

The Looks of a Winner: Beauty, Gender and Electoral Success*

Niclas Berggren[§]

Henrik Jordahl[#]

Panu Poutvaara[□]

Abstract. We study the role of beauty in politics. For the first time, focus is put on differences in how women and men evaluate female and male candidates and how different candidate traits relate to success in real elections. We have collected assessments of photos of 1,929 Finnish political candidates from 2,772 respondents from outside of Finland and 3,698 Finnish respondents. It turns out that beauty evaluations are positively related to electoral success both in the national and in municipal elections. The effect is always statistically significant for female candidates, and often also for male candidates. Extensive sensitivity analysis confirms the results.

JEL Codes. D72, J45, J7

Keywords. Beauty, gender, elections, political candidates, beauty premium

* The authors wish to thank Bryan Caplan, Daniel Hamermesh and Mikael Priks and participants at the 2006 IIPF conference on Cyprus and at seminars at the Center for Study of Public Choice, George Mason University; Institute of International Economic Studies (IIES), Stockholm University; Department of Economics, Uppsala University; Department of Psychology, University of Helsinki; Department of Political Science, University of Helsinki; Department of Economics, Lund University; the Centre for Economics and Business Research (CEBR), Copenhagen Business School; and the Ratio Institute for helpful comments and suggestions, all respondents and those who helped us attract them (especially our colleagues, who advertised the study to their students in several countries, and several bloggers), as well as Otto Kässi for excellent research assistance, Karl Bengtsson for outstanding technical assistance, and the Torsten and Ragnar Söderberg Foundations (Berggren and Jordahl), the Jan Wallander and Tom Hedelius Foundation (Jordahl) and the Yrjö Jahnsson Foundation (Poutvaara) for financial support.

[§] The Ratio Institute, P.O. Box 5095, SE-102 42 Stockholm, Sweden. niclas.berggren@ratio.se

[#] Research Institute of Industrial Economics, Box 55665, SE-102 15 Stockholm, Sweden. henrik.jordahl@riie.se

[□] Department of Economics, University of Helsinki, Arkadiankatu 7 (P.O. Box 17), 00014 University of Helsinki, Finland. panu.poutvaara@helsinki.fi

1. Introduction

Some note a tendency in elections to put focus on candidates rather than on political platforms.¹ Others point out that voters are not very well informed.² Hence, they may not base their voting decisions on a thorough examination of positions and programs but may choose which candidates to support on the basis of various heuristics or on the basis of “thin slices” of information (Ambady and Rosenthal, 1992). Looks may be such a heuristic or thin slice of information.

We investigate to what extent evaluations of photos of political candidates can explain election outcomes. In particular, we ask whether there is a political “beauty premium”, such that better-looking candidates have a higher chance of becoming elected to public office. The advantage of beauty could, according to the logic of expressive voting, be substantial in elections: in the absence of opportunity cost people might as well vote for the most beautiful candidate.³

We are the first to study how evaluations of male and female political candidates, in particular of their beauty, by men and women differ.⁴ Our use of photos representing real political candidates allows us to outline how gender differences in evaluations are related to the electoral performance of male and female candidates.

We have carried out a web survey with over 1,900 facial photos of Finnish political candidates, collecting evaluations from about 2,800 respondents who were asked to assess their beauty, as well as perceived competence and trustworthiness, on a five-point scale. The survey was carried out outside of Finland, to minimize the risk that respondents would recognize the persons in the photos. We then collected evaluations of the same photos from about 3,800 Finnish respondents, to serve as extensive sensitivity analysis. Each candidate’s beauty score, on average based on about nine assessments when focusing on evaluations outside of Finland, has then been related to the number of votes they obtained in the 2003 and 2004 elections. Respondents were shown the same photos that Finnish voters had seen in the advertisements by political parties. Each municipality is obliged to provide outdoor ad-

¹ See e.g. Wattenberg (1991), Stewart and Clarke (1992) and Mughan (2000) for such views. Cf. King (2002) for some dissent.

² See e.g. Downs (1957), Bartels (1996) and Caplan (2002).

³ On expressive voting, see e.g. Brennan and Lomasky (1993).

⁴ Throughout the paper, we use the terms “men” and “women” to denote respondents, i.e. those who participated in our study by evaluating political candidates, and “male” and “female” to denote political candidates.

vertising space to all political parties, implying that almost all voters have been exposed to the photos of the candidates in their electoral district repeatedly.

Our main result is that a candidate's beauty (relative to the beauty of competing candidates) is conducive to electoral success, but far from a necessary condition for it. In real elections, when candidates are chosen in competition with others on the same party list, beauty is in most specifications more strongly correlated with success than perceived competence and trustworthiness for female candidates, and in several specifications also for male candidates. While the estimated effect of beauty on electoral success is not large, nor is it trivial and it might tip the scale for candidates at the cut-off point on a list.

A central contribution of this study to the literature on the role of beauty in politics consists of a systematic investigation of the role of gender. If beauty matters for electoral success, then the question is if it confers differential advantages on male and female candidates. Furthermore, we investigate whether men and women differ in their evaluation of candidates' beauty and other traits. There are a few other aspects of our research design that we think comprise a contribution, in addition to the main focus on gender, beauty and electoral success as such. The Finnish system is unusually suited for gender analysis, since there is a sizable number of female candidates in all districts competing against male candidates.

Through the choice of Finland, our study is the first of the effects of facial appearance on the success of politicians in a wholly proportional electoral system, enabling us to focus on within-party competition. There are at least four benefits of focusing on this type of competition.

First, many electoral systems are proportional and hence political candidates in many places do compete against other candidates from the same party. This makes our results applicable to such settings, unlike previous results from countries that do not have a proportional system.

Second, plurality-vote systems like the American one also contain intraparty competition between candidates in the primary-election stage, and for that reason our study has clear implications for American politics, and for politics in countries with similar systems, as well.

Third, there could be a problem of reverse causality or a missing underlying variable in all previous studies of one-member districts with between-party competition. If party A is stronger than party B in a given district, party A is also likely to have more potential candidates to choose from, and thus it is likely to be able to attract candidates that score high on various traits. Also, such candidates are likely to prefer districts with higher chances of win-

ning. So, it could be that the causality is not from, say, perceived competence and beauty to electoral success, but that districts in which one party has an advantage in a given election is more successful in recruiting a candidate who is beautiful or perceived as more competent. We largely avoid this problem by our focus on intraparty competition.

Fourth, when studying within- rather than between-party competition, like most previous studies do, it is much easier to isolate the effect of beauty on electoral success, as the choice between candidates of different parties can be expected to be influenced by many other factors, such as ideology, which we are able to effectively control for with our approach.

Aside from the benefits of being able to study intraparty competition, the Finnish system allows us to focus on non-incumbent candidates, about whom voters can be expected to have little or no factual information, and to see whether thin slices of information are used by voters to inform their choice of a political candidate. Previous studies almost always feature competition between an incumbent and challenger. Political competition in Finland also features a sizable number of female candidates in each district, which enables us to study the gender issue in more detail.

Furthermore, we use more respondents than other studies – almost 2,800 in one specification and 3,700 in the other compared to four (Hamermesh, 2006), 50 (Mobius and Rosenblat, 2006), 100 (Banducci et al., 2003), 264 (Benjamin and Shapiro, 2006) and 843 (Todorov et al., 2005).⁵ This is a kind of sensitivity test. We also use a larger number of photos (to reduce the risk that results are driven by outliers), the photos actually shown to voters in the elections (as the parties distribute one official photo of each of their candidates), other respondents than students (e.g. from blogs and from postings to colleagues and university alumni), and respondents from many different countries. Our cross-country analysis allows us to rule out that a single country accounts for the observed patterns.

2. The Literature

We relate our study to four strands of literature: general studies of the effects of beauty on individuals, labor-market studies, election studies (the strand to which we belong) and studies on cognitive bases for voter decision-making (the use of “thin slices” of information, in-

⁵ Todorov et al. (2005) collected evaluations of beauty from only 34 respondents.

formation shortcuts and heuristics). The latter strand supports the approach taken in our study, in showing that people often base evaluations and decisions on brief visual cues in the absence of other forms of information. The beauty literature overall has so far paid scant attention to our main focus, the gender issue in elections.

The first strand establishes that being beautiful confers many advantages on a person. Langlois et al. (2000), in a meta-analysis of 102 studies, report that the looks of people influence how they are perceived and treated by others, even by those who know them.⁶ As for the role of gender, Langlois et al. (2000: 399) report the following:

The meta-analyses showed that, both within and across cultures, people agreed about who is and is not attractive. Furthermore, attractiveness is an advantage in a variety of important, real-life situations. *We found not a single gender difference and surprisingly few age differences, suggesting that attractiveness is as important for males as for females and for children as for adults.* (Our italics.)

The second strand demonstrates that the benefits of having good looks, as perceived by others, extend to and are substantial in the labor market. Beautiful people receive higher wages (a so-called beauty premium). According to Hamermesh and Biddle (1994), workers of above-average beauty earn about 10 to 15 percent more than workers of below-average beauty. Other studies obtain qualitatively similar results: see e.g. Frieze et al. (1991), Biddle and Hamermesh (1998), Harper (2000), Pfann et al. (2000), Hamermesh et al. (2002), French (2002) and Mocan and Tekin (2006). Experimental studies confirm this picture, e.g. Andreoni and Petri (2004) and Mobius and Rosenblat (2006).⁷ As for specific gender effects relating to beauty in the labor market, Hamermesh and Biddle (1994: 1187) conclude that there is an

absence of significantly larger penalties and premia, especially the latter, for women than for men. If anything, the evidence goes in the opposite direction: men's looks may have slightly larger effects on their earnings than do women's.⁸

⁶ . Cf. Feingold (1992a) and Eagly et al. (2001).

