Centre Emile Bernheim
Research Institute in Management Sciences

Solvay Brussels School
Economics & Management

Partisan stereotypes

Carmelo Licata and Pierre-Guillaume Méon

Using two surveys, we study how respondents process visual cues to identify the political orientation (left- vs. right-wing) of members of the French National Assembly (referred to as “deputies”), based on official photographs only, to test the type of heuristic that they use. We first confirm that respondents outperform random guesses. Second, we find that their categorizations correlate with observable characteristics (gender, tie color, jewelry) and subjective assessments of deputies’ personality traits (attractiveness, competence, trustworthiness). Third, the objective visual cues that respondents use are consistent with the actual characteristics of left- and right-wing deputies, and respondents mistakenly react to subjective personality traits that differ little across the two groups of deputies. Fourth, left- and right-wing respondents use the same cues in the same way, attractiveness being the only exception. Fifth, the magnitude of the marginal impact of a characteristic on the probability of a respondent categorizing a photograph as left- or right-wing increases strictly with the representativeness of that characteristic. Finally, we find evidence that some characteristics correlate with categorization errors. Findings 1, 2, 4, and the finding that respondents use cues in the correct way are consistent with both Bayesian behavior and the representativeness heuristic. Findings 5, 6, and the finding that respondents react to subjective cues that do not differ across groups are at odds with Bayesian inference but consistent with the representativeness heuristic suggested by Kahneman and Tversky (1972) and recently modelled by Gennaioli and Shleifer (2010), and Bordalo et al. (forthcoming).

Keywords: Stereotypes, Appearance, Political candidates, Party affiliation.
JEL Classifications: D72, D83, D84, J45, J7.

CEB Working Paper N° 16/037
July 2016

Université Libre de Bruxelles - Solvay Brussels School of Economics and Management
Centre Emile Bernheim
ULB CP114/03 50, avenue F.D. Roosevelt 1050 Brussels BELGIUM
e-mail: ceb@admin.ulb.ac.be Tel.: +32 (0)2/650.48.64 Fax: +32 (0)2/650.41.88

ULB

Partisan stereotypes

July 2016

Carmelo Licata
Université libre de Bruxelles (ULB)
CP-114/03
Avenue F.D. Roosevelt, 50
1050 Brussels, Belgium
Phone: +32 2 650 65 99
e-mail: carmelo.licata@gmail.com

Pierre-Guillaume Méon
Université libre de Bruxelles (ULB)
Centre Emile Bernheim
CP-114/03
Avenue F.D. Roosevelt, 50
1050 Brussels, Belgium
Phone: +32 2 650 65 99
e-mail: pgmeon@ulb.ac.be

Abstract: Using two surveys, we study how respondents process visual cues to identify the political orientation (left- vs. right-wing) of members of the French National Assembly (referred to as "deputies"), based on official photographs only, to test the type of heuristic that they use. We first confirm that respondents outperform random guesses. Second, we find that their categorizations correlate with observable characteristics (gender, tie color, jewelry) and subjective assessments of deputies' personality traits (attractiveness, competence, trustworthiness). Third, the objective visual cues that respondents use are consistent with the actual characteristics of left- and right-wing deputies, and respondents mistakenly react to subjective personality traits that differ little across the two groups of deputies. Fourth, left- and right-wing respondents use the same cues in the same way, attractiveness being the only exception. Fifth, the magnitude of the marginal impact of a characteristic on the probability of a respondent categorizing a photograph as left- or right-wing increases strictly with the representativeness of that characteristic. Finally, we find evidence that some characteristics correlate with categorization errors. Findings 1, 2, 4, and the finding that respondents use cues in the correct way are consistent with both Bayesian behavior and the representativeness heuristic. Findings 5, 6, and the finding that respondents react to subjective cues that do not differ across groups are at odds with Bayesian inference but consistent with the representativeness heuristic suggested by Kahneman and Tversky (1972) and recently modelled by Gennaioli and Shleifer (2010), and Bordalo et al. (forthcoming).

JEL Codes: D72, D83, D84, J45, J7.

Keywords: Stereotypes, Appearance, Political candidates, Party affiliation.

Acknowledgments: We thank Niclas Berggren, Marco Gianni, Christophe Hurlin, Henrik Jordahl, Rachel Kranton, Krisztina Kis-Katos, Grégory Levieuge, James Rockey, Sandra Rothenberger, and Claudia Toma for helpful comments. We are also grateful to the participants of the Silvaplana Workshop in Political Economy in Pontresina, to the participants of the European Public Choice Society Meeting in Groningen, to the participants of the annual conference of the Society for Institutional and Organizational Economics in Paris, and to seminar participants at Centre Emile Bernheim, at the Research Institute of Industrial Economics (IFN), Stockholm, and at the University of Orléans. The authors might nitpick about the distribution of responsibility for remaining errors and omissions between them, but agree that none of it rests on anyone but themselves.

1. Introduction

Any social group can be stereotyped, and researchers from various fields have investigated stereotypes about groups defined using characteristics as diverse as race, gender, occupation, age, nationality, region of origin, or college major (Jussim et al., 2016). Political parties are no exception, and partisan stereotypes have been documented. Westfall et al. (2015) and Bordalo et al. (forthcoming) are but recent examples of studies of the stereotypes held about Republicans and Democrats.

The literature on stereotyping and categorization provides conflicting views on the way in which people form and use stereotypes to categorize others. One class of theories assumes that people are rational and correctly follow Bayesian principles. Thus, people should form beliefs about groups that accurately reflect all the characteristics of those groups. They should then on average correctly infer group membership from available cues. Models of statistical discrimination inspired by Phelps (1973) follow those principles. A second class of theories, spurred by Kahneman and Tversky (1972, 1973), has emphasized and modelled various types of heuristics resulting in biased judgments. Studies on categorization such as Mullainathan (2002) or Fryer and Jackson (2008) assume that people use coarse categories to save on the cost of processing information. More to the point, work by Kahneman and Tversky (1972, 1973), Gennaioli and Shleifer (2010), and Bordalo et al. (forthcoming) describes how people use the representativeness heuristic to form beliefs about groups and categorize individuals across groups. In a nutshell, people do not use all the characteristics of groups when making category judgments, but only the characteristics that are the most representative. Specifically, they focus on characteristics whose frequencies differ the most across groups, even if those characteristics are not the most frequent in each group considered separately. That heuristic simplifies the judgment problem but may lead to biases and judgement errors.

The aim of this paper is to test the two classes of theories by investigating how respondents to two online surveys categorized anonymous photographs of members of the French National Assembly (referred to as “deputies”) as left- or right-wing. In doing so, our study elaborates on the recent literature on the ability of people to identify the political orientation of anonymous politicians. Samochowiec et al. (2010), for example, observed that survey respondents could outperform chance when guessing the political orientation of unknown Swiss and German politicians from their photographs. Olivola et al. (2012) found that survey respondents could

distinguish Republican from Democrat candidates with above-chance accuracy simply by looking at their photographs. This literature suggests that people hold stereotypes about the appearance of left- and right-wing politicians, and can use those stereotypes.

However, the research done so far has not investigated *how* visual cues are processed to categorize politicians. We fill that gap by relating this recent literature to the theoretical literature on stereotypes and categorization. In doing so, we therefore contribute to both strands of literature. On the one hand, we suggest a theoretical interpretation for the behavior reported in the literature devoted to the visual identification of left- and right-wing politicians. On the other hand, we empirically test theories of stereotypes and categorization. Our paper is the first to bring the two literatures together.

The classic left-right distinction makes it possible to test the theories of how people categorize others, because it provides what Bordalo et al. (forthcoming) refer to as a "natural definition of comparison groups". It is a key dimension of politics that results from a long historical process and is therefore fairly consensual. This natural definition of the two groups allows us to study the partisan stereotypes prompting respondents to classify politicians as left- or right-wing, and how they are processed.

Specifically, we use two complementary surveys asking respondents to classify actual French deputies as left- or right-wing based on their anonymous photographs. In the two surveys, respondents had to make decisions based on the stereotypes that they held about the left and the right, and on the visual cues appearing on photographs. The first survey elicited nearly 70,000 categorizations and the second more than 3,000. The two surveys therefore resulted in the largest samples used so far to assess the ability of respondents to categorize politicians. In the second survey, respondents additionally assessed subjective traits of deputies, namely competence, attractiveness, and trustworthiness. We complemented the information on those characteristics by coding objective characteristics, such as gender, ethnicity, tie color, or jewelry. Our dataset therefore contains at the same time information about the beliefs that respondents hold about what left- and right-wing deputies look like, and information about what left- and right-wing deputies truly look like. We use the dataset to address a series of embedded questions in order to test whether respondents are rational or use the representativeness heuristic.

First, we confirm that respondents outperform random guesses when categorizing anonymous deputies as left- or right-wing. Our findings thus extend the findings of Samochowiec et al. (2010) and Olivola et al. (2012) to another country. Second, we test whether categorizations

correlate with specific observable cues (gender, tie color, jewelry) and subjective assessments of deputies' personality traits (attractiveness, competence, trustworthiness). We find that they do. While Rule and Ambady (2010) relate categorizations to subjective assessments of warmth and power, our paper is the first to determine a set of objective visual cues that influence categorizations. Third, we test whether the cues that respondents use are consistent with the actual characteristics of left- and right-wing deputies. We find that objective visual cues, such as gender or race, clearly are. Subjective perceptions of deputies' competence and trustworthiness are used in the correct direction by respondents. Respondents correctly associate competence with the right and trustworthiness with the left, while actual left- and right-wing deputies differ little in terms of those two dimensions. Our paper is therefore the first to show that visual stereotypes contain a kernel of truth, but can also exaggerate differences, in line with Bordalo et al.'s (forthcoming) modelling of the representativeness heuristic. Fourth, we test whether left- and right-wing respondents use the same cues in the same way to categorize deputies. We find that they do. Attractiveness is the only exception, as left-wing respondents categorize attractive deputies as left-wing while right-wing respondents categorize them as right-wing. Fifth, we investigate the magnitude of the marginal impact of a characteristic on the probability of a respondent categorizing a photograph as left- or right-wing and observe that it increases with the representativeness of that characteristic. Finally, we test whether categorization errors are randomly distributed. We find that representative characteristics tend to correlate with categorization errors. Whereas the first test is an extension of an existing result to a new country, all the other tests are original and have not been performed previously.

While findings 1, 2, and 4 are consistent with both Bayesian behavior and the representativeness heuristic, findings 5 and 6 are at odds with Bayesian behavior but consistent with the representativeness heuristic. Finding 3 is more ambiguous. Finding that respondents use cues in the correct direction is consistent with both Bayesian behavior and the representativeness heuristic. However, finding that they tend to react to subjective characteristics that differ little across groups suggests they resort to the representativeness heuristic, which results in stereotypes that amplify differences across groups. Overall, our findings therefore lend support to the contention that respondents resort to the representativeness heuristic.

Beside the implications for theories of stereotypes and our understanding of how people behave when categorizing unknown deputies, investigating how people form stereotypes about political parties and infer party membership matters because appearances have been shown to affect

political success, for instance by Rosenberg et al. (1986), Todorov et al. (2005), Ballew and Todorov (2007), or Banducci et al. (2008). Antonakis and Dalgas (2009) even find that the preference for candidates of children aged 5 to 13 years is a predictor of electoral success, suggesting that the capacity to evaluate appearance is acquired early in life. Berggren et al. (2010) show that beauty is associated with the number of votes non-incumbents receive in Finnish parliamentary elections.

More to the point, appearance interacts with electoral outcomes through partisan affiliation. Specifically, Berggren et al. (2011) find that if beauty increases the number of votes that a candidate receives in local elections for all candidates, its effect is larger for right-wing than for left-wing candidates. Olivola et al. (2012) also find that Republican candidates who look more Republican, insofar as they are more accurately identified as Republicans by survey participants, perform better in Republican constituencies.

Furthermore partisan stereotypes may interfere with political discourse and campaigns. Rahn (1995) provides experimental evidence that people discard the information about a candidate's favored policies contained in the candidate's speech if the candidate's party affiliation is known and the information does not conform to the party affiliation of the candidate. As a result, if appearance reveals party affiliation and party affiliation trumps discourse, then appearance may interfere with party platforms. Conversely, if politicians manipulate their discourse but people are able to correctly read visual cues, stereotypes may reveal politicians' true preferences, and attempts to misrepresent views may become futile.[1] If appearance reveals political affiliation and affects electoral outcomes, determining the cues that people use to classify unknown politicians as left- or right-wing, and how they use them, becomes a key research question.

With those ends in view, the rest of the paper is organized as follows. The next section provides a theoretical framework. Section 3 describes our data. Section 4 tests the ability of respondents to categorize deputies as left- or right-wing. Section 5 investigates the stereotypes used by respondents by investigating the visual cues that respondents associate with left- and right-wing deputies. Section 6 refines the results of Section 5 by investigating separately the stereotypes held by respondents with different political preferences. Section 7 investigates the categorization errors that respondents make when categorizing deputies. Section 8 concludes.

[1] Anecdotal evidence is provided by the media buzz caused during the French 2007 presidential election by the footage of Pierre Bourdieu declaring about socialist candidate in the presidential election Ségolène Royal "You instantly know she's not left-wing. […] She has a way of being, a way of speaking that tells you 'she's right-wing' […] even if she utters left-wing words" (see e.g. Le Monde, 2006, or the following link to the footage: http://www.dailymotion.com/video/x1g7vj_bourdieu-segolene-royal-est-de-droi_news).

2. Theoretical framework

The basic premise necessary to consider that people can categorize left- and right-wing politicians is that left- and right-wing politicians must look different. In other words, political affiliations must correlate with observable characteristics or visual cues. Otherwise, respondents would be unable to categorize deputies based on their photographs only. The first sub-section therefore discusses the mechanisms that may prompt such differences. However, visual cues only matter to the extent that people can use them to categorize politicians. The second sub-section therefore surveys the cues that the literature has so far shown to operate. The final sub-section then discusses how such information is processed by agents and the models that have been put forward to describe their behavior.

2.1. Why should left- and right-wing politicians look different?

One can consider two broad sets of reasons why left- and right-wing politicians may look different. First, appearances may reveal personality attributes that directly affect preferences and political affiliations. Second, observable characteristics may affect people's outcomes on the labor market, thereby prompting them to favor different political programs.

A first personality trait that may be revealed by appearance is social dominance. Samochowiec et al. (2010) argue that conservatism correlates with facial appearance through the impact of male testosterone on social dominance, i.e. the tendency to support group-based hierarchy in the psychology literature. Their argument rests on the finding that people who rank higher on social dominance tend to endorse conservative, hierarchy-enhancing, political stances, as argued by Sidanius and Pratto (2001), while social dominance tends to correlate with testosterone (see Sellers et al. 2006 for a survey and evidence of a stable relationship). Because higher testosterone also results in a more masculine appearance (Penton-Voak and Cheng, 2004), it may indirectly provide a visual cue on ideological preferences. Similarly, Price et al. (2011) observe that more muscular men tend to report a lower preference for egalitarianism, a trait that one would associate with more redistributive policies endorsed by left-wing parties.

Gender is another easily observable cue that correlates with traits that can be indicative of political preferences. For instance, Croson and Gneezy (2009) survey experimental evidence suggesting that, although there are variations across experiments, women tend to favor more equal income distributions. Price et al. (2011) similarly observe that men report higher social dominance than women.

Beside visual cues that are inherited or immutable, people's styles may correlate with their political preferences. Carney et al. (2008) found that self-reported liberals tend to smile more than self-reported conservatives. Clothing, fashion accessories, hairstyle, or facial hair can also provide information on political preferences. Samochowiec et al. (2010) thus argue that politically more conservative people may emphasize a high social status through their appearance and favor more old-fashioned and less experimental styles. In comparison, political liberals may look more unconventional.

In addition to directly revealing politically-relevant personality traits, observable cues may also affect the performance of individual agents on the labor market, thereby influencing their policy preferences. Evidence of race- and gender- inequality abounds, as surveys by Fryer (2011) or Bertrand (2011) emphasize. Groups that are discriminated against may thus rationally support left-wing parties, because they more visibly endorse policies aimed at reducing discrimination and support redistribution.

By the same token, evidence surveyed by Hamermesh (2013) pointing to a beauty premium on the labor market has accumulated since Hamermesh and Biddle's (1994) classic paper. If agents perceived as more attractive fare better on the labor market and get higher incomes, they will have an incentive to oppose more distributive policies, according to the classic model of Meltzer and Richard (1981) and the related literature surveyed in Alesina and Giuliano (2011). As distributive policies are usually advocated by left-wing parties, more attractive agents have an incentive to be more right-wing, as Berggren et al. (2011) point out.

