

THE LOOKS OF A WINNER: BEAUTY, GENDER, AND ELECTORAL SUCCESS

Niclas Berggren

The Ratio Institute

Panu Poutvaara

University of Helsinki and CEBR

Henrik Jordahl

Research Institute of Industrial Economics (IFN)

Abstract

We study the role of beauty in politics using candidate photos that figured prominently in electoral campaigns. Our investigation is based on visual assessments of 1,929 Finnish political candidates from 10,011 respondents (of which 3,708 were Finnish). As Finland has a proportional electoral system with a relatively large number of female candidates, we are able to perform a systematic study of gender differences, and to compare the electoral success of non-incumbent candidates representing the same party. An increase in beauty by one standard deviation is associated with an increase of 17–20% in the number of votes for the average non-incumbent candidate. The relationship is virtually always statistically significant for female candidates, and in most specifications also for male candidates. When interpreting our results, we also evaluate alternative explanations of why beauty matters.

Acknowledgments: The authors wish to thank Bryan Caplan, Mikael Elinder, Justina Fischer, Daniel Hamermesh, Daniel Klein, Claus Thustrup Kreiner, Markku Lanne, Mikael Priks and Roope Uusitalo; participants at the 2006 IIPF conference, the 2007 Annual Meeting of Finnish Economists, the 2007 World Meeting of the Public Choice Society, the 2007 CESifo Public Sector Area Conference and the 22nd congress of the European Economic Association; and participants at seminars at Bocconi University, Copenhagen Business School, ETH Zürich, George Mason University, HUI Stockholm, KU Leuven, Lund University, the Ratio Institute, Stockholm University, TU Dresden, Umeå University, University of Gävle, University of Helsinki and Uppsala University 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.

E-mail addresses: Berggren: niclas.berggren@ratio.se; Jordahl: henrik.jordahl@ifn.se; Poutvaara: panu.poutvaara@helsinki.fi

1. Introduction

Are good looks an advantage in politics? For several reasons they could be. If good-looking people are more persuasive, are treated better in social interaction and achieve higher occupational success – as evidenced in a meta-study by Langlois et al. (2000) – they might do better also in politics. In a recent contribution, Mobius and Rosenblat (2006) find that beautiful people receive higher wages in an experimental labor market, even though performance was not correlated with beauty. Part of the difference is explained by higher confidence and better communication skills by good-looking people, while part of the explanation relates to employers perceiving good-looking people as more skilled than they really are.

In the vocabulary of Ambady and Rosenthal (1992), good looks could function as a “thin slice” of information or as a heuristic in decision-making. Already Downs (1957) proposed that many voters are rather uninformed about the details of politics, a view that is further supported by Bartels (1996). A consequence of this could be that voters focus on personal characteristics of the candidates rather than on political programs, as Wattenberg (1991) argues is the case. Or for that matter, people might just prefer to look at beautiful people as suggested by the importance of looks in the entertainment industry. Against this background, we investigate if visual assessments of political candidates can explain election outcomes.

Our main result is that beauty helps. We find that an increase in beauty by one standard deviation is associated with a 17- to 20-percent increase in the number of votes for the average non-incumbent. Beauty is more strongly correlated with success than either perceived competence or trustworthiness. Our empirical analysis also suggests that beauty matters more for female candidates.¹

Our study is based on four web surveys with over 1,900 facial photos of Finnish political candidates. Altogether, we collected assessments from 10,011 respondents. About 2,800 non-Finnish and about 3,700 Finnish respondents were told that the persons in photos are political candidates. About 3,500 respondents from outside of Finland were not told anything

¹ 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.

about the persons in photos. In these three surveys with a large number of respondents, each respondent was shown a random selection of photos and was asked to assess the candidates' beauty, as well as perceived competence, trustworthiness, likability and intelligence. In the fourth survey with 16 respondents, each respondent assessed *all* 504 photos of candidates in the Helsinki municipal election. For each survey, we investigate to what extent the candidates' individual beauty scores (relative to the average beauty of competing candidates) are associated with their electoral success in the 2003 parliamentary or 2004 municipal elections.

Our main contributions can be summarized in three points. First, we are the first to study the effects of facial appearance on the success of political candidates who compete against other candidates from the same party. Second, we are able to focus on competition between a large number of non-incumbent candidates – about most of whom voters can be expected to have little or no information other than party, occupation, education and visual cues. Both of these advances are made possible by the proportional electoral system in Finland.² And while we focus on competition between non-incumbent candidates, we can also analyze incumbency effects. Finnish voters who are unsatisfied with an incumbent can vote for a challenger from the same party. Our third contribution consists of a systematic investigation of the role of gender. If beauty matters for electoral success, then an important question is if it confers differential advantages on male and female candidates. Finnish elections are unusually suited for gender analysis, since there is a sizable number of both male and female candidates in all districts. We also investigate whether men and women differ in their assessment of candidates' beauty and other traits.

A major benefit of focusing on within-party competition is that we avoid problems of reverse causality which may plague studies based on between-party competition in one-member districts. Political parties are more likely to attract more popular (e.g. more beautiful) candidates in districts in which they have an electoral advantage. This problem could confound the intriguing finding by Todorov et al. (2005), that quick photo assessments of competence help predict the outcomes of elections to the U.S. Congress. Unlike studies of between-

² According to Reynolds, Reilly, and Ellis (2005), there are proportional electoral systems with party lists in 68 countries including Finland.

party competition, we are able to construct our electoral-success variable in such a way – basically as the vote share on a list featuring competition against candidates from the same party – that the relationship between expected electoral outcomes of various parties and candidate selection is unlikely to influence the results.

In studying within- rather than between-party competition, we also automatically control for the effect of ideology on voter choice, as candidates of the same party in Finland are ideologically quite homogeneous, unlike candidates of different parties. In the Finnish election study from the 2003 parliamentary election, most voters said that political opinions and party were crucial for their choice of candidate. Even so, personal appearance and style was important for one third of the voters. As more than half of the voters considered a candidate's political experience important and as more than a third valued a candidate's education, most voters clearly also think that a candidate's personal characteristics and expected competence, and not just party, matter (see Bengtsson and Grönlund 2005).

Studying within-party competition in Finland offers interesting insights also for countries with one-member electoral districts, like the United States. Most obviously, party primaries are an important stage in American federal and state-level elections. Our study provides reliable estimates on the relative importance of several aspects of candidate appearance at this stage of the electoral process. As ideological considerations are more important in general elections, our results arguably give an upper bound for the effects of various aspects of candidate appearance in between-party competition. However, there is no reason to expect the relative importance of various aspects of personal appearance to differ between within-party and between-party competition. With or without ideological competition, voters prefer competent, trustworthy and likable politicians.³

We also think that some aspects of our research design form a contribution. By having respondents from Finland and from many other countries and by studying their assessments separately, we are able to say that the results hold irrespective of the nationality of the respon-

³ For example, Besley (2004), Caselli and Morelli (2004), Messner and Polborn (2004), and Poutvaara and Takalo (2007) study electoral competition between candidates who differ in their competence or honesty; two non-ideological traits that voters care about.

dents (who, in the case of Finns, may recognize the candidates). Our survey where 16 respondents assessed all photos of candidates for the Helsinki municipal election shows that surveys with a small number of respondents may produce unstable results when the respondents come from the same country as the persons whose appearance they assess. In addition, our survey where respondents were not told that the photos depict political candidates provides information about whether knowing this affects the assessment.

Importantly, we use photos that the political parties displayed on their campaign posters. Finnish municipalities are obliged to provide each political party with the same number of stands on which the parties can present their posters. All parties make extensive use of posters that display the names and photos of all candidates in the district. As all parties have a large number of such posters outdoors during the electoral campaign and as the same photos are also displayed in newspaper ads, it is likely that a large majority of voters have seen most or all of the candidate photos from the parties that they consider voting for. In order to guarantee a uniform style on the electoral posters, it is common practice for Finnish parties to use one photographer who takes photos of all candidates on a given list. Using official candidate photos from electoral campaigns, rather than photos from the press or photos supplied by candidates individually, reduces empirical problems of reverse causality, omitted variables and measurement error substantially.

A tentative interpretation of our findings is that beauty helps either because good-looking people are more successful in social interaction, or because voters like to watch good-looking politicians. The interpretation is based on three pieces of evidence. First, assessments of four other traits (among them competence and intelligence) allows us to demonstrate that beauty has an independent effect, rather than working as a signal through other trait assessments. Second, the candidates' education and occupation, as reported on electoral lists, also serve as signals of competence, and by including this information we demonstrate in another way that beauty has an effect that is independent of its signaling competence. Third, the fact that beauty appears to be as important for incumbents as for challengers also supports the interpretation that social interaction or the pleasure of looking at beautiful people explain why better-looking candidates are more successful.

In order to see how reliable our findings are, we conduct an extensive sensitivity analysis along several dimensions. The sensitivity analysis confirms the basic result that beauty is positively related to electoral success, and more so for female candidates.