⁷ There are more general experimental studies of the effects of beauty, e.g. Mulford et al. (1998) on the Prisoners' Dilemma, Solnick and Schweitzer (1999) on the Ultimatum Game, and Eckel and Wilson (2006) on trust games.

⁸ Similar qualitative results are obtained in other labor market studies, e.g. on lawyers (Biddle and Hamermesh, 1998), on young U.S. workers (Mocan and Tekin, 2006), on Chinese workers (Hamermesh et al., 2002) and on U.K. workers (Harper, 2000). There are also studies that look at different treatments of men and women in the labor market that do not focus on beauty, e.g. Neumark et al. (1996) on restaurants and Goldin and Rouse (2000) on symphony orchestras.

The third strand, to which we most directly belong, studies the role of beauty in politics, the initial results of which can generally be said to be in need of verification and extension, not least in terms of whether they are characterized by gender differences. Banducci et al. (2003) find that attractive candidates to British community boards are at an advantage – moving from the lowest rating to the highest increases the candidate’s share of the votes by 14 percentage points. Female candidates are evaluated as less attractive, on average, than male candidates, but it is not reported whether the electoral effect of beauty differs between men and women. Klein and Rosar (2006), in a study of a German state election, find that variations in a candidate’s attractiveness can cause a change in his or her vote share of two to four percentage points. Being male influences a candidate’s vote share positively, but only by about 0.5 percentage point. They do not investigate if there is a gender difference in beauty assessments. King and Leigh (2006), in a study of Australian elections, report that beauty is related to electoral success (a one standard deviation increase in beauty raises the average vote share by 1.5-2 percentage points). Female candidates are considered more beautiful than male candidates but the marginal effect of beauty on electoral success is greater for the latter. The authors interpret this as an instance of the “dumb blonde syndrome”. We report below our results when testing this hypothesis. Todorov et al. (2005) find that inferences of competence from photos predict the outcomes of actual elections to the U.S. Congress quite well (71.6 percent of the Senate races and 66.8 percent of the House races).⁹ Benjamin and Shapiro (2006) report that about 20 percent of the variation of the actual vote shares in U.S. gubernatorial elections can be explained by assessments of video clips. Hamermesh (2006) looks at elections to the high offices of the American Economic Association, and his results indicate that there is a large and almost statistically significant effect of beauty on the electoral success of a male candidate; but also that there is virtually no such effect for a female candidate.¹⁰

⁹ Todorov et al. (2005) have virtually nothing to say about the gender aspects of their findings, aside from the observation that candidates who were perceived as more competent only had a small advantage from this in races between one male and one female candidate.

¹⁰ Hamermesh and Schmidt (2003), in a study of gender discrimination in elections to the Econometric Society, find that men and women are treated identically, if their qualifications are identical. Donald and Hamermesh (2006), in a study of elections to the American Economic Association, find that women are at an advantage: their chance of victory is much better than for men, and even more so when other determinants are accounted for. Dillingham et al. (1994), in another study of elections in a professional society, found that women tend to vote for other women to a disproportionate degree, whereas men were not guided by candidate gender in their voting. These three studies do not explicitly look at beauty, however.

The fourth strand, lastly, looks into the role of heuristics, information shortcuts, stereotyping and thin slices of information for voters' decisions. It thereby underpins many empirical studies in the other strands, including the present study, where it is seen to what extent observations of photos and video clips can influence perceptions and behavior. Downs (1957) is a precursor, stressing the uncertainty of voter decision-making and regarding, on this basis, parties and ideologies as devices used to attract voters who are not all that familiar with detailed policies. Lau and Redlawsk (2001) study the effects of using five different cognitive heuristics (party, ideology, endorsements, polls and candidate appearance) that many voters appear to use at times, especially when decisions are perceived to be complex. Candidate appearance seems particularly important for those low in political sophistication. In this vein, Redlawsk and Lau (2003) show that voters in their experiment preferred dislikeable candidates they agree with on the issues over much more likeable candidates they disagree with on the issues.¹¹ Schubert and Curran (2003), in studying three groups with varying degrees of knowledge about politics, find that appearance stereotyping is widely used by people with minimal political knowledge and less used by experts. Barrett and Barrington (2005) study experimentally what impact there is from using different photos of the same politician in the same newspaper article on readers' propensity to vote for him. While 20 percent in the worse picture group indicated that they would have voted for the candidate, 46 percent gave the same reply in the better-picture group.¹²

In all, as shown in the articles discussed and in the metastudy by Ambady and Rosenthal (1992), people often form evaluations and act on the basis of thin slices of information.¹³

¹¹ Cf. Budesheim and DePaola (1994: 339), who look at image- vs. issue-based evaluations of candidates and who find that "(a) physical appearance influenced evaluations even when individuating personality information was provided, (b) subjects' evaluations were less influenced by their agreement with the candidates' issue positions when image information was presented than when it was not ..."

¹² There are other heuristic devices in low-information election as well, such as race and gender, as shown by McDermott (1998).

¹³ This conclusion is supported further by Nisbett and Ross (1980), Kahneman et al. (1982), Simon (1985), Lau and Sears (1986), Fiske and Taylor (1991), Riggle et al. (1992), Lupia (1994), Macrae et al. (1994), Piattelli-Palmarini (1994), Bartels (1996), Caprara et al. (1997), Keating et al. (1999) and Willis and Todorov (2006).

3. Institutional Facts, Survey and Data

3.1 Institutional Facts

The political setting for this study is Finland, and its electoral system is proportional.¹⁴ Finland has a one-chamber legislature, and the country is divided into fourteen mainland districts electing in total 199 legislators and the autonomous province of Åland (Ahvenanmaa in Finnish) electing one. Elections are held every four years. Currently, eight parties are represented in the Parliament.

In each district, parties¹⁵ present lists of their candidates, typically in alphabetical order but sometimes incumbents listed first, and each voter can and must choose one candidate on one list in the elections. The number of candidates that a party can present equals the number of representatives elected from the district if this is 14 or more, and otherwise 14. The number of seats in the 14 mainland districts varies between seven and 32. The legislature seats of a given district are allocated based on party vote shares to the candidates in accordance with their “competitive indices”. The candidate in each party with the highest number of votes receives as his or her competitive index the total number of votes obtained by his or her party, the candidate with the second highest number of votes obtains an index calculated as half of the party votes, the third candidate gets an index equal to a third of the party votes, etc. Then all candidates are ranked on the basis of their indices, and from this list, there will be elected as many candidates as there are seats in the electoral district. In the municipal elections, competitive indices are calculated in a similar way, with each municipality forming a district.

In the 2003 national election, the participation rate among eligible citizens was 69.7 percent. Female candidates received 42.6 percent of all votes, and 75 of the elected 200 members of parliament were women (Statistics Finland, 2006). Finnish voters also elect a President. At present, Social Democrat Tarja Halonen serves her second six-year term.

¹⁴ See Raunio (2005) and the Election Act of 1998 at <<http://www.finlex.fi/pdf/saadkaan/E9980714.PDF>>

¹⁵ Or electoral alliances or joint lists. For simplicity, only parties are mentioned forthwith.

3.2 The Survey

In order for beauty to be a meaningful variable for social scientists to study, perceptions of it need to be quantified as well as reflect somewhat of a stable consensus. Langlois et al. (2000) in fact find that there is considerable and statistically significant agreement about who is and who is not attractive, both within and across cultures. As Hamermesh and Biddle (1994: 1175) put it: “within a culture at a point in time there is tremendous agreement on standards of beauty, and these standards change quite slowly.”¹⁶ On this basis, we have conducted a web survey.¹⁷ We did not only ask about beauty but also about possibly related traits in order to pinpoint more precisely how the beauty results are to be interpreted and what determines electoral success.¹⁸

We have carried out three surveys. In the first pilot survey in 2005-2006, we showed respondents from outside of Finland photos without giving any background information on the persons in photos. This survey serves as an extensive sensitivity analysis to show that average evaluations of beauty and related traits were not affected by us telling that the persons in photos are political candidates. This information was revealed in our second survey, conducted in the spring and summer of 2006 outside of Finland. With the help of dozens of colleagues, students in various universities were invited to participate, either in lectures or by e-mail. As for traditional student recruitment, the biggest participant numbers, more than 100 from each, came from Sciences Po in France and Uppsala University in Sweden. To attract also non-students, invitations to participate in our study were sent to Uppsala University alumni as well as to members of two professional associations (International Institute of

¹⁶ The same point is made by e.g. Adams (1977), Feingold (1992b), Jones and Hill (1993), Perrett et al. (1994), Cunningham et al. (1995) and Aharon et al. (2001). We also find remarkably small differences between countries. For example, the average beauty rating on a five-point scale by Swedes was 2.7, by Americans 2.8, by French 2.6, by Germans 2.8, by Danes 2.7 and by others 2.7. Since respondents from different countries may not have evaluated the same photos the same share of times, one can also compare the mean deviation of national evaluations from the mean evaluations of each photo among respondents from all nationalities. Then we get the following results: Swedes -0.0002, Americans 0.0164, French -0.1092, Germans 0.0288, Danes -0.0246 and others 0.0151. It is clear that respondents in different countries make very similar evaluations of the same photos (with the French possibly finding candidates a little less beautiful than others).

¹⁷ For an excerpt of questions asked, including reply alternatives, see Box A1 in the Appendix.