The impact of attractiveness on party affiliation may be amplified for politicians if its impact on political success differs across parties. Berggren et al. (2011) observe that right-wing candidates in Finnish elections are not only better looking but more importantly enjoy a larger beauty premium in terms of votes. Career concerns would therefore give an incentive to attractive candidates to join a right-wing party, where they would have a greater chance of success. Similar mechanisms may be at work for other physical characteristics, resulting in more stereotypical parties.

Finally, visual cues may be manipulated to maximize electoral success. Whether such manipulation dampens or amplifies existing differences is beyond the scope of the present paper. What our study allows us to do is to test whether perceivers can categorize politicians, and how they do so, regardless of the origin of politicians' differences. Perceivers must therefore be able to interpret existing visual cues. In the next sub-section, we therefore survey the results of the literature investigating what perceivers can infer from the photographs of politicians.

2.2. What do people perceive?

As Samochowiec et al. (2010) argue, differences between left- and right-wingers are ubiquitous, and people must have experienced many opportunities in their lives to distinguish the two groups. In fact, their finding, and that of Olivola et al. (2012) that survey respondents could classify candidates with above-chance accuracy by looking at anonymous photographs, suggest that people indeed use visual cues.

The question then becomes whether one can identify the cues that people use. The two most visible characteristics are probably gender and race. Unsurprisingly, there is evidence that people use those characteristics to infer political preferences. Using two surveys, McDermott (1998) finds that US voters tend to assume that female and black candidates are more liberal than white male candidates and adjust their voting behavior accordingly. In their study, Samochowiec et al. (2010) explicitly asked survey respondents to list the cues they used to categorize anonymous photographs of Swiss politicians. Respondents mentioned gender among the cues that they used.

Borkenau and Liebler (1995) asked strangers to assess the personality of other participants in an experiment where the only available information was a video. They find that viewers analyze a broad spectrum of mannerisms, including the propensity to smile, when inferring a person's personality. Moreover, the cues that viewers analyze and relate to specific personality traits match the cues that do indeed correlate with those traits. In other words, they find that viewers correctly use visual cues to infer personality. Furthermore, viewers are particularly good at inferring conscientiousness from visual cues, which is a personality trait that Carney et al. (2008) have found to strongly correlate with conservatism.

The role of clothing is less clear. Participants in the study by Samochowiec et al. (2010) mentioned clothes and hairstyle among the visual cues that they used. However, in a particular variant of their study, Samochowiec et al. (2010) presented two versions of the photographs of the same pool of politicians, one where clothing cues were removed and one where faces were removed. While viewers could categorize photographs showing faces only with above-chance accuracy, their answers were indistinguishable from random answers when categorizing politicians from their clothes only. Clarifying the role of clothing is therefore one of the aims of the present study.

In any case, the evidence overall suggests that faces are perceived holistically, and that viewers do not isolate specific features.[2] Unsurprisingly therefore, perceivers seem to react to more

[2] See Gold et al. (2012) and Shen and Palmeri (2014) for a survey and a discussion.

general features of the photographs that they observe. Electoral success has thus been found to relate to attractiveness, by Hamermesh (2006) or Bergren et al. (2010, 2011), perceived competence, and perceived trust, by Todorov et al. (2005) and Ballew and Todorov (2007). Rule and Ambady (2010) observe that perceivers rank photographs along two main dimensions and categorize politicians accordingly. The first dimension, called warmth, features high loadings on likeability and trustworthiness. The second dimension, called power, features high loadings on dominance and facial maturity. Photographs that are higher on warmth tend to be categorized as left-wing while those higher on power tend to be categorized as right-wing.

Those results show that people do use visual cues to categorize anonymous photographs. However, they may use those cues in many different ways. The next section surveys how those cues are processed.

2.3. How do people process observable cues?

When asked to categorize photographs, a rational agent would process all observable cues in a Bayesian way. That agent would consider all the alternatives and their probabilities, and make guesses that would be on average accurate. Such behavior would be in line with the model of statistical discrimination put forward by Phelps (1973). A key implication is that classification errors would be random and on average zero.

However, people may not always behave in a Bayesian way, especially when they have to make a quick guess. Instead, they use stereotypes and heuristics. According to Hilton and von Hippel (1996, page 240), the standard definition of stereotypes in psychology is "beliefs about the characteristics, attributes, and behaviors of members of certain groups […] also theories about how and why certain attributes go together". Rahn (1986) more specifically defines partisan stereotypes as "cognitive structures that contain citizens' knowledge, beliefs, and expectancies about the two major political parties."

Kahneman and Tversky (1972, 1973) describe how people in practice categorize individuals across groups, a task they refer to as category prediction or nominal prediction. They argue that people depart from normative Bayesian decision rules in systematic ways and use simpler rules called heuristics. In particular, people seem to categorize individuals using the representativeness heuristic, whereby they categorize individuals on the basis of the characteristics that are the most diagnostic of the group to which those variables belong. Specifically, a characteristic is

representative of a group if the relative frequency of this characteristic is much higher in that group than in the reference group (Kahneman and Tversky, 1972).[3]

As Kahneman and Tversky (1972) point out, the representativeness heuristic sometimes yields reasonable outcomes, but may result in biases. Gennaioli and Shleifer (2010) and Bordalo et al. (forthcoming) explore this heuristic. They assume that people making a decision in a probabilistic environment will overweight the characteristics that are the most representative. Bordalo et al. (forthcoming) argue that the characteristics that "come to mind" when thinking about a particular group are the most representative of that group with respect to the rest of the population, as opposed to the most likely. Because stereotypes are based on the true distributions of characteristics of the groups, they contain a "kernel of truth". However, they may exaggerate differences when the characteristics that are representative of a group are not the most likely in that group.

Gennaioli and Shleifer (2010) model a task that is the closest to the task that respondents in our surveys were asked to perform. They model a decision maker who must determine whether an individual has a given characteristic based on information on a set of other characteristics. They assume that the decision maker only considers the most representative scenarios when categorizing the individual. They show that such a heuristic can lead to over- or under-reaction to information with respect to a Bayesian processing of the available information. The bias will be smaller if the characteristics that are representative of the groups are also likely.

To illustrate Gennaioli and Shleifer's (2010) model in the context of Survey 2, let us consider the reaction of respondents to tie color. Wearing a red tie will be distinctive of the left if the share of deputies wearing a red tie is larger on the left than on the right. According to Gennaioli and Shleifer's (2010) model, respondents would consider two scenarios, specifically left-wing deputies with a red tie and right-wing deputies with another tie. They will overlook the possibilities that a left-wing politician might wear another tie, and that a right-wing politician might wear a red tie. People will therefore systematically classify deputies wearing a red tie as left-wing. Tie color will indeed be informative if a vast majority of left-wing deputies wear such ties while right-wing deputies do not. In that case, the representativeness heuristic will result in a minor bias. However, if

[3] Stereotypes presuppose the existence of "groups". How people define the boundaries of groups is the focus of a specific literature. The interested reader may refer to Mullainathan (2002) or Fryer and Jackson (2008). Participants in the two surveys were asked to consider two groups: left- and right-wing deputies. The definition of those two groups is grounded in a long historical process. In the next section, we argue that the two groups could be objectively defined in our sample because of the composition and workings of the French parliament at the time of the surveys.

a majority of deputies wear red ties in both groups, a red tie will not be very informative, therefore leading to large biases.

The upshot of this section is threefold. First, both statistical discrimination and the representativeness heuristic will prompt respondents to react to characteristics that actually distinguish the groups. Both types of stereotypes do indeed contain a "kernel of truth". Second, if survey respondents use the kind of heuristic described by Kahneman and Tversky (1972, 1973) and Gennaioli and Shleifer (2010), they will react more to more representative characteristics. Determining whether the visual cues that elicit the largest reactions, specifically that affect categorizations the most, are also the most distinctive is a first way to test the representativeness heuristic. Third, if respondents use the representativeness heuristic, they may make systematic mistakes when reacting to observable cues. Conversely, errors will be random if respondents use Bayesian decision rules.

By testing the accuracy of predictions, finding which visual cues are used by respondents to classify deputies as left- or right-wing, determining whether they match the true characteristics of the two groups and/or whether they are representative, and looking at the distribution of errors, we can test the two families of behaviors. This is the aim of the rest of this paper.[4]

3. The data

In this section we describe the data used in this study. We first introduce the two surveys from which the data come. We then describe the characteristics of the deputies that were shown to respondents and assess the representativeness of the visual cues in their photographs.

3.1. The surveys

The data used in this study stem from two online surveys. The first was designed as part of a press event and was circulated in France prior to the 2012 presidential election, eliciting almost 70,000 responses. To gain more insight in the behavior of respondents, we designed a second survey specifically for the present study.

[4] One should bear in mind that while the existence of systematic biases can be interpreted as evidence against Bayesian behaviour, their absence is not necessarily evidence against the representativeness heuristic, because the heuristic may lead to moderate biases if the most representative characteristics are also the most likely, as Gennaioli and Shleifer (2010) show.

The first survey was launched prior to the 2012 election by an anonymous developer known by the nickname “wax-o”. It was initially developed as part of an event called “Hack the Press 2”, held in Paris in January 2012, where journalists and developers competed to design applications to collect data that could be used by journalists. We will refer to this survey as Survey 1. Shortly after it was launched, the survey went viral and attracted a considerable number of respondents. When we downloaded the collected data, which can be freely retrieved from the website, there were nearly 70,000 observations.

Respondents were presented with randomly selected photographs of actual French deputies and asked to classify them as left- or right-wing. The photographs used were those that appeared on the official website of the French Parliament and were of deputies in office during the presidential term of Nicolas Sarkozy (2007-2012). The photographs showed the face and upper chest. No personal information such as name and party affiliation was given about any deputy. Respondents therefore had to rely on partisan stereotypes to guess deputies’ political orientations. The site featured the photographs of the 554 deputies in office at the time of the survey, 333 right-wing and 221 left-wing, 447 male and 107 female. Photographs were randomly assigned to respondents. We therefore have a balanced number of respondents for each deputy. Respondents were asked to specify if they were themselves left- or right-wing.

Respondents were asked no other question. Moreover, they were immediately told if they had correctly guessed the deputy’s true party affiliation and were provided with the share of correct guesses they had so far made. Respondents could then choose to move to the next photograph. There was no time limit to classify politicians. The sole motivation to keep going was respondents’ interest in the game.

We complement Survey 1 with the results of an online survey designed specifically for the present study, Survey 2. Like in Survey 1, respondents were shown the anonymous photographs of French deputies, whom they were asked to classify as left- or right-wing by clicking on an icon.

Unlike in Survey 1, results were only reported to respondents after they had assessed a batch of ten photographs. Two key arguments support this feature of Survey 2. First, it was a way to collect more data, because the sole motivation to keep going was again respondents’ interest in the game. Reporting results after batches of ten photographs elicited interest and encouraged respondents to assess more photographs. Second, and more importantly, reporting results on a single screen after an assessment of ten photographs limited the ability of respondents to learn from

the photographs they had been shown, as this could have affected their behaviors and contaminated the stereotypes they held prior to the study with information acquired during the study. After each batch, respondents were invited to assess another batch of ten politicians. There was no time limit to answer the questions and classify deputies.

In addition to guessing political affiliations, respondents were asked to assess the deputies in terms of attractiveness, competence, and trustworthiness on a one-to-five scale.[5]

Before entering the survey page, respondents were asked to specify their gender and political orientation. Disclosing political orientation was not mandatory to participate in the survey, because in pre-testing some respondents had refused to do so and then left the survey. We therefore left respondents the option of not answering that question. Respondents to Survey 2 can thus be classified as left-wing, right-wing, or undeclared.

The survey was promoted through the newsletter of the authors' research center and through their personal networks, reaching an audience consisting predominantly of non-French respondents. It was launched on 25 February 2013. Data collection was stopped on 4 June 2013. We obtained six assessments per politician.

A key point if we want to interpret the responses of survey respondents as revealing their stereotypes is that they must indeed have interpreted their task as being to correctly guess deputies' affiliations, rather than to make a judgment about how left- or right-wing deputies are. In the two surveys that we use, we can safely assume that respondents did not let their own assessment of the policies supported by deputies interfere with the classification of individuals as left- or right-wing.

The first reason for this is that French political parties can in general be fairly unambiguously classified as left- or right-wing, because the French political system is bipolar. Godbout and Foucault (2013) show that the French National Assembly during the period 1958-2012 was one-dimensional, with a majority/opposition dimension corresponding to the left-right distinction. This was particularly true in 2012, because parties represented in the parliament had formed an unambiguous majority and opposition. Specifically, left-wing parties represented in the parliament at the time of the survey were the communist party ("Parti communiste"), the environmentalist party ("les Verts"), the socialist party ("Parti socialiste"), and a slightly left-of-center party ("Parti radical socialiste"). Some deputies though not formally members of those

[5] The exact wording was simply: "Is this person attractive?", "Is this person competent?", and "Is this person trustworthy?". The first points of the five-point scale were "Unattractive", "Incompetent", "Untrustworthy", while the fifth points were "Very attractive", "Very competent", and "Very trustworthy". For our data analysis, the replies were coded from 1 to 5, but those numbers did not appear in the survey. There was no option to abstain.

parties were declared as "various left" ("Divers gauche") and participated in the same group as members of the socialist party and the left-of-center party. Right-wing parties were the right-of-center party ("Nouveau centre") and the right-wing party ("Union pour un mouvement populaire"). Even unaffiliated candidates, of whom there were only six in 2012, have to declare whether they support the majority or the opposition and can therefore be classified too. There was no extreme right deputy. At the time of the survey, the ruling coalition consisted of the two right-wing parties and supported a right-wing government led by François Fillon, a member of "Union pour un mouvement populaire". The president, Nicolas Sarkozy, was a member of the same party.

Moreover, a way to check the interpretation of the task by survey respondents is to look at the share of correct identifications of well-known politicians. At the time of the survey, François Hollande, who was later elected president and appears in Survey 1, was the socialist party's official candidate for the presidential election. He could therefore be identified by virtually anyone. However, the socialist party, and François Hollande in particular, is often accused of not being left-wing enough by parties to its left. Respondents who classified François Hollande as right-wing could not possibly have made a mistake about his party affiliation, but probably signaled their disagreement with the stance of his party. However, François Hollande was correctly classified as left-wing by 99 percent of respondents, suggesting that very few respondents let their own assessment of his actual orientation interfere with their answers. We can therefore interpret the answers of respondents as genuine attempts at guessing deputies' affiliations.

3.2. Visual cues

In addition to the data collected in the two surveys, we coded objective characteristics of the candidates that were directly observable from their photographs. Specifically, we coded dummies capturing their gender and whether they were non-Caucasian, in line with the findings of McDermott (1997, 1998) and Samochowiec et al. (2010). Following Carney et al. (2008), we coded whether deputies are smiling.[6] Finally, we created a dummy equal to one if the deputy is wearing glasses, which is an easily observable cue.

We also coded gender-specific dummy variables. For male deputies, we created a dummy taking the value one if the deputy wears a red tie, a dummy taking the value one if he wears a tie that is neither blue nor red, and a dummy taking the value one if he wears no tie at all. The

[6] The smile dummy variable is set to one if the teeth appear when an upward curving of the corners of the mouth is observed and zero otherwise.

reference group consists of deputies wearing a blue tie.[7] We also created a dummy capturing whether the deputy wears a beard, and one capturing whether he wears a moustache.

For female deputies, we coded dummies capturing whether the deputy wears a suit, whether she wears jewelry, and three dummies coding hair-color and hair-length. Specifically, we coded a dummy set to one if the deputy's hair is blond, a dummy capturing whether her hair is short, and a dummy if her hair is long. To be classified as blond, the hair had to be entirely light-colored. Streaked hair did not count. To be classified as short, the hair had to let the ears be visible. Conversely, the deputy's hair was classified as long if it touched her shoulders.[8] The reference category for hair-length is medium length.

Finally, we created a dummy variable taking the value one if the deputy had held national office prior to the survey. National office is defined as being a minister, a party leader, or having chaired one of the two assemblies.[9]

Because the two surveys are based on the same set of photographs, we can match their results and the objective characteristics that we coded ourselves.