2. The Literature

A large body of research has established that it is good to be beautiful. In a meta-analysis of 102 studies, Langlois et al. (2000) report that the looks of people influence how they are perceived and treated by others, even by those who know them.⁴ Furthermore, good looks could serve as a signal of better health. Jackson, Hunter, and Hodge (1995) and Kanazawa and Kovar (2004) argue that they could also be a signal of higher intelligence. As for gender, Langlois et al. (2000, p. 399) say:

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.

Economic research has demonstrated similar substantial benefits in the labor market. Beautiful people receive higher wages, a beauty premium. According to Hamermesh and Biddle's (1994) seminal study, workers of above-average beauty earn about 10 to 15% more than workers of below-average beauty. Other studies obtain similar results: see e.g. Biddle and Hamermesh (1998), Harper (2000), Pfann et al. (2000), Hamermesh, Meng, and Zhang (2002), French (2002), and Mocan and Tekin (2006). As for gender, Hamermesh and Biddle (1994, p. 1187) conclude that there is an "absence of significantly larger penalties and premia, especially the latter, for women than for men."

In a recent contribution, Mobius and Rosenblat (2006) find a beauty premium in an experimental labor market involving a maze-solving task, despite the fact that performance was

⁴ Cf. Eagly et al. (1991) and Feingold (1992a).

uncorrelated with beauty. Mobius and Rosenblat decompose the beauty premium into three parts. First, good-looking workers are more confident and higher confidence increases wages. Second, for a given level of confidence, good-looking workers are wrongly considered more able by employers. Third, good-looking workers benefit from having better communication and social skills. As communication, social skills and confidence are all important in politics, one could expect a substantial beauty premium in elections.

The role of beauty in politics has attracted academic interest only recently. 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 male candidates; but that there is virtually no effect for female candidates. Rosar, Klein, and Beckers (forthcoming) study the role of attractiveness evaluations in a German state election in one-member districts. Their results indicate that an increase in beauty can increase a candidate's vote share by at most 4 percentage points. King and Leigh (2007) study beauty in Australian elections and report that it matters: a one standard deviation increase in beauty raises a candidate's vote share by 0.7–1.8 percentage points. (Note that these numbers cannot readily be compared with our findings since the settings of electoral competition are different and since we define electoral success relative to list size.) While Rosar, Klein, and Beckers (forthcoming) and King and Leigh (2007) study competition between candidates representing different parties in one-member districts, we focus on within-party competition in a proportional electoral system with several candidates being elected from each district. Furthermore, as we ask our respondents to assess the competence, intelligence, likability and trustworthiness of the candidates we are able to test whether beauty has an independent role of its own, in addition to its potential role as a signal for other traits.

Our work is also related to voting research on the role of heuristics, information shortcuts, stereotyping, and thin slices of information. Downs (1957) stresses the uncertainty of voter decision-making and regards parties and ideologies as devices used to attract voters who are not all that familiar with detailed policies. Lau and Redlawsk (2001) find that voters low in political sophistication use candidate appearance as a heuristic.⁵ Among more recent stud-

⁵ Cf. Budesheim and DePaola (1994, p. 339) and Redlawsk and Lau (2003).

ies, Todorov et al. (2005) find that inferences of competence from photos help predict the outcomes of elections to the U.S. Congress (71.6% of Senate races and 66.8% of House races). Benjamin and Shapiro (2006) report that about 20% of the variation of the actual vote shares in U.S. gubernatorial elections can be explained by predictions based on video clips. While these authors analyze photos or video clips as the only thin slice of information, we also study occupation and education, as reported on electoral lists. The meta-study by Ambady and Rosenthal (1992) further confirms that people – whether right or wrong – often form assessments and act on the basis of thin slices of information.⁶

3. Institutional Facts, Surveys, and Data

3.1. Institutional Facts

The political setting for this study is Finland, and its electoral system is proportional.⁷ 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 electing one. Elections are held every four years. The number of MPs elected from the 14 mainland districts varies between seven and 32.

In each parliamentary district, parties present lists of their candidates, typically in alphabetical order but sometimes with incumbents listed first, and each voter chooses one candidate on one list. The number of candidates that a party can present equals the number of representatives elected from the district, if this is 14 or more. In small districts with less than 14 seats, parties can present 14 candidates. The legislature seats of a given district are allocated based on party vote shares to the candidates in accordance with their “competitive in-

⁶ This conclusion is supported further by Kahneman, Slovic, and Tversky (1982), Simon (1985), Lupia (1994), Macrae, Milne, and Bodenhausen (1994), Bartels (1996), Caprara, Barbaranelli, and Zimbardo (1997), and Willis and Todorov (2006).

⁷ See Raunio (2005).

dices”. In each party, the candidate 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. The number of elected municipal councilors depends on municipality size, reaching a maximum of 85 in Helsinki. In the municipal elections each party is allowed to present one and a half as many candidates on its list as the number of seats in the municipal council. The maximum number of candidates that each party can present in Helsinki is thus 127.

In the 2003 parliamentary election, turnout was 69.7%. Female candidates received 42.6% of all votes, and 75 of the 200 elected members of parliament were women (Statistics Finland, 2006).

3.2. *The Surveys*

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 agreement about who is and who is not attractive, both within and across cultures. As Hamermesh and Biddle (1994, p. 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.”⁸ On this basis, we have conducted four web surveys based on the same questionnaire, but with some modifications in each treatment. In addition to asking about beauty, we also included questions about four other traits in order to find out more precisely

⁸ The same point is made by e.g. Feingold (1992b), Cunningham et al. (1995), and Aharon et al. (2001).

what determines electoral success and how the results are to be interpreted.⁹ By collecting responses from several countries we are also able to check for cross-cultural differences. We find, in our main survey with non-Finnish respondents, that respondents in different countries make very similar assessments of the same photos (with the French possibly finding candidates a little less beautiful than Americans, Swedes, Germans, Danes and others).

The four surveys are described briefly in Table 1.

TABLE 1. The four surveys.

| Name of survey | Nationality of respondents | Information to respondents that the photos depict political candidates | Selection of photos shown to respondents | Number of respondents | Number of responses | Time when carried out |
|-------------------------------------|----------------------------|--|--|-----------------------|---------------------|-------------------------|
| Survey 1: The main survey | Non-Finnish | Yes | Random (four per round) | 2,772 | 16,218 | Spring-summer 2006 |
| Survey 2: The survey of Finns | Finnish | Yes | Random (four per round) | 3,698 | 26,477 | Fall 2006 |
| Survey 3: The small survey | Swedish and Finnish | Yes | All (504 per round) | 16 | 8,064 | Winter 2007 |
| Survey 4: The no-information survey | Non-Finnish | No | Random (ten per round) | 3,525 | 38,985 | Autumn-winter 2005/2006 |

Notes: In the columns with the number of respondents and responses, only respondents who assessed at least four photos (and their responses) are reported.

Our main survey, survey 1, was conducted in the spring and summer of 2006 outside of Finland. The main reason for using non-Finnish respondents is that they can be expected not to recognize any of the candidates, which is an advantage when analyzing whether visual images function as thin slices of information. With the help of dozens of colleagues, students in various universities were invited to participate, either in lectures or by e-mail. The biggest

⁹ We do not claim that the assessments represent true characteristics of the political candidates. This study is about perceptions and none of the relationships reported should be interpreted as claims of a relationship among any underlying true characteristics.

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 Public Finance and European Public Choice Society). We also cooperated with several blogs 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.

After replying to some personal background questions, each respondent was shown four photos, one at a time, randomly chosen from the database of photos, in total two of each gender. In connection with each photo, several questions were asked (see Box A1 in the Appendix for further details). There was an option, after having assessed four photos, to assess additional rounds of four photos, this time with a choice as to whether to assess only females, only males or a continued mixture. There was no time limit for looking at the photos.¹⁰ The size of the photos was approximately 5 x 3.5 centimeters (2 x 1.4 inches), and they depicted faces only. No other information than the photo was given about any candidate. The candidates come from four parties with 63% of the elected members of parliament in the 2003 election: the Social Democratic Party, the National Coalition Party (a center-right party), the Left Alliance and the Green League.

Finnish political parties advertise their candidates on posters with individual photos of all candidates in a district. Since the participating political parties provided us with these photos, our respondents assessed the same photos as the voters were exposed to. There are two potential problems related to the use of candidate photos – in this and in other studies on the role of candidate appearance in politics. The first one is reverse causality: successful politicians could have access to stylists and better photographers. The second one is omitted va-

¹⁰ 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 observation do not yield greater predictive accuracy, something which seems to hold, not least, with regard to faces (cf. Todorov et al. 2005, pp. 1623–1624, and Willis and Todorov 2006).

riables, if some politicians both “dress for success” and do other unobserved things, like visit large numbers of voters, which help them getting elected. However, we expect both problems to be smaller when using official candidate photos. Our investigation does not suffer from the problem that more successful or better financed candidates hire better photographers: official candidate photos taken by the same photographer offers a more equal playing ground. Moreover, a “bad hair day” would produce measurement error for a candidate if photos from the press were used, whereas with official candidate photos, one expects an unflattering picture exposed in numerous posters to be detrimental for electoral success. In any case, Hamermesh, Meng, and Zhang (2002) find that attempts to improve one’s looks, in the realm of clothing and cosmetics, only have a small impact on how beautiful one is perceived by others.

