates with better facial characteristics. In either case, better candidate faces would obtain in districts with the most party effort of recruitment.

That candidate face is non-randomly correlated with the partisanship of a district and incumbency is a testable proposition if the faces of candidates can be compared across districts. We next present a method that allows us to do this by separately measuring individual candidate facial competence. This measure allows us to test more complete models of the relationship between face, incumbency, partisanship, and election outcomes.

Measuring Individual Candidate Facial Competence

We created a survey to measure the perceived facial competence of candidate faces for the U.S. House in 2004, and the U.S. Senate 1990-2006 on a common scale. Todorov et al. (2005) asked respondents to compare the two faces of opposing candidates in a given election. This method measures the competence of the two faces relative to each other, head to head, but does not indicate how the two faces would compare to other faces from outside of that election.

We wrote a computer-based survey that presented to each participant two randomly-drawn faces from the pool of all candidate faces. After a one-second exposure to the image pair, the participant was asked to choose which of the two faces was more competent. The text of the question and experimental design followed as closely as possible that used by Todorov et al. (2005). Each participant evaluated hundreds of face pairs. An example of the survey can be found at `http://sjhill.bol.ucla.edu/faces`.

We conducted two separate surveys. In the first, 296 students in a lower division political science class at UCLA evaluated images of white male candidates from 2004 House elections.[4] In the second survey, 349 students from an upper division UCLA political science class evaluated images of 1990-2006 Senate candidates, of all races and genders, and, separately, the 2004 House candidates from the first survey.[5] We used images of candidate faces provided to us by Todorov et al. (2005); for other candidates, we followed the methodology described in Todorov et al. (2005) of obtaining pictures from CNN.com and supplementing them as necessary with pictures from other internet sites. We standardized these photos in size and

[4]We limited our initial survey to this subset of candidates because we were unsure of the number of evaluations needed to get a precise measure of competence. When we determined the effectiveness of the survey and estimation procedure, we were able to add more faces into the second survey.

[5]Before the first survey, participants were asked to identify the photo of the Member of Congress for the UCLA area from a lineup of photos containing members of the California Assembly as a test of candidate recognition. The participants recognized Henry Waxman at levels barely better than chance. Following the survey of Senator faces, participants were asked to identify the faces of Senators from the current Senate that they recognized. Prior to estimation, we removed the evaluations in which a participant claimed to recognize the face, as did Todorov et al. (2005). In both surveys, images of individuals that we felt had a high probability of recognition, such as members of the leadership, presidential candidates, and those with high-profile scandals were not included.

pixel count, turned all to black and white, and added a standard gray background. See the Appendix for a discussion of our efforts to purge the pictures of "quality" that could be associated with candidate traits.

Each participant evaluated hundreds of face pairings, all randomly-drawn. We used more than 167,000 binary choices by participants to build competence scores for each face in the candidate pool.[6] We make three assumptions to calculate these scores.

First, we assume that there is a latent continuum of facial competence on which each face can be placed that drives the perceptions of all potential raters. Second, we assume that participant evaluations have a probabilistic, not deterministic, relationship with the latent facial competence dimension; in the pairwise evaluations performed by our participants, the face chosen as more competent is not always the face that is higher in latent facial competence because of both measurement error and idiosyncratic preferences of different raters. Third, we assume that participant evaluations are transitive.

Based on these assumptions, facial competence scores can be estimated by resolving the following question: What relative positions of the faces on the competence continuum would have been most likely to produce the choices made by our participants? Using maximum likelihood, we estimate each face's position on the continuum of competence. This provides a numerical estimate of the latent facial competence for each face in our pool. Note also that with the estimated locations on the continuum, we can calculate the probability that any given face will be chosen over any other face in a pairwise evaluation by our participants based upon the distance between the two faces. We present the technical details of our estimation model in the Appendix.

[6]We estimated the scores with a variety of robustness checks based upon recognition, respondent consistency (we repeated the same face pairs within respondents, varying left-right status of the repeat pair, to measure their consistency), dropping early and late evaluations for fatigue and learning. None of the alternatively estimated scores substantively affected our results.

Replication

We use the estimated latent competence scores generated from our survey to compare our estimates to those of Todorov et al. (2005). The Todorov et al. (2005) facial competence measure is the proportion of times one candidate's face is chosen more competent than their opponent's. Given our estimated competence scores, we can construct the predicted probability that any face will be chosen more competent than any other by reversing the estimation model described in the Appendix.

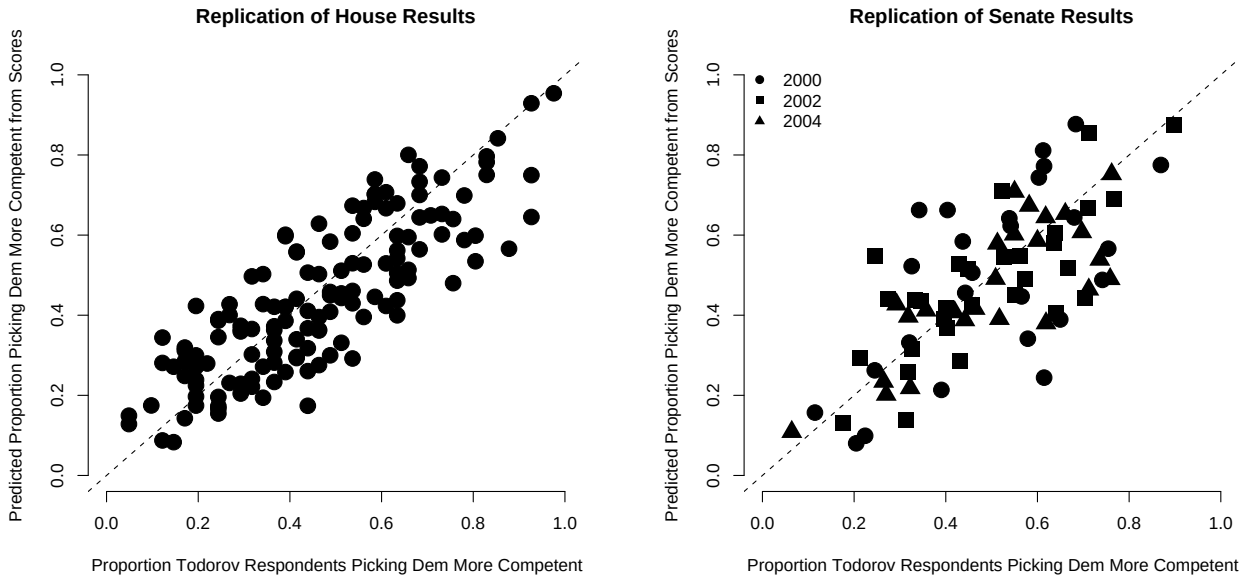
We were able to closely replicate Todorov et al.'s (2005) measure of facial competence using evaluations from our survey. We present our replication of House and Senate evaluations in Figure 2. In each frame, the x-axis plots the proportion of the Todorov et al. (2005) participants who chose the Democratic candidate more competent, and the y-axis the predicted proportion of pairwise evaluations in which the Democrat would be picked the more competent by our participants. The dashed line is a 45 degree line indicating perfect correspondence. Our method effectively replicates the choices of the Todorov et al. (2005) participants. Our scores also reproduce the relationship between facial competence and the vote (not presented).