3.3. Deputies

In this section, we describe the actual characteristics of left- and right-wing deputies. We therefore describe the prototypes of left- and right-wing deputies. Figure 1 focuses on the objective characteristics of deputies. It reports the share of deputies of both groups displaying each characteristic. It shows that right-wing deputies are more often male, or non-Caucasian, and smile more often than left-wing deputies. Conversely, right-wing deputies wear glasses less often than left-wing deputies. If we look at gender-specific characteristics, we observe that male right-wing deputies wear red ties, ties that are neither red nor blue, no tie at all, a beard or a moustache less often than their left-wing counterparts. We observe that right-wing female deputies wear a suit less often than their left-wing counterparts. Finally, we observe almost no difference in the propensity to wear jewelry between left- and right-wing female deputies.

[7] In pre-tests, respondents cited tie colour as a key visual cue.
[8] We coded hair colour for female deputies only, because dyeing is uncommon among men. We did not code hair-length for men, because it displays almost no variation in our sample.
[9] We also created a dummy capturing whether the photograph's background was not neutral and computed the deputy's age. We found some evidence that older deputies were perceived as more right-wing in Survey 1. However, controlling for those variables never affected our results for other variables. To save on space, we therefore do not report regression results controlling for those variables. They are available on request.

Figure 1: Actual characteristics of deputies: Objective characteristics, as a percentage of the left-wing and right-wing populations

Fig. 1a: Whole sample

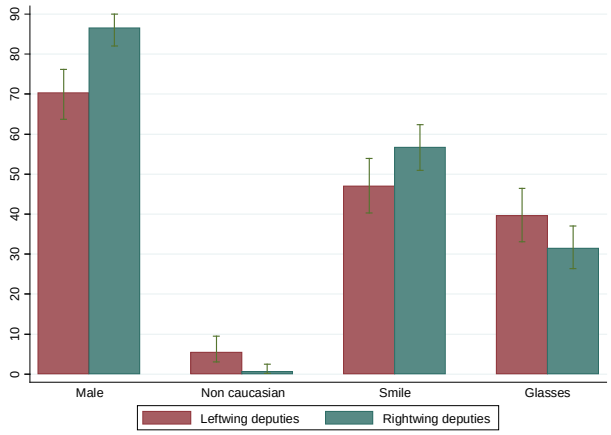


Fig. 1b: Male deputies

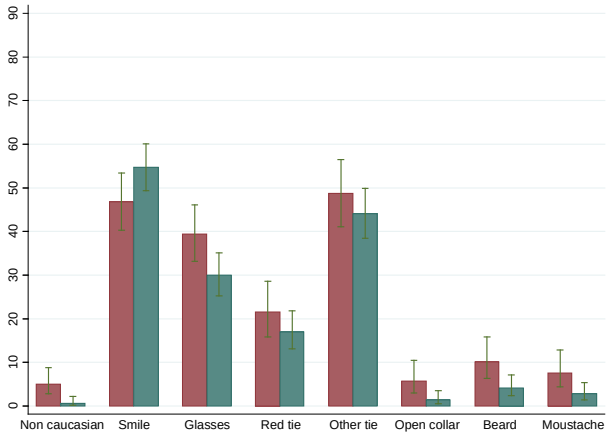
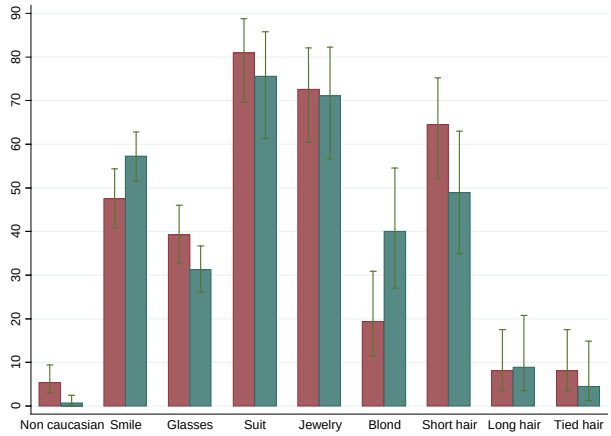


Fig. 1c: Female deputies



Source: Authors' calculations. Deputies who have held national office are dropped.
Confidence intervals are computed for a five-percent level of confidence.

*** insert Table 1a around here ***

*** insert Table 1b around here ***

A Chi-squared test reported in Column 3 of Table 1a confirms that the share of male deputies is different between the left and the right. The difference is statistically significant at the one-percent level. Chi-squared tests also confirm that the shares of non-Caucasian deputies differ

between the left and the right. The difference is significant at the one-percent level in the whole sample, at the ten-percent level in the sample of male deputies, and at the five-percent level in the sample of male deputies. The share of deputies who smile also significantly differs across the two groups. However, if the difference is significant at the five-percent level in the whole sample and at the one-percent level in the sample of male deputies, it is insignificant in the sample of female deputies. Finally, the shares of deputies who wear glasses significantly differ between left- and right-wing deputies in the whole sample as well as in both gender-specific sub-samples.

Figure 2: Actual characteristics of deputies: Average subjective scores

Fig. 2a: Whole sample

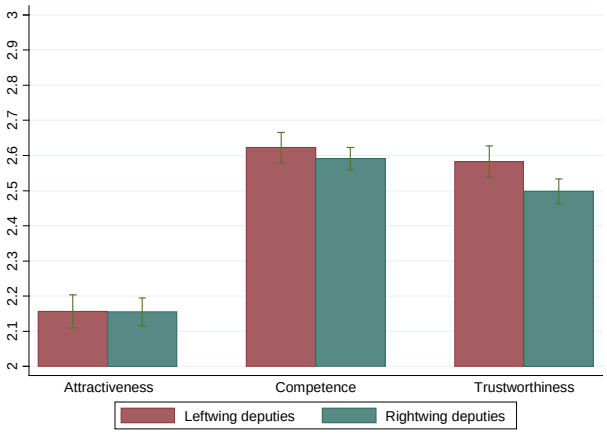


Fig. 2b: Male deputies

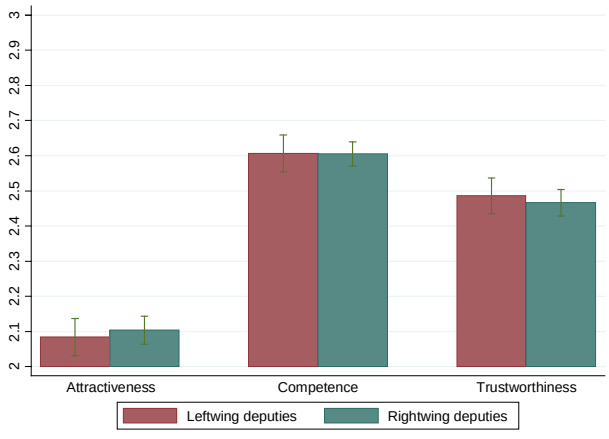
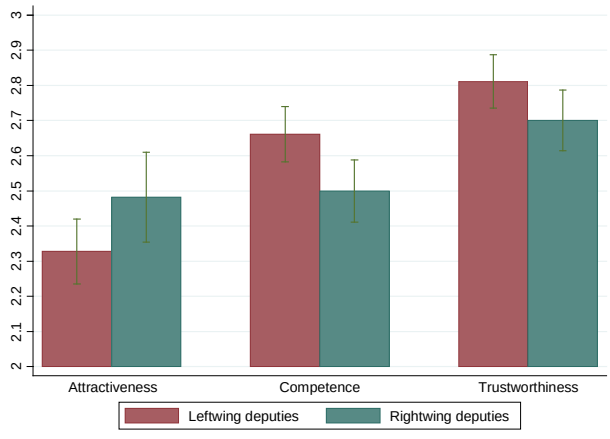


Fig. 2c: Female deputies



Source: Authors' calculations. Deputies who have held national office are dropped.
Confidence intervals are computed for a five-percent level of confidence.

If we turn to male-specific objective characteristics, Chi-squared tests reveal that the shares of deputies with a beard and with a moustache are different between the left and the right, and that the difference is significant at the five-percent level for both characteristics.

Chi-squared tests suggest that most female-specific characteristics do not significantly differ between left- and right-wing deputies. The shares of female deputies wearing suits, jewelry, short hair, long hair, or tied hair seem to be indistinguishable across the two groups. However, we find that the difference between the two groups is significant at the ten-percent level for the blond hair dummy.

Figure 2 reports the subjective evaluations of actual left- and right-wing deputies. There appears to be little difference in attractiveness, trustworthiness, and competence between the two groups of deputies. Table 1b complements the figure by reporting the means and the variances of subjective characteristics. For each deputy, the subjective characteristics correspond to their average assessments by respondents to Survey 2. They are therefore continuous variables, and their representativeness must be assessed by looking at their means and their variances, as Bordalo et al. (forthcoming) show. For each sub-sample, the characteristics are reported by order of decreasing difference between the means of left-wing and right-wing deputies.

Table 1b confirms that the means of subjective characteristics differ little between left-wing and right-wing deputies. The t-tests reported in Column 3 confirm that the mean evaluations of subjective characteristics are in general statistically indistinguishable between groups. The tests, however, signal two statistically significant differences. In the whole sample, left-wing deputies are perceived as slightly more trustworthy than right-wing deputies, and the difference is significant at the ten-percent level. In the sub-sample of female deputies, the difference between left- and right-wing deputies in average perceived competence is also significant at the ten-percent level. Female left-wing deputies are thus perceived as slightly more competent than their right-wing counterparts.

We complement Figures 1 and 2 by a series of logit regressions where deputies' actual orientations are regressed on their objective characteristics and on their average subjective evaluations across all respondents. This not only summarizes the characteristics of those deputies but provides a rational benchmark against which to weigh the behavior of survey respondents. If an econometrician was asked to guess deputies' affiliations from their official photographs, he or she would estimate those models. We restrict the sample to deputies who have not held national office, because stereotypes will be assessed on this sample in the following sections.

The results of those regressions are reported in Tables 2a to 2c. As the dependent variable is set to one if the deputy is right-wing, positive coefficients signal that a characteristic is associated with a larger probability of the deputy being right-wing.

*** insert Table 2a around here ***

*** insert Table 2b around here ***

*** insert Table 2c around here ***

Table 2a reports the results of the regressions run on the whole sample of deputies, both male and female. Chi-squared tests all reject the null hypothesis that coefficients are jointly zero. Specifically, the coefficient of the male deputy dummy is significantly positive, signaling that being male increases the probability of being right-wing, while the coefficients of the non-Caucasian and glasses dummies are negative, signaling characteristics that increase the probability of being left-wing.

Table 2b reports the results of the regressions run on the sub-sample of male deputies. It confirms that wearing a red tie, any other tie that is not blue, an open collar, a beard, a moustache, or glasses is associated with a lower probability of actually being right-wing. It confirms that a deputy who smiles is more likely to be right-wing. Table 2b also shows that no subjective characteristic is associated with deputies' orientations at standard levels of significance.

Finally, Table 2c focuses on the sample of female deputies. It shows that very few of the cues that we coded relate to the orientation of female deputies.[10] We find some evidence that blond deputies are more likely to be right-wing and that deputies wearing glasses are more likely to be left-wing, but those cues are only significant at the ten-percent level and in the parsimonious specification of Column 1. The coefficients become insignificant when subjective characteristics are added to the set of explanatory variables. We observe that the probability of being right-wing increases with attractiveness. The effect is significant at the five-percent level in both regressions where attractiveness is controlled for.

3.4. Representativeness of characteristics

To test the relevance of the representativeness heuristic, we must first assess the representativeness of the visual cues across the two groups of deputies. To do so, Table 1a reports

[10] Note that the non-Caucasian dummy does not appear in Table 1c, because there is no non-Caucasian right-wing deputy in the sample of female deputies.

the frequencies of visual cues separately for left- and right-wing deputies, in the whole sample, the sample of male deputies, and the sample of female deputies. The third column of Table 1a reports the likelihood ratio, i.e. the ratio of the frequency of right-wing deputies exhibiting a characteristic to the frequency of left-wing deputies exhibiting the same characteristic. For each sample, characteristics are reported by order of increasing likelihood ratio. As a result, appearing at the top are the characteristics that are the most representative of the left, while those at the bottom are the most representative of the right.

Table 1a reports frequencies and likelihood ratios computed on our sample. The top panel of Table reports those figures for the whole sample pooling male and female deputies. The two characteristics that appear as the most distinctive of left-wing deputies versus right-wing deputies are gender and being non-Caucasian.

The middle panel of Table 1a focuses on the sub-sample of male deputies. In that panel, the most representative feature of left-wing deputies is being non-Caucasian. Conversely, smiling is distinctive of right-wing male deputies.

Finally, the bottom panel of Table 1a reports frequencies and likelihood ratios for female deputies. It shows that being non-Caucasian and wearing glasses are the most representative features of the left. In fact, being non-Caucasian is perfectly diagnostic of being left-wing for female deputies, as there was no non-Caucasian right-wing female deputy in our sample.

As Bordalo et al. (forthcoming) argue, the representativeness of a continuous variable must be assessed by looking at both its mean and its variance. Recall that the only statistically significant differences in means between left- and right-wing deputies are observed for trustworthiness in the whole sample, and for competence in the female-only sub-sample. Left-wing deputies are perceived as more trustworthy in the whole sample, and left-wing female deputies are perceived as more competent than their right-wing counterparts in the female-only sample.

The right-hand panel of Table 1b compares the standard deviations of the subjective scores of left- and right-wing deputies. Again, most standard deviations are very close and statistically indistinguishable according to F-tests of equality of standard deviations. Competence is the only exception. In the whole sample and the sub-sample of male deputies, the standard deviations of the competence score are larger for left- than for right-wing deputies, and the differences are statistically significant at the ten-percent level.

Taken together those findings allow us to specify what the stereotypes of the two groups of deputies would look like according to the representativeness heuristic. Left-wing deputies in the whole sample are perceived as more trustworthy on average. The variances of trustworthiness in the two groups are however statistically indistinguishable. According to Bordalo et al. (forthcoming), the stereotype about left-wing deputies should lead them to be perceived as on average more trustworthy than they really are, and more trustworthy than right-wing deputies. The stereotype should also underestimate the variance of the trustworthiness of left-wing deputies. By the same token, the stereotype of left-wing female deputies should lead them to be described as more competent than they really are, and with less variance in competence. They should also be perceived as more competent than right-wing female deputies.

In the whole sample, average perceived competence is the same for the two groups of deputies, but the variance in competence is larger for left-wing deputies. The stereotype should therefore overemphasize extreme cases of competence for left-wing deputies. In other words, the competence of left-wing deputies should be perceived as more heterogeneous than it really is. The same conclusion applies to the competence of male deputies.

4. Respondents' accuracy

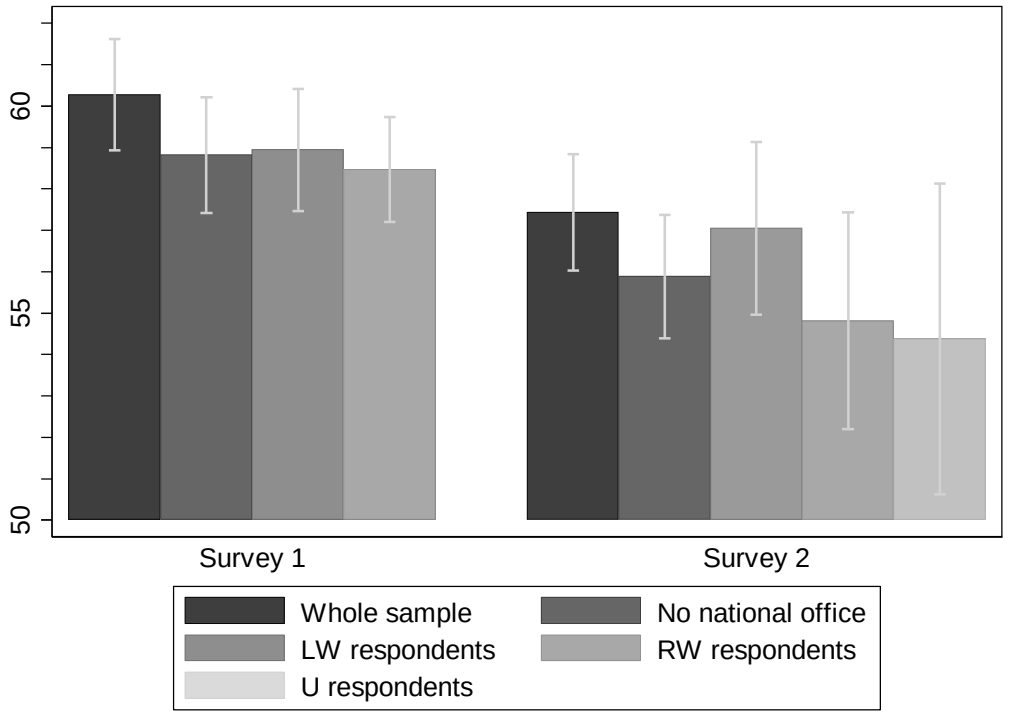
Samochowiec et al. (2010) and Olivola, et al. (2012) find that survey respondents could outperform chance when guessing the political orientation of anonymous politicians. In this sub-section, we check whether the finding holds in the two surveys that we use in this study.