Survey 2, the survey of Finns, was carried out in the fall of 2006 in Finland. This time, we attracted mainly student participants. This survey allows us to investigate how recognition of candidates affects assessments and to verify that assessments by Finnish respondents are broadly in line with patterns of non-Finnish respondents. The biggest participant numbers, more than 300 from each, came from the University of Jyväskylä, the University of Helsinki, and the University of Oulu. Respondents could participate in a lottery of 30 movie tickets.

Survey 3, the small survey, took place in early 2007 in Finland and Sweden with 16 respondents of varying age and gender. This time, each respondent assessed all 504 photos of candidates in the Helsinki municipal election. The main reason was to see whether this way of assessing candidates – used in labor market studies – yields similar results as our large-scale surveys where each one of a large number of respondents assesses a small number of randomly selected photos.

Survey 4, the no-information survey, was conducted in the autumn-winter of 2005/2006. Respondents from outside of Finland were shown photos without any information on the persons appearing. This allows us to test whether assessments of beauty and other traits were affected by us telling that the persons in photos are political candidates.

We focus our investigation on the main survey with non-Finnish respondents who knew that they were assessing political candidates, and discuss results from the three other surveys in Section 7.

3.3. Data

Our database contains 1,929 photos of Finnish political candidates – 1,009 of men and 920 of women, from the municipal (57%) and parliamentary level (43%). We only include assessments by respondents who assessed at least four photos. We only include photos with at least three assessments. This gives us 1,786 photos. In Section 5, we divide the photos into two groups – those of non-incumbents (1,555 photos) and those of incumbents (231 photos). By “incumbents” is meant political candidates who served in the office in question, or as members of the national or the European parliaments at the time of the election. On average, each photo was assessed by nine respondents in the main survey.

As indicated in Table 2, Americans and Swedes make up a majority of our 2,772 respondents. Large groups of respondents also come from France, Germany and Denmark.

TABLE 2. Respondents by country.

| Country | Number | Percent |
|---------------|--------|---------|
| USA | 859 | 31.0 |
| Sweden | 850 | 30.7 |
| France | 261 | 9.4 |
| Germany | 220 | 7.9 |
| Denmark | 156 | 5.6 |
| Other country | 426 | 15.4 |
| Total | 2,772 | 100 |

Notes: Respondents denote those who assessed at least four photos (one full round). 66% were men, 34% women. 32% were undergraduate students, and 14% were graduate students. Average age: 31 (32 for men and 30 for women).

Through our four web surveys, we use more respondents than other studies of beauty or competence: 6,303 from outside of Finland and 3,708 from Finland, compared to four (Hamermesh 2006), five (King and Leigh 2007), 50 (Mobius and Rosenblat 2006), 264 (Benjamin and Shapiro 2006), 843 (Todorov et al. 2005), and 903 (Rosar, Klein, and Beckers forthcoming).¹¹

¹¹ Todorov et al. (2005) collected assessments of beauty from only 34 respondents.

4. Perceptions of Beauty and Other Traits

Each photo was assessed in the five dimensions beauty, competence, trustworthiness, likability, and intelligence using five reply options, which we have converted to a five-number scale.¹² The lowest possible beauty rating corresponds to 1, and the highest possible to 5, etc. In assessing each trait, respondents had an option to abstain. In our main survey, the share of those who abstained varied between 0.5% for beauty and 7.9% for trustworthiness. There is substantial agreement among respondents; if we concentrate on two groups of beauty assessments — 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% level.

However, men and women did not always agree on their assessments (Table 3). There is a clear tendency for men, on average, to give photos of female candidates less positive assessments than women do. There are smaller differences in the assessments of photos of male candidates; the only statistically significant difference is that men find male candidates more handsome or beautiful compared to what women find.

¹² Using a cardinal scale of this kind is standard fare in the literature: see e.g. Hamermesh and Biddle (1994). As reported more fully in section 6.1, we have also used alternative variables based in ordinal assessments: the share of responses where a candidate was evaluated as the most beautiful, most competent and most trustworthy among four photos.

TABLE 3. Assessments of five traits.

| Variable | Men assessing male candidates | Women assessing male candidates | Men assessing female candidates | Women assessing female candidates |
|-------------------------|-------------------------------|---------------------------------|---------------------------------|-----------------------------------|
| Average beauty | 2.64 (0.90) | 2.57 (0.91) | 2.79 (1.06) | 3.01 (0.97) |
| Average competence | 3.30 (0.88) | 3.27 (0.88) | 3.21 (0.84) | 3.39 (0.85) |
| Average trustworthiness | 3.04 (0.86) | 3.02 (0.89) | 3.29 (0.82) | 3.42 (0.83) |
| Average likability | 3.07 (0.92) | 3.06 (0.95) | 3.23 (0.93) | 3.37 (0.94) |
| Average intelligence | 3.38 (0.83) | 3.35 (0.82) | 3.23 (0.79) | 3.37 (0.79) |

Notes: Standard deviations in parentheses. The figures are from our main survey.

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 female candidates more positive assessments of all traits, even though the difference in the assessment of intelligence is small and not statistically significant. There is, lastly, no indication of a “dumb blonde syndrome,” which King and Leigh (2007) suggest as an interpretation of their results. There is a strong positive relationship, both for female and for male candidates, between beauty and perceived competence and between beauty and perceived intelligence. This holds irrespective of the gender of the respondents or the age of the candidates. A general pattern is that assessments of any pair of traits are positively correlated with each other, but correlations are far from perfect.¹³

¹³ For correlation coefficients, see Table A1 in the Appendix.

5. Beauty and Electoral Success

5.1. *The Empirical Setting*

In this section we investigate the relationship between beauty and electoral success. Given that assessments by Finnish voters could be influenced by their knowledge of the candidates, there is a risk that using Finnish respondents would create systematic measurement error. To avoid this, the results in this and the following section are based on assessments by non-Finnish respondents in our main survey.¹⁴ We present results for other respondent groups, including Finns, in Section 7.

Like Hamermesh (2006), we first look at the share of the elected candidates who receive above-average assessments. In the case of beauty, about 62% of the elected non-incumbent candidates were assessed as being above average on their list. This indicates that although beauty may be an asset in politics, it is by no means a necessary requirement for being elected. However, again we find that there is a clear gender gap: whereas only 43% of the elected male candidates had a beauty rating above average, the corresponding number for female candidates is 74%. Compared to other non-incumbent candidates of their own gender, 57% of elected male candidates and 70% of elected female candidates were thought to be of above average beauty on their list. This gender gap suggests that it may be fruitful to analyze the effects of beauty for each gender separately.¹⁵

A more detailed picture emerges if we look at average assessments and also take the gender of the respondents into account. Both men and women assess elected and non-elected male candidates similarly. One difference is that perceived competence is a bit higher among elected compared to non-elected male candidates. For beauty, the assessments of elected and

¹⁴ None of the non-Finnish 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.”

¹⁵ We have done this throughout the paper but in general only report statistically significant gender differences.

non-elected male candidates are very close to each other.¹⁶ For female candidates the picture is quite different. Both men and women find elected female candidates more good looking than non-elected ones. 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 assessments than they give to non-elected female candidates.¹⁷

Next we investigate to what extent beauty and other traits can be related to the relative success of candidates in the 2003 and 2004 elections. Unlike other studies we focus first on the large group of non-incumbent candidates (defined as political candidates who were not elected to the office in question and who were not members of the national or European parliaments at the time of the election) and then look at the full set of candidates, including incumbents. One reason for making this division is that incumbency is a very strong predictor of electoral success (see e.g. Lee, forthcoming), and if a dummy variable fails to capture all of its effects, other estimates risk being biased. Another reason is that appearance and other thin slices of information may be more important for less well-known candidates.¹⁸ Furthermore, Andreoni and Petrie (forthcoming) report that individual contributions in a public goods game are higher in the presence of beautiful players as long as the individual contributions are unknown, but that this beauty premium turns into a beauty penalty once the individual contributions are revealed. This suggests that beauty could be more important for non-incumbents.

The trait variables are constructed in two steps. First we compute the mean of all assessments of a particular photo. From this measure we then subtract the mean assessment for each trait for the candidates on the same list. That is, we use *relative* measures of the different traits, capturing how beautiful, competent and trustworthy a candidate is perceived to be in relation to his or her competitors on the list.

¹⁶ See Figure A1 in the Appendix. However, incumbent candidates are seen as slightly better-looking than non-incumbent candidates (an average of 2.82 vs. an average of 2.73).

¹⁷ See Figure A2 in the Appendix.