¹⁸ We do not claim that the evaluations represent true characteristics of the persons in the photos. This study is about perceptions and how these are related. None of the relationships reported should be interpreted as claims of a relationship in any underlying true characteristics.

Public Finance and European Public Choice Society). We also cooperated with several blogs¹⁹ that advertised our study. Our data collection method allows us to study separately traditional student respondents and respondents recruited in other ways. The respondents had the option to participate in a lottery of 100 euros and could also order a future summary of the results.

Each respondent was shown four photos, one at a time, two of females and two of males, randomly chosen from the database of photos. In connection with each of them, several questions were asked. There was an option, after having evaluated four photos, to evaluate additional rounds of four photos, this time with a choice as to whether to evaluate only females, only males or a continued mixture. There was no time limit for looking at the photos.²⁰ The size of the photos was approximately 5 x 3.5 centimeters (2 x 0.7 inches), and they depicted faces only. No other information than the photo was given about anyone. In Finland, political parties have in each district a poster that has a photo of each candidate. These photos were provided to us by the participating political parties. Thus, our respondents were shown the same photos as the voters. Our usage of photos that have actually been shown to the voters means that the quality of the photos or what the candidates wear etc. are less important issues than if other photos had been used. Even more, one could argue that using other photos than those actually shown to voters – which is the standard in other studies – could bias estimates of the effects of appearance on electoral success. After all, most voters do not participate in events in which they could see the candidates in person and therefore have to rely on the photos provided in advertisements.²¹ Lastly, the candidates on the photos come from four parties: the Social Democratic Party, the National Coalition Party,

¹⁹ A “blog” is a weblog, i.e. a website with personal commentary, often on current affairs and topics.

²⁰ Presumably, respondents have used different periods of time when looking at the photos, but this need not be a problem. Ambady and Rosenthal (1992) document that studies using longer periods of behavioral observation do not yield greater predictive accuracy, something which seems to hold, not least, with regard to faces (cf. Todorov et al., 2005: 1623–1624, and Willis and Todorov, 2006).

²¹ We also think that the usage of actual photos shown to voters, and that these photos all portray only faces, makes it unlikely that beauty can be much affected by spending and thereby be endogenous. Furthermore, since we use photos of non-incumbents, who have not been able to take advantage of the possibly beauty-enhancing effects of elected office, reverse causality is even more improbable. In the labor market, Hamermesh et al. (2002) look at whether one can affect one’s perceived beauty through spending on clothing and makeup. They find that such spending has a positive marginal impact on a woman’s perceived beauty – but also that it pays back no more than 15 percent of additional spending in the form of higher income.

the Left Alliance and the Green League. These parties represent 63 percent of elected members of parliament in the 2003 election.

Our third survey was carried out in fall 2006 in Finland. This time, we attracted mainly student participants. Most respondents were recruited by contacting student organizations with a web page. Also several departments at the University of Helsinki distributed the survey. Respondents could participate in a lottery of 30 movie tickets.

In the remainder of this section and in the next two sections we focus on the second survey, discussing results from the two other surveys in section 6.

3.3 Data

Number of photos

The database contained 1,929 photos of Finnish political candidates – 1,009 of men (52 percent) and 920 of women (48 percent), from the municipal (57 percent) and national level (43 percent). We only include evaluations by respondents who evaluated at least a full round of four photos. Except when studying hypothetical elections (part of our sensitivity analysis) we only include photos with at least three evaluations. This gives us 1,789 photos. In the section on the effect of beauty in real elections, we confine most of the analysis to 1,569 photos of non-incumbents. We then only exclude incumbents, not elected non-incumbents.

Number of responses

The number of photo evaluations from respondents evaluating at least four photos (one full round) was 16,218. On average, each photo was evaluated by nine respondents.

Number of respondents (per country and in total)

Table 1. Respondents by country

| Country | Country of residence | | Home country | |
|---------------|----------------------|---------|--------------|---------|
| | Number | Percent | Number | Percent |
| US | 859 | 31.0 | 800 | 28.9 |
| Sweden | 850 | 30.7 | 863 | 31.1 |
| France | 261 | 9.4 | 230 | 8.3 |
| Germany | 220 | 7.9 | 219 | 7.9 |
| Denmark | 156 | 5.6 | 154 | 5.6 |
| Other country | 426 | 15.4 | 506 | 18.2 |
| Total | 2,772 | 100 | 2,772 | 100 |

Note: Respondents denote those who assessed at least four photos (one full round). 66 percent were men, 34 percent women. 32 percent were students, an additional 14 percent doctoral students. Average age: 31.

In order to guarantee that respondents did not recognize politicians in the photos, respondents who lived in Finland were shown photos of Swedish politicians, and their evaluations are not part of this study. However, as detailed below, we later undertook a follow-up study as a sensitivity test, in which Finns evaluate the same Finnish politicians as respondents from other countries.

4. Perceptions of Beauty and Other Traits

Each photo was evaluated using five reply options, which we have converted to a five-number scale. The lowest possible beauty rating corresponds to 1, the highest possible beauty rating to 5, etc. In evaluating each trait, respondents had an option to abstain. The share of those who abstained varied between 0.5 percent for beauty and 7.9 percent for trustworthiness. There is substantial agreement among respondents; if we concentrate on two groups of beauty evaluations — above average (4 and 5) and below average (1 and 2) — the kappa coefficient of inter-rater agreement is 0.48 and highly statistically significant. The corresponding coefficients for the other four traits range from 0.18 to 0.23, all of them statistically significant at the 1 percent level.

However, men and women did not always agree on their evaluations (Table 2). There is a clear tendency for men, on average, to give photos of female candidates less positive evaluations than women do. There are smaller differences in the evaluations of photos of male candidates; the only statistically significant difference at the 5 percent level is that men find male candidates more handsome or beautiful compared to what women find.

Table 2. Average evaluations

| Variable | Men evaluating male candidates | Women evaluating male candidates | Men evaluating female candidates | Women evaluating female candidates |
|-----------------|--------------------------------|----------------------------------|----------------------------------|------------------------------------|
| Beauty | 2.644 (0.899) | 2.573 (0.912) | 2.790 (1.063) | 3.009 (0.971) |
| Competence | 3.304 (0.875) | 3.272 (0.875) | 3.205 (0.844) | 3.389 (0.845) |
| Likability | 3.069 (0.916) | 3.060 (0.951) | 3.223 (0.927) | 3.371 (0.942) |
| Trustworthiness | 3.043 (0.856) | 3.015 (0.894) | 3.292 (0.820) | 3.418 (0.826) |
| Intelligence | 3.382 (0.831) | 3.352 (0.816) | 3.229 (0.786) | 3.367 (0.794) |

Note: Standard deviations in parentheses.

On average, men perceive male candidates to be more intelligent and competent than female candidates, and female candidates to be more beautiful, likable and trustworthy. Women give more positive evaluations of female candidates in all respects, even though the difference in the evaluation of intelligence is small and not statistically significant.²² There is, lastly, no indication of a “dumb blonde syndrome”: there is a strictly positive relationship for both female and male candidates between beauty and perceived average competence and between beauty and perceived average intelligence. This holds irrespective of gender combinations or age of the candidates.

5. The Effect of Beauty in Real Elections

5.1 The Empirical Setting

We estimate the importance of beauty and other perceived traits on electoral success. Given that evaluations by Finnish voters could be influenced by their knowledge of the candidates, there is a risk that using Finnish respondents would create a systematic measurement error. To avoid this, the results here are based only on the evaluations of non-Finnish respondents.²³ In the next section, we present results with Finnish respondents, as a sensitivity analysis.

Like Hamermesh (2006), we first looked at the share of the candidates elected in real elections who score above average on their list for the traits.²⁴ In the case of beauty, a little more than half of non-incumbent candidates elected had an above-average rating. This indicates that although beauty may be an asset in politics, it is by no means a necessary requirement for getting elected. However, again we find that there is a clear gender gap: whereas only 40 percent of the elected male candidates had a beauty rating above average, the corresponding number for female candidates is 63 percent. Compared to other non-incumbent

²² For correlation coefficients, see Table A1 in the Appendix.

²³ None of the respondents correctly recognized anyone of the candidates. In 17 cases the respondent mistook a candidate for another politician. Tarja Halonen was the only Finnish politician that anyone, incorrectly, claimed to recognize. Ten answers were of the kind “I recognize her but don’t remember her name.”

²⁴ See Table A2 in the Appendix.

candidates of their own gender, 57 percent of elected men and 65 percent of elected women were evaluated to be of above-average beauty.

A more detailed picture emerges if we look at average evaluations and also take the gender of the respondents into account. Both men and women evaluate elected and non-elected male candidates similarly. One notable difference is that perceived competence is a bit higher among elected male candidates compared to non-elected male candidates. For beauty, the evaluations of elected and non-elected male candidates are very close to each other.²⁵ For female candidates the story is quite different. Both men and women assign higher beauty scores to elected than to non-elected female candidates. Other differences are smaller, but not as small as for male candidates. Here one can mention that men seem to give elected female candidates higher competence evaluations than they give to non-elected female candidates.²⁶

Next we test to what extent beauty and other traits can be related to the relative success of candidates in actual elections in 2003 and 2004. In this we pay particular attention to gender differences. We focus first on the large group of non-incumbent candidates (elected and non-elected) and then look at the full set of candidates, including incumbents. The reason that we make this division (unlike the previous literature) is that incumbency is a strong predictor of electoral success (see e.g. Lee, 2001) and the processes that determine electoral outcomes appear to be different for incumbents and non-incumbents.²⁷ Using incumbency to explain electoral success in this kind of study would also dodge the question why the incumbent politicians were elected in the first place. Because of multicollinearity between the five traits, only beauty, competence, and trustworthiness are used in our regressions: we exclude likability, which is quite highly correlated with beauty, and intelligence, which is strongly correlated with competence.²⁸

²⁵ See Figure A1 in the Appendix. However, incumbent candidates are seen as slightly more beautiful than non-incumbent candidates (an average of 2.82 vs. an average of 2.73).