The Competence of Politician Faces

Our survey allows us to compare the perceived facial competence of Members of Congress and their challengers. The numbers measure "inferred facial competence" based upon the evaluations of our survey participants, and we standardize the scores to have mean zero and unit variance. The scales for each house were estimated separately, so comparing the numeric values across chambers is not meaningful.

In Figure 3 we present a graphical display of the distributions of estimated facial competence for all candidates in our pool for each chamber. The white bars for each histogram are the distributions for all candidates – both challengers and incumbents – while the shaded

Figure 2: Replication of Todorov et al. Experimental Results



bars are the distributions for challengers only. In both the House and Senate, the challenger face distribution is shifted to the left of the incumbent face distribution, indicating lower challenger facial competence. However, the difference is smaller in the Senate, perhaps due to the higher profile Senate campaigns attracting, on average, better challengers. In both houses, challengers make up the bulk of the lower tail of the distribution.

To meaningfully illustrate the competence scores, we have placed a set of faces above each distribution. The median 2004 House incumbent face (denoted 'Median Incumbent' in the upper panel) was Representative Cliff Stearns (R-FL 6th). Stearns handily defeated his opponent David Bruderly despite having only a relatively small advantage in facial competence. The median challenger was Warren Redlich who challenged Representative Michael R. McNulty (D-NY 21st) and only won 30 percent of the two party vote.

The most competent looking incumbent for the House was Representative Alan B. Mollohan (D-WV 1st). Mollohan defeated his opponent Alan Parks and had a commanding competence advantage. In contrast, the most competent looking challenger was Greg Dority, who ran for Congress in North Carolina's 1st District as a Republican and, despite

his advantage, lost to Democrat G.K. Butterfield. Butterfield was well below the median competence for incumbents.

Gary Ackerman (D-NY 5th) was evaluated to have the least competent incumbent face, yet overcame a deficit of inferred facial competence, defeating his more competent-appearing opponent Stephen Graves. Not as fortunate was the least competent challenger face, Robert Lucas, who lost to, now-Senator, Sherrod Brown in Ohio's 13th.

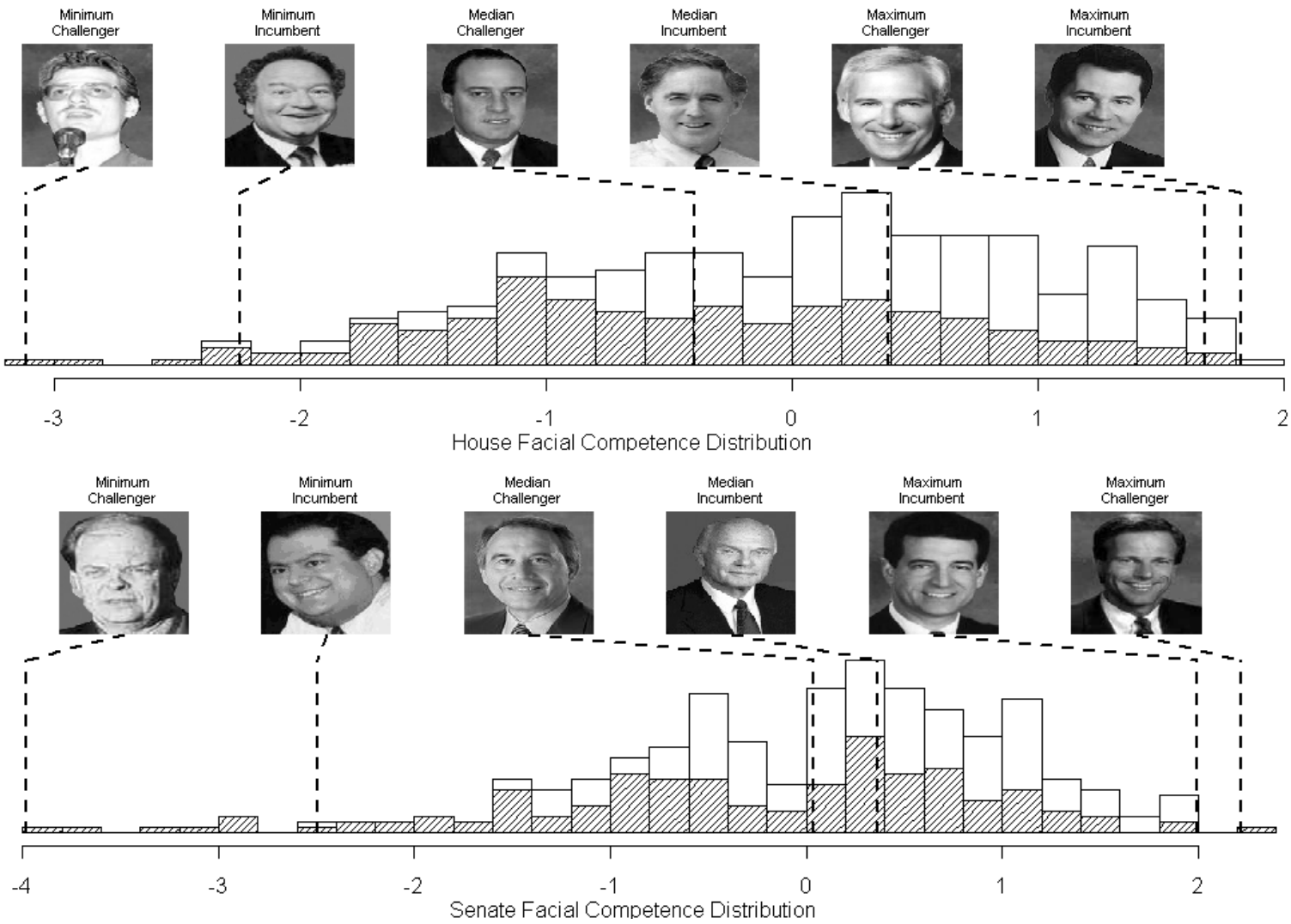
In the Senate we included candidates regardless of race or gender, except high-profile senators likely to be recognized by survey participants (excluding interesting candidates such as Hillary Clinton). Our pool includes Senate candidates from 1990-2006. The median Senate challenger was David Walters who lost to David Inhofe (R-OK) despite having an advantage in facial competence. The median incumbent face in our Senate pool was John Glenn (D-OH). Glenn easily defeated his opponent in 1992, Michael DeWine, over whom he had a substantial advantage in competence.[7] The most competent challenger was John Thune (R-SD), who defeated Tom Daschle in 2004, despite having narrowly lost to Tim Johnson (D-SD) in 2002, over whom he had a considerable advantage in face. John Thune also has the distinction of being the most competent looking current Senator. The most competent incumbent at the time of the campaign was Russ Feingold (D-WI) who was first elected in 1992.