¹⁸ We are able to study non-incumbents separately as Finland has a proportional electoral system with personal votes determining the order in which candidates are elected, resulting in within-party competition. A plurality-vote system, like that of the United States, typically features competition between an incumbent and a challenger from different parties. Benjamin and Shapiro (2006), Rosar, Klein, and Beckers (forthcoming), and King and Leigh (2007) use a dummy for incumbency.

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

$$\text{relative success}_{i,j} = (p_i / v_j) * 100 \quad (1)$$

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 .¹⁹ When studying non-incumbents in section 5.2 we calculate both the trait measures and relative success based on non-incumbent candidates only. In section 5.3 the same measures are calculated for incumbent and non-incumbent candidates together. Each candidate's vote share would be a simpler and more direct choice of dependent variable. We use that measure in the sensitivity analysis, but the advantage of the relative success measure is that it makes election outcomes comparable, as list sizes differ (especially between parliamentary and municipal elections).

As regressors, we use the three trait variables beauty, competence and trustworthiness. These three were selected to keep the analysis simple by focusing on dissimilar traits.²⁰ In our preferred specification we also include the age dummies young, which denotes an age under 30, and old, which denotes an age over 60. Our data show that both men and women find younger candidates more beautiful than older candidates.

5.2. Non-Incumbent Candidates

We begin by looking at the effects in the parliamentary election for female and male non-incumbent candidates. Most notably, as reported in Table 4, we find that beauty is clearly our most important explanatory variable of relative success both for female and for male candidates, and the only regressor that consistently attains statistical significance.

¹⁹ The mean of relative success is 100, capturing that on average each candidate must receive a share of the votes equal to 1 / list size. The average of relative success for elected candidates (incumbents and non-incumbents) is 338. That is, they receive 3.38 times the votes of the average candidate.

²⁰ Beauty and likability showed a high correlation and intelligence and competence showed a high correlation. In section 6.2 we describe results from a specification that includes all five traits.

TABLE 4. Relative success in the parliamentary election, non-incumbents.

| | (1) | (2) | (3) | (4) | (5) |
|----------------------|--|--|--|---|---|
| | Relative success all non-incumbents | Relative success all non-incumbents | Relative success all non-incumbents | Relative success female non- incumbents | Relative success male non-incumbents |
| Beauty | 34.89*** (6.31) | | 31.17*** (6.55) | 33.43*** (8.58) | 29.85*** (11.25) |
| Competence | | 23.08*** (8.34) | 10.95 (8.61) | 5.441 (15.6) | 11.70 (9.88) |
| Trustworthiness | | 9.94 (9.30) | 6.07 (8.89) | 15.27 (14.2) | -1.61 (12.3) |
| Male candidate | 3.77 (6.37) | -0.05 (6.77) | 4.72 (6.74) | | |
| Young (age<30) | -18.93** (9.45) | -3.93 (9.54) | -16.23* (9.70) | -18.47 (12.4) | -17.15 (14.9) |
| Old (age>60) | 11.59 (22.5) | 0.74 (21.8) | 8.19 (22.3) | -28.21 (20.3) | 48.26 (38.5) |
| Number of candidates | 641 | 641 | 641 | 343 | 298 |
| Adjusted R-squared | 0.06 | 0.02 | 0.06 | 0.09 | 0.04 |

Notes: Robust standard errors in parentheses. The regressions include a constant term. * significant at 10%; ** significant at 5%; *** significant at 1%.

In column 1, beauty is the only of the three traits that is included, and it is found to be highly statistically significant. The coefficient of beauty becomes marginally smaller when competence and trustworthiness are included as well (in columns 3–5). When we exclude beauty in column 2, the size of the estimated coefficient for perceived competence is substantially higher than in columns 3–5 and also attains statistical significance. This suggests that as perceptions of beauty and competence are positively correlated, the claim in Todorov et al. (2005) that voting preferences are anchored on inferences of competence from facial appearance may need to be reconsidered.

The three last columns include all three traits. A higher beauty score of one standard deviation implies an increase in the number of personal votes, relative to the average number of votes for the non-incumbents on the list, by 20.3% for all candidates, 24.1% for female candi-

dates, and 16.4% for male candidates.²¹ The gender difference is however not statistically significant (which generally holds true for regression results based on this main dataset). To facilitate the interpretation of the estimated impact of beauty, note that an increase of one unit in relative success means a one-percentage point increase in the number of votes, relative to the average number of votes of all candidates on the same list. Accordingly, an increase in the beauty assessment by one standard deviation is associated with a 20-percent increase in the number of votes for the average non-incumbent. One can also note that being young may be a disadvantage.

Table 5 reveals that the point estimate of beauty is only marginally smaller for the municipal elections. A higher beauty score of one standard deviation implies an increase in the number of personal votes, relative to the average number of votes for the non-incumbents on the list, by 16.6% for all candidates, 21.4% for female candidates and 19.4% for male candidates. Except among male candidates, the estimates for competence are statistically significant and larger than in the parliamentary election.

²¹ The standard deviation is 0.65 for all candidates, 0.72 for female candidates and 0.55 for male candidates.

TABLE 5. Relative success in the municipal elections, non-incumbents.

| | (1) | (2) | (3) |
|----------------------|--|---|---|
| | Relative success all non-incumbents | Relative success female non-incumbents | Relative success male non-incumbents |
| Beauty | 25.58*** (6.74) | 27.16** (11.30) | 19.44*** (6.03) |
| Competence | 18.54** (8.15) | 33.27** (15.7) | 7.278 (7.99) |
| Trustworthiness | -15.60* (8.17) | -14.20 (12.4) | -15.01 (10.8) |
| Male candidate | -27.82*** (6.53) | | |
| young (age<30) | -22.82*** (7.86) | -26.58* (13.5) | -17.01** (7.88) |
| old (age>60) | -3.50 (12.8) | -20.76 (14.3) | 11.69 (18.9) |
| Number of candidates | 914 | 460 | 454 |
| Adjusted R-squared | 0.05 | 0.04 | 0.02 |

Notes: Robust standard errors in parentheses. The regressions include a constant term. * significant at 10%; ** significant at 5%; *** significant at 1%.

5.3. All Candidates (Incumbents and Non-Incumbents)

The previous literature has focused on plurality-vote systems and has not studied competition between non-incumbents. We now investigate what the effect would be, as shown in Table 6, of adding incumbents and an incumbency dummy.²²

²² To economize, in the tables reporting regression results from here on, we generally only report results corresponding to column 3 in Table 4, i.e. for female and male candidates together in a specification that includes a dummy for male candidates and age dummies. The reason for this choice is that when comparing the estimated beauty coefficients for female and male candidates, the difference is not statistically significant in regressions based on data from our main survey.

TABLE 6. Relative success in the parliamentary and municipal elections, incumbents and non-incumbents.

| | (1) | (2) |
|----------------------|--|---|
| | Relative success parliamentary election | Relative success municipal elections |
| Beauty | 19.13*** (5.82) | 17.36** (7.74) |
| Competence | 11.57 (8.09) | 5.49 (10.48) |
| Trustworthiness | 6.41 (6.59) | -0.25 (12.16) |
| Incumbent | 190.86*** (19.35) | 352.91*** (35.40) |
| Male candidate | -2.57 (6.79) | -18.33** (9.12) |
| Young (age<30) | -19.27** (7.61) | -5.49 (10.08) |
| Old (age>60) | -14.72 (18.16) | -9.51 (17.63) |
| Number of candidates | 743 | 1,043 |
| Adjusted R-squared | 0.36 | 0.39 |

Notes: Robust standard errors in parentheses. The regressions include a constant term. * significant at 10%; ** significant at 5%; *** significant at 1%.

For the parliamentary election, reported in column 1, the beauty coefficient is smaller than its counterpart in the regression with non-incumbents only and implies that a one standard deviation increase in beauty is associated with an increase of relative success of 12.4 units. Perceived competence does not attain statistical significance. For the municipal elections, reported in column 2, beauty has a coefficient of almost the same size as in the regression with non-incumbents only, and the statistical significance of perceived competence that appeared in that regression vanishes.²³

Finally, we have carried out some hypothetical and purely mechanical calculations in order to roughly see how many non-elected candidates that could have been elected if they had had better looks. On each list, this was done by an imaginary reduction of the beauty as-

²³ Results without age dummies for Tables 4–6 are very similar and are available upon request.

assessment of all elected candidates by one standard deviation combined with an equally large imaginary increase in the beauty assessment of the same number of non-elected candidates. Using the estimated beauty coefficients in Table 6, this hypothetical procedure adds to the relative success of non-elected candidates at the expense of the elected ones. This crude experiment shows that 15% of the candidates elected in the parliamentary election would be replaced by competitors who were made more beautiful through this procedure. The corresponding figure in the municipal elections is 11%.