²⁶ See Figure A2 in the Appendix.

²⁷ Because of the proportional electoral system with party lists and intraparty competition we are able to study non-incumbents only, whereas a plurality-vote system like that of the U.S. typically features an incumbent and a non-incumbent from different parties facing each other. Benjamin and Shapiro (2006) use a dummy for incumbency, but it is unclear to what extent it captures all effects of incumbency, which might appear in other variables or in other ways. Hence, we consider our setting an advantage.

²⁸ This is in line with factor analysis carried out by Todorov et al. (2005: 1624), who identify competence, trustworthiness and likability as the main factors in predicting electoral outcomes.

The trait variables are constructed in two steps. First we compute the mean over all respondents who evaluated a particular photo. From this measure we then subtract its mean over the non-incumbent candidates on the same list. That is, we use *relative* measures of the different traits, capturing how beautiful, competent and trustworthy a candidate is in relation to his or her competitors on the list.

The dependent variable, *relative success*, is defined in the following way for candidate i on list j :

$$\text{relative success}_{i,j} = (p_i / v_j) * 100$$

where p_i is candidate i 's number of personal votes and v_j is all votes for candidates on list j divided by the number of candidates on list j .²⁹ In the next section, we use a candidate's vote share as the dependent variable in a sensitivity analysis, but the advantage of the relative success measure is that it makes election outcomes comparable, as list sizes differ (especially between national and municipal elections).

As regressors, we use the three trait variables *beauty*, *competence* and *trustworthiness*, measuring for each candidate the average trait evaluation minus the average evaluation of that trait over all candidates on the same list. In an extended regression we also include the age dummies *young*, which denotes an age under 30, and *old*, which denotes an age over 60. This is a way to control for possible age effects – as it is e.g. clear that both women and men consider younger candidates of both genders more beautiful, on average, than older candidates.³⁰

5.2 Non-Incumbent Candidates

We begin by looking at the effects in the national election for female and male non-incumbent candidates. This means that the variables defined above are calculated by exclusively including non-incumbent candidates. Most notably, as reported in Table 3, we find that relative beauty is clearly the most important explanatory variable of relative success both for female and for male candidates, and the only regressor that consistently attains sta-

²⁹ The mean of relative success is 100, capturing that on average each candidate must receive a 1 / list size share of the votes. The average of relative success for elected candidates is 348. That is, they receive 3.48 times the votes of the average candidate.

³⁰ Detailed figures are available upon request.

tistical significance. A higher relative beauty score of one unit (which is very close to the standard deviation for both female and male candidates) implies an increase of between 19.4 and 23.6 units of relative success for female candidates and between 17.3 and 20.9 units for male candidates. To ease interpretation, note that an increase of 1 unit in relative success means a one-percentage increase in the number of voters, relative to the average number of votes of all candidates in the same list. For example, column 2 reports that an increase by one step in the average beauty evaluation is associated with a 19-percent increase in the number of voters for female candidates. For male candidates, specification (3) indicates that relative perceived competence is statistically significant and almost as large as relative beauty. When controlling for beauty, competence and trustworthiness evaluations, being young is a disadvantage especially for female candidates. Old age is an advantage for male candidates only.

Table 3. Relative success in the national election, non-incumbents

| | (1) Relative success All candidates | (2) Relative success Female candi- dates | (3) Relative success Male candidates | (4) Relative success All candidates | (5) Relative success Female candi- dates | (6) Relative success Male candidates |
|------------------|---|---|--|---|---|--|
| beauty | 18.30*** (4.79) | 19.38*** (6.78) | 17.31** (6.74) | 21.55*** (4.96) | 23.60*** (7.10) | 20.86*** (6.84) |
| competence | 9.932 (6.66) | 5.301 (11.4) | 13.36* (7.40) | 7.246 (6.73) | 0.959 (11.5) | 9.155 (7.54) |
| trustworthiness | 6.408 (6.92) | 8.626 (11.4) | 4.751 (7.99) | 5.936 (6.92) | 11.25 (11.4) | 1.934 (8.01) |
| male candidate | -1.289 (6.19) | | | -2.073 (6.18) | | |
| young (age < 30) | | | | -18.34** (7.60) | -22.95** (10.9) | -13.92 (10.4) |
| old (age>60) | | | | 4.406 (15.8) | -31.76 (24.3) | 43.18** (19.4) |
| Observations | 646 | 348 | 298 | 646 | 348 | 298 |
| R-squared | 0.04 | 0.03 | 0.05 | 0.05 | 0.05 | 0.07 |

Note: Standard errors in parentheses. The regressions include a constant term. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Compared to the national election, Table 4 reveals that in the municipal elections, relative beauty is somewhat less important for female candidates but still statistically significant. For male candidates, statistical significance vanishes and the size of the estimates is reduced considerably. A higher relative beauty score of one unit (again, the approximate standard deviation for both female and male candidates) implies an increase of between 15.0 and 21.7 units of relative success for female candidates and between 6.5 and 8.9 units for male candidates. Here, relative perceived competence is statistically significant only for male candidates and the effect is larger than for relative beauty.

If we exclude beauty, the effect of competence is substantially higher and statistically more significant. This would reconcile our findings with those of Todorov et al. (2005), who

do not include beauty in most of their analysis and find that competence plays a primary role. Our findings suggest that as perceptions of beauty and competence are positively correlated, excluding beauty as a control results in an overestimation of the effect of perceived competence.

Table 4. Relative success in municipal elections, non-incumbents

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------|------------------------------------|--|-------------------------------------|------------------------------------|--|-------------------------------------|
| | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates |
| beauty | 12.48*** (4.17) | 14.96** (6.29) | 6.465 (5.26) | 17.23*** (4.44) | 21.71*** (7.01) | 8.874 (5.39) |
| competence | 19.28*** (6.49) | 32.52*** (11.6) | 9.844 (6.76) | 15.00** (6.62) | 25.99** (11.9) | 6.572 (6.84) |
| trustworthiness | -14.57** (6.87) | -22.52* (12.4) | -7.588 (7.11) | -15.89** (6.86) | -24.32* (12.4) | -8.385 (7.08) |
| male candidate | -19.92*** (5.35) | | | -20.61*** (5.34) | | |
| young (age < 30) | | | | -16.05** (6.29) | -22.51** (10.3) | -8.571 (7.13) |
| old (age > 60) | | | | 12.26 (8.78) | 5.170 (16.3) | 20.34** (8.93) |
| Observations | 921 | 464 | 457 | 921 | 464 | 457 |
| R-squared | 0.04 | 0.04 | 0.01 | 0.05 | 0.05 | 0.03 |

Note: Standard errors in parentheses. The regressions include a constant term. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

To further pinpoint the relationship between beauty and electoral success, and to see whether the relationship is driven by outliers, we have computed Spearman rank correlations for the 449 non-incumbent candidates in the Helsinki municipal elections.³¹ The Spearman rank correlation between *beauty* and relative success is especially strong for female candidates, for whom Spearman's rho is 0.278.³²

The analysis of Spearman rank correlations also allows us to compare the relationship between electoral success and the evaluations of the five different traits one at a time and to implement a horse race between these as explanatory variables for electoral success. For both females and males, the Spearman rank correlation between electoral success and beauty is larger and has a higher statistical significance than the rank correlation between electoral success and perceived competence, intelligence, trustworthiness and likability.

³¹The Helsinki municipal elections are best suited for this, thanks to the large number of candidates and all four parties having about the same number of non-incumbents.

³² See Table A3 in the Appendix.

5.3 All Candidates (Incumbents and Non-Incumbents)

The previous literature has, probably because of the focus on plurality-vote systems, not studied non-incumbents separately. We now investigate what the effect would be, as shown in Tables 5 and 6, of adding incumbents and an incumbency dummy.

Table 5. Relative success in the national election, incumbents and non-incumbents

| | (1) | (2) | (3) | (4) | (5) | (6) |
|-----------------|--------------------|------------------------|--------------------|--------------------|------------------------|--------------------|
| | Relative success | Relative success | Relative success | Relative success | Relative success | Relative success |
| | All candidates | Female candi- dates | Male candidates | All candidates | Female candi- dates | Male candidates |
| beauty | 17.08*** (6.14) | 20.30** (9.11) | 12.10 (7.83) | 20.13*** (6.33) | 25.65*** (9.50) | 13.74* (7.99) |
| competence | 15.09* (8.44) | 7.874 (15.1) | 19.78** (8.62) | 12.45 (8.51) | 2.576 (15.2) | 17.94** (8.79) |
| trustworthiness | 5.090 (8.77) | 12.21 (15.6) | 0.856 (9.07) | 5.410 (8.78) | 15.92 (15.6) | -0.425 (9.16) |
| incumbent | 188.0*** (10.4) | 197.8*** (17.7) | 179.1*** (11.2) | 185.7*** (10.6) | 194.2*** (17.9) | 176.9*** (11.4) |
| male candidate | -4.637 (7.81) | | | -5.208 (7.80) | | |
| young (age< 30) | | | | -22.13** (10.1) | -30.90** (15.0) | -11.84 (12.9) |
| old (age>60) | | | | -15.19 (17.9) | -45.62 (30.4) | 11.01 (19.3) |
| Observations | 743 | 394 | 349 | 743 | 394 | 349 |
| R-squared | 0.33 | 0.27 | 0.44 | 0.33 | 0.28 | 0.44 |

Note: incumbent is a dummy variable for incumbent candidates.