The least competent incumbent Senate face across all races and years was Spencer Abraham (R-MI), who lost to the more competent but still below average Debbie Stabenow (D-MI) in 2000. The least competent challenger was Charles M. Oberly, III who lost to the late Bill Roth (R-DE) in 1994.

Finally, Republicans typically have more competent faces than Democrats. The median Republican candidate is more competent looking than the median Democrat in the House by one-third standard deviation (-0.040 to 0.312) and the Senate by one-tenth standard deviation (0.129 to 0.234).

[7]DeWine was elected to the Senate in the following election.

Figure 3: Example Faces and Competence Distributions by Chamber



Shaded regions are distribution of challengers.

House Candidates (from Left to Right): Robert Lucas, Rep. Gary Ackerman ($D-NY5th$), Warren Redlich, Rep. Cliff Stearns ($R-FL6th$), Greg Dority, Rep. Alan B. Mollohan ($D-WV1st$).

Senate Candidates (from left to right): Charles M. Oberly, III, Sen. Spencer Abraham ($R-MI$), David Walters, Sen. John Glenn ($D-OH$), Sen. Russ Feingold (D-WI), Sen. John Thune ($R-SD$).

Scores are not comparable across chambers.

Competition and Faces in Congressional Elections

With estimates of the facial competence of individual candidates in hand, we return now to assessing the mechanism that connects face to political success. Having separate estimates for each candidate allows us to estimate the separate effect of each candidate's face on vote. It also allows us to look for evidence of the selection of higher quality faces into more competitive and partisan-advantaged districts.

We first regress incumbent vote share on the incumbent and challenger's facial competence scores as separate variables. The importance of incumbency, challenger quality, and variation in spending effects in congressional elections (Jacobson 2004) suggests that the effect of challenger and incumbent faces on vote may vary.

We present these results in the left column for each chamber in Table 2. The effect of challenger and incumbent faces do differ. Moving from the least competent challenger face to the most competent challenger face is estimated to decrease the incumbent's vote share by almost 6 percentage points in the House and by almost 20 percentage points in the Senate.[8] Oddly, in both chambers a more competent incumbent face is estimated to *decrease* the incumbent's vote share. In the House this estimate is statistically different from zero, with an estimated decrease across incumbent faces almost as large as that for the challenger face.

We suggest that the counter-intuitive notion that a better incumbent face causes lower incumbent vote share is unlikely to be true. It is more likely that we are observing the effects of strategic party recruitment.

If strategic party recruitment were driving candidate selection, parties would recruit the most competent-looking faces to the most closely contested elections. If parties engage in this strategic allocation of recruitment effort, then incumbents from marginal districts (i.e., expected incumbent vote share of about 50 percent) would tend to have more competent faces than incumbents from safe districts (i.e., expected incumbent vote share of 80 percent). This would induce the negative coefficient on incumbent facial competence because the best

[8]Including year fixed effects does not affect the results.

Table 2: Using Facial Competence and District Competitiveness to Predict Incumbent Vote Share.

	House 2004		**Senate 1992-2006**	
	Faces Only	With Expectation	Faces Only	With Expectation
Intercept	64.51	57.42	60.14	49.66
	(0.54)	(2.76)	(0.80)	(6.66)
Challenger Facial Competence	-1.30	-0.57	-3.37	-2.05
	(0.47)	(0.41)	(0.63)	(0.50)
Incumbent Facial Competence	-1.17	-0.72	-0.45	-0.30
	(0.56)	(0.48)	(0.88)	(0.66)
Cook: Incumbent Lean		-0.39		5.08
		(2.98)		(6.77)
Cook: Incumbent Likely		1.53		8.32
		(3.20)		(6.76)
Cook: Incumbent Safe		8.72		16.89
		(2.79)		(6.73)
Cook: Tossup				-0.37
				(6.88)
N	145	145	147	147
R^2	0.08	0.38	0.17	0.56
Adjusted R^2	0.06	0.35	0.16	0.54
Std. Error of Regression	5.62	4.67	8.97	6.65

Ordinary least squares with standard errors in parentheses. Dependent variable is incumbent vote share.

faces are in the districts with lower incumbent vote share.