5.4. Occupation and Education as Alternative Thin Slices of Information

Finnish candidates are allowed to report their education and occupation on the official party lists that are placed in voting booths. Almost all candidates – 98% in our sample – report at least one of these pieces of information on their party list. This information on the candidates' education and occupation is also listed in most electoral ads. Therefore, voters have access to at least two other thin slices of information, in addition to photos.²⁴

Regression results indicate that the beauty coefficient is virtually unaffected, both in terms of size and statistical significance, when we include our battery of occupational and educational dummy variables. Listing oneself as a worker, artist or student is associated with lower electoral success when both occupational and educational dummies are included. Likewise, reporting upper-secondary education or comprehensive school or less is negatively related to electoral success. Details are reported in Table A2 in the Appendix; column 3 in Table 4 can be consulted for comparison.

To summarize our findings, beauty emerges as an asset in politics.

²⁴ In fact, 37% of the voters in the Finnish election study stated that a candidate's education had a considerable impact on their choice (Bengtsson and Grönlund 2005, p. 245).

5.5. Interpretation

Having established a link between beauty assessments and electoral success, we next bring together our results to evaluate alternative explanations of why beauty matters. Recall that Mobius and Rosenblat (2006) found that beauty can influence payoffs even if it is uncorrelated with productivity and that Todorov et al. (2005) found that competence was the most important predictor of electoral success. With our data we are able to discriminate between, on the one hand, the explanation that voters favor good-looking candidates because they enjoy watching them or because good-looking politicians are more successful in social interaction, and, on the other hand, the explanation that beauty matters as a signal related to competence or similar traits.

We have already reported in columns 1 to 3 of Table 4 that beauty assessments are a more robust explanatory variable of electoral success than competence or trustworthiness assessments. Beauty assessments are likewise a more robust explanatory variable than intelligence and likability assessments (as reported in Section 6.2). This is our first piece of evidence suggesting that beauty plays a role of its own, rather than just serving as a signal of competence. Also our finding that beauty retains its statistical significance when occupation and education are included in the regressions suggests that beauty does not only matter as a signal of competence.

Our inclusion of several trait evaluations and of information about occupation and education thus suggests that voters favor good-looking candidates because they enjoy watching good-looking politicians, or because good-looking politicians are more successful in social interaction. Another possibility – which we cannot rule out with our data – is that beauty serves as a signal for some other characteristics than competence, intelligence, likability or trustworthiness.

As our third piece of evidence we conjecture that if beauty helps because voters like to watch good-looking politicians or because good-looking people are more successful in social interaction, then it should matter similarly for incumbents and challengers. If, on the other hand, beauty serves as a signal (of competence, health or other traits) then one would expect

its effect to be larger for challengers, as voters do not have as much information about them as they have about incumbents. We find that an interaction term between beauty and incumbency is statistically insignificant, both in the parliamentary and in the municipal elections. Similarly, if we include interaction terms between incumbency and the three variables beauty, competence and trustworthiness, they are all statistically insignificant. Since beauty appears to be as important for incumbents as for challengers, also our third test supports the interpretation that social skills or the pleasure of looking at beautiful people explain why better looking candidates are more popular.

6. Sensitivity Analysis

We will now investigate to what extent the results reported so far are sensitive to various alternative ways of examining the relationship between beauty and electoral success. We report the results briefly, but in each case, all details are available upon request. Our finding that beauty is strongly associated with electoral success is maintained in each alternative specification.

6.1. Measures Based on Ordinal Assessments

Beauty, competence and trustworthiness have so far been measured cardinally. We have also used alternative measures based on ordinal assessments. Like our previous trait variables, these variables, *beautyshare*, *competenceshare* and *trustshare*, are constructed in two steps. First we compute the share of assessments where a candidate was found to be the most beautiful, most competent and most trustworthy, when presented with three other randomly chosen candidates. From this measure we then subtract its mean over the non-incumbents on the same list. The results reveal that the previous qualitative results of Tables 4 and 5 hold, as beauty dominates and retains statistical significance. An increase in the beauty share by one standard deviation is associated with an increase in the number of votes by 39% for the average non-

incumbent candidate. These results indicate that the positive relation between beauty and electoral success is not just a consequence of the question used or the way we construct the explanatory variables.

6.2. *Sensitivity in Other Dimensions*

We have also made a number of minor changes in our main empirical specifications (in Tables 4 and 5). Here we report results from specifications where we exchange the dependent variable, redefine incumbency, include perceptions of additional traits, check for outliers, separate students and non-students, and use perceived instead of real age.

We begin by replacing 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 ,

$$vote\ share_{i,j} = (p_i / w_j) * 100 \quad (2)$$

where p_i is non-incumbent candidate i 's number of personal votes and w_j is the number of all votes for non-incumbent candidates on list j . The relationship between this measure and relative success is that $w_j = v_j * \text{the number of non-incumbent candidates on list } j$. 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.²⁵ It turns out that the results are qualitatively very similar to those of Table 4. Again and most notably, we find that beauty is by far our most important explanatory variable. Competence does not attain statistical significance. A higher beauty score of one standard deviation implies an increase of 1.61 percentage points in the vote share in the parliamentary election. In the municipal election, with more candidates on the lists, the corresponding figure is 0.15. Although these numbers may appear small, note that the average vote share among all non-

²⁵ See Berggren, Jordahl, and Poutvaara (2006) for a version of this study where vote share was used as the main dependent variable.

incumbents is 4.47% in the parliamentary election and 0.57% in the municipal elections. The corresponding averages for elected non-incumbents are 11.75% and 2.21%.

The second change we make to test the sensitivity of our results is to redefine incumbency. Above, incumbents were defined as political candidates who served in the office in question or as members of the national or European parliaments at the time of the election. A more common definition of incumbency is to include only the candidates who served in the office *in question* (hence, regarding candidates who had been elected to some other office as non-incumbents). Using this definition, Tables 4–6 have been reproduced and no big differences appear, neither for the parliamentary nor for the municipal elections.

In the empirical models reported so far, we have included three of the five traits that were assessed by our respondents: beauty, competence and trustworthiness. We excluded likability and intelligence in order to simplify the analysis and keep the focus on three dissimilar traits (e.g., intelligence can be expected to be conceptually quite similar to competence). We have conducted the analysis with all five traits included, and it shows that the exclusion is an innocuous one. Beauty retains its statistical significance and remains about as important in terms of coefficient size compared to Table 4 (the coefficient is 27.3 for the parliamentary election for all non-incumbent candidates, compared to 31.12 in Table 4); whereas the likability and intelligence coefficients do not attain statistical significance.²⁶

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 444 non-incumbent candidates in the Helsinki municipal elections. The Helsinki municipal elections are best suited for this, since all four parties have a large number of candidates and about the same number of non-incumbents on their lists. The rank correlation between beauty and relative success is especially strong for female candidates, for whom Spearman's rho is 0.285. For male candidates, the correlation is 0.103 but not statistically significant. Combining female and male candidates we get a statistically significant correlation of 0.232.

²⁶ Including five traits instead of three does not result in multicollinearity problems according to variance inflation factors.

The analysis of Spearman rank correlations also allows us to compare the relationship between electoral success and the assessments 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 level of statistical significance than the rank correlation between electoral success and perceived competence, trustworthiness, likability, or intelligence.

Unlike several other studies, we have substantial numbers of both students and non-students among our respondents. It turns out that the assessments by (undergraduate and graduate) students and other respondents are remarkably similar, with the only statistically significant differences being that students assess 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). In terms of regression results, looking at relative success, non-incumbent candidates and confining analysis to photos with at least three student assessments, beauty attains statistical significance and the size of the coefficient is 23.3 in the parliamentary and 21.0 in the municipal elections. Beauty remains important, even when just using this group of respondents.

Finally, as we asked respondents to estimate the age of each candidate, we have also exchanged the real age used in the regressions above with the age perceived by respondents. It turns out that the estimated coefficient of beauty is almost identical when perceived age is used.

7. Three Additional Surveys

In addition to the sensitivity analysis in the preceding section, with an investigation of the results derived from our main survey, we have also carried out three additional surveys, as was reported in Section 3.2. We have done this in order to study the effects of using respondents from Finland (who may recognize candidates), in order to compare our approach of having many respondents each of whom assesses a small number of photos with that of most pre-

vious studies (which use few respondents each of whom assesses a large number of photos) and, lastly, in order to see whether knowing that the photos depict political candidates affects the assessments.

7.1. Finnish Respondents

We have undertaken a survey based on the same set of political candidates with only Finnish respondents (survey 2). The results indicate only small differences compared to our main survey with non-Finnish respondents.

As we asked the Finnish respondents to indicate if they recognized candidates, we are able to study how results differ in the degree of recognition. In Table 7, we report estimated beauty and competence coefficients stemming from regressions using the same set of variables as in Table 4, column 3 – i.e. beauty, competence, trustworthiness, male candidate, young and old. As before, we restrict ourselves to non-incumbents.

Column 1 contains regression results for all candidates. Column 2 contains results from regressions where we exclude individual assessments of candidates that the respondents recognized (by giving a first name, a family name or both). In column 3 we exclude photos of candidates recognized by at least one respondent. Lastly, column 4 contains results from non-Finnish respondents based on the same sample of candidates.²⁷ Hence, as one moves to the right from column 1 to 4, the probability of candidate recognition is gradually diminished.