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

For the national election, this boosts the explanatory power of the model significantly. The results for female candidates are similar to when only non-incumbents were studied, but for male candidates beauty is less important and perceived competence more important (and clearly statistically significant). Youth is a disadvantage for female candidates. The strong benefits of incumbency are very clearly confirmed.

Table 6. Relative success in municipal elections, incumbents and non-incumbents

| | (1) | (2) | (3) | (4) | (5) | (6) |
|-----------------|------------------------------------|--|-------------------------------------|------------------------------------|--|-------------------------------------|
| | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates |
| beauty | 17.72** (8.11) | 22.81** (11.5) | 8.049 (11.6) | 20.06** (8.68) | 30.20** (12.8) | 6.727 (12.0) |
| competence | 10.22 (12.6) | 11.81 (20.8) | 11.12 (15.2) | 8.205 (12.9) | 4.968 (21.3) | 12.01 (15.5) |
| trustworthiness | -2.895 (13.2) | -12.30 (22.4) | 4.325 (15.4) | -3.439 (13.2) | -14.51 (22.5) | 4.562 (15.5) |
| incumbent | 339.6*** (14.7) | 359.9*** (22.9) | 318.7*** (18.4) | 337.9*** (14.9) | 353.9*** (23.4) | 319.3*** (18.8) |
| male candidate | -19.97* (10.3) | | | -20.09* (10.4) | | |
| young (age< 30) | | | | -10.07 (12.7) | -27.02 (19.4) | 8.306 (16.4) |
| old (age>60) | | | | 0.842 (15.9) | 4.363 (27.9) | 2.886 (18.4) |
| Observations | 1043 | 524 | 519 | 1043 | 524 | 519 |
| R-squared | 0.35 | 0.33 | 0.38 | 0.35 | 0.33 | 0.38 |

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

For the municipal elections, comparing again with the non-incumbency results, beauty has a larger effect for female candidates, whereas the previous effect of perceived competence vanishes. For male candidates, only incumbency attains statistical significance.

5.4 Summary

To summarize, is perceived beauty an asset in politics? Our findings indicate that it is, especially in the national context and to a higher extent for female than for male candidates. However, although it generally appears to be a bigger asset than e.g. competence, as inferred from photos, it does not seem to be a *dominant* explanation of relative success in the Finnish, proportional electoral system. First of all, the R²s reported are quite low, indicating that the fit of the empirical model is not all that precise (although the inclusion of incumbents boosts the R²s quite a bit). Second, the size of the effect is rather modest, even where it matters the most (i.e. for female candidates). However, even a small difference in relative success may make a difference inside a party list if the candidates — like they are in Finland — are elected in the order determined by the number of personal votes.

6. Sensitivity Analysis

We will now investigate to what extent the results reported thus far are sensitive to various alternative ways of investigating the relationship between beauty and electoral success. We report briefly on the results, but in each case, the detailed results are available upon request.

6.1 Finnish Respondents

For all results reported thus far, we have only used respondents from other countries than Finland in order to avoid the risk that respondents recognize the Finnish political candidates. Such recognition could severely bias the evaluations. However, it could be that Finns and non-Finns – in the cases where candidates are not recognized by either category – evaluate the candidates differently on the traits. If there are large differences, trait evaluations from non-Finns could not be expected to predict electoral success for Finnish candidates particularly well. To investigate this, we have undertaken a follow-up study based on the same set of political candidates with only Finnish respondents. The results, from 3,698 respondents who completed at least a full round of evaluating four photos and 25,598 responses, indicate small differences.

First, average evaluation can be compared, as reported for Finns in Table 7 and for non-Finns in Table 2. The overall pattern is the same. Unlike in our main study, Finnish women give more positive evaluations than Finnish men of male candidates' competence ($p=0.01$). The differences for beauty, trustworthiness and intelligence are not statistically significant.³³

³³ Correlations between the traits are also similar to the ones reported for non-Finns in Table A1 in the Appendix; and the Kappa coefficients of inter-rater agreement are virtually identical.

Table 7. Average evaluations by Finnish respondents³⁴

| Variable | Men evaluating male candidates | Women evaluating male candidates | Men evaluating female candidates | Women evaluating female candidates |
|-----------------|--------------------------------|----------------------------------|----------------------------------|------------------------------------|
| Beauty | 2.70 (0.88) | 2.68 (0.85) | 2.87 (1.00) | 3.15 (0.84) |
| Competence | 3.20 (0.85) | 3.25 (0.79) | 3.09 (0.81) | 3.34 (0.76) |
| Likability | 2.97 (0.89) | 3.03 (0.88) | 3.16 (0.92) | 3.35 (0.86) |
| Trustworthiness | 3.02 (0.87) | 3.02 (0.82) | 3.25 (0.81) | 3.40 (0.77) |
| Intelligence | 3.30 (0.81) | 3.32 (0.73) | 3.17 (0.75) | 3.34 (0.68) |

Note: Standard deviations in parentheses. Excluding recognized candidates.

Second, we have done regressions for the national and municipal elections using responses from Finns only. Here, we only report results for the national election for reasons of brevity, but the results are qualitatively similar. As before, we restrict ourselves to non-incumbents. In Table 8, we exclude evaluations of candidates that respondents indicated were recognized (by stating the candidate's first or second name or both). In Table 7, we include evaluations of recognized candidates to see what the effect is. These results are to be compared with those reported for non-Finns in Table 3.

Such a comparison reveals quite substantial differences.³⁵ For one, the effect of beauty on relative success is much larger for female candidates and smaller and statistically insignificant for male candidates. The effect of perceived competence is, on the other hand, much larger for male candidates as well as more statistically significant, whereas it is smaller and statistically insignificant for female candidates.

³⁴ A corresponding table has been calculated for all evaluations, including those of recognized candidates, and the differences are minimal.

³⁵ The comparison is complicated by the fact that the share of female respondents is 73 percent in the Finnish study and only 32 percent in the study of non-Finns.

Table 8. Relative success in the national election, non-incumbents, non-recognized candidates only

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------|------------------------------------|--|-------------------------------------|------------------------------------|--|-------------------------------------|
| | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates |
| beauty | 26.61*** (7.56) | 41.73*** (10.6) | 9.887 (11.0) | 32.31*** (8.05) | 47.18*** (11.3) | 16.93 (11.6) |
| competence | 37.27*** (11.8) | 9.812 (19.9) | 54.63*** (14.3) | 29.88** (12.3) | 2.354 (20.5) | 45.78*** (15.1) |
| trustworthiness | -22.54* (11.9) | -11.89 (19.3) | -26.32* (15.0) | -21.29* (11.9) | -8.168 (19.4) | -26.69* (14.9) |
| male candidate | -2.578 (8.00) | | | -1.696 (8.00) | | |
| young (age < 30) | | | | -17.67* (9.10) | -21.00* (12.3) | -16.70 (13.7) |
| old (age >60) | | | | 8.772 (17.0) | -32.73 (26.0) | 42.68* (21.9) |
| Observations | 664 | 326 | 338 | 664 | 326 | 338 |
| R-squared | 0.05 | 0.06 | 0.05 | 0.05 | 0.07 | 0.07 |

Note: Evaluations of recognized candidates are excluded. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

However, taking a look at Table 9 indicates that recognition of candidates affects the results. Including evaluations of recognized candidates reduces the effect of beauty but spectacularly raises the effect of competence (which also becomes statistically significant) for female candidates. Since recognition in reality is not a dichotomous but a continuous variable, it is probable that the results of Table 8 are also affected by recognition.

Table 9. Relative success in the national election, non-incumbents, including recognized candidates

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------|------------------------------------|--|-------------------------------------|------------------------------------|--|-------------------------------------|
| | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates | Relative success All candidates | Relative success Female candi- dates | Relative success Male candidates |
| beauty | 20.57*** (7.87) | 28.79** (11.4) | 9.980 (11.2) | 25.13*** (8.41) | 33.35*** (12.2) | 15.17 (11.9) |
| competence | 60.46*** (12.4) | 64.38*** (21.3) | 57.14*** (14.9) | 54.21*** (13.0) | 57.16** (22.2) | 50.65*** (15.8) |
| trustworthiness | -25.91** (12.6) | -19.34 (20.6) | -27.55* (15.8) | -24.91** (12.6) | -15.52 (20.6) | -28.09* (15.8) |
| male candidate | -4.448 (8.40) | | | -3.600 (8.41) | | |
| young (age < 30) | | | | -14.47 (9.59) | -18.46 (13.3) | -11.04 (14.0) |
| old (age > 60) | | | | 5.048 (18.3) | -36.67 (29.1) | 36.86 (23.2) |
| Observations | 671 | 326 | 345 | 671 | 326 | 345 |
| R-squared | 0.06 | 0.07 | 0.05 | 0.06 | 0.08 | 0.06 |

Note: Evaluations of recognized candidates are included. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In general, we think that this finding indicates that the results of previous studies should be interpreted with some caution, as they are based on evaluations by respondents of the same nationality as the political candidates. This entails a risk for non-reported recognition.

As a sensitivity analysis, the follow-up study with Finnish respondents confirms that beauty has an effect on relative success for female candidates in the national election. Unlike the study with non-Finns, male candidates's beauty does not have any effect, but perceived competence clearly does. A further difference is that female candidates' perceived competence is very important when evaluations of recognized candidates are included.