We find more evidence of party recruitment in the second and fourth columns of Table 2. When we add a measure of district competitiveness to the model, the coefficient on incumbent face moves towards zero, which we would expect to occur if incumbent face were negatively correlated with expected probability of incumbent victory. We measure the competitiveness of a contest by coding each race according to the classifications provided by the *Cook Political Report* (Cook 1992-2006).[9] Cook classifies each campaign as "Tossup", "Lean", "Likely", or "Safe" for each party. In an attempt to keep our measure untainted by the challenging candidate's competence, we use Cook publications from at least one year before each election so that the challenger is unlikely to have yet been selected. For example, our measure of competitiveness for the 2004 elections are taken from the August 2003 *Cook Political Report* newsletter. We recode the measure of competitiveness in the direction of the expected probability of incumbent victory, and include dummy indicator variables for each category. For the House results, the excluded category is "Tossup", and for the Senate, the excluded category is "Challenger Lean".[10]

In both chambers, adding the district competitiveness variable increases the amount of explained vote share variance. For the House, the perverse coefficient on incumbent face is cut in half and loses statistical significance, as does the estimated effect of challenger face. In the Senate, the effect of challenger face is decreased by a third but maintains a confidence interval well outside of zero. Moving from the least competent Senate challenger to the most competent Senate challenger is estimated to decrease incumbent vote share by 12 points. The effect of the incumbent's face remains uncertainly-estimated yet maintains its perverse sign. We suspect that incumbent face coefficients remain negative for both chambers due to

[9]For examples of scholarly work employing Cook's report, see Gimple, Karnes, McTague, & Pearson-Merkowitz (2008) or Vavreck (2001). Models using the Cook variable exclude the 1990 Senate contests because we were unable to obtain a *Cook Political Report* for 1989.

[10]It is possible that even the Cook measure is endogenous to the characteristics of the challenger. We ran the same models with the presidential vote share of the district or state from the previous election in lieu of Cook as a measure of competitiveness. This results in little change in the substance of what we describe here. We are more satisfied, however, that Cook's measure better captures the competitiveness of district, which is based on more holistic information than just the result of the previous presidential election.

our inability to fully capture the strategic political calculations at play.

It is not surprising that the effect of challenger face is stronger in the Senate. Senate races are higher salience elections and we presume voters are more likely to receive information about the appearance of the challenging candidate. We will explore the effect that faces have on individual voters in Senate elections in more detail following an exploration of the allocation of candidate faces.

The Distribution of Faces Across Districts

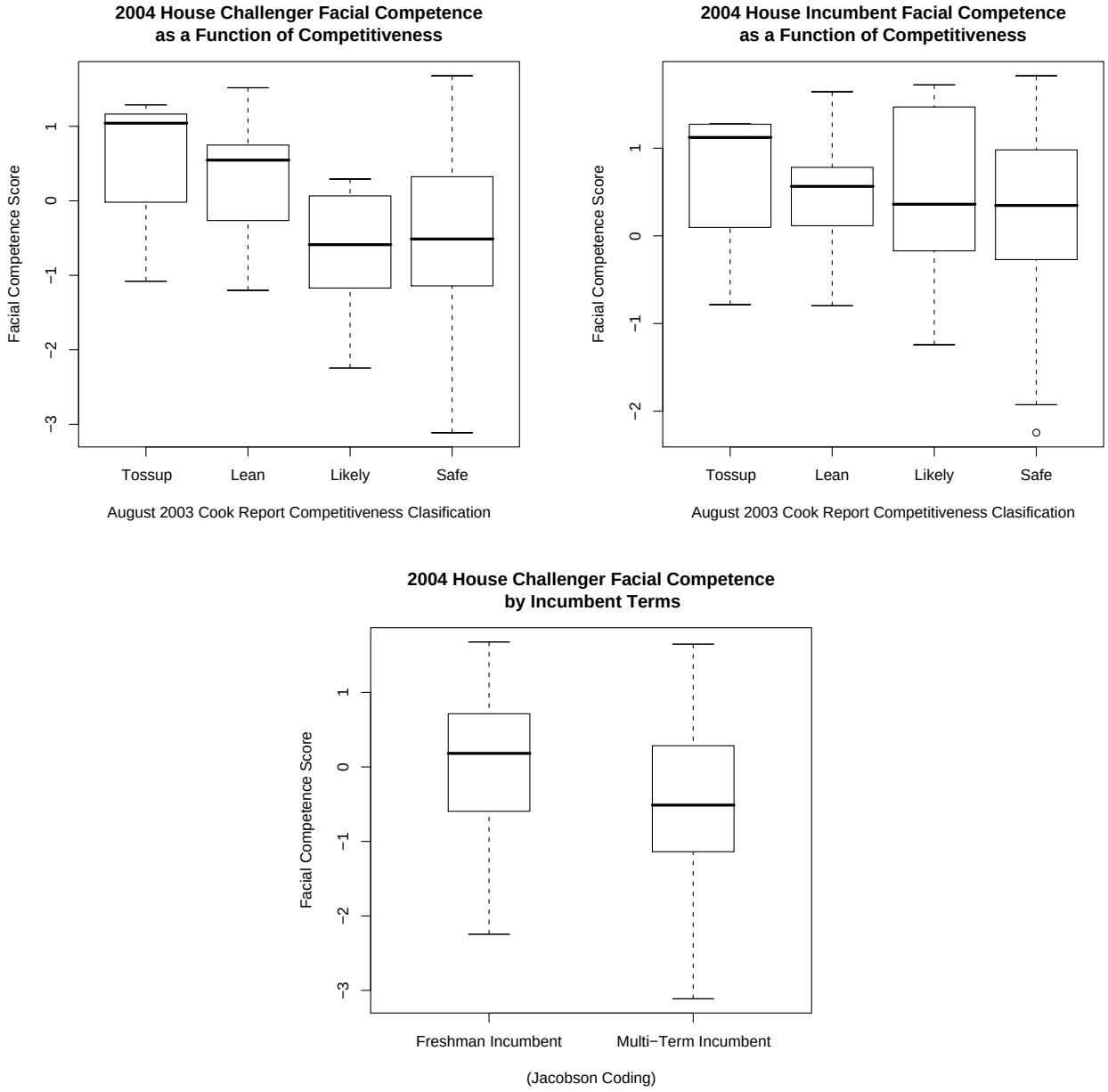
We turn now to a graphical exploration of the distribution of faces by competitiveness. Our regression results indicate that faces are correlated with district and election characteristics, perhaps due to strategic behavior. In Figure 4, we present the distribution of candidate facial competence by competitiveness measures. The top two panels are boxplots of House challenger and incumbent competence as a function of year-prior Cook district classification. Both indicate that more competent faces appear to enter the more competitive elections.