²⁷ Since coefficients do not change much when we adjust the sample of candidates (compare column 4 with column 3 in Table 4 and column 1 in Table 5), the differences that we do observe seem not to be driven by sample composition, but instead by recognition.

TABLE 7. Relative success, non-incumbents.

| | (1) | (2) | (3) | (4) |
|--|--|---|--|-------------------------|
| | Finnish respondents, including recognized candidates | Finnish respondents, individual assessments of recognized candidates are excluded | Finnish respondents, photos of candidates recognized by at least one respondent are excluded | Non-Finnish respondents |
| Beauty, parliamentary election | 30.37*** | 32.54*** | 21.48** | 29.10*** |
| Beauty, municipal elections | 27.05*** | 32.50*** | 31.15*** | 24.25*** |
| Competence, parliamentary election | 39.62*** | 28.33** | 52.49*** | 11.74 |
| Competence, municipal elections | 31.90** | 6.03 | 9.50 | 13.29* |
| Number of candidates, parliamentary election | 704 | 704 | 559 | 559 |
| Number of candidates, municipal elections | 965 | 965 | 799 | 799 |

Notes: The regression model used is that of Table 4, column 3, and Table 5, column 1. To facilitate comparability, the sample in columns 3 and 4 is adjusted to contain the same set of candidates. This table only reports the beauty and competence coefficients. * significant at 10%; ** significant at 5%; *** significant at 1% (based on robust standard errors).

The beauty coefficients are rather stable. The competence coefficients are in contrast quite unstable. Previous studies have either just excluded individual assessments of recognized candidates (Benjamin and Shapiro 2006), excluded “well-known” candidates from the set of photos (King and Leigh 2007), or both (Todorov et al. 2005).²⁸ Since recognition can be partial and unconscious, we think that the results of previous studies should be interpreted with some caution, as they are based on assessments by respondents of the same nationality as the political candidates and do not systematically test if the use of foreign respondents produces similar results. This entails a risk for non-reported recognition which we avoid in our main study with non-Finnish respondents. In particular, the unstable competence coefficients point at a possible problem with the results of Todorov et al. (2005), who find that perceived competence is a good predictor of electoral success.

²⁸ Benjamin and Shapiro (2006) did not ask their participants to evaluate candidates from Massachusetts, the state in which almost all of their participants resided, or to evaluate candidates from the state where they grew up. King and Leigh (2007) also use one non-Australian respondent to evaluate photos in a sensitivity test.

7.2. Respondents Assessing All Photos

We have also conducted survey 3, with a small number of respondents who each assessed *all* 504 photos of Helsinki municipal candidates.²⁹ The reason was to see whether this way of assessing photos – used in labor market studies – gives rise to different overall assessments and results compared to the approach taken in our other surveys, where a much greater number of respondents each assessed a randomly drawn small number of photos. We have ten Finnish and six Swedish respondents in this survey. For both nationalities, one half of the respondents are men and the other half women. The youngest respondent is 22 and the oldest 70, with 36 as the mean age.³⁰

When looking at regression results in Table 8, three new comparisons can be made (all of them restricted to the Helsinki municipal election): between results based on this small survey’s Swedish respondents (see column 1) and results from our main survey with non-Finnish respondents (column 3); between results based on this survey’s Finnish respondents (column 2) and results from our survey of Finns (column 4); and between results based on Swedish and results based on Finnish respondents in the small survey (columns 1 and 2)

TABLE 8. Relative success in the Helsinki municipal elections, all candidates, non-incumbents.

| | (1) | (2) | (3) | (4) |
|------------|---------------------|---------------------|-------------------------|---------------------|
| | Swedish respondents | Finnish respondents | non-Finnish respondents | Finnish respondents |
| | small survey | small survey | main survey | survey of Finns |
| Beauty | 26.68** | 24.77** | 28.71** | 28.69** |
| Competence | 2.86 | 45.76** | 21.36 | 43.61* |

Notes: The regression model used is that of Table 4, column 3. This table only reports the beauty and competence coefficients. * significant at 10%; ** significant at 5%; *** significant at 1%.

²⁹ The reason for using only this subset of all photos is that it would be too time-consuming for a respondent to evaluate 1,929 photos.

³⁰ The pairwise correlations of beauty assessments among our Swedish respondents range from 0.42 to 0.61, with an average of 0.52, compared to a range from 0.12 to 0.62 with an average of 0.42 for the Finnish respondents.

The differences as far as beauty is concerned are very small irrespective of which comparison is made – and notably, beauty retains statistical significance throughout. Once again, there is a difference in the competence coefficients between the non-Finnish respondents in columns 1 and 3 and the Finnish respondents in columns 2 and 4. This difference, as we argue in Section 7.1, plausibly depends on the occurrence of recognition.

Thus, the two methods – using a small number of respondents who assess all photos and using a large number of respondents who assess a random selection of photos – seem to yield quite similar results. We have investigated this tentative conclusion further, in order to see to what extent the results are sensitive to the composition of respondents. To do this we have composed different groups of four respondents and estimated regressions based on their assessments, along the lines of Hamermesh (2006) and King and Leigh (2007). In doing this we kept the groups balanced in terms of age and gender of included respondents. When studying Swedish respondents, this was done by letting each group be composed of two men and two women, including the oldest person of each gender. This restriction gives rise to four groups of Swedish respondents. We find that beauty coefficients are quite stable (ranging from 25.8 to 29.3) and always statistically significant. However, when combining the ten Finnish respondents in 16 different groups of four, the results are not as clear-cut, again plausibly reflecting the importance of recognition. The beauty coefficients range from 10.4 to 20.3 and are only statistically significant in six of the 16 regressions. This fluctuation suggests that surveys with a small number of respondents – four in this case – may produce unstable results when respondents come from the same country as the persons that they assess.

7.3. Respondents without Information about the Photos

We have furthermore conducted survey 4, the no-information survey, in which it was *not* revealed that the photos depict political candidates or that we are studying politics.³¹ Once

³¹ In this survey, respondents had to evaluate at least ten photos. Another difference was that there was no option of choosing “Do not know / Do not want to answer” when evaluating the photos.

more, the same photos are used as in the main survey and in the survey with only Finnish respondents.

In terms of average assessments of the traits, we find small differences, typically in the order of 0.1–0.2 units on the five-point scale, compared to when respondents knew that the photos depicted political candidates. We have also carried out regressions for the parliamentary and municipal elections for non-incumbents. In the parliamentary election, beauty retains statistical significance for all candidates, but it is somewhat less important in terms of estimated coefficient size, compared to the results in Table 4. In the municipal elections, beauty is somewhat more important for female candidates and less important for male candidates (compared with Table 5). Furthermore, tests of statistical significance indicate a difference between the beauty coefficients of female and male candidates, the latter being smaller and not statistically significantly different from zero.

In all, these findings indicate that assessments may be modestly affected by the knowledge that the photos depict political candidates. But beauty comes through as a strong explanatory variable here too, especially for female candidates.

8. Conclusions

We investigate how beauty is related to electoral success in Finland and find that candidates who look better than their list competitors are more successful. In the parliamentary election, an increase in beauty of one standard deviation is associated with a 20% increase in the number of votes for the average non-incumbent candidate. In the municipal elections, the figure is 17%. The figures are based on assessments by non-Finns in order to make sure that candidates were not recognized.

The Finnish electoral system provides an ideal testing ground. It is proportional and each voter has to vote for one candidate on a party list, which makes it possible to look at the effect of beauty in within-party competition. Studying within-party competition holds several advantages. First, studies of between-party competition may face a reverse-causality problem

if a party is more successful in recruiting good-looking candidates in districts where it enjoys strong support. Second, within-party competition allows us to control for ideology very effectively. Third, we can study non-incumbent candidates separately in addition to a sample of both incumbents and non-incumbents.

Why does beauty matter? We are able to discriminate between two alternative explanations. On the one hand, voters may favor good-looking candidates either because they enjoy watching them or because good-looking politicians are more successful in social interaction. On the other hand, beauty may matter as a thin slice of information used to infer competence or similar traits. This explanation accords well with the results of Todorov et al. (2005), indicating that competence assessments predict electoral success.

Evidence assembled by studying beauty assessments together with assessments of competence and other traits, by including information about education and occupation as alternative thin slices of information, and by testing whether the effects of beauty differ between incumbents and non-incumbents all point in the same direction: voters favor good-looking candidates either because they enjoy watching them, or because good-looking politicians are more successful in social interaction. There are at least two potential reasons for why Todorov et al. obtain a different result: unreported candidate recognition and reverse causality.

Extensive sensitivity analysis confirms our main results. Using measures based on ordinal assessments of beauty, competence and trustworthiness, as well as more minor specification changes – none of these modifications alter the qualitative findings. Furthermore, our three additional surveys, using Finnish respondents, using respondents who assessed all as opposed to a random selection of photos and using respondents who did not know that the photos depict political candidates, confirm the main result. Beauty matters.