6.2 Vote Share as Dependent Variable

We have replaced relative success with vote share as the dependent variable in the regressions reported in section 5, to see whether the results are qualitatively affected. *Vote share* is defined in the following way for candidate i on list j ,

$$vote\ share_{i,j} = (p_i / w_j) * 100$$

where p_i is non-incumbent candidate i 's number of personal votes and w_j is all votes for non-incumbent candidates on list j . That is, $w_j = v_j * \text{the number of non-incumbent candidates on list } j$. This variable is easier to interpret intuitively than relative success, but since the number of candidates differ between lists, the estimated coefficients for different lists are not readily comparable.

Table 10. Vote shares in the national election, non-incumbents

| | (1) Vote share All candidates | (2) Vote share Female candi- dates | (3) Vote share Male candidates | (4) Vote share All candidates | (5) Vote share Female candi- dates | (6) Vote share Male candidates |
|------------------|-------------------------------------|---|--------------------------------------|-------------------------------------|---|--------------------------------------|
| beauty | 2.169*** (0.47) | 2.560*** (0.62) | 1.583** (0.73) | 2.447*** (0.48) | 2.966*** (0.65) | 1.798** (0.74) |
| competence | 0.877 (0.65) | 0.226 (1.05) | 1.353* (0.80) | 0.644 (0.65) | -0.186 (1.06) | 1.100 (0.82) |
| trustworthiness | 0.512 (0.67) | 0.852 (1.04) | 0.325 (0.86) | 0.511 (0.67) | 1.097 (1.04) | 0.166 (0.87) |
| male candidate | 0.165 (0.60) | | | 0.0983 (0.60) | | |
| young (age < 30) | | | | -1.773** (0.74) | -2.198** (1.00) | -1.442 (1.12) |
| old (age > 60) | | | | -0.805 (1.53) | -2.876 (2.22) | 1.154 (2.10) |
| Observations | 646 | 348 | 298 | 646 | 348 | 298 |
| R-squared | 0.05 | 0.06 | 0.04 | 0.06 | 0.08 | 0.05 |

Note: Standard errors in parentheses. The regressions include a constant term. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

As revealed in Table 10, the results are qualitatively very similar to those of Table 3. Again and most notably, we find that relative beauty is by far the most important explana-

tory variable of vote share both for female and for male candidates. Competence attains statistical significance in one specification for male candidates as well. A higher relative beauty score of one unit (which is very close to the standard deviation for both female and male candidates) implies an increase of between 2.6 and 3.0 percentage points in the vote share for female candidates and between 1.6 and 1.8 percentage points for male candidates.³⁶ In the municipal elections, for which the results are not reported here for reasons of brevity, beauty is again statistically significant and again more important for female candidates. Interestingly, in local elections we get the result that competence is more important than beauty for female candidates, but it seems rather unimportant for male candidates. The main qualitative difference, when compared to Table 4, is that beauty here does not attain statistical significance for male candidates in any specification.

6.3 Hypothetical Election

To further see if there is a relationship between evaluations of beauty and a propensity to choose a political candidate, we asked respondents to vote for one of four candidates (the ones they had just evaluated) in a hypothetical election, or to abstain from voting.³⁷ If one looks at the share of the thus “elected” candidates that were also picked as the most extreme one (positively so) in the evaluated traits, one actually finds that beauty gets the lowest score: only 45 percent of the respondents thought that the candidate they chose to vote for was also the most beautiful one. Competence seems to be the most important trait in this regard: 60 percent of the respondents thought that the candidate they chose to vote for was the most competent one. It does seem clear that the traits play a role.

In Table 11, we report results of a linear probability model with the respondents’ choices of candidates in the hypothetical election as the dependent variable. The explanatory variables are dummy variables for choices as the most beautiful, the most competent and the most trustworthy candidate. Male candidate is a dummy for male candidates. The table is only based on choices between two male and two female candidates. It turns out that both men and women prefer candidates of their own gender, a pattern which is especially strong

³⁶ The results are similar if we instead use each candidate’s vote share among non-incumbents minus $1/n$, where n is the number of non-incumbents on each list.

³⁷ The instruction reads: “Sometimes people have to vote in an election with only a little information. Let us assume that you would have to either vote for one of these persons as a member of Parliament [non-US respondents]/the House of Representatives [US respondents], or abstain from voting. Which would be your choice?”.

for women. Beauty is less important than competence for both men and women, but still, the probability of getting elected increases by 16 percent for a female candidate if she was ranked as the most beautiful.³⁸

Table 11. Linear probability model of the hypothetical election

| | (1) | (2) |
|------------------|----------------------|---------------------|
| | Elected by women | Elected by men |
| beautychoice | 0.161*** (0.012) | 0.143*** (0.009) |
| competencechoice | 0.294*** (0.012) | 0.295*** (0.009) |
| trustchoice | 0.210*** (0.012) | 0.210*** (0.009) |
| male candidate | -0.063*** (0.010) | 0.024*** (0.008) |
| Observations | 5172 | 8650 |
| R-squared | 0.27 | 0.24 |

Standard errors in parentheses. The regressions include a constant term.

* significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent

Furthermore, to see whether the effect of beauty is non-linearly related to being elected, we have replaced the beauty variable with dummies for each of the five beauty levels. Overall, the effect is approximately linear.

6.4 Students and Non-Students

Unlike most previous studies, we not only use students as respondents. It turns out that the evaluations by (undergraduate and graduate) students and other respondents are remarkably similar, with the only statistically significant differences being that students evaluate the candidates somewhat more negatively in beauty (average of 2.69 vs. 2.79 for non-students) and somewhat more positively in trustworthiness (average of 3.23 vs. 3.17 for non-students).

6.5 Survey without Information about the Photos

We have conducted another web survey with the same photos, getting almost 39,000 responses from about 3,500 participants from outside of Finland. We focus on the current survey as we did not report in the earlier survey that the photos depict political candidates or that we are studying politics, and thus did not ask who the respondents would vote for. Nonetheless, the earlier survey serves as an extensive sensitivity analysis of the results and

³⁸ The results are confirmed in a probit model.

indicates high test-retest reliability. When both surveys are combined, each photo has received almost 30 ratings on average from non-Finns alone, rendering the results statistically very robust. Importantly, we found that the results were relatively similar whether the respondents knew that the persons in photos are politicians or not.

7. Concluding Remarks

People's visual appearance has been shown to exert an independent influence on how they are perceived and treated in many settings. For example, good looks seem to be a distinct advantage in the labor market. An emerging literature indicates that beauty may play an important role in politics as well (see e.g. Banducci et al., 2003, and Klein and Roser, 2006). We add to this literature by studying whether there are differences in how men and women are assessed by men and women. Our study is the first one to analyze the role that beauty and other perceived traits play in real political elections taking into account gender differences among both electoral candidates and respondents.

As voters in Finland vote for a person on a party list in multi-member jurisdictions, our analysis especially sheds light on the role of beauty for electoral success in a proportional electoral system, but the results are valuable also for majoritarian electoral systems with primary elections, where there are also elements of intraparty competition. One advantage of focusing on such competition is that it allows us to control for ideology very effectively and look purely at the effects of beauty and other perceived traits.

Our results, based on a much bigger dataset than has previously been used, imply that relative beauty does play a role for electoral success – candidates who are more beautiful than their list competitors are more successful. In national elections, an increase in beauty of one unit is associated in average with 17 to 24 percent more votes. In municipal elections, the corresponding figure is 15 to 22 percent for female candidates, but it turns statistically insignificant for male candidates as well as much smaller.

Extensive sensitivity analysis confirms the results. Using Finnish respondents, exchanging the dependent variable, exploring choice in a hypothetical election, separating students and non-students, using respondents that were not informed that they evaluated politicians – all these changes do not alter the qualitative findings. Beauty matters in politics, but it is not a necessary nor a sufficient condition for success.

Appendix

Box A1. Excerpt from the web survey

What is your evaluation of the physical appearance or attractiveness of this person compared to the average among people living in your country of residence?
 Very unattractive
 Below average
 Average
 Above average
 Very handsome or beautiful
 Cannot say/Prefer not to answer

What is your evaluation of the competence of this person compared to the average among people living in your country of residence?
 Very incompetent
 Below average
 Average
 Above average
 Very competent
 Cannot say/Prefer not to answer

What is your evaluation of the likability of this person (i.e. how nice, pleasant, and agreeable do you find this person) compared to the average among people living in your country of residence?
 Very unlikable
 Below average
 Average
 Above average
 Very likable
 Cannot say/Prefer not to answer

What is your evaluation of the trustworthiness of this person (i.e. how ethical, honest, and responsible do you find this person) compared to the average among people living in your country of residence?
 Very untrustworthy
 Below average
 Average
 Above average
 Very trustworthy
 Cannot say/Prefer not to answer

What is your evaluation of the intelligence of this person compared to the average among people living in your country of residence?
 Very unintelligent
 Below average
 Average
 Above average
 Very intelligent
 Cannot say/Prefer not to answer

What is your evaluation of the age of this person? Use your keyboard to fill in the age in the box below.

Table A1. Correlation matrix

| | beauty | competence | likability | trustworthiness | Intelligence |
|-----------------|--------|------------|------------|-----------------|--------------|
| beauty | 1.0000 | | | | |
| competence | 0.3154 | 1.0000 | | | |
| likability | 0.4077 | 0.3185 | 1.0000 | | |
| trustworthiness | 0.2169 | 0.3801 | 0.5058 | 1.0000 | |
| intelligence | 0.2778 | 0.6531 | 0.2803 | 0.3632 | 1.0000 |

Note: All of the reported traits exhibit positive (and statistically significant) correlations with each other. The strongest correlation is the one between competence and intelligence. The weakest is the one between beauty and trustworthiness. Beauty is most strongly correlated with likability. Correlations are similar for students and non-students, with the only statistically significant differences being that students associate trustworthiness less strongly with beauty (correlation 0.199 for students and 0.235 for non-students) and with competence (correlation 0.360 for students and 0.396 for non-students).