The bottom panel shows the distribution of House challenger faces by the tenure of the incumbent, a measure likely more robust to potential endogeneity in the Cook measure. The graphic demonstrates more competent faces challenge freshmen incumbents, generally the most vulnerable of incumbents. The polyserial correlation between challenger face and a freshman incumbent indicator is $r = .28$. Similar analysis for the Senate (not presented) reveals a similar but, not surprisingly, less distinct pattern.

Individual Voters and Faces

In the preceding analysis we found evidence that the effect of face on vote share is not solely a voter-level process, but rather that the allocation of faces to districts is related to incumbency and district partisanship. In this section, we extend our investigation to the analysis of individual vote choice to see how incumbency and district partisanship structure

Figure 4: Facial Competence by District Competitiveness



individual response to candidate faces.

There are two avenues through which candidate faces might influence the vote choice. On the one hand, the characteristics of each face may add or subtract some constant probability of candidate vote across all voters. On the other hand, the face may interact with another variable such as partisanship so that the response to the face is conditional on the intersection of voter and candidate partisanship.

How face functions as a cue to different types of partisans is a subject of theoretical importance to understanding the relationship between faces and election outcomes. If voters were making comparisons of faces independent of partisanship or incumbency, we would expect the influence of facial competence on vote choice to be most pronounced among nonpartisan voters. Those voters that lack a partisan cue to structure their vote choice would be more influenced by candidate characteristics such as face than voters whose vote choice is structured by their partisan identification. If, on the other hand, face comparisons are not made naively but instead interact with other voter dispositions such as partisanship, we would expect face to affect the choices of partisans as well.

We estimate the interaction of facial cues and partisanship by combining the facial competence measures from our survey with election exit poll surveys.[11] We chose to use exit polls, rather than other common surveys such as the American National Election Study, because of the advantages of a relatively large sample size and the temporal proximity to the actual vote choice.

In Table 3 we present the results from a probit regression analysis of the effect on vote choice produced by both the additive effect of facial competence and the interaction of competence and respondent partisanship.[12] Because the results are probit coefficients and include interactions, we will discuss the results of Table 3 with reference to predicted values using the estimates from the full House and Senate models. Our predicted values are

[11] We use network exit polls for the 2004 House elections and for the 1992-2002 and 2006 Senate elections. The *National Election Pool* exit poll questionnaires did not include a question on Senate vote choice in 2004.

[12] Including year fixed effects does not affect the results.

calculated by holding one candidate's facial competence at the 50th percentile and holding the Cook report of district competitiveness at "likely" going to the incumbent party. For each type of voter (independent, incumbent co-partisan, and challenger co-partisan), we estimate the change in the predicted probability that a respondent votes for the incumbent candidate that is produced by moving the incumbent candidate's facial competence from the 25th percentile to the 75th percentile. We then estimate the same change in incumbent vote probability produced by moving the challenging candidate's face from the 25th to 75th percentile. In Figure 5 we present these estimated first difference effects with 95 percent confidence intervals.

The three bars furthest to the left in Figure 5 represent — for independents, challenger co-partisans and incumbent co-partisans, respectively — the predicted first difference effect on the probability of voting for the incumbent produced by changing incumbent facial competence from the 25th to the 75th percentile. For example, if better incumbent face increased incumbent vote all the bars would be in the positive region (upper left of the panel). However, consistent with the incumbent facial competence coefficient reported in the aggregate-level analysis in Table 2, the first difference effect of increasing the House incumbent's level of facial competence is statistically insignificant for all three voter categories, and two of the three point estimates exhibit the estimate that better incumbent face decreases incumbent vote.

The three subsequent bars represent the estimated effects of increasing the House challenger's level of facial competence. If better challenger face decreased incumbent vote all the bars would be in the negative region (the bottom right of the panel). Here only the effect of challenger face on challenger co-partisan vote is statistically distinguishable from zero. Thus for the House overall, challenger face appears to have a small effect on vote choice, but in general the effect of candidate faces on House elections appears to be negligible.

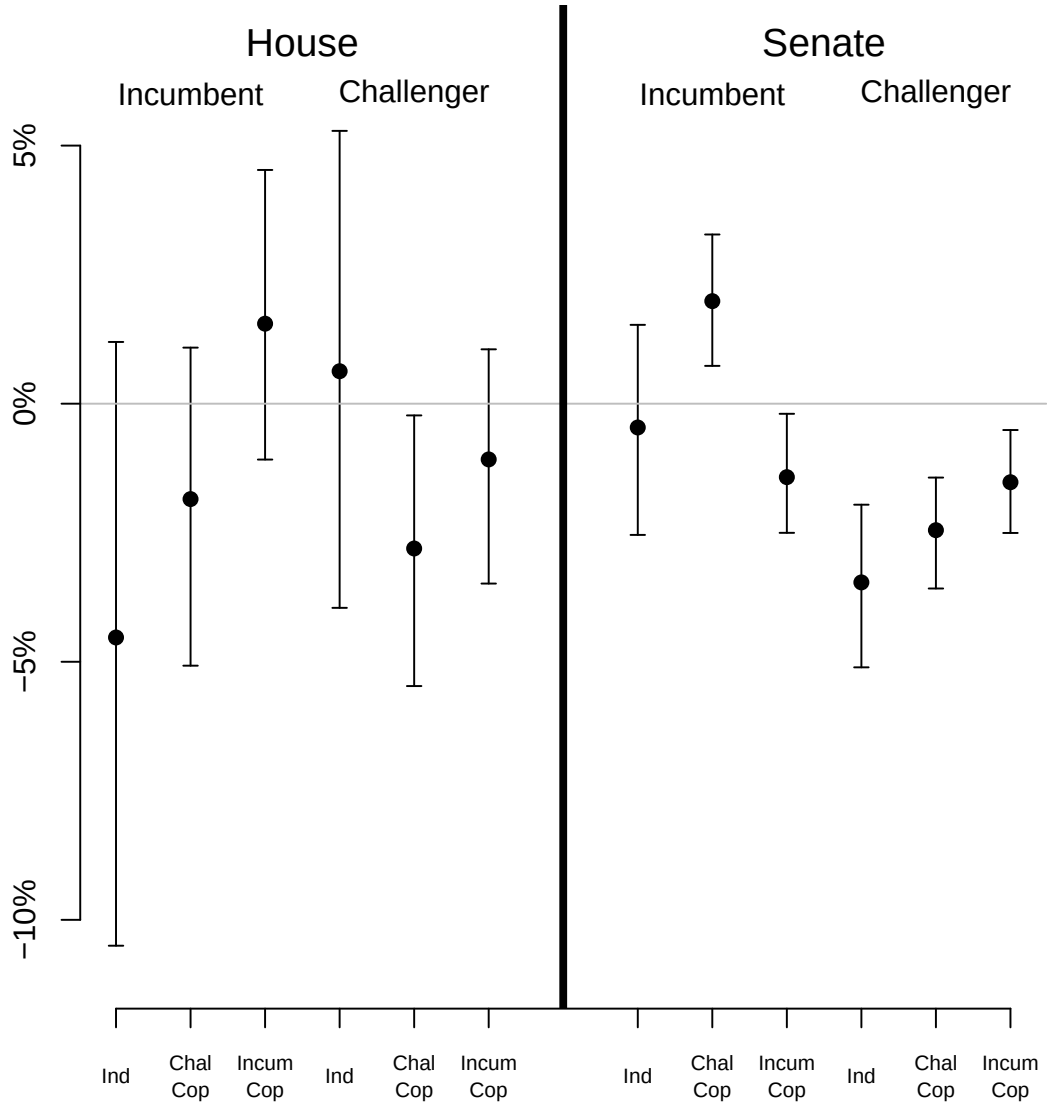
In the Senate, however, more than just challenger co-partisans are estimated to respond to face. A better incumbent face is predicted to increase the probability of defection to the