Lastly, in Finland, all parties have several male and female candidates on each list. Although the estimated effect of beauty is about the same for male and female candidates in our main regressions, there are some signs of beauty being more important for female candidates. First, beauty matters only for female candidates in some specifications in the municipal elections. Second, the Spearman rank correlation between beauty and our measure of electoral success is statistically significant only for female candidates (in the Helsinki municipal elec-

tion). Third, the standard deviation of the candidates' beauty is higher for female than for male candidates, meaning that a larger share of female candidates can be found in the upper tail of the beauty distribution. In consequence, beauty seems more important for female candidates. This stands in contrast to labor market studies, where the beauty premium has been found to benefit males more than females.

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 | Trustworthiness | Likability | Intelligence |
|-----------------|--------|------------|-----------------|------------|--------------|
| Beauty | 1.00 | | | | |
| Competence | 0.32 | 1.00 | | | |
| Trustworthiness | 0.22 | 0.38 | 1.00 | | |
| Likability | 0.41 | 0.32 | 0.51 | 1.00 | |
| Intelligence | 0.28 | 0.65 | 0.36 | 0.28 | 1.00 |

Notes: All of the reported traits exhibit statistically significant correlations with each other.

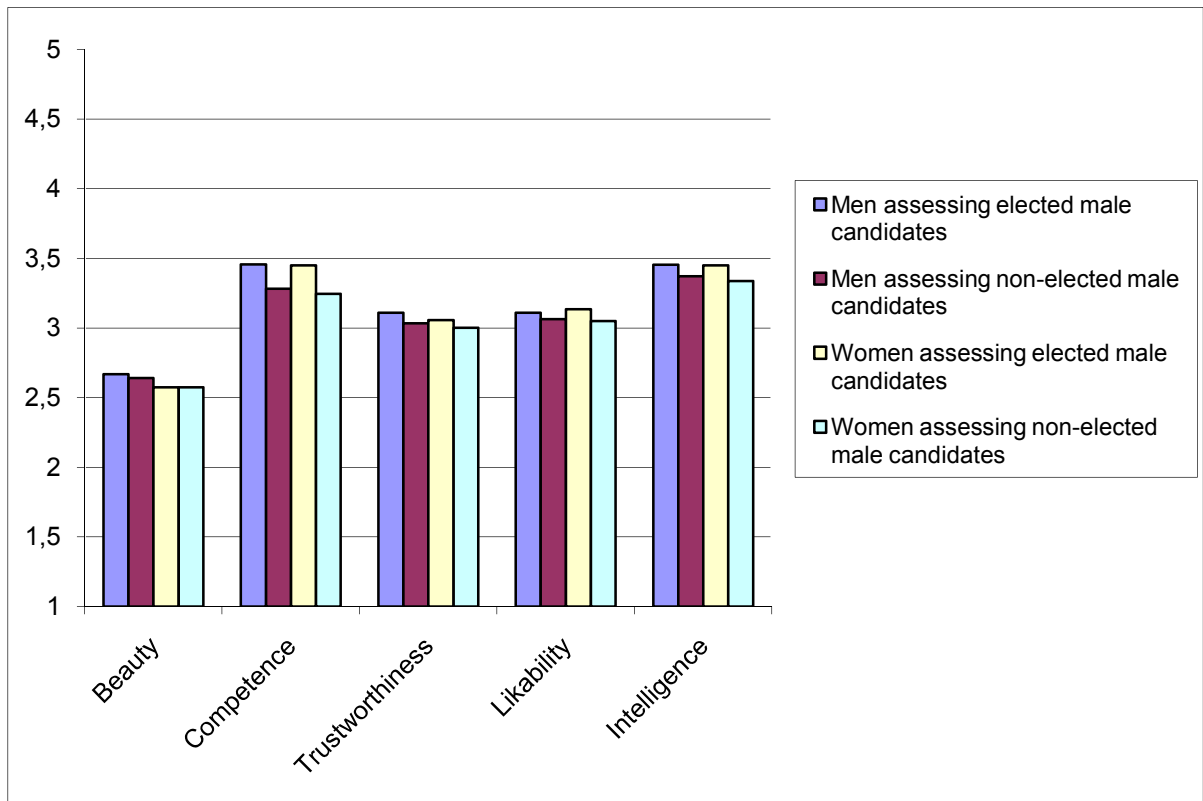


FIGURE A1. Assessments of elected and non-elected male candidates.

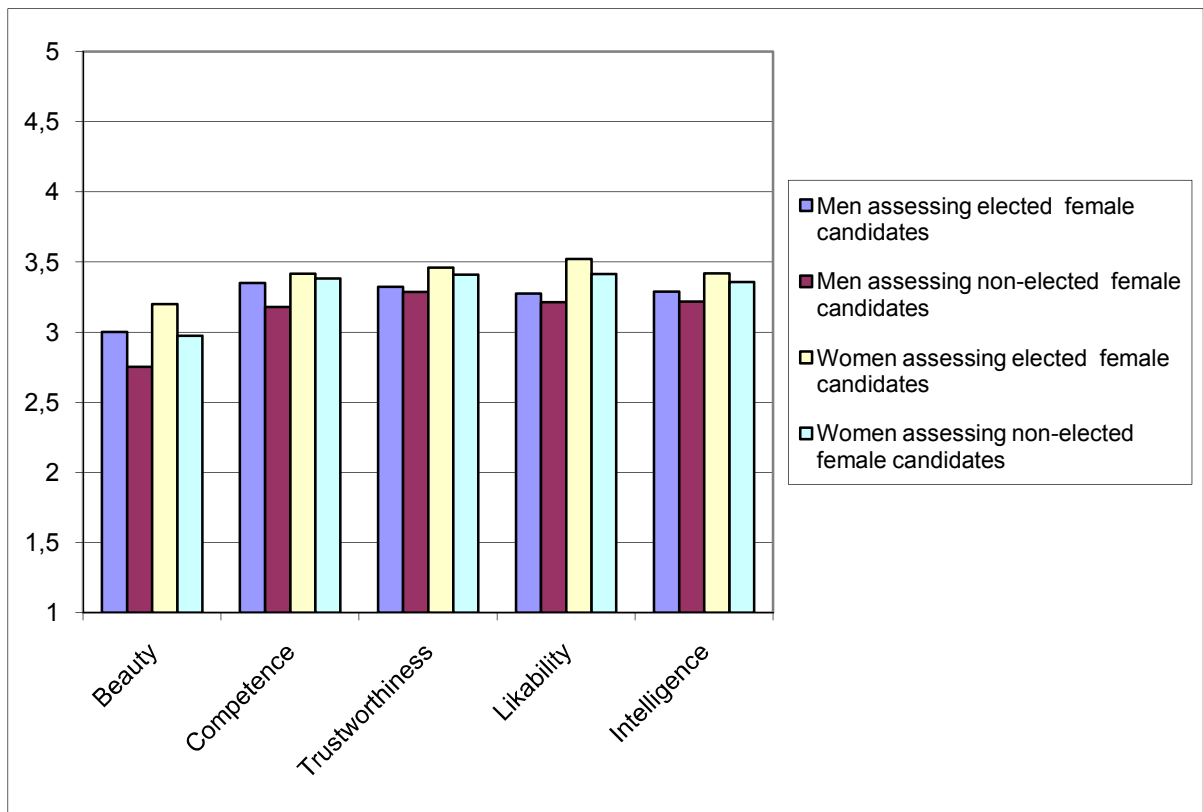


FIGURE A2. Assessments of elected and non-elected female candidates.

TABLE A2. Relative success in the parliamentary election, with occupational and educational dummies.

| | Relative success non-incumbents | Relative success non-incumbents | Relative success non-incumbents |
|------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Beauty | 29.58*** (6.22) | 29.16*** (6.42) | 28.79*** (6.22) |
| Competence | 9.56 (8.22) | 10.51 (8.58) | 10.18 (8.25) |
| Trustworthiness | 2.56 (8.52) | 6.54 (8.72) | 3.53 (8.54) |
| Party worker | 16.12 (30.8) | | 14.77 (30.0) |
| Management | 6.82 (18.5) | | -0.01 (18.9) |
| Researcher | 36.61 (27.8) | | 26.79 (28.1) |
| Teacher | -15.87 (14.9) | | -25.54 (16.3) |
| Upper white collar | -8.38 (15.6) | | -15.21 (16.4) |
| Medical doctor | -2.17 (17.7) | | -14.43 (19.4) |
| Nurse | -19.58 (16.1) | | -14.26 (17.5) |
| Lower white collar | -23.15 (14.9) | | -23.31 (15.3) |
| Worker | -34.92** (14.1) | | -30.07** (13.9) |
| Entrepreneur | -15.09 (17.8) | | -16.32 (17.7) |
| Artist | -36.73** (16.2) | | -39.49** (16.0) |
| Student | -53.25*** (17.8) | | -34.23** (15.6) |
| Not employed | -38.37* (20.4) | | -29.67 (19.6) |
| University education | | 17.14** (8.62) | 13.15 (10.5) |
| Vocational education | | -9.98 (8.86) | -4.70 (10.3) |
| Upper-secondary education | | -35.18*** (12.5) | -25.33* (14.6) |
| Comprehensive school or less | | -49.06*** (10.4) | -44.18*** (10.7) |
| Male dummy | 3.80 (6.92) | 6.14 (6.76) | 4.83 (7.00) |
| Young (age<30) | -5.77 (12.2) | 3.16 (12.7) | 0.54 (13.1) |
| Old (age>60) | 10.67 (22.5) | 9.62 (21.5) | 9.79 (21.9) |
| Number of candidates | 641 | 641 | 641 |
| Adjusted R-squared | 0.09 | 0.09 | 0.10 |