A first insight into that question is provided by Figure 3, which reports the shares of correct guesses in Survey 1 and Survey 2. The left panel of Figure 3 focuses on Survey 1 while the right panel focuses on Survey 2. For each survey, the first bar reports the share of correct guesses for all deputies together, the second for deputies who have not held national office, while the following ones then distinguish respondents according to their declared orientation, still considering only deputies who have not held national office. Survey 2 allows us to measure the share of correct guesses for undeclared respondents, reported in the final bar.

Figure 3 shows that the shares of correct guesses are always larger than 50 percent in both surveys. The difference is moreover statistically significant at the five-percent level at least. The result holds even when deputies who have held national office are excluded from the sample.

In Survey 1, the shares of correct guesses of left- and right-wing respondents are very close and statistically indistinguishable. In Survey 2, we can distinguish undeclared respondents in addition to left- and right-wing respondents. Again, it appears that the shares of correct guesses of the three groups are statistically indistinguishable.

Figure 3: Share of correct guesses in the two surveys



Confidence intervals are computed for a five-percent level of confidence.

Further evidence is provided by a Chi-squared test testing the independence of actual and guessed political orientations. We performed that test on the dataset resulting from our own survey, because it provides individual observations as opposed to aggregate classifications. When computed on the whole sample, the Chi-squared statistic amounts to 64.56, which allows us to comfortably reject the null hypothesis of independence at the one-percent level of significance. To make sure that the result was not driven by identifiable well-known politicians, we performed the same test on a sample restricted to deputies who had not held national office. The Chi-squared statistic for that sample is lower but still equals 37.04, which still allows us to reject the null hypothesis of independence at the one-percent level of significance. Those tests confirm that survey respondents

did not randomly classify the photographs of deputies. On the contrary, their guesses at least partly reflected the actual orientations of deputies appearing in the photographs.

We performed a series of t-tests to see if categorization errors were independent of the political affiliation of respondents, and a series of Chi-squared tests to compare average accuracy across groups. None of those tests allowed us to reject the null hypothesis of no differences across groups of respondents. Groups of respondents therefore appear equally able to categorize deputies. These results are in line with those of Rule and Ambady (2010), who could find no difference in accuracy across Republican and Democrat perceivers asked to categorize politicians.

We also tested whether respondents were more or less accurate when categorizing deputies from their own side of the political spectrum as compared with deputies from the other side, because Samochowiec et al. (2010) observed that perceivers were more accurate when categorizing politicians of their out-group. We did this by performing for each group of respondents in the second survey a t-test on the shares of correct guesses for left- and right-wing deputies. For no group could we find a significant difference in the means of correct guesses across left- and right-wing deputies. We complemented those tests by performing, for each group of respondents, Chi-squared tests of independence of correct guesses with the true political preferences of deputies. Again, we could reject the null hypothesis of independence for no group of respondents. In contrast with the results of Samochowiec et al. (2010), these findings suggest that respondents were not more accurate at classifying deputies whose attitudes were opposite to their own position, i.e. out-group deputies, than deputies whose attitudes they shared, i.e. in-group deputies.

*** insert Table 3 around here ***

We provide additional evidence that guessed orientations were not random by testing whether a deputy is indeed more likely to be right-wing when he/she is classified as right-wing. To do so, we regress deputies' actual orientations on guessed orientations. We estimate two series of binary logit models where the dependent variable is a dummy variable capturing the actual political orientation of deputies. In the first series of regressions, the key explanatory variable is the share of respondents who classified the deputy as right-wing in Survey 1. In the second series of regressions, the key explanatory variable is respondents' guesses in Survey 2.

In the first model, observations are individual deputies, and the explanatory variable is the share of respondents who perceived a deputy as right-wing, which we refer to as the deputy's right-wing score.

$$Prob(Actually\ right\text{-}wing_i) = F(a_0 + a_1\ RWscore_i + u_i) \quad (1a)$$

Where $Actually\ right\text{-}wing_i$ is a dummy variable taking the value one if deputy i is indeed right-wing, and $RWscore_i$ is the share of respondents who classified deputy i as right-wing in Survey 1.

In the second model, observations are individual answers by survey respondents in Survey 2. The dependent variable is the actual orientation of the deputy assessed by a respondent:

$$Prob(Actually\ right\text{-}wing_{ij}) = F(a_0 + a_1\ Assumed\ right\text{-}wing_{ij} + u_{ij}) \quad (1b)$$

Where $Actually\ right\text{-}wing_{ij}$ is a dummy variable taking the value one if the actual orientation of deputy i assessed by respondent j is indeed right-wing, $Assumed\ right\text{-}wing_{ij}$ is a dummy variable taking the value one if respondent j classified deputy i as right-wing, and u_{ij} is a dyad-specific error term. The only difference with Model 1a is that observations now pertain to deputy-respondent dyads as opposed to individual deputies.

The estimation of both models is performed using cluster-robust standard errors, with clusters defined over deputies, since each deputy was classified by several respondents.

Columns 1 to 4 of Table 3 report the outcome of estimating a model where the key explanatory variable is the deputy's right-wing score. In Column 1, the model is estimated on the whole sample. We drop deputies who have held national office in Column 2. In Column 3, we look at the deputy's right-wing score among left-wing respondents. In Column 4, we look at the deputy's right-wing score among right-wing respondents. In all models, the coefficient of the right-wing score is positive and significant at the one-percent level. Again, those results mean that the more survey respondents have classified a deputy's photograph as right-wing, the more likely that deputy is to be right-wing. Quantitatively, the estimated marginal impact implies that a one-percentage point increase in a deputy's share of right-wing guesses results in a 0.99 to 1.15 percentage-point increase in the probability of the deputy actually being right-wing. The effect is unsurprisingly

larger when deputies having held national office are included in the sample, which confirms that respondents were better able to recognize them.

The outcome of estimating Model 1b is reported in Columns 5 to 9 of Table 3. Column 5 reports the result of the model estimated on the whole sample of deputies. We observe that the coefficient of the guessed right-wing deputy dummy is positive and significant at the one-percent level. The dummy remains positive and significant at the one-percent level in Column 6, where deputies having held national office are dropped from the sample. In Columns 7 to 9, the coefficient appears significantly positive for left-wing, right-wing, and undeclared respondents, though only at the five-percent level for right-wing respondents and at the ten-percent level for undeclared respondents.

Again, the positive sign of the guessed right-wing dummy means that a deputy's actual orientation is more likely to be right-wing if respondents classified him/her as right-wing. The probability of a deputy being right-wing is 7.65 to 12.7 percentage points larger if a respondent has classified him/her as right-wing.

The results reported in Table 3 imply that the blind guesses of survey respondents strongly correlate with deputies' true orientations. In other words, respondents are indeed able to outperform random guesses when trying to find deputies' true orientations. This finding confirms those of Samochowiec et al. (2010), Olivola and Todorov (2010), Rule and Ambady (2010), and Olivola et al. (2012), and extend them to an additional country.

5. What stereotypes look like

If respondents can categorize deputies with above-chance accuracy, they must be able to use visual cues that appear on photographs. In this section, we try to identify those cues. In other words, we look at the content of the stereotypes that respondents use to categorize deputies. To do so, we look at the characteristics of deputies that correlate with being categorized as left- or right-wing. We then compare those characteristics with those that a rational viewer would use, and with those that the representativeness heuristic predicts.

5.1. Empirical strategy

Like in the previous section, we estimate two series of models, to make the most of the two surveys. The first model uses data from Survey 1 as its dependent variable. The dependent variable

is deputy *i*'s right-wing score, i.e. the share of respondents who classified that deputy as right-wing. Each deputy's score was matched to the deputy's objective characteristics as coded by us, and to his/her average score on the three subjective questions of Survey 2. The model's specification is therefore the following:

$$RWscore_i = a_0 + A_1\ MP_i + A_2 \overline{MP}_i + \varepsilon_i \quad (2a)$$

where $RWscore_i$ is deputy *i*'s right-wing score. MP_i is the vector of objective characteristics of deputy *i* that we coded ourselves. $\overline{MP}_i$ is a vector containing the evaluation of deputy *i*'s subjective characteristics averaged over all the respondents who assessed him/her in Survey 2. a_0 is a scalar, A_1 and A_2 are vectors of coefficient, and ε_i is the deputy-specific error term.

The dependent variable of Model 2b is not bounded, because no right-wing score in our sample reaches the zero or one bounds. The model was therefore estimated using OLS.

In the second model, each observation is a deputy-respondent dyad contained in Survey 2. The dependent variable is a dummy variable taking the value one if deputy *i* was classified as right-wing by respondent *j*. The probability that that variable assumes the value one is related to deputy *i*'s objective characteristics that we coded ourselves, to the deputy's subjective characteristics as assessed by all respondents in Survey 2 except respondent *j*, and to respondent *j*'s characteristics. The specification of the model thus reads:

$$\Pr\left(Assumed\ right-wing_{ij} = 1\right) = \mathrm{F}\left(b_0 + B_1 MP_i + B_2\ \overline{MP}_{ij} + B_3\ R_j + u_{ij}\right) \quad (2b)$$

where MP_i is defined as in equation (2a). $\overline{MP}_{ij}$ is a dyad-specific vector containing the evaluation of deputy *i*'s subjective characteristics averaged over all the respondents who assessed him/her in Survey 2 except respondent *j*. R_j the vector of characteristics of respondent *j*. A_1, A_2, and A_3 are vectors of coefficients, and u_{ij} is a dyad-specific error term.

We average the assessments of deputy *i*'s characteristics while excluding respondent *j*'s assessment so as to minimize the risk of reverse causality. For instance, one concern might be that left- or right-wing respondents assume that deputies that they perceive as politically closer look

more attractive, competent, or trustworthy, due to a halo effect (Herrmann and Shikano, 2016).[11] Also, a respondent who, for whatever reason, has categorized a deputy as left-wing could apply to the deputy the stereotypical characteristics of left-wing deputies. As we are interested in how respondents use visual cues to categorize deputies, and not how they use the deputy's guessed affiliation to infer his/her characteristics, we want to rule out that possibility. We therefore use averages, and as respondent *j*'s assessment may influence the average assessment, we exclude that assessment from the average. As a result, the respondents who assessed the trustworthiness, competence, and attractiveness of deputies are not those whose categorization is the dependent variable of Equation 2a.[12]

The model is estimated using a logit estimate. We cluster standard errors by deputy to take into account the fact that the same photograph was shown to several respondents.

The advantage of Model 2a is that it aggregates the results of hundreds of attempts at guessing each deputy's political affiliation. It may therefore be considered as a better measure of the deputy's general left- or right-wing appearance. Moreover, it uses two independent datasets. The evaluations of attractiveness, competence, and trustworthiness are those that were obtained in our own survey, which reduces the risk of reverse causality. Model 2b uses more observations. Moreover, it uses individual answers, and can control for the characteristics of respondents.

As each model has its advantages, we estimate them both in turn. This allows us to conduct a systematic robustness check of the measured impact of a given characteristic of a deputy on his/her propensity to be perceived as left- or right-wing. Most importantly of all, the respondents in Model 2a and Model 2b are two different groups of people. If the results of the two models concur, this will be evidence that the stereotypes that we unveil are not limited to a particular group of survey respondents.

To make sure that the results were not driven by the more prominent deputies, we dropped all observations pertaining to deputies having held national office. In the remainder of the paper what we refer to as the whole sample is the sample of both male and female deputies from which those deputies were dropped.

[11] Section 5.4 below directly assesses the presence of a halo effect.

[12] Using the score granted by respondent *j* to deputy *i* instead of deputy *i*'s average score, however, essentially leads to the same results.

5.2. Baseline results

Table 4a reports the results of the estimation of Model 2a on data from Survey 1 on the whole sample of French deputies, specifically both left- and right-wing and both male and female, from which deputies who have held national office were dropped. The sample thus consists of 491 deputies, none of whom have held national office.

We include explanatory variables by group. The first group consists of objective observable characteristics of the deputy, namely gender, being non-Caucasian, whether the deputy smiles, and whether he/she wears glasses. The second group features the deputy's subjective characteristics, namely attractiveness, competence, and trustworthiness. The third group is made up of the respondent's self-declared characteristics: his/her political orientation, or whether he/she refused to provide it, and his/her gender. We then consider all the independent variables together.

F-tests reject the null hypothesis that all coefficients are jointly zero in all columns. The first column of Table 4a reports the outcome of estimating Model 2a while controlling for objective variables. The adjusted R-squared shows that the model explains more than 20 percent of the variance of right-wing scores. The regression shows that being male increases the right-wing score by 15.3 percentage points. Being non-Caucasian decreases it by 30 percentage points, while wearing glasses reduces the score by 6.96 percentage points. All three coefficients are significant at the one-percent level of statistical significance. In the same regression, the dummy variable capturing whether the deputy smiles in his/her photograph exhibits a positive sign significant at the five-percent level, suggesting that smiling politicians were perceived as more right-wing. However, the effect is quantitatively small, as it amounts only to 3.36 percentage points of the right-wing score.

The second column of Table 4a reports the outcome of estimating Model 2b when the set of explanatory variables consists of deputies' subjective characteristics. While the adjusted R-squared is only 7.2 percent, two subjective characteristics appear significantly. First, deputies perceived as more competent are also perceived as more likely to be right-wing. The effect is significant at the one-percent level. A one-point increase on the competence scale results in a 10.5 percentage points increase in the right-wing score. Table 4a also shows that deputies perceived as more trustworthy are less likely to be perceived as right-wing. The effect is also significant at the one-percent level. It implies that a one-point increase in the trustworthiness score decreases the right-wing score by 15.4 percentage points. Conversely, we find that the impact of looking attractive does not on average correlate with deputies' perceived political orientations. When the two sets of variables are

considered jointly, in Column 3 of Table 4a, the results reported in the first two columns do not change qualitatively, and the coefficients change very little.

The results reported in Table 4b complement those reported in Table 4a by reporting the result of estimating Model 2b using data from Survey 2. The sets of explanatory variables are the same in both tables. The key differences are that each observation is now a deputy-respondent dyad, and that the dependent variable is a dummy variable taking the value one if the respondent has classified a deputy as right-wing in Survey 1. Moreover, we can control for respondents' characteristics. All regressions were performed on a sample of 2,983 assessments. The reported coefficients are marginal effects and therefore lend themselves to a quantitative interpretation.

*** insert Table 4a around here ***

*** insert Table 4b around here ***

Although the pseudo-R-squared is low, the Chi-squared test rejects by a large margin the null hypothesis that all coefficients are zero. The only exception is Estimation 3, where the set of independent variables is restricted to the respondent's characteristics.

The first column of Table 4b reports the outcome of estimating Model 2a while controlling for the first set of explanatory variables. The male deputy dummy exhibits a positive sign that is significant at the one-percent level. The probability of being classified as right-wing was therefore 15.2 percentage points larger for male deputies than for female ones. Conversely, the non-Caucasian dummy exhibits a negative sign significant at the five-percent level. Non-Caucasian deputies therefore had a 16.7 percentage point lower probability of being classified as right-wing. We also observe that wearing glasses decreased the probability of being classified as right-wing by 12.2 percentage points, and that the effect is statistically significant at the one-percent level. Finally, smiling on the photograph had no statistically significant effect on the probability of being classified as left- or right-wing.

The second column of Table 4b reports the outcome of estimating Model 2b when the set of explanatory variables is restricted to the average assessment of the deputy's subjective characteristics. We observe that deputies scoring one more point on the average competence score had a 6.73 percentage points larger probability of being classified as right-wing. This effect is moreover significant at the five-percent level. Conversely, deputies scoring one more point on the

trustworthiness score were 12.5 percentage points less likely to be classified as right-wing, and the effect was significant well beyond the one-percent level. The marginal effect of being perceived as attractive is statistically insignificant in Regression 2, suggesting that respondents on average did not take attractiveness into account when assessing deputies' orientations.

The third column controls for respondents' self-declared characteristics. None of those characteristics exhibits a significant effect. Respondents' gender and political orientation therefore seem unrelated to their propensity to classify deputies on the political spectrum.