Table 3: Using Facial Competence and Partisanship to Predict Individual-Level Vote Choice.

	House 2004		**Senate 1992-2006**	
	Faces Only	With Expectation	Faces Only	With Expectation
Intercept	0.20	0.34	0.19	-0.04
	(0.05)	(0.24)	(0.02)	(0.03)
Respondent Shares Challenger Party	-1.29	-1.29	-1.01	-1.03
	(0.08)	(0.08)	(0.02)	(0.02)
Respondent Shares Incumbent Party	1.32	1.31	1.03	1.02
	(0.08)	(0.08)	(0.03)	(0.03)
Challenger Facial Competence	-0.01	0.02	-0.09	-0.07
	(0.10)	(0.11)	(0.02)	(0.02)
Incumbent Facial Competence	-0.25	-0.20	0.00	-0.01
	(0.13)	(0.13)	(0.02)	(0.02)
Respondent Shares Challenger Party	-0.32	-0.31	-0.01	-0.00
*Challenger Facial Competence	(0.16)	(0.16)	(0.02)	(0.02)
Respondent Shares Incumbent Party	-0.14	-0.14	0.02	0.01
*Challenger Facial Competence	(0.17)	(0.17)	(0.02)	(0.02)
Respondent Shares Challenger Party	0.02	0.02	0.05	0.07
*Incumbent Facial Competence	(0.20)	(0.20)	(0.03)	(0.03)
Respondent Shares Incumbent Party	0.36	0.38	-0.05	-0.04
*Incumbent Facial Competence	(0.21)	(0.20)	(0.03)	(0.03)
Cook: Incumbent Lean		-0.42		0.18
		(0.24)		(0.04)
Cook: Incumbent Likely		-0.26		0.15
		(0.24)		(0.04)
Cook: Incumbent Safe		-0.12		0.42
		(0.23)		(0.03)
N	4350	4350	21974	21619
AIC	3460.51	3453.03	21655.22	21113.43

Probit regression coefficients with standard errors in parentheses. Dependent variable is respondent voted for incumbent candidate. The omitted category is Cook: Tossup.

Figure 5: First Difference Effects on Incumbent Vote of Candidate Facial Competence, by Chamber, Incumbency Status and Voter Partisan Category.



The first difference effect reported here with 95% confidence intervals is the percentage point change in the probability that respondent supports the incumbent candidate, when the incumbent or challenger candidate face moves from the 25th percentile to 75th percentile (Imai, King, & Lau 2007a, Imai, King, & Lau 2007b). Using the full models in Table 3, predictions are calculated holding competitiveness at "likely incumbent victory" and holding the rival candidate's face at the median of all candidate faces for the applicable chamber.

incumbent by challenger party voters by 2.0 percentage points while, once again, a better incumbent face is associated with a *decrease* of 1.4 points in the probability of incumbent party voter loyalty. The least ambiguous candidate face effect in our exit poll analysis occurs among Senate challengers. As the three rightmost bars in Figure 5 indicate, the predicted effect of changing a Senate challenger's face is significant for all three subpopulations, with increased probability of challenger vote of 3.5 points for independents, 2.5 points for challenger-party voters, and 1.5 points for incumbent-party voters.

There are two important points that we can conclude from this above analysis. First, the facial competence of challenging candidates appears to have a much more sizeable effect on vote choice than the facial competence of incumbents. This is not surprising given that voters can make use of other readily available cues in evaluating incumbents. Second, the predicted effects of candidate faces on vote choice are not limited to political independents. Even individuals who identify with a political party can be influenced by faces under the right circumstances. In fact, voters who share the party label of the challenger candidate are the only group demonstrating a statistically significant response to candidate face in both chambers. That the most distinctive effects occur among this group suggests that voter response to candidate faces is structured by political attitudes. Therefore, we suggest it likely that the process producing the causal effect of candidate faces on vote choice is more complex than voters making decisions based on a naive and rapid comparison of faces.

Discussion

In this paper, we have measured the effect of candidate faces on the vote. We show that the considerable bivariate relationship of face to vote is reduced in the presence of observational controls. Holding incumbency and partisanship constant, the effect of facial competence is small. However, we do find an effect of face to aggregate vote when the two candidate faces are accounted for separately, and also at the level of the individual voter. We find evidence

for the effect of candidate face on individual voters among partisans as well as independents.

We remain concerned, despite our efforts at control described in the Appendix, that some important part of the measure of facial competence is due to variation in image quality. If our survey participants picked up from clothing, haircuts, professional lighting, or any other potential cue something about the quality or funding of the candidate's campaign, part of our measure of facial competence is in fact just a proxy for campaign competence. Were this true, it would actually support the contention that quality of candidate is not distributed randomly across districts. But it would further attenuate the effect of face on the vote choice.