Notes: The occupational classification follows Statistics Finland (2001), though we have merged certain occupational categories with a small number of candidates and listed party workers as a group of their own. The reference group for occupation is candidates who did not list their occupation. Comprehensive school or less corresponds to at most 10 years of schooling. Upper-secondary education corresponds to 12 years of schooling, and vocational education 10–12 years. Upper-secondary education serves usually as preparing for university-level education, and many of the candidates with upper-secondary education listed as highest education have started, but not completed, university studies. Vocational education includes, e.g., basic nurses, nurses, commercial school graduates, clerks, and artisans. The reference group for education is candidates who did not list their education. Robust standard errors in parentheses. The regressions include a constant term. * significant at 10%; ** significant at 5%; *** significant at 1%.

References

- Aharon, Itzhak, Nancy Etcoff, Dan Ariely, Christopher F. Chabris, Ethan O'Connor, and Hans C. Breiter (2001). "Beautiful Faces Have Variable Reward Value: fMRI and Behavioral Evidence." *Neuron*, 32, 537–551.
- Ambady, Nalini, and Robert Rosenthal (1992). "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis." *Psychological Bulletin*, 111, 256–274.
- Andreoni, James, and Ragan Petrie (forthcoming). "Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments." *Journal of Economic Psychology*.
- Bartels, Larry M. (1996). "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science*, 40, 194–231.
- Bengtsson, Åsa, and Kimmo Grönlund (2005). "Ehdokasvalinta." ["Candidate Selection."] In *Vaalit ja demokratia Suomessa [Elections and Democracy in Finland]*, edited by Heikki Paloheimo. WSOY.
- Benjamin, Daniel J., and Jesse M. Shapiro (2006). "Thin-Slice Forecasts of Gubernatorial Elections." NBER Working Paper No. 12660. November.
- Berggren, Niclas, Henrik Jordahl, and Panu Poutvaara (2006). "The Looks of a Winner: Beauty, Gender and Electoral Success." IZA Discussion Paper No. 2311. September.
- Besley, Timothy (2004). "Paying Politicians: Theory and Evidence." *Journal of the European Economic Association*, 2, 193–215.
- Biddle, Jeff E., and Daniel S. Hamermesh (1998). "Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics*, 16, 172–201.
- Budesheim, Thomas Lee, and Stephen J. DePaola (1994). "Beauty or the Beast? The Effects of Appearance, Personality, and Issue Information on Evaluations of Political Candidates." *Personality and Social Psychology Bulletin*, 20, 339–348.
- Caprara, Gian Vittorio, Claudio Barbaranelli, and Philip G. Zimbardo (1997). "Politicians' Uniquely Simple Personalities." *Nature*, 385, 493.

- Caselli, Francesco, and Massimo Morelli (2004). "Bad Politicians." *Journal of Public Economics*, 88, 759–782.
- Cunningham, Michael R., Alan R. Roberts, Anita P. Barbee, Perry B. Druen, and Cheng-Huan Wu (1995). "Consistency and Variability in the Cross-Cultural Perception of Female Physical Attractiveness." *Journal of Personality and Social Psychology*, 68, 261–279.
- Downs, Anthony (1957). *An Economic Theory of Democracy*. HarperCollins.
- Eagly, Alice H., Richard D. Ashmore, Mona G. Makhijani, and Laura C. Longo (1991). "What Is Beautiful Is Good, But...: A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype." *Psychological Bulletin*, 110, 109–128.
- Feingold, Alan (1992a). "Good-Looking People Are Not What We Think." *Psychological Bulletin*, 111, 304–341.
- Feingold, Alan (1992b). "Gender Differences in Mate Selection Preferences: A Test of the Parental Investment Model." *Psychological Bulletin*, 112, 125–139.
- French, Michael T. (2002). "Physical Appearance and Earnings: Further Evidence." *Applied Economics*, 34, 569–572.
- Hamermesh, Daniel S. (2006). "Changing Looks and Changing 'Discrimination': The Beauty of Economists." *Economics Letters*, 93, 405–412.
- Hamermesh, Daniel S., and Jeff E. Biddle (1994). "Beauty and the Labor Market." *American Economic Review*, 84, 1174–1194.
- Hamermesh, Daniel S., Xin Meng, and Junsen Zhang (2002). "Dress for Success – Does Priming Pay?" *Labour Economics*, 9, 361–373.
- Harper, Barry (2000). "Beauty, Stature and the Labour Market: A British Cohort Study." *Oxford Bulletin of Economics and Statistics*, 62, 771–800.
- Jackson, Linda A., John E. Hunter, and Carole N. Hodge (1995). "Physical Attractiveness and Intellectual Competence: A Meta-Analytic Review." *Social Psychology Quarterly*, 63, 108–122.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky (1986). *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press.

- Kanazawa, Satoshi, and Jody L. Kovar (2004). "Why Beautiful People Are More Intelligent." *Intelligence*, 32, 227–243.
- King, Amy, and Andrew Leigh (2007). "Beautiful Politicians." Available at SSRN: <http://ssrn.com/abstract=961138>.
- Langlois, Judith H., Lisa Kalakanis, Adam J. Rubenstein, Andrea Larson, Monica Hallamm, and Monica Smoot (2000). "Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review." *Psychological Bulletin*, 126, 390–423.
- Lau, Richard R., and David P. Redlawsk (2001). "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science*, 45, 951–971.
- Lee, David S. (forthcoming). "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics*.
- Lupia, Arthur (1994). "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review*, 88, 63–76.
- Macrae, C. Neil, Alan B. Milne, and Galen V. Bodenhausen (1994). "Stereotypes as Energy-Saving Devices: A Peek Inside the Cognitive Toolbox." *Journal of Personality and Social Psychology*, 66, 37–47.
- Messner, Matthias, and Matthias K. Polborn (2004). "Paying Politicians." *Journal of Public Economics*, 88, 2423–2445.
- Mobius, Markus M., and Tanya S. Rosenblat (2006). "Why Beauty Matters." *American Economic Review*, 96, 222–235.
- Mocan, Naci, and Erdal Tekin (2006). "Ugly Criminals." NBER Working Paper No. 12019.
- Pfann, Gerard A., Jeff E. Biddle, Daniel S. Hamermesh, and Ciska M. Bosman (2000). "Business Success and Businesses' Beauty Capital." *Economics Letters*, 67, 201–207.
- Poutvaara, Panu, and Tuomas Takalo (2007). "Candidate Quality." *International Tax and Public Finance*, 14, 7–27.
- Raunio, Tapio (2005). "Finland: One Hundred Years of Quietude." In *The Politics of Electoral Systems*, edited by Michael Gallagher and Paul Mitchell. Oxford University Press.

- Redlawsk, David P., and Richard R. Lau (2003). "Do Voters Want Candidates They Like or Candidates They Agree With? Affect vs. Cognition in Voter Decision Making." Mimeograph, University of Iowa.
- Reynolds, Andrew, Ben Reilly, and Andrew Ellis (2005). *Electoral System Design: The new International IDEA Handbook*. International Institute for Democracy and Electoral Assistance.
- Rosar, Ulrich, Markus Klein, and Tilo Beckers (forthcoming). "The Frog Pond Beauty Contest: Physical Attractiveness and Electoral Success of the Constituency Candidates at the North Rhine-Westphalia State Election 2005." *European Journal of Political Research*.
- Simon, Herbert A. (1985). "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review*, 79, 293–304.
- Statistics Finland (2001). "Classification of Occupations 2001." http://tilastokeskus.fi/tk/tt/luokitukset/lk_en/ammatti_index.html.
- Statistics Finland (2003). "Parliamentary Elections, 2003." http://www.tilastokeskus.fi/tup/suoluk/suoluk_vaalit_en.html.
- Todorov, Alexander, Anesu N. Mandisodza, Amir Goren, and Crystal C. Hall (2005). "Inferences of Competence from Faces Predict Election Outcomes." *Science*, 308, 1623–1626.
- Wattenberg, Martin P (1991). *The Rise of Candidate-Centered Politics: Presidential Elections of the 1980s*. Harvard University Press.
- Willis, Janine, and Alexander Todorov (2006). "First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face." *Psychological Science*, 17, 592–598.