The fourth column of Table 4b controls simultaneously for the first two sets of explanatory variables. Doing so in general affects neither the statistical significance nor the magnitude of estimated coefficients. This regression therefore confirms that being male increases a deputy's probability of being classified as right-wing. Conversely, being non-Caucasian, wearing glasses, and being perceived as more trustworthy increases the probability of being classified as left-wing. In Regression 4, the coefficient of being perceived as competent is insignificant, in contrast to the previous regression. We obtain the same results in the fifth column, when controlling in addition for respondents' characteristics. In this column, respondents' characteristics remain statistically insignificant at accepted levels of significance.

The results of Models 2a and 2b therefore concur. Being female, non-Caucasian, wearing glasses, and looking trustworthy are features that survey respondents tend to associate with left-wing deputies. Being male and looking competent are perceived as right-wing features.

The results of Tables 4a and 4b are based on regressions using a sample pooling of male and female deputies. They therefore rely on rather general characteristics. Splitting the sample across genders allows us to code gender-specific characteristics. This is why we now turn to gender-specific regressions in Tables 5a, 5b, 6b, and 6b.

*** insert Table 5a around here ***

*** insert Table 5b around here ***

Tables 5a and 5b restrict the sample to male deputies. We can thus control for tie color (or absence of a tie), and for facial hair, both of which are male-specific features.

Table 5a reports the results of estimating Model 2a on the sub-sample of male deputies. It by and large confirms the qualitative impact of the variables that were significant in previous regressions. In particular, being non-Caucasian exhibits a negative coefficient that is statistically significant at the one-percent level (Columns 1 and 3). Similarly, wearing glasses bears a negative coefficient that is statistically significant at the one-percent level (Columns 1 and 3). The smile dummy bears a positive sign significant at the ten-percent level. We also find again that being perceived as competent increases the probability of being classified as right-wing (Columns 2 and 3). The effect is significant at the one-percent level. Conversely, trustworthiness remains correlated with a lower probability of being classified as right-wing, and the effect is significant at the one-percent level (Columns 2 and 3).

However, the key new information appearing in Table 5a is provided by the coefficients of male-specific variables that now appear in Columns 1 and 3. Specifically, wearing a red tie reduces the right-wing score by nearly 12 percentage points, and the effect is significant at the one-percent level. We also observe that wearing a tie that is neither blue nor red also reduces the score. The effect exceeds 6 percentage points, and is significant at the one-percent level. The effect of having an open collar is always significant at the one-percent level and exhibits a negative coefficient with a magnitude revolving around 25 percentage points.

Variables coding facial hair also affect the perception of deputies' political orientations. In particular, wearing a beard has a significantly negative effect on a deputy's right-wing score. A beard reduces the score by approximately 30 percentage points, and the effect is significant at the one-percent level. A moustache reduces the score by around 23 percentage points, and the effect is also significant at the one-percent level.

Table 5b reports the outcome of estimating Model 2b on the sub-sample of male deputies. It broadly confirms the findings of Table 5a. In more detail, we observe that wearing glasses (Columns 1, 4, and 5) and being perceived as trustworthy (Columns 2, 4, and 5) correlate with being categorized as left-wing. Conversely, being perceived as competent correlates with being categorized as right-wing, although the coefficient is only significant in Column 1 and at the ten-percent level. Being non-Caucasian and smiling have a significant effect in none of the regressions.

If we focus on male-specific characteristics (Columns 1, 4, and 5), we observe that the red tie dummy exhibits a negative coefficient, implying that male deputies wearing a red tie are around 12.8 percentage points less likely to be classified as right-wing than deputies wearing a blue tie. The

effect is always significant at the one-percent level. We also observe that wearing a tie that is neither blue nor red exhibits a negative coefficient. The effect is smaller (around 4.6 percentage points) and significant at the ten-percent level. The open collar dummy exhibits a negative sign significant at the one-percent level, implying that not wearing a tie at all increases the probability of being classified as left-wing. The effect exceeds 33 percentage points.

Variables coding facial hair also affect the perception of deputies' political orientations. In particular, wearing a beard has a significantly negative effect on the probability of being classified as right-wing. The effect is significant at the one-percent level, and its magnitude is approximately 18 percentage points. Wearing a moustache also reduces the probability of being classified as right-wing. The effect is always significant at the one-percent level. Wearing a moustache reduces the probability of being classified as right-wing by around 23 percentage points.

All in all, the results of Tables 5a and 5b reveal stereotypes associated with male deputies. Wearing a red tie, no tie, a beard, or a moustache is perceived as evidence that the deputy is left-wing. Conversely, wearing a blue tie and being clean-shaven are perceived as right-wing characteristics. Like in the set of regressions pooling deputies of both genders, we observe that perceived competence is associated with a right-wing orientation, and perceived trustworthiness with a left-wing orientation. Attractiveness seems insignificant in the sample of male deputies.

*** insert Table 6a around here ***

*** insert Table 6b around here ***

We now turn to the sample of female deputies to determine female stereotypes. The results are reported in Tables 6a and 6b. Table 6a reports the results of estimating Model 2a on the sub-sample of female deputies, measuring the perception of a deputy's orientation by her right-wing score in Survey 1. Although the sample size shrinks due to the low share of women in the French parliament, the model highlights some of the variables that were significant in previous regressions.

The results confirm the robust effect of being non-Caucasian and wearing glasses. Both are significant beyond the one-percent level in all the regressions where they appear (Columns 1 and 3). We find some evidence that being attractive raises the right-wing score while being perceived as competent decreases it, but those variables are only significant in Regression 2. The coefficient of attractiveness is only significant at the ten-percent level, and that of competence at the five-percent level. We also find evidence that being perceived as trustworthy decreases the right-wing score in

Regression 3, which includes both objective and subjective characteristics. In that regression the coefficient of trustworthiness is significant at the five-percent level.

If we turn to female-specific characteristics, we find that the suit dummy variable bears a positive sign. However that sign is only significant in the regression that controls both objective and subjective characteristics, and only at the ten-percent level. Wearing a suit increases the deputy's right-wing score by 6.9 percentage points. We also observe that the jewelry dummy exhibits a positive sign, but only in the first regression and only at the ten-percent level. Wearing jewelry increases the right-wing score by nearly 7 percentage points.

We find no evidence of a significant effect from hair length or style (Columns 1 and 3). The short hair, long hair, and tied hair dummies are all statistically insignificant at conventional levels of significance. However, we find robust evidence that being blond increases the right-wing score. The coefficient of the blond dummy is significant at the one-percent level in both Regression (1) and (3). Its point estimate suggests that being blond increases the right-wing score by 13 to 14 percentage points.

Table 6b reports the outcome of estimating Model 2b on the sub-sample of female deputies. The results of Model 2b pertaining to female-specific variables both lend additional support to and qualify the results of Model 2a. In particular, being non-Caucasian leads to a lower probability of being classified as right-wing, and the coefficient is significant at the five-percent level (Columns 1, 4, 5). Wearing glasses also bears a negative coefficient that is statistically significant at the five- or one-percent level (Columns 1, 4, 5). Like in previous regressions, the smile dummy is never significant at standard levels of significance (Columns 1, 4, 5). And as in the tables devoted to male politicians, we find that attractiveness is not statistically correlated with deputies' guessed orientations (Columns 2, 4, 5).

In the regression controlling only for deputies' subjective characteristics, we find that the competence dummy exhibits a negative sign significant at the ten-percent level, suggesting that female deputies perceived as more competent are more likely to be categorized as left-wing (Column 2), which contrasts with results obtained for the whole sample and for male politicians, but is in line with Table 6a. However, the coefficient turns insignificant in subsequent regressions (Columns 4, 5). Trustworthiness is positively related to the probability of being categorized as right-wing (Columns 2, 4, 5), which contrasts not only with results obtained for the whole sample and for male deputies, but also with regressions on female deputies using data from Survey 1, but is

significant only in two regressions (Columns 2 and 5) and only at the ten-percent level. Again, we can see no statistically significant effect from respondents' characteristics (Columns 3 and 5).

We can now turn to female-specific characteristics. Firstly, we observe that hair length, hair color, and wearing jewelry exhibit insignificant coefficients in all regressions (Columns 1, 4, 5). Two dummies however seem to be taken into account by survey respondents. The suit dummy bears a positive coefficient (Columns 1, 4, 5). It is significant at the five-percent level in Column 1 and at the ten-percent level in Column 4. This provides additional evidence that wearing a suit increases respondents' propensity to classify a female deputy as right-wing. Specifically, wearing a suit increases that probability by 9.7 to 12 percent. Having tied hair also bears a positive coefficient significant beyond the ten-percent level (Columns 1, 4, 5). It increases the propensity to be categorized as right-wing by 21 to 23 percent.

The results provided in this section provide a consistent description of partisan stereotypes. Basically, being male, Caucasian, and looking competent are features that survey respondents associate with the right. Conversely, being female, non-Caucasian, wearing glasses, and looking trustworthy are features that survey respondents associate with the left.

For the sub-group of male deputies, wearing a blue tie, and being clean-shaven are associated with the right. In contrast, wearing a red tie, a tie of any color but red or blue, or no tie at all, and wearing a beard or a moustache are features that survey respondents associate with the left. The results are remarkably stable across the two surveys.

For the sub-group of female deputies, we find evidence that wearing a suit, wearing jewelry, having tied hair, or being blond leads survey respondents to classify deputies as right-wing. However, the results seem less stable across the two surveys. The signs of characteristics are in general the same in the two surveys, but their significances sometimes differ. Having tied hair only appears significant in Survey 2. Conversely, being blond only appears significant in Survey 1. We find some evidence that female deputies perceived as competent are associated with the left, which stands in contrast with the results for the whole sample and for the sample of male deputies, but the evidence is overall fragile. The evidence for trustworthiness differs across the two surveys, and is also fragile.

5.3. Do stereotypes look like deputies?

The rational model of behavior implies that agents use a correct and complete distribution of group characteristics to categorize deputies. The representativeness heuristic implies that stereotypes can exaggerate differences between groups, but that they contain a kernel of truth, insofar as they amplify differences that exist between groups.

Section 5.2 above shows that survey respondents systematically related certain characteristics to being left- or right-wing, and therefore describes the stereotypes that they have in mind. Section 3.3 describes prototypical deputies. We now compare stereotypes and prototypes as a test of the accuracy of beliefs. If respondents used cues in a way that was completely at odds with the true characteristics of left- and right-wing deputies, we would have to reject both the rational model and the representativeness heuristic.

Regressions on the whole sample of deputies show that respondents associate being female, non-Caucasian, and looking trustworthy with the left. Conversely, they associate being male, Caucasian, and looking competent with the right. This is in line with the descriptive statistics of left- and right-wing politicians reported in Table A1 in the appendix. They show that left-wing deputies are indeed more often female and non-Caucasian and wear glasses more often than right-wing deputies. Those differences are statistically significant.

The coefficients of objective characteristics reported in Tables 4a and 4b, where the dependent variable is deputies' guessed orientations in Survey 2, are in general in line with those obtained in regressions where the dependent variable was deputies' true orientations and reported in Table 1a. An exception is the smile dummy, which is positive and significant at the one-percent level in Table 1a, indicating that right-wing deputies indeed smile more, while it is insignificant in Table 4b, suggesting that respondents did not take it into account in Survey 2. However, the smile dummy was also significantly positive in Table 4a where the dependent variable was deputies' right-wing scores in Survey 1.

As regards subjective characteristics, we find that trustworthiness bears a positive and significant sign in Tables 4a and 4b, while Table 2a reports weak evidence that trustworthiness actually correlates with being left-wing. This suggests that the respondents' association of looking trustworthy with the left is somewhat founded. The finding that perceived competence correlates with respondents categorizing a deputy as right-wing in Tables 4a and 4b is however at odds with the finding that perceived competence does not significantly correlate with actual orientation, as

shown in Table 1a. We remark however that competence loses significance in Table 5b when objective characteristics are controlled for.

We moreover ran a series of t-tests to test the significance of the difference between the coefficients of Table 4b and those of Table 1a.[13] Focusing on the most complete models, reported in Column 4 of Table 4b and Column 3 of Table 1a, we ran a t-test for each explanatory variable to test whether its coefficient was the same in the two regressions. None of those tests could reject the null hypothesis that coefficients were equal. This suggests that, overall, the reaction of respondents to visual cues is in line with actual characteristics of deputies.

When categorizing male deputies, respondents associated wearing a red tie, a tie of any color but red or blue or no tie at all, and wearing a beard or a moustache with the left. Again, the stereotypes used by respondents seem in line with the true characteristics of deputies. Table 1a shows that left-wing deputies do indeed wear red ties, ties of any color but red or blue or no tie at all, and a beard or a moustache significantly more often than right-wing deputies.

Moreover, the signs of the coefficients of objective characteristics in Table 5a and 5b are in general the same as in Table 2b. The smile dummy is an exception, as it is significant in Table 5b but not in Table 1b.

The way in which subjective characteristics are used contrasts with the true differences in the assessment of the subjective characteristics of deputies of the two groups. Table 2b shows that average subjective characteristics do not correlate with the true orientations of deputies. Yet, Tables 5a and 5b suggest that respondents in both Survey 1 and 2 associate competent-looking deputies with the right and trustworthy-looking deputies with the left.

This suggests that respondents used the correct objective visual cues to categorize deputies, but over-reacted when associating subjective perceptions with political orientation. A series of t-tests of the significance of the difference between the coefficients of Table 5b and those of Table 2b qualify that remark. We could never reject the null hypothesis that coefficients were equal in the two sets of regressions.

[13] Note that we have to focus on Model 2b here, because its dependent variable is discrete, like the dependent variable in Tables 2a, 2b, and 2c, while Model 2a's dependent variable is the deputy's right-wing score, which is continuous. The coefficients of Model 2a cannot therefore be compared with those of regressions where the dependent variable is deputies' true orientations.

When categorizing female deputies, respondents associated being non-Caucasian, wearing a suit, wearing jewelry, having tied hair, or being blond with the right. We found weak evidence that they associated perceived competence with the left and attractiveness with the right, but the evidence was overall fragile, as was the evidence for trustworthiness. Table 2a shows that stereotypes partly follow actual differences between the two groups, as female left-wing deputies are indeed significantly more often non-Caucasian, less often blond than their right-wing counterparts, and wear glasses more often. It therefore seems that survey respondents correctly picked up being non-Caucasian and wearing glasses as left-wing cues, and blondness and attractiveness as right-wing cues. However, the admittedly fragile evidence that respondents react to suits, jewelry, and tied hair seems at odds with the actual characteristics of the two groups of deputies.

When we ran t-tests to test the significance of the difference between the coefficients of Table 6b and those of Table 2c, we could find no significant difference in the coefficients of the two tables. This may be due to the fact that the sample size is small, and that coefficients are therefore estimated with large standard errors.

The upshot of this sub-section is that the direction in which respondents use objective visual cues is generally in line with the actual characteristics of the two groups of deputies. This is particularly true for the whole sample and for the sample of male deputies. This finding suggests that stereotypes overall look like deputies. Stereotypes therefore contain a kernel of truth. The finding rules out behaviors that would be completely irrational, but does not allow us to discriminate between Bayesian behavior and the representativeness heuristic.

Similarly, respondents react in the right direction to the subjective characteristics of deputies, although those characteristics differ little across the two groups. Reacting in the correct direction to a subjective cue is consistent with both rational behavior and the representativeness heuristic. Finding that respondents over-react to minor differences suggests the representativeness heuristic is at work, as it results in stereotypes that amplify actual differences between groups.

The results for the sample of female deputies are more mixed. Respondents correctly picked up being non-Caucasian and wearing glasses as left-wing cues, and blondness and attractiveness as right-wing cues, but reacted to suits, jewelry, and tied hair in a way that is at odds with the actual characteristics of deputies. When interpreting these findings, one should bear in mind that female deputies account for less than a fifth of deputies, and their numbers were even fewer in previous

parliaments. The scarcity of female deputies may explain why stereotypes of female deputies are partly disconnected from the prototypical female deputy.

5.4. The role of representativeness

If survey respondents use the representativeness heuristic described by Kahneman and Tversky (1972, 1973) and Gennaioli and Shleifer (2010), they will react more to more representative characteristics. In this section, we compare the marginal effects of objective visual cues against their likelihood ratios.[14] Specifically, we plot the marginal effects estimated in Section 5.2 against the likelihood ratios of the two characteristics reported in Table 1a. We do so for both Model 2a, estimated with data from Survey 1, and Model 2b, estimated with data from Survey 2. We first look at the outcome of regressions on the whole sample, then on male deputies, and finally on female deputies. The results are reported in Figure 4.