We have also constructed an original measure of candidate quality. To the extent that others have demonstrated that appearance is correlated with other success (Rule & Ambady 2008, Mobius & Rosenblat 2006, Hamermesh & Biddle 1994, Biddle & Hamermesh 1998), this measure is not inconsistent with the standard measures of candidate quality (e.g. Jacobson 2004). If this is an accurate measure of candidate quality, we have found that higher quality candidates are more likely to enter the subset of congressional elections in which the contest is competitive. How these candidates come to enter these contests has important implications for understanding the dynamics of party and candidate influence in shaping the choices available to voters. On the one hand, higher quality candidates with more to lose could be selecting into contests which they are more likely to win. On the other hand, parties could be recruiting the best candidates into the most competitive contests (Gibson, Cotter, Bibby, & Huckshorn 1985, Gibson, Cotter, Bibby, & Huckshorn 1983, Herrnson 1986, Mann & Ornstein 1981).

Our evidence suggests that candidates or parties, or likely both, are making decisions which limit the range of faces and candidate quality from which voters choose. Because many congressional contests in the United States are not competitive and because candidates with high competence are more likely to enter the contests in which they have a reasonable chance of success, we find high competence candidates defeating their usually low competence challengers in the majority of contests. This dynamic produces a high correlation between

facial competence and election outcomes. That candidate faces are not distributed randomly across contests and that it is likely that parties and candidates are making decisions which affect the allocation of candidates to races suggests avenues for future research. For example, our unique measure of candidate quality may be useful in research adjudicating between candidate and party centered models of electoral politics.

The attribution of election outcomes to potentially-objectionable voter decision processes also has implications for normative political theory. We have demonstrated, however, that even if voters do sometimes make decisions based on candidate appearance, the effect does not appear large enough to decide any but the closest election outcomes, and what effect there is appears to be conditioned by voter partisanship and candidate incumbency status.

Some of the media attention surrounding the research by Todorov et al. (2005) was probably generated by the sense that the finding demonstrates that the voting public is uninformed. We have demonstrated that appearance plays a much smaller role in election outcomes than one might infer from a casual reading of Todorov et al. (2005) or its representation in the media. However, a system in which most contests pit a good-looking incumbent who shares partisanship with the vast majority of the district against a challenger with little chance of victory might also raise concerns about the vibrancy of democracy.

Appendix

Details of Facial Competence Estimation

Facial competence scores for each candidate are calculated using the binary choices made by survey participants between two candidate faces. The scores are estimated by a method used to model congressional committee choice (Groseclose & Stewart III 1998). We assume each candidate face i has a location c_i on the latent competence scale. Each participant evaluates two faces i and j with some amount of measurement error ϵ_i and ϵ_j. Participants choose face i over face j if and only if $c_j + \epsilon_j < c_i + \epsilon_i$ which is the same as $\epsilon_j - \epsilon_i < c_i - c_j$. Without loss

of generality, we assume that the ϵ_i are identically and independently distributed according to a mean-zero normal distribution with standard deviation σ. Given these assumptions, the probability that the respondent reports candidate i more competent than candidate j is

$$\Phi\Big(\frac{c_i - c_j}{\sigma\sqrt{2}}\Big)$$

where $\Phi(\cdot)$ is the cumulative normal probability function.

To estimate the c_i we let each observation be the evaluation by one participant of two faces. We define an indicator matrix V, with K rows of observations and I columns of faces, where each element v_{ki} takes the value 1 if the kith face were selected the more competent, -1 if the kith face were not selected the more competent, and zero if the kith face was not evaluated. For each observation k, therefore, the probability that the respondent evaluated the face pair randomly presented in the way that they did is

$$\Phi\Big(\frac{\sum_{i=1}^{I} c_i v_{ki}}{\sigma\sqrt{2}}\Big).$$

This implies a likelihood function for parameters c given data V

$$L(c|V) = \prod_{k=1}^{K} \Phi\Big(\frac{\sum_{i=1}^{I} c_i v_{ki}}{\sigma\sqrt{2}}\Big).$$

Again following Groseclose & Stewart III (1998) we implement this estimation with an intercept-free probit model, with the number of explanatory variables equal to the number of candidates. As with standard probit estimation, we set σ to 1, which means that our estimates of the c_i are in units of σ. For identification, one face is set to the value of 1 on the competence dimension. After implementing the estimation, the probit coefficients are utilized as facial competence scores.

With the estimated competence scores, we can calculate the probability that any one face will be chosen the more competent over any other face by our participants by reversing

the estimation model. Specifically,the probability that a participant will choose candidate i more competent than candidate j given estimated competence scores c_i and c_j is

$$\Phi\Big(\frac{c_i - c_j}{\sqrt{2}}\Big).$$

The quantity σ is ignored as it is fixed at 1 in the estimation.

Details of Image Quality Estimation

A casual look at the candidate images will reveal that some are of higher quality than others. That some photos are taken on more high quality cameras or simply produced by better photographers could be a reflection of the quality of the candidate and the campaign. It is possible then that participant responses are not based solely on qualities of the candidate's face but also to the qualities of the candidate's image. This would present a difficulty for accurate estimates of the role of facial competence alone in campaigns if one were to argue that facial competence has a direct effect on voters.

To ascertain how much of election results are directly attributable to face, we attempt to control for image quality. We constructed a unique measure of photo quality derived from the variance of pixel color at different points in an image. We find that image quality is related to individual's evaluations of the competence of the faces in the images. However, we find that even when controlling for image quality, evaluations are still related to election outcomes and our substantive results are unchanged.

We detail here our efforts to objectively-measure image quality. We considered hand-coding by eye each image for quality, but were concerned the facial characteristics would influence such an effort. Instead, we implemented an objective measurement based on spatial statistics using each pixel in the image.

We first constructed a *variogram* of each image. The variogram is a common concept in spatial statistical studies that measures variance γ between the value of a process k at an

arbitrary point s_i in a space S and every other point $s_{i+1}, ..., s_n$ in S that is at each distance $d_0, d_1, ..., d_m$ from s_i, where m is usually 1/3 the maximum possible distance between points (for details on variogram estimation, see Cressie 1993, Ripley 2004). In this case, S is the pixel matrix of the image.