Figure A1. Evaluations of elected and non-elected male candidates

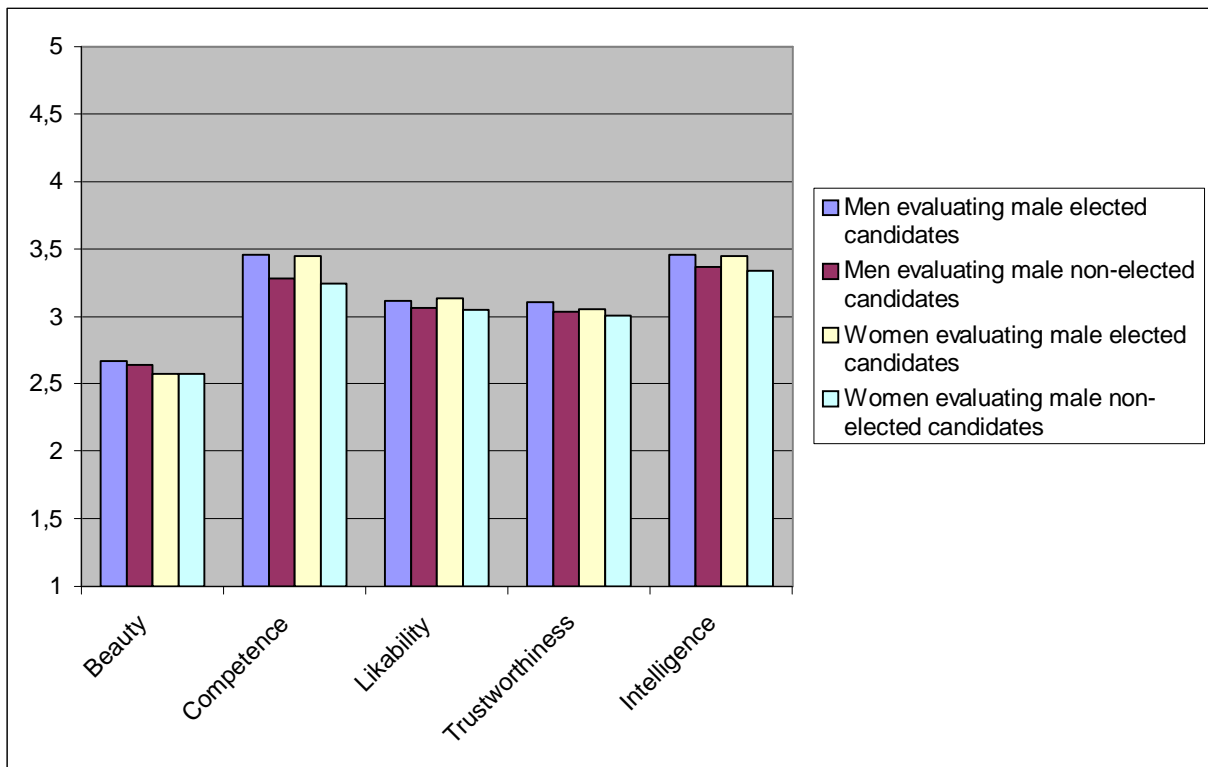


Figure A2. Evaluations of elected and non-elected female candidates

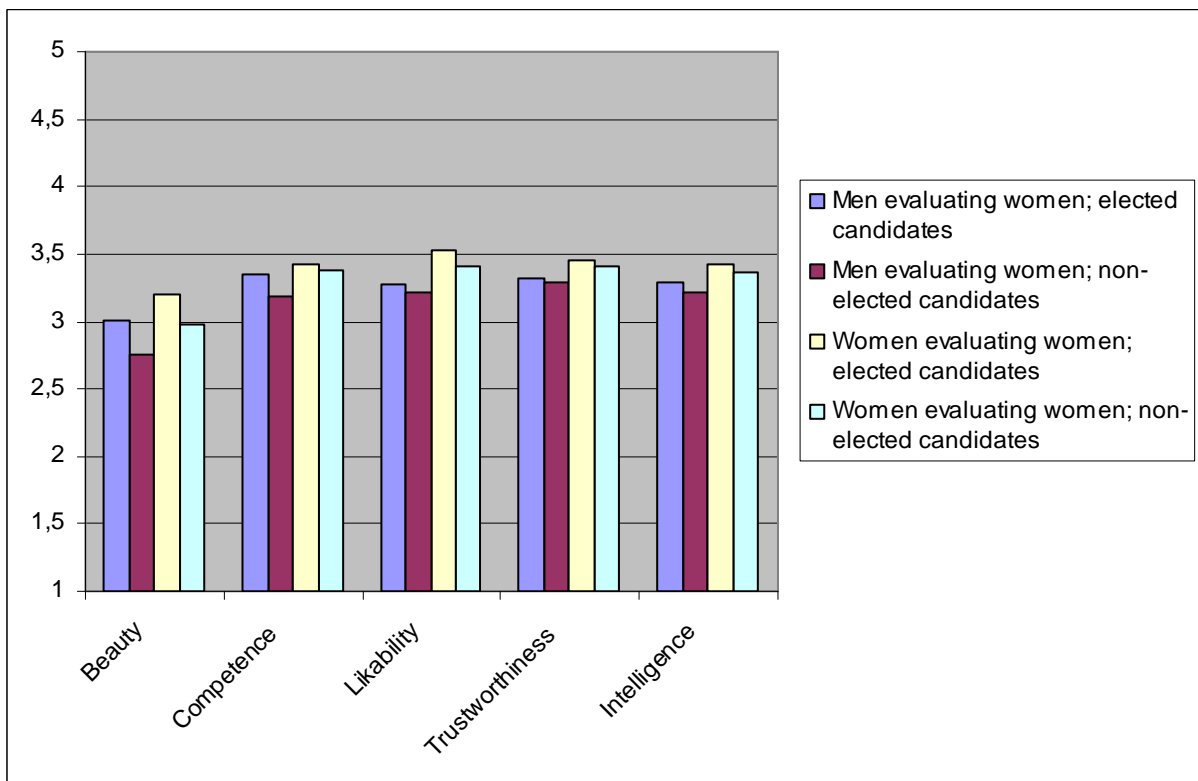


Table A2. Share of elected non-incumbent candidates with scores above average, real elections

| Trait | All elected candidates | Elected male candidates | Elected female candidates |
|----------------------|------------------------|-------------------------|---------------------------|
| beauty | 52.2 % | 40.0 % | 62.8 % |
| competence | 63.3 % | 65.6 % | 61.4 % |
| likability | 57.8 % | 40.8 % | 72.4 % |
| trustworthiness | 56.3 % | 40.0 % | 70.3 % |
| intelligence | 58.5 % | 64 % | 53.8 % |
| # elected candidates | 270 | 125 | 145 |

Table A3. Spearman rank correlations for Helsinki, non-incumbent candidates only

| | | All candidates | Male candidates | Female candidates |
|-----------------|--|--------------------|--------------------|---------------------|
| beauty | Spearman's ρ | 0.247 | 0.138 | 0.278 |
| | Test of H0: beauty and vote share are independent | Prob > t = 0.000 | Prob > t = 0.041 | Prob > t = 0.000 |
| competence | Spearman's ρ | 0.134 | 0.069 | 0.181 |
| | Test of H0: competence and vote share are independent | Prob > t = 0.004 | Prob > t = 0.311 | Prob > t = 0.0061 |
| likability | Spearman's ρ | 0.148 | 0.062 | 0.132 |
| | Test of H0: likability and vote share are independent | Prob > t = 0.002 | Prob > t = 0.357 | Prob > t = 0.047 |
| trustworthiness | Spearman's ρ | 0.095 | -0.063 | 0.006 |
| | Test of H0: trustworthiness and vote share are independent | Prob > t = 0.044 | Prob > t = 0.353 | Prob > t = 0.926 |
| intelligence | Spearman's ρ | 0.051 | 0.013 | 0.167 |
| | Test of H0: intelligence and vote share are independent | Prob > t = 0.284 | Prob > t = 0.843 | Prob > t = 0.011 |
| | # candidates | 449 | 220 | 229 |

References

- Adams, G. (1977). "Physical Attractiveness Research: Toward a Developmental Social Psychology of Beauty." *Human Development* 20: 217–239.
- Aharon, I., Etcoff, N., Ariely, D., Chabris, C., O'Connor, E. and Breiter, H. (2001). "Beautiful Faces Have Variable Reward Value: fMRI and Behavioral Evidence." *Neuron* 32: 537–551.
- Ambady, N. and Rosenthal, R. (1992). "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis." *Psychological Bulletin* 111(2): 256–274.
- Andreoni, J. and Petri, R. (2004). "Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments." Working Paper. University of Wisconsin, Madison, WI.
- Banducci, A. A., Thrasher, M., Rallings, C. and Karp, J. A. (2003). "Candidate Appearance Cues in Low-Information Elections." Paper presented at the Annual Conference of the American Political Science Association, Philadelphia.