The top panel of Figure 4 plots the marginal effects of gender-neutral characteristics estimated in Models 2a and 2b on the whole sample of deputies except those who have held national office, against the likelihood ratio computed so as to increase when a characteristic is more representative of the right. Both figures show that the marginal effect of a characteristic is larger, the larger the likelihood ratio. Being male is the most distinctive characteristic of right-wing deputies, with a likelihood ratio of 1.23, and exhibits the largest marginal effect in both models. Conversely, being non-Caucasian is the most distinctive characteristic of the left, with a likelihood ratio of 0.13, and exhibits the most negative marginal effect on the right-wing score and the propensity to categorize a deputy as right-wing. More generally, the relationship between a characteristic and its likelihood ratio is strictly increasing. The scatterplots therefore show that the most representative characteristics elicit the largest reactions from respondents.

The middle panel of Figure 4 plots the marginal effect of deputies' characteristics estimated in regressions restricted to male deputies, which allows us to increase the number of characteristics. Again, in both models, we observe a clear positive association between a characteristic's likelihood ratio and its representativeness, as both scatterplots are clearly upward sloping.

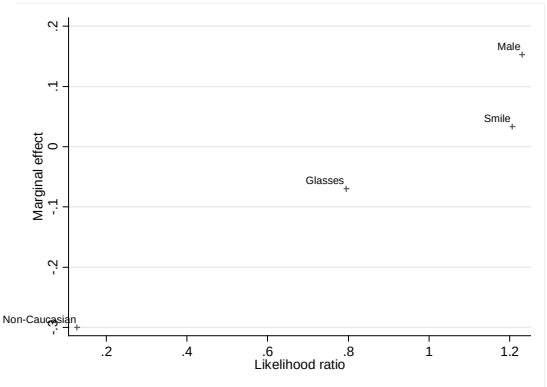
[14] We focus on objective characteristics here, because differences in the means of those characteristics are quantitatively small and seldom statistically significant. One may however note that the absolute marginal effect of trustworthiness is larger than the absolute marginal effect of competence in regressions run on the whole sample (Table 4a and 4b), which is in line with the finding that average trustworthiness of left-wing deputies is statistically significantly larger than that of right-wing deputies, while the difference between the two groups in terms of mean perceived competence is smaller and statistically insignificant. By the same token, the absolute marginal effect of competence in regressions on female deputies is larger than the absolute marginal effect of competence (Tables 6a and 6b), which is in line with the finding that female left-wing deputies are on average perceived as more competent than their right-wing counterparts.

Finally, the bottom panel of Figure 4 focuses on the sample of female deputies. Both scatterplots are again clearly upward sloping. Female characteristics that are more representative of the right elicit larger reactions.

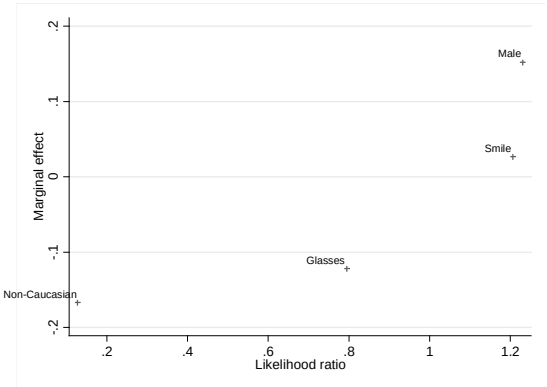
Figure 4: Estimated marginal effects of objective characteristics on the propensity to categorize a deputy as right-wing vs. likelihood ratios

Whole sample

Survey 1 (Reg 4a.1)

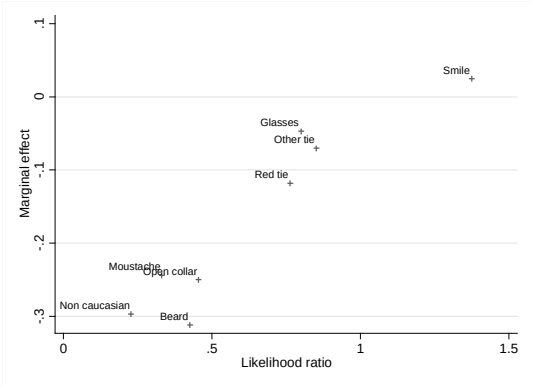


Survey 2 (Reg 4b.1)

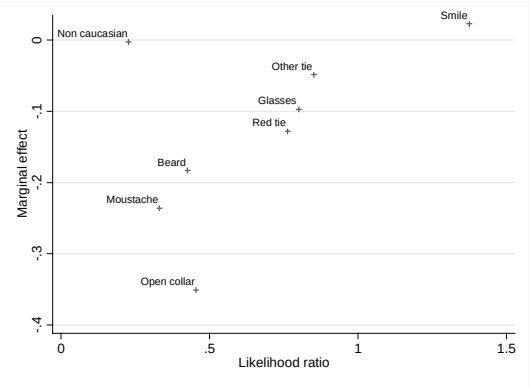


Male deputies

Survey 1 (Reg 5a.1)

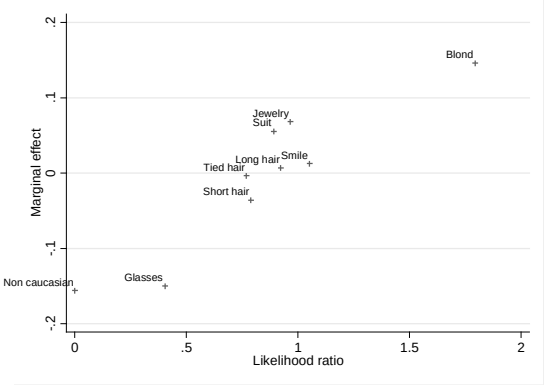


Survey 2 (Reg 5b.1)

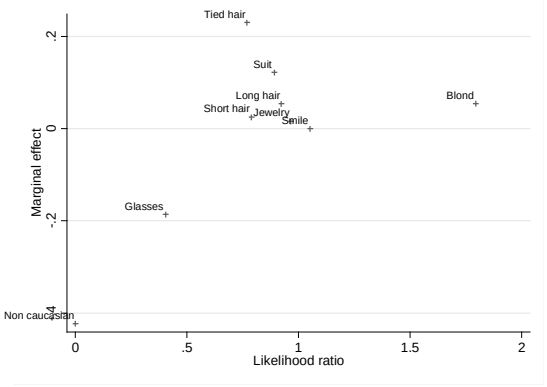


Female deputies

Survey 1 (Reg 6a.1)



Survey 2 (Reg 6b.1)



All scatterplots signal a strong association between a characteristic's likelihood ratio and its marginal effect on the propensity to classify a deputy as right-wing. The same finding applies regardless of the survey used and of the sub-sample considered. Although we cannot test its statistical significance, because there are too few coefficients, the association shows that the more distinctive of the right a characteristic is, the more respondents to the two surveys reacted to it by classifying a deputy as right-wing. Conversely, characteristics that are more representative of the left result in larger propensities to classify deputies as left-wing.

What is key in our sample is that the most representative characteristics are not necessarily those that are the most prevalent in each group. For instance, being non-Caucasian is the characteristics that elicits the strongest association with the left. This is in line with the representativeness of that characteristic. According to Table 1a, it is indeed representative of the left, as the likelihood ratio is only 0.13 in the whole sample and 0.23 in the sample of male deputies. It is even strictly diagnostic of being left-wing for female deputies, as there is simply no female non-Caucasian right-wing deputy in our sample. Yet, only 5.45 percent of left-wing deputies are non-Caucasian. The proportion falls to 3.52 percent among male left-wing deputies and is 10 percent among female deputies. This finding and the association between the marginal effects of characteristics and their representativeness suggest that respondents used the representative heuristic to categorize deputies.

6. Are partisan stereotypes partisan?

Both Bayesian classification and the representativeness heuristic imply that cues used by respondents should be the same across groups, because the two models only take into account the characteristics of the groups to be categorized and not those of categorizers. Yet, categorizers may differ, in particular left- and right-wing respondents.

Westfall et al. (2015) find that American citizens who identify strongly as Republicans or Democrats perceive a greater polarization between the two parties, and that they perceive the degree of polarization of the other group as larger than the degree of polarization of their own. That exaggeration of polarization may extend to visual cues. The judgements of respondents may in addition be biased by a halo effect, whereby respondents infer specific attributes about a person

from their overall impression of that person.[15] In the case of deputies' political orientations, a halo effect would result in respondents who have a positive overall impression about a deputy attributing their favored ideological position to that deputy (Herrmann and Shikano, 2016). They could for instance attribute positive personality traits to deputies who share their views. In addition, respondents may perceive more polarization among deputies that they perceive as members of the opposite group, as reported by Westfall et al. (2015), which would amplify the halo effect.

Testing whether the cues used by left- and right-wing respondents differ is therefore a further test of the two models, although it does not allow us to discriminate between them. In this sub-section, we estimate the same models as in the previous section, but distinguish left-wing respondents, right-wing respondents, and respondents who have not declared their political orientation.[16]

To save on space, we focus on three specifications. The first considers objective explanatory variables. The second considers subjective explanatory variables. The third considers both subjective and objective variables, and controls for deputies' actual orientations. We no longer control for respondents' characteristics, because they were never significant in all previous regressions except one.

*** insert Table 7a around here ***

*** insert Table 7b around here ***

Table 7a estimates Model 2a while distinguishing the responses of left- and right-wing respondents when the sample pools male and female deputies. The results corresponding to left-wing respondents are reported in odd-numbered columns while those for right-wing respondents are reported in even-numbered columns. The F-statistics indicate that the null hypothesis that all coefficients are zero can be rejected for all models.

[15] In line with this possibility, we observe a positive correlation of attractiveness, competence, and trustworthiness scores. However, the coefficient of correlation is moderate even if the correlation is significant at the one-percent level. In the sample of deputies who have not held national office, the coefficient of correlation is 0.70 between trustworthiness and competence, 0.51 between trustworthiness and attractiveness, and only 0.49 between attractiveness and competence. Moreover, our results are virtually unaffected if those variables are introduced one by one.

[16] We also distinguished observations according to deputies' true political affiliations to check if the characteristics used to classify deputies differ across left- and right-wing deputies. We found no systematic difference across the two groups. Such a non-result is unsurprising, as survey respondents are precisely asked to guess the orientation of deputies and can by construction not condition their classification on a characteristic that they do not observe. Those results are available on request.

The first striking result of Table 7a is that left- and right-wing respondents use the same characteristics to categorize politicians. As before, the share of right-wing classifications increases in both groups if a deputy is male, smiles, and looks competent. It decreases if the deputy is non-Caucasian, wears glasses, or looks trustworthy. Coefficients are moreover quantitatively very similar, suggesting that stereotypes are very similar in the two groups.

T-tests show that the coefficients of objective variables are in general statistically indistinguishable at standard levels of confidence. The only exception is the non-Caucasian dummy, whose coefficients are statistically different at the one-percent level across the two groups. However, the difference between the coefficients remains quantitatively small, at around five percentage points of the right-wing score.

The findings regarding subjective characteristics are more contrasted. The coefficients of the competence score are indistinguishable across the two groups of respondents at standard levels of significance. The coefficients of the trustworthiness score are statistically different at the five-percent level, but the difference does not exceed two percentage points of the trustworthiness score.

The second striking finding of Table 7a is to be found in the coefficient of the attractiveness score. Its coefficient is statistically insignificant in the sample of left-wing respondents (Columns 3 and 5), but significantly positive up to the one-percent level among right-wing respondents (Columns 4 and 6). A t-test moreover confirms that the two coefficients are significantly different at the one-percent level. Left- and right-wing respondents therefore differed in that right-wing respondents considered that attractive deputies were more likely to be on their side of the political spectrum.

In Survey 2, respondents had the option of refusing to specify their own political orientation and can be categorized in three categories. Table 7b therefore compares the outcomes of estimating Model 2b separately for left-wing, right-wing, and undeclared respondents. The Chi-squared statistics indicate that the null hypothesis that all coefficients are zero can be rejected for all models.

When comparing the results for the three groups of respondents, we again see that most explanatory variables exhibit the same sign across groups of respondents, suggesting that respondents with different political preferences use visual cues in a similar way to guess deputies' political orientations. Thus, the male deputy dummy exhibits a positive sign in all regressions and is significant at the one- or five- percent level. Being non-Caucasian exhibits a negative sign in all regressions, confirming that it is used as a signal of being left-wing by survey respondents. It is

significant at the five- or ten-percent level. Wearing glasses bears a negative sign throughout Table 7b and is significant at the one- or five-percent level. All types of respondents therefore use it as a left-wing cue. The smile dummy is positive but insignificant at standard levels of significance in nearly all regressions of Table 7b. It is marginally significant at the ten-percent level for undeclared respondents in Column 3, but becomes insignificant once subjective personality traits are controlled for.

The results for subjective personality traits are more contrasted. The coefficient of the competence score is now only significant for undeclared respondents (Columns 6 and 9). In line with previous results, its sign is positive for those respondents, suggesting that they associate perceived competence with right-wing deputies. In line with previous findings, the coefficient of the trustworthiness score is negative for all groups of respondents. However, it is never significant for right-wing respondents (Columns 5 and 8). For left-wing respondents it is significant at the one-percent level in Column 4, but marginally fails to be significant when objective cues are controlled for (Column 7).

Again, the attractiveness score is used differently by different groups. It bears a negative and statistically significant coefficient among left-wing respondents (Columns 4 and 7). Conversely, in regressions restricted to right-wing voters, the coefficient is positive and significant at the ten-percent level (Columns 5 and 8). Finally, the coefficient is statistically insignificant in regressions restricted to undeclared respondents. A t-test confirms that the difference between the coefficients of the group of left-wing respondents and right-wing respondents is significant well beyond the one-percent level. The difference is also significant at the ten-percent level between right-wing and undeclared respondents. It is statistically insignificant between left-wing and undeclared respondents. Those findings therefore confirm the difference between left- and right-wing respondents in Survey 1.

In a nutshell, those findings suggest that respondents with a declared political affiliation tend to associate attractiveness with their own side of the political spectrum, while undeclared respondents do not use attractiveness as a cue to guess the orientation of deputies. In other words, left-wing respondents consider attractiveness as a left-wing characteristic, whereas right-wing respondents consider it as right-wing.[17] This finding suggests a halo effect in the way respondents react to the attractiveness of deputies.

[17] It could be argued that beauty is in the eye of the beholder, and that the result means that respondents tend to find deputies attractive because they perceive them to be on their side of the political spectrum. However, one must bear in mind that attractiveness here is measured by each deputy's average attractiveness score computed over all the other

*** insert Table 8a around here ***

*** insert Table 8b around here ***

Table 8a reports the results of distinguishing left- and right-wing respondents when estimating Model 2a on the sub-sample of male politicians. It confirms that respondents of both groups react in a similar way to visual cues. Glasses and being non-Caucasian are perceived as left-wing cues, while smiling is perceived as right-wing by both groups of respondents. Deputies perceived as competent are categorized as right-wing while those who are perceived as trustworthy are categorized as left-wing. Finally, like in previous tables, attractiveness is the only characteristic that results in opposite categorizations by the two groups. It is significant in the sub-sample of right-wing respondents and bears a positive sign, suggesting that right-wing respondents categorize attractive deputies as right-wing, in line with a halo effect. On the contrary, the coefficient of the attractiveness score is insignificant in the sub-sample of left-wing respondents.

The new results concern male-specific characteristics, which all bear coefficients that are significant at the one-percent level and for which we find a great deal of consensus. The red tie, other tie, and open collar dummies bear a negative sign for both groups of respondents and are therefore associated with being left-wing. In both groups, the beard and moustache dummies exhibit a negative sign. The difference in the coefficients is in general statistically insignificant, except for the other tie and beard dummies, but even in those two cases, the magnitude of the difference is small. It does not exceed two points of the right-wing score for the other tie dummy and four for the beard dummy.

Table 8b reports the results of distinguishing left- and right-wing respondents when estimating Model 2b on the sub-sample of male politicians. Again, the Chi-squared of all regressions show that the hypothesis that all coefficients are zero can be rejected at the one-percent level of significance. Again, we observe little difference in the way that objective visual cues correlate with the propensity to be categorized as left- or right-wing. Being non-Caucasian and smiling are insignificant in all regressions.[18] The glasses dummy exhibits a negative sign significant

respondents and not by the evaluation of the respondent. It is accordingly a personal characteristic of the deputy. The finding therefore suggests that respondents tend to classify attractive deputies as belonging to their own group, and not that they find deputies that they perceive as belonging to their group more attractive.