A black and white digital image has for each pixel a number representing the place on the gray scale of that point on the image. This number can range from 0 to 255. Each photograph was standardized to a matrix of 105x147 pixels. In constructing the variogram, we first median-polished the matrix so that each pixel was the residual after the median tendency of each row and column of the matrix had been removed.

Using this variogram, we constructed an image quality measure, I, for each photograph such that

$$I = \frac{\gamma d_1}{\gamma d_{49}}.$$

This image quality measure is designed to capture the average variance of any two adjacent pixels given the average variance of two pixels at a distance in which the pixels would be expected to have no autocorrelation. This is based on the assumption that adjacent pixels are autocorrelated because they are capturing the same feature of an image. For example, in a photograph of a man, two adjacent pixels might both be capturing the man's tie and should be the same color. As pixels are further apart, the autocorrelation goes to zero and pixels have no dependence on each other. We assume that images of higher quality have more autocorrelation between pixels because of higher original density of the photograph. However, this is relative to the overall variance of the photograph, for example the man might wear a solid tie (low variance) or a patterned tie (high variance), so the denominator of the image quality measure represents the variance at a point in which no autocorrelation is expected to exist.

References

Ballew II, Charles C., & Alexander Todorov. 2007. "Predicting Political Elections From Rapid and Unreflective Face Judgments." *Proceedings of the National Academy of Sciences*.

Benjamin, Daniel J., & Jesse M. Shapiro. 2006. "Thin-Slice Forecasts of Gubernatorial Elections." *National Bureau of Economic Research Working Paper Series* (Working Paper No. 12660).

Biddle, Jeff E., & Daniel S. Hamermesh. 1998. "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics* 16 (1): 172–201.

Campbell, Angus, Philip E. Converse, Warren E. Miller, & Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Cook, Charles E. 1992-2006. "The Cook Political Report.". Washington, D.C.: National Political Review, Inc.

Cressie, Noel A. C. 1993. *Statistics for Spatal Data*. New York: Wiley-Interscience.

Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3 (3): 395–405.

Fox, John. 2007. *polycor: Polychoric and Polyserial Correlations*. R package version 0.7-5.

Gelman, Andrew, & Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34 (4): 1142–1164.

Gibson, James L., Cornelius P. Cotter, John F. Bibby, & Robert J. Huckshorn. 1983. "Assessing Party Organizational Strength." *American Journal of Political Science* 27 (2): 193–222.

Gibson, James L., Cornelius P. Cotter, John F. Bibby, & Robert J. Huckshorn. 1985. "Whither the Local Parties?: A Cross-Sectional and Longitudinal Analysis of the Strength of Party Organizations." *American Journal of Political Science* 29 (1): 139–160.

Gimple, James G., Kimberly A. Karnes, John McTague, & Shanna Pearson-Merkowitz. 2008. "Distance-Decay in the Political Geography of Friends-and-Neighbors Voting." *Political Geography* 27: 231-252.

Groseclose, Tim, & Charles Stewart III. 1998. "The Value of Committee Seats in the House, 1947-91." *American Journal of Political Science* 42 (2): 453–474.

Hamermesh, Daniel S., & Jeff E. Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84 (5): 1174–1194.

Hamilton, John. 2005. "Scientists Search for that Winning Look." *National Public Radio All Things Considered* (June).

Herrnson, Paul S. 1986. "Do Parties Make a Difference? The Role of Party Organizations in Congressional Elections." *Journal of Politics* 48 (3): 589–615.

Imai, Kosuke, Gary King, & Olivia Lau. 2007a. *probit: Probit Regression for Dichotomous Dependent Variables*. In Kosuke Imai, Gary King, and Olivia Lau, "Zelig: Everyones Statistical Software," http://gking.harvard.edu/zelig.

Imai, Kosuke, Gary King, & Olivia Lau. 2007b. *sim: Simulating Quantities of Interest*. In Kosuke Imai, Gary King, and Olivia Lau, "Zelig: Everyones Statistical Software," http://gking.harvard.edu/zelig.

Jacobson, Gary C. 1996. "The 1994 House Elections in Perspective." *Political Science Quarterly* 111 (2): 203–223.

Jacobson, Gary C. 2004. *The Politics of Congressional Elections*. 6 ed. New York: Longman.

Jacobson, Gary C., & Samuel Kernell. 1983. *Strategy and Choice in Congressional Elections*. New Haven: Yale University Press.

Little, Anthony C., Robert P. Burriss, Benedict C. Jones, & S. Craig Roberts. 2007. "Facial Appearance Affects Voting Decisions." *Evolution and Human Behavior* 28 (1): 18–27.

Mann, Thomas E., & Norman J. Ornstein. 1981. "The Republican Surge in Congress." In *The American Elections of 1980*, ed. Austin Ranney. American Enterprise Institute.

Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven: Yale.

Mobius, Markus M., & Tanya S. Rosenblat. 2006. "Why Beauty Matters." *American Economic Review* 96: 222–235.

R Development Core Team. 2007. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Ripley, Brian D. 2004. *Spatal Statistics*. New York: Wiley-Interscience.

Rule, Nicholas O., & Nalini Ambady. 2008. "The Face of Success: Inferences From Chief Executive Officers' Appearance Predict Company Profits." *Psychological Science* 19 (2): 109-111.

Skloot, Rebecca. 2007. "Faces Decide Elections." *New York Times Magazine*.

Todorov, Alexander, Anesu N. Mandisodza, Amir Goren, & Crystal C. Hall. 2005. "Inferences of Competence from Faces Predict Election Outcomes." *Science* 308 (5728): 1623-1626.

Vavreck, Lynn. 2001. "The Reasoning Voter Meets the Strategic Candidate: Signals and Specificity in Campaign Advertising, 1998." *American Politics Research* 29 (5): 507-529.

Zebrowitz, Leslie A., & Joann M. Montepare. 2005. "PSYCHOLOGY: Appearance DOES Matter." *Science* 308 (5728): 1565-1566.

Zebrowitz, Leslie A., Judith A. Hall, Nora A. Murphy, & Gillian Rhodes. 2002. "Looking Smart and Looking Good: Facial Cues to Intelligence and Their Origins." *Personality and Social Psychology Bulletin* 28 (2): 238-249.