- Barrett, A. W. and Barrington, L. W. (2005). "Is a Picture Worth a Thousand Words? Newspaper Photographs and Voter Evaluations of Political Candidates." *The Harvard International Journal of Press/Politics* 10(4): 98-113.
- Bartels, L. M. (1996). "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40(1): 194-231.
- Benjamin, D. J. and Shapiro, J. M. (2006). "Thin-Slice Forecasts of Gubernatorial Elections." NBER Working Paper No. 12660. National Bureau of Economic Research (NBER), Cambridge, MA.
- Biddle, J. E. and Hamermesh, D. S. (1998). "Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics* 16(1): 172-201.
- Brennan, G. and Lomasky, L. (1993). *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge: Cambridge University Press.
- Budesheim, T. L. and DePaola, S. J. (1994). "Beauty or the Beast? The Effects of Appearance, Personality, and Issue Information on Evaluations of Political Candidates." *Personality and Social Psychology Bulletin* 20(4): 339-348.
- Caplan, B. (2002). "Systematically Biased Beliefs about Economics: Robust Evidence of Judgemental Anomalies from the Survey of Americans and Economists on the Economy." *Economic Journal* 112(April): 1-12.
- Caprara, G. V., Barbaranelli, C. and Zimbardo, P. (1997). "Politicians' Uniquely Simple Personalities." *Nature* 385(February): 493.
- Cunningham, M., Roberts, A. R., Barbee, A. P., Cruen, P..B. and Wu, C.-H. (1995). "Consistency and Variability in the Cross-Cultural Perception of Female Physical Attractiveness." *Journal of Personality and Social Psychology* 68: 261-279.
- Dillingham, A., Ferber, M. and Hamermesh, D. S. (1994). "Gender Discrimination by Gender: Voting in a Professional Society." *Industrial and Labor Relations Review* 47(4): 622-633.
- Donald, S. G. and Hamermesh, D. S. (forthcoming). "What Is Discrimination? Gender in the American Economic Association, 1925-2004." *American Economic Review*.
- Downs, A. (1957). *An Economic Theory of Democracy*. New York, NY: HarperCollins.
- Eagly, A. H., Ashmore, R.D., Makhijani, M. G. and Longo, L. C. (2001). "What Is Beautiful Is Good, But: A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype." *Journal of Management* 27: 363-381.
- Eckel, C. C. and Wilson, R. K. (2006). "Judging a Book by Its Cover: Beauty and Expectations in the Trust Game." *Political Research Quarterly* 59(2): 189-202.

- Feingold, A. (1992a). "Good-Looking People Are Not What We Think." *Psychological Bulletin* 111: 304–341.
- Feingold, A. (1992b). "Gender Differences in Mate Selection Preferences: A Test of the Parental Investment Model." *Psychological Bulletin* 112: 125–139.
- Fiske, S. T. and Taylor, S. E. (1991). *Social Cognition*. New York, NY: McGraw-Hill.
- French, M. (2002). "Physical Appearance and Earnings: Further Evidence." *Applied Economics* 34: 569–572.
- Frieze, I., Olson, J. and Russell, J. (1991). "Attractiveness and Income for Men and Women in Management." *Journal of Applied Social Psychology* 21(3): 1039–1057.
- Goldin, C. and Rouse, C. (2000). "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians." *American Economic Review* 90(4): 715–741.
- Hamermesh, D. S. (2006). "Changing Looks and Changing 'Discrimination': The Beauty of Economists." *Economics Letters* 93(3): 405–412.
- Hamermesh, D. S. and Biddle, J. E. (1994). "Beauty and the Labor Market." *American Economic Review* 84 (5): 1174–1194.
- Hamermesh, D. S., Meng, X. and Zhang, J. (2002). "Dress for Success – Does Primping Pay?" *Labour Economics* 9(3): 361–373.
- Hamermesh, D. S. and Schmidt, P. (2003). "The Determinants of Econometric Society Fellows Elections." *Econometrica* 71(1): 399–407.
- Harper, B. (2000). "Beauty, Stature and the Labour Market: A British Cohort Study." *Oxford Bulletin of Economics and Statistics* 62: 771–800.
- Jones, D. M. and Hill, K. (1993). "Criteria of Facial Attractiveness in Five Populations." *Human Nature* 4: 271–296.
- Kahneman, D., Slovic, P. and Tversky, A. (Eds.) (1982). *Judgement Under Uncertainty: Heuristics and Biases*. New York, NY: Cambridge University Press.
- Keating, C. F., Randall, D. and Kendrick, T. (1999). "Presidential Physiognomies: Altered Images, Altered Perceptions." *Political Psychology* 20(3): 593–610
- King, A. (Ed.) (2002). *Leaders' Personalities and the Outcomes of Democratic Elections*. Oxford: Oxford University Press.
- King, A. and Leigh, A. (2006). "Beautiful Politicians." Unpublished paper. Australian National University, Canberra, Australia.
- Klein, M. and Rosar, U. (2006). "The Frog Pond Beauty Contest: Physical Attractiveness and Electoral Success of the Constituency Candidates at the North Rhine-Westphalia State Election 2005." Unpublished paper. University of Köln, Köln, Germany.

- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallamm, M. and Smoot, M. (2000). "Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review." *Psychological Bulletin* 126(3): 390-423.
- Lau, R. R. and Redlawsk, D. P. (2001). "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(October): 951-971.
- Lau, R. R. and Sears, D. O. (Eds.) (1986). *Political Cognition: The 19th Annual Carnegie Symposium on Cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- Lee, D. S. (2001). "The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of the Elections to the U.S. House." NBER Working Paper No. 8441. National Bureau of Economic Research (NBER), Cambridge, MA.
- Lupia, A. (1994). "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88(1): 63-76.
- Macrae, C. N., Milne, A. B. and Bodenhausen, G. V. (1994). "Stereotypes as Energy-Saving Devices: A Peek Inside the Cognitive Toolbox." *Journal of Personality and Social Psychology* 66(1): 37-47.
- McDermott, M. L. (1998). "Race and Gender Cues in Low-Information Elections." *Political Research Quarterly* 51: 895-918.
- Mobius, M. M. and Rosenblat, T. S. (2006). "Why Beauty Matters." *American Economic Review* 96(1): 222-235.
- Mocan, N. and Tekin, E. (2006). "Ugly Criminals." IZA Discussion Paper No. 2048. Institute for the Study of Labor (IZA), Bonn, Germany.
- Mughan, A. (2000). *Media and the Presidentialization of Parliamentary Elections*. Basingstoke: Palgrave.
- Mulford, M., Orbell, J., Shatto, C. and Stockard, J. (1998). "Physical Attractiveness, Opportunity and Success in Everyday Exchange." *The American Journal of Sociology* 103: 1565-1593.
- Neumark, D., Bank R. J. and Nort, K. D. van (1996). "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics* 111(3): 915-941.
- Nisbett, R. E. and Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

- Perrett, D. I., Burt, M. D., Penton-Voak, I. S., Lee, K. J., Rowland, D. A. and Edwards, R. (1999). "Symmetry and Human Facial Attractiveness." *Evolution and Human Behavior* 20: 295-307.
- Pfann, G. A., Biddle, J. E., Hamermesh, D. S. and Bosman, C. M. (2000). "Business Success and Businesses' Beauty Capital." *Economics Letters* 67(2): 201-207.
- Piattelli-Palmarini, M. (1994). *Inevitable Illusions: How Mistakes of Reason Rule Our Minds*. New York, NY: John Wiley & Sons.
- Raunio, T. (2005). "Finland: One Hundred Years of Quietitude." In Gallagher, M. and Mitchell, P. (Eds.) *The Politics of Electoral Systems*. Oxford: Oxford University Press.
- Redlawsk, D. P. and Lau, R. R. (2003). "Do Voters Want Candidates They Like or Candidates They Agree With? Affect vs. Cognition in Voter Decision Making." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.
- Riggle, E. D., Ottati, V., Wyer, R. S., Kuklinski, J. and Schwartz, N. (1992). "Bases of Political Judgments: The Role of Stereotypic and Nonstereotypic Information." *Political Behavior* 14(March): 67-87.
- Rosenberg, S. W., Bohan, L., McCafferty, P. and Harris, K. (1986). "The Image and the Vote: The Effect of Candidate Presentation on Voter Preference." *American Journal of Political Science* 30(1): 108-127.
- Rosenberg, S. W. and McCafferty, P. (1987). "The Image and the Vote: Manipulating Voters' Preferences." *The Public Opinion Quarterly* 51(1): 31-47.
- Schubert, J. N. and Curran, M. A. (2003). "Political Sophistication and Appearance Stereotyping in Candidate Evaluation." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.
- Siegelman, L., Siegelman, C. K. and Fowler, C. (1987). "A Bird of a Different Feather? An Experimental Investigation of Physical Attractiveness and the Electability of Female Candidates." *Social Psychology Quarterly* 50(1): 32-43.
- Simon, H. A. (1985). "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review* 79(March): 293-304.
- Solnick, S. J. and Schweitzer, M. E. (1999). "The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions." *Organizational Behavior and Human Decision Processes* 79(3): 199-215.
- Statistics Finland (2006). "Parliamentary Elections, 2003." <http://www.tilastokeskus.fi/tup/suoluk/suoluk_vaalit_en.html>

- Stewart, M. and Clarke, H. (1992). "The (Un)importance of Party Leaders: Leader Images and Party Choice in the 1987 British Election." *Journal of Politics* 54: 447-470.
- Todorov, A., Mandisodza, A. N., Goren, A. and Hall, C. C. (2005). "Inferences of Competence from Faces Predict Election Outcomes." *Science* 308(5728): 1623-1626.
- Wattenberg, M. (1991). *The Rise of Candidate-Centred Politics: Presidential Elections of the 1980s*. Cambridge, MA: Harvard University Press.
- Willis, J. and Todorov, A. (2006). "First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face." *Psychological Science* 17(7): 592-598.