[18] The non-Caucasian dummy does not appear at all in regressions involving only undeclared respondents, because no undeclared respondent was matched to a non-Caucasian deputy in the survey. This is a reminder that some deputy characteristics are rare in our sample, which makes it difficult to assess their effect in smaller samples.

at the one-percent level in all regressions except in those focusing on undeclared respondents, where it is still negative but insignificant. Table 8b also broadly confirms the results of Table 7b concerning perceived competence, which is associated with being right-wing by undeclared respondents, and trustworthiness, which is associated with being left-wing by left-wing and undeclared respondents. Again, attractiveness is associated with being left-wing by left-wing respondents and with being right-wing by right-wing respondents, while undeclared respondents do not use it in a statistically significant way.

We again find a great deal of consensus for male-specific characteristics. The red tie dummy bears a negative sign for all groups and is therefore associated with being left-wing, but is insignificant at standard levels of significance among undeclared respondents. An open collar is associated with the left by left- and right-wing respondents, but its effect cannot be assessed for undeclared respondents. The other tie dummy exhibits a negative sign in all groups, although it is only significant, at the ten-percent level, for right-wing respondents. In all groups, wearing a beard or a moustache exhibits a negative sign, but the moustache dummy is insignificant at standard levels of significance in the group of undeclared respondents.

*** insert Table 9a around here ***

*** insert Table 9b around here ***

Table 9a reports results obtained when estimating Model 2a on the sub-sample of female deputies. Again, we observe a great deal of consensus among left- and right-wing respondents. They concur in categorizing non-Caucasian female deputies and those who wear glasses as left-wing, and in categorizing blond female deputies and those who wear a suit as right-wing. Both dummy variables bear coefficients that are significant at the one-percent level. Left- and right-wing respondents also concur in categorizing trustworthy-looking deputies as more left-wing. The competence score bears a negative coefficient in both groups of respondents, but is only significant among right-wing respondents. Finally, we still observe that a deputy's attractiveness score is positively significant in the group of right-wing respondents and insignificant in the group of left-wing respondents, in line with a halo effect.

Few female-specific dummies are statistically significant. The blond dummy is significantly positive at the one-percent level in both groups, signaling that they both associate blondness with being right-wing. Similarly, the suit dummy exhibits a positive coefficient in all regressions where

it appears. It is insignificant for both groups of respondents in Columns (1) and (2), but significant at the ten-percent level for both groups in the comprehensive specification (Columns 5 and 6). Finally, the jewelry dummy bears a positive sign throughout Table 9a. It is only significant in the group of left-wing respondents, at the ten-percent level, but is marginally insignificant at the same level in the right-wing group.

Table 9b reports estimations of Model 2b on the sub-sample of female deputies. Likely due to the more limited number of observations, fewer coefficients appear significant in that sample. Still, we remark that the signs of objective cues in general do not differ across groups of respondents. Moreover, we obtain results that are reminiscent of previous tables. Specifically, we observe that being non-Caucasian exhibits a significantly negative sign and is therefore considered as a left-wing feature by the three groups of respondents (Columns 1, 2, 3, 7, 8, 9). Similarly, wearing glasses bears a negative coefficient in all regressions (Columns 1, 2, 3, 7, 8, 9) and is significant at the one-percent level among left-wing respondents (Columns 1, 7).

Like in Tables 4a and 4b, trustworthiness exhibits a positive sign (Columns 1, 2, 3, 7, 8, 9). The coefficient is statistically significant only among right-wing respondents, and at the five-percent level (Columns 5, 8). As in previous regressions, competence bears a negative sign, but is insignificant at standard levels in all regressions. Unlike in the sample of male deputies, attractiveness never exhibits a positive coefficient.

Female-specific characteristics provide little additional insight. The jewelry dummy never appears significant at accepted levels of significance. We however observe that the suit dummy exhibits a positive sign significant at the one-percent level in all regressions run on right-wing respondents (Columns 2, 8). Hair dummies are never significant, except the tied hair dummy, which exhibits a positive sign and is significant at the ten-percent level in the sample of right-wing respondents (Columns 2, 8).

Distilling the results of this section, we can say that respondents with different political preferences concur in associating some characteristics with the left- or the right-wing of the political spectrum. The consensus is nearly perfect on objective characteristics. The characteristics that are consensually associated with the left are being female, being non-Caucasian, and wearing glasses. When it comes to male deputies, all respondents also concur in relating ties that are not blue, an open collar, and any form of facial hair to the left rather than to the right. In the female sample, we

find some evidence that deputies wearing suits, tying their hair, and who are blond are considered as more right-wing.

We report some evidence that trustworthiness is consensually associated with the left, while competence is associated with the right in regressions on all deputies and on male deputies. The association of trustworthiness and competence with political orientation seems to be reversed for female deputies, though the sample is smaller. Respondents therefore seem not to be subject to a halo effect when evaluating competence and attractiveness. We therefore find little evidence of a halo effect for competence and trustworthiness.

We however do find a halo effect for attractiveness. Attractiveness stands out, because it is the only characteristic about which left- and right-wing respondents disagree. When all deputies are pooled or when the sample is restricted to male deputies, we observe that attractiveness bears opposite signs in the samples of left- and right-wing respondents, while no such pattern is observable among undeclared respondents in Survey 2. Furthermore, both left- and right-wing respondents associate attractiveness with their own group.

The absolute magnitude of the effect of attractiveness is also larger for right-wing than for left-wing respondents. Specifically, the absolute coefficient of the attractiveness score is one-third larger in the sample of right-wing respondents than in the sample of left-wing respondents. The attractiveness dummy is insignificant in the sample of female deputies. In Survey 1, the coefficient of the attractiveness dummy is always significant in the sub-sample of right-wing respondents while it is always insignificant in the left-wing sub-sample.

Those findings may provide an explanation for the finding of Berggren, et al. (2011), who observe a larger beauty premium for right-wing politicians in Finnish municipal elections. Our results suggest that right-wing voters are more prone to perceive attractiveness as a sign of being right-wing. Right-wing voters may accordingly cast their ballot in favor of attractive candidates, because they interpret good looks as a signal that candidates' preferences are aligned with their own.

7. Are stereotypes misguiding?

While stereotypes are qualitatively in line with the true characteristics of the two groups of deputies, they do not lead to perfect categorization. Indeed, respondents do make categorization errors. In this section, we use those errors to provide additional evidence on the way in which respondents used stereotypes. Our starting point is that if respondents process information in a rational way and therefore use all the relevant information available, they should not make systematic errors. In other words, errors should be random. Conversely, if respondents process information using the representativeness heuristic, some visual cues could correlate with a larger or lower propensity to make a categorization error. Testing whether errors relate to observable cues in a systematic way is therefore a way to test the rationality of stereotypes.

To do so, we estimate Models 2a and 2b but change the dependent variables. In Model 2a, we replace the dependent variable by the share of respondents who have incorrectly categorized deputy *i* as left- or right-wing. Likewise, in Model 2b, the dependent variable is replaced by a dummy variable that takes the value one if respondent *j* has incorrectly categorized deputy *i* as left- or right-wing.

The rationale of the test is straightforward. If respondents are rational, there should be no correlation between errors made in classifying deputies and visual cues. The coefficients of independent variables coding visual cues in Models 2a and 2b should therefore be statistically indistinguishable from zero. Conversely, if we observe that one of those variables exhibits a coefficient that significantly differs from zero, then we can conclude that the errors that respondents make are not randomly distributed. This will signal a bias that is inconsistent with rationality, because some variables will lead to worse or better categorizations than others.

*** insert Table 10a around here ***

*** insert Table 10b around here ***

*** insert Table 10c around here ***

Tables 10a to 10c report the results of estimating Models 2a and 2b when dependent variables code the propensity of respondents to incorrectly classify deputies. Table 10a reports the results obtained when pooling all deputies together regardless of gender.

In the first five columns of Table 10a, we report the outcome of regressions where the dependent variable is the share of respondents who have incorrectly classified each deputy in

Survey 1. We observe that the coefficients of the gender and non-Caucasian dummies are negative and statistically significant at the ten-percent level, implying that female and non-Caucasian deputies are classified with more accuracy than other deputies (Columns 1 and 3). When distinguishing left- and right-wing respondents (Columns 4 and 5), we find that the gender and non-Caucasian cues are used in a different way by the two groups of respondents. The gender dummy is only significant in the sample of left-wing respondents while the non-Caucasian dummy is only significant in the sample of right-wing respondents. Note however that when one of the two dummies is insignificant it marginally misses the ten-percent threshold. Finally, we can see no correlation between subjective evaluations and categorization errors, as the coefficients of attractiveness, competence, and trustworthiness are all statistically insignificant (Columns 2, 3, 4 and 5).

The last seven columns of the table report the outcome of estimating Model 2b, where the dependent variable is a dummy taking the value one if respondent *j* has incorrectly categorized deputy *i*. Columns 6 to 9 report the results of estimating the models when all respondents are pooled together. They exhibit no sign of systematic biases, insofar as the coefficients of all the variables coding the characteristics of deputies appear statistically insignificant at standard levels of significance. Columns 10 to 12 run separate regressions for each group of respondents. Again, they show little sign of bias. The only exception is the coefficient of attractiveness in the group of undeclared respondents, which is statistically significant at the five-percent level (Column 12). It bears a positive and significant sign, implying that more attractive deputies tend to be more often misclassified by undeclared respondents. Finally, the coefficient of the respondent's gender in the sample of right-wing respondents is negative and significant at the five-percent level, suggesting that right-wing male respondents are better at guessing deputies' orientations than right-wing female respondents. As this variable concerns respondents and not deputies, its exhibiting a significant sign does not help to distinguish the models of behavior. It suggests that male respondents may be better informed, but not more or less rational.

The striking feature of Table 10a is that the two characteristics that significantly correlate with a smaller share of categorization errors in Survey 1 are gender and being non-Caucasian, which are precisely the two characteristics that are the most distinctive of left-wing deputies versus right-wing deputies. Accordingly, respondents seem to react the most to the two most distinctive features of the two groups of deputies. This supports the notion that respondents use the representativeness heuristic to categorize politicians.

Table 10b reports results pertaining to male deputies. Here we observe that four objective characteristics bear a statistically significant coefficient at least in some regressions, namely the red tie dummy, the other tie dummy, the non-Caucasian dummy, and the glasses dummy. Moreover, the results obtained with the two surveys concur. When significant, the red tie and other tie dummies both exhibit a positive sign implying that they are associated with more categorization errors. The red tie dummy is significant in almost all regressions (Columns 1 to 9), except in the regressions on the samples of right-wing and undeclared respondents in Survey 2 (Columns 11 and 12). The other tie dummy is significant in the whole sample and in the sample of right-wing respondents in Survey 2 (Columns 6, 9, and 11). It is insignificant in Survey 1 (Columns 1, 3, 4, and 5). The non-Caucasian dummy appears positive and significant in Survey 2 (Columns 6 and 9), but not when respondents are split across groups (Columns 10, 11, and 12). The glasses dummy bears a negative coefficient in the first regression using Survey 1 (Column 1). We find little evidence that subjective characteristics are related to correct categorizations, except for attractiveness in the regression focusing on undeclared respondents in Survey 2, where it exhibits a positive sign significant at the ten-percent level (Column 12). Finally, we still find some evidence that male right-wing respondents are more accurate than female right-wing respondents, since the gender dummy exhibits a negative sign in Column 11.

The results of Table 10b add some additional support to the representativeness heuristic model. Being non-Caucasian correlated with a lower propensity to correctly guess a deputy's orientation in Survey 2 according to Table 10b. At the same time, Table 2a shows that while being non-Caucasian is the most representative feature of left-wing male deputies, the share of non-Caucasians is low among both left- and right-wing deputies. This is a typical configuration where Gennaioli and Shleifer's (2010) model predicts a large categorization bias. Specifically, respondents over-reacted to observing a non-Caucasian deputy because this trait is representative of left-wing deputies, but thereby neglected the possibility that some non-Caucasian deputies are right-wing.

Similarly, red ties, and to some extent ties of other colors, resulted in more mis-categorizations. With an odd ratio of 0.76, red ties are indeed distinctive of the left. However, only 22.54 percent of left-wing deputies wear one. By reacting to a red tie by categorizing a deputy as left-wing, respondents overlooked the fact that a majority of left-wing deputies wear another type of tie. They also overlooked the fact that a non-negligible portion of right-wing deputies, 17.20 percent, also wear a red tie.

Finally, while smiling is distinctive of right-wing male deputies, it does not correlate with the propensity to make mistakes. This suggests that a distinctive characteristic may not be used by respondents. Admittedly, smiling is a less permanent characteristic than the other visual cues that we have coded, which may explain why respondents did not take it into account. Kahneman and Tversky (1972) stress that the representativeness heuristic is not the only mechanism that shapes recall. They also consider the availability of characteristics, defined as the ease with which a characteristic can be recalled by respondents. It may be argued that because smiling is a transient characteristic, it is less easily recalled by respondents than other characteristics.

Table 10c reports results pertaining to female deputies. In that sample, we find that being non-Caucasian leads to more accurate classifications of deputies (all columns except Column 10). The finding appears in both surveys. In Survey 1, it holds for both left- and right-wing respondents. In Survey 2, it holds when all respondents are pooled together (Columns 6 and 9) and in the sub-samples of right-wing (Column 11) and undeclared respondents (Column 12), but not in the sub-sample of left-wing respondents (Column 10). Wearing glasses is also associated with better categorizations (Columns 1, 3, 4, 5, and 10). The finding appears in all regressions using Survey 1 (Columns 1 to 5). It only appears in Survey 2 in the regression focusing on left-wing respondents (Column 10).

We moreover observe some evidence that the long hair and tied hair dummies exhibit a positive sign in the regression restricting the sample to right-wing respondents of Survey 2 (Column 11), implying that those dummies result in fewer correct guesses. Also, attractiveness bears a positive sign and trustworthiness a negative sign that are respectively significant at the five- and ten-percent levels in the regression focusing on undeclared respondents of Survey 2 (Column 12).

Table 1a shows that being non-Caucasian and wearing glasses are the most representative features of the left. Being non-Caucasian is even perfectly diagnostic of being left-wing for female deputies, as there was no non-Caucasian right-wing female deputy in our sample. Finding that the two characteristics stand out in Table 10c is therefore in line with the representativeness heuristic.

Finally, while being blond is the most representative characteristic of right-wing female deputies, it does not correlate with the probability of being correctly categorized. It however correlates with the propensity of respondents in Survey 1 to categorize female deputies as left-wing, according to Table 6a.

8. Conclusion

We used two complementary surveys asking respondents to classify actual French deputies as left- or right-wing based on their anonymous photographs. The results obtained with the two surveys concurred. We first confirmed that respondents outperform random guesses when categorizing deputies.

Second, we found that categorizations correlated with observable visual cues and with subjective assessments of deputies' photographs. For instance, male deputies tend to be categorized as right-wing and female deputies tend to be categorized as left-wing, while deputies perceived as competent tend to be categorized as right-wing, and those perceived as trustworthy are categorized as left-wing.

Third, we found that objective visual cues are used in the correct direction by survey respondents, at least when deputies of both genders are included in the sample and for male deputies. For instance, respondents tend to classify non-Caucasian deputies as left-wing and non-Caucasian deputies are indeed more likely to be left-wing. Stereotypes therefore contain a kernel of truth. Furthermore, respondents react to deputies' perceived competence and trustworthiness in the correct direction, despite actual left- and right-wing deputies being similar in terms of those dimensions. This shows that respondents may amplify minor differences across groups, which is in line with their use of the representative heuristic.

Strikingly, the cues that correlate with categorizations are independent of the political affiliation of survey respondents, which is our fourth result. Partisan stereotypes are consensual. Attractiveness is the only characteristic that prompts different reactions from respondents with different political preferences, as respondents who are affiliated with one group associate attractive deputies with that group, while unaffiliated respondents do not react to attractiveness.

Fifth, we observe that the magnitude of the marginal impact of a characteristic on the probability of a respondent categorizing a photograph as left- or right-wing tends to increase with the representativeness of that characteristic.

Finally, we find that representative characteristics tend to correlate with categorization errors. While findings 1, 2, and 4 are consistent with both Bayesian behavior and the representativeness heuristic applied to objective characteristics, findings 5 and 6 are at odds with Bayesian behavior but consistent with the representativeness heuristic. Finding 3 is more ambiguous. Finding that respondents use cues in the correct direction is consistent with both Bayesian behavior and the representativeness heuristic, in particular for objective characteristics. However, finding that they

tend to react to subjective characteristics while those characteristics differ little across groups is in line with the notion that the representativeness heuristic results in stereotypes that amplify differences across groups.

In addition, we observed differences in the accuracy of stereotypes for male and female deputies. Those differences could be due to the scarcity of female deputies preventing respondents from forming accurate beliefs. They could also be due to an interaction of political stereotypes with gender stereotypes. Finally, because the results of the present paper rely on photographs of actual deputies they consider the appearance of deputies as given. Yet, deputies may manipulate their looks in some sort of signaling game with voters. What our results show is that in equilibrium, left- and right-wing deputies look different and that respondents can distinguish them. Our results do not allow us to determine how looks could be manipulated nor to what effect. In fact, whereas some characteristics to which respondents react are inherited, like race or gender, others can be chosen, like tie color or facial hair. Differences in the prevalence of inherited characteristics and characteristics that can be consciously altered across left- and right-wing politicians may signal manipulation. Providing a model of how respondents differently use and form beliefs on objective and subjective cues, how gender interacts with political affiliation, and whether and how appearances are manipulated is food for further research.

9. References

Alesina, A. and Giuliano, P., 2011. Preferences for Redistribution. In: A. Bisin and J. Benhabib, eds. *Handbook of Social Economics*. North Holland, pp. 93-132.

Antonakis, J. and Dalgas, O., 2009. Predicting Elections: Child's Play!. *Science*, 323, pp. 1183-1183.

Ballew, C. C. and Todorov, A., 2007. Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Science*, 104(46), pp. 17948–17953.

Banducci, S. A., Karp, J. A., Thrasher, M. and Rallings, C., 2008. Ballot Photographs as Cues in Low-Information Elections. *Political Psychology*, 29(6), pp. 903-917.

Berggren, N., Jordahl, H. and Poutvaara, P., 2010. The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, 94(1), pp. 8-15.

Berggren, N., Jordahl, H. and Poutvaara, P., 2011. The Right Look: Conservative Politicians Look Better and their Voters Reward it. Centre Emile Bernheim Working Paper N° 11/004.

Bertrand, M., 2011. New Perspectives on Gender. In: O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics*, 4, Elsevier, pp. 1545-1592.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., forthcoming. Stereotypes. *Quarterly Journal of Economics.*

Borkenau, P. and Liebler, A., 1995. Observable Attributes as Manifestations and Cues of Personality and Intelligence. *Journal of Personality*, 63(1), pp. 1-25.

Carney, D. R., Jost, J. T., Gosling, S. D. and Potter, J., 2008. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, 29(6), pp. 807-840.

Croson, R. and Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic Literature*, 47(2), pp. 448-474.

Fryer, R. G., 2011. Racial inequality in the 21st century: The declining significance of discrimination. In: O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics*, 4, Elsevier, pp. 855-971.

Fryer, R. G. and Jackson, M., 2008. A Categorical Model of Cognition and Biased Decision-Making. *B.E. Journal of Theoretical Economics*, 8(1).

Gennaioli, N. and Shleifer, A., 2010. What comes to mind. *Quarterly Journal of Economics*, 125(4), pp. 1399-1433.

Godbout, J.-F. and Foucault, M., 2013. French legislative voting in the Fifth Republic. *French Politics*, 11(4), pp. 307-331.

Gold, J. M., Mundy, P. J. and Tjan, B. S., 2012. The perception of a face is no more than the sum of its parts. *Psychological science*, 23(4), pp. 427-434.

Hamermesh, D. S., 2006. Changing looks and changing "discrimination": The beauty of economists. *Economics Letters*, 93(3), pp. 405-412.

Hamermesh, D. S., 2011. *Beauty pays: Why attractive people are more successful.* Pinceton NJ: Princeton University Press.

Hamermesh, D. S. and Biddle, J. E., 1994. Beauty and the labor market. *American Economic Review*, 84(5), pp. 1174-1194.

Herrmann, M. and Shikano, S., 2016. Attractiveness and facial competence bias face-based inferences of candidate ideology. *Political Psychology*, 37(3), pp. 401-417.

Hilton, J. L. and Von Hippel, W., 1996. Stereotypes. *Annual review of psychology*, 47(1), pp. 237-271.

Jussim, L. et al., 2016. Stereotype Accuracy: One of the Largest and Most Replicable Effects in All of Social Psychology. in: T. Nelson, ed. *Handbook of Prejudice, Stereotyping, and Discrimination.* Mahwah, NJ: Lawrence Erlbaum Publishing, pp. 31-64.

Kahneman, D. and Tversky, A., 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), pp. 430-454.

Kahneman, D. and Tversky, A., 1973. On the psychology of prediction. *Psychological Review*, 80(4), pp. 237-251.

Le Monde, 2006. Ségolène Royal "est de droite"... selon Pierre Bourdieu. 5 October.

McDermott, M. L., 1997. Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary US Elections. *American Journal of Political Science*, 41(1), pp. 270-278.

McDermott, M. L., 1998. Race and Gender Cues in Low-Information Elections. *Political Research Quarterly*, 51(4), pp. 895-918.

Meltzer, A. H. and Richard, S. F., 1981. A rational theory of the size of government. *Journal of Political Economy*, 89(5), pp. 914-927.

Mullainathan, S., 2002. *Thinking through Categories,* Mimeo Harvard University.

Olivola, C. Y. and Todorov, A., 2010. Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, 46(2), pp. 315–324.

Olivola, C. Y., Sussman,A. B., Tsetsos, K., Kang, O. E., and Todorov A., 2012. Republicans Prefer Republican-Looking Leaders: Political Facial Stereotypes Predict Candidate Electoral. *Social Psychological and Personality Science*, 3(5), pp. 605-613.

Penton-Voak, I. S. and Chen, J. Y., 2004. High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, 25(4), pp. 229-241.

Phelps, E. S., 1972. The statistical theory of racism and sexism. *American Economic Review*, 62(4), pp. 659-661.

Price, M. E., Kang, J., Dunn, J. and Hopkins., S., 2011. Muscularity and attractiveness as predictors of human egalitarianism. *Personality and Individual Differences*, 50(5), pp. 636-640.

Rahn, W. M., 1993. The Role of Partisan Stereotypes in Information Processing about Political Candidates. *American Journal of Political Science*, 37(2), pp. 472-496.

Rosenberg, S. W., Lisa, B., 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), pp. 108-127.

Rule, N. O. and Ambady, N., 2010. Democrats and Republicans Can Be Differentiated from Their Faces. *PLoS ONE*, 5(1), pp. 1-7.

Samochowiec, J., Wänke, M. and Fiedler, K., 2010. Political Ideology at Face Value. *Social Psychological and Personality Science*, 1(3), pp. 206-213.

Sellers, J. G., Mehl, M. R. and Josephs, R. A., 2007. Hormones and personality: Testosterone as a marker of individual differences. *Journal of Research in Personality*, 41(1), pp. 126-138.

Shen, J. and Palmeri, T. J., 2014. The perception of a face can be greater than the sum of its parts. *Psychonomic Bulletin and Review*, 22(3), pp. 1-7.

Sidanius, J. and Pratto, F., 2001. *Social dominance: An intergroup theory of social hierarchy and oppression.* Cambridge: Cambridge University Press.

Todorov, A., Mandisodza, A. N. and Goren, A., 2005. Inferences of Competence from Faces Predict Election Outcomes. *Science*, 308, pp. 1623-1926.

Westfall, J., Boven, L. V., Chambers, J. R. and Judd, C. M., 2015. Perceiving Political Polarization in the United States Party Identity Strength and Attitude Extremity Exacerbate the Perceived Partisan Divide. *Perspectives on Psychological Science*, 10(2), pp. 145-158.

Westfall, J., Van Boven, L., Chambers, J. R. and Judd, C. M., 2015. Perceiving Political Polarization in the United States Party Identity Strength and Attitude Extremity Exacerbate the Perceived Partisan Divide. *Perspectives on Psychological Science*, 10(2), pp. 145-158.

Appendix

***Insert Table A1 about here ***

***Insert Table A2 about here ***

***Insert Table A3 about here ***

Tables

Table 1a: Representativeness of visual cues

	Left-wing deputies (in %)	Right-wing deputies (in %)	Chi-squared	Likelihood ratio (2)/(1)
	(1)	(2)	(3)	(4)
Whole sample				
Non-Caucasian	5.45	0.69	10.42 ***	0.13
Glasses	39.60	31.49	3.45 *	0.80
Smile	47.03	56.75	4.50 **	1.21
Male deputy	70.30	86.51	19.40 ***	1.23
Male deputies				
Non-Caucasian	3.52	0.80	3.82 *	0.23
Moustache	8.45	2.80	6.27 **	0.33
Beard	11.27	4.80	5.71 **	0.43
Open collar	3.52	1.60	1.49	0.45
Red tie	22.54	17.20	1.67	0.76
Glasses	42.96	34.40	2.83 *	0.80
Other tie	50.70	43.20	2.05	0.85
Smile	40.14	55.20	8.22 ***	1.38
Female deputies				
Non-Caucasian	10.00	0.00	4.15 **	0.00
Glasses	31.67	12.82	26.15 ***	0.40
Tied hair	6.67	5.13	0.098	0.77
Short hair	65.00	51.28	1.85	0.79
Suit	83.33	74.36	1.18	0.89
Long hair	8.33	7.69	0.01	0.92
Jewelry	71.67	69.23	0.068	0.97
Smile	63.33	66.67	0.11	1.05
Blond	20.00	35.90	3.08 *	1.79

*** p<0.01, ** p<0.05, * p<0.1. Deputies who have held a national office are not included in the sample.

Table 1b: Representativeness of subjective visual cues

	Means				Standard deviations			
	Left-wing deputies	Right-wing deputies	t-test	LW-RW	Left-wing deputies	Right-wing deputies	F-test	LW-RW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Whole sample								
Trustworthiness	2.582	2.498	1.919*	0.084	0.490	0.465	1.112	0.025
Competence	2.623	2.591	0.767	0.032	0.477	0.425	1.262*	0.052
Attractiveness	2.156	2.155	0.027	0.001	0.522	0.523	0.993	-0.002
Male deputies								
Attractiveness	2.084	2.104	-0.388	-0.020	0.494	0.490	1.013	0.003
Competence	2.606	2.605	0.029	0.001	0.481	0.424	1.285*	0.057
Trustworthiness	2.486	2.466	0.393	0.019	0.474	0.465	1.041	0.009
Female deputies								
Attractiveness	2.328	2.482	-1.277	-0.154	0.549	0.611	0.810	-0.061
Trustworthiness	2.811	2.700	1.251	0.111	0.453	0.414	1.194	0.038
Competence	2.662	2.499	1.788*	0.163	0.470	0.423	1.234	0.047

*** p<0.01, ** p<0.05, * p<0.1. *** p<0.01, ** p<0.05, * p<0.1. Deputies who have held a national office are not included in the sample.

Table 2a: Dependent variable: Actual orientation. Whole sample.

	(1)	(2)	(3)
Male deputy	0.216		0.228
	(5.253)***		(4.930)***
Non-Caucasian	-0.364		-0.383
	(-2.519)**		(-2.508)**
Smile	0.0996		0.0969
	(3.064)***		(2.913)***
Glasses	-0.0980		-0.0866
	(-2.839)***		(-2.385)**
Attractiveness		0.0494	0.0614
		(0.957)	(1.374)
Competence		0.0428	-0.0402
		(0.583)	(-0.624)
Trustworthiness		-0.146	-0.0258
		(-2.093)**	(-0.424)
Observations	491	491	491
Pseudo R-squared	0.0614	0.00808	0.0648
Log likelihood	-312.2	-329.9	-311.0
Chi-squared	37.51	5.045	37.68

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2b: Dependent variable: Actual orientation. Male deputies.

	(1)	(2)	(3)
Red tie	-0.157		-0.157
	(-3.023)***		(-2.986)***
Other tie	-0.125		-0.126
	(-3.117)***		(-3.113)***
Open collar	-0.235		-0.244
	(-1.912)*		(-1.945)*
Beard	-0.158		-0.161
	(-2.194)**		(-2.209)**
Moustache	-0.183		-0.189
	(-1.793)*		(-1.802)*
Non-Caucasian	-0.233		-0.240
	(-1.548)		(-1.544)
Smile	0.105		0.105
	(2.223)**		(2.208)**
Glasses	-0.0762		-0.0762
	(-1.916)*		(-1.897)*
Attractiveness		0.0341	0.0220
		(0.568)	(0.428)
Competence		0.0171	-0.0211
		(0.200)	(-0.294)
Trustworthiness		-0.0495	0.0142
		(-0.627)	(0.217)
Observations	392	392	392
Pseudo R-squared	0.0652	0.00123	0.0658
Log likelihood	-239.9	-256.3	-239.8
Chi-squared	30.43	0.587	30.99

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2c: Dependent variable: True orientation. Female deputies.

	(1)	(2)	(3)
Suit	-0.0734		-0.0497
	(-0.564)		(-0.389)
Jewelry	0.0730		0.0508
	(0.582)		(0.444)
Blond hair	0.210		0.139
	(1.848)*		(1.265)
Short hair	-0.182		-0.146
	(-1.459)		(-1.203)
Long hair	-0.236		-0.250
	(-0.935)		(-0.958)
Tied hair	-0.0855		0.168
	(-0.329)		(0.676)
Smile	0.110		0.0559
	(0.880)		(0.477)
Glasses	-0.269		-0.0564
	(-1.671)*		(-0.379)
Attractiveness		0.285	0.278
		(2.491)**	(2.298)**
Competence		-0.273	-0.191
		(-1.475)	(-1.014)
Trustworthiness		-0.132	-0.170
		(-0.662)	(-0.925)
Observations	93	93	93
Pseudo R-squared	0.0761	0.0760	0.120
Log likelihood	-58.43	-58.44	-55.68
Chi-squared	9.796	7.756	13.00

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Dependent variable: Actual orientation.

	(1) Whole	(2) No national office	(3) LW	(4) RW	(5) Whole	(6) No national office	(7) LW	(8) RW	(9) U
Right-wing score	1.147	1.067	0.990	1.214					
	(9.943)***	(8.664)***	(8.437)***	(9.073)***					
Guessed right-wing					0.127	0.105	0.125	0.0858	0.0763
					(6.652)***	(4.953)***	(4.761)***	(2.560)**	(1.788)*
Observations	553	491	491	491	3,357	2,983	1,527	976	480
Pseudo R-squared	0.148	0.118	0.109	0.136	0.0143	0.00917	0.0134	0.00610	0.00468
Log likelihood	-316.8	-293.2	-296.3	-287.5	-2225	-2003	-1021	-655.8	-324.5
Chi-squared	94.26	70.64	66.17	79.84	37.68	21.86	19.41	5.940	2.978

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4a: Dependent variable: Share of right-wing classification in Survey 1. Whole sample.

	(1)	(2)	(3)
Male deputy	0.153		0.124
	(7.959)***		(5.827)***
Non Caucasian	-0.300		-0.310
	(-8.546)***		(-9.271)***
Smile	0.0336		0.0309
	(2.187)**		(2.037)**
Glasses	-0.0696		-0.0596
	(-4.340)***		(-3.616)***
Attractiveness		-0.000203	0.0114
		(-0.0110)	(0.652)
Competence		0.105	0.0650
		(4.417)***	(2.652)***
Trustworthiness		-0.154	-0.0997
		(-6.445)***	(-4.039)***
Constant	0.482	0.704	0.562
	(22.88)***	(15.08)***	(11.00)***
Observations	491	491	491
R-squared	0.214	0.078	0.240
Adj. R-squared	0.207	0.0720	0.229
F-test	45.66	15.56	32.91

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4b: Dependent variable: Guessed orientation in Survey 2. Whole sample.

	(1)	(2)	(3)	(4)	(5)
Male deputy	0.152			0.129	0.129
	(5.686)***			(4.372)***	(4.380)***
Non Caucasian	-0.167			-0.170	-0.169
	(-2.328)**			(-2.367)**	(-2.356)**
Smile	0.0268			0.0268	0.0266
	(1.201)			(1.212)	(1.202)
Glasses	-0.122			-0.118	-0.118
	(-5.630)***			(-5.385)***	(-5.397)***
Attractiveness		-0.0157		-0.0181	-0.0180
		(-0.669)		(-0.797)	(-0.791)
Competence		0.0673		0.0346	0.0346
		(1.981)**		(1.026)	(1.031)
Trustworthiness		-0.125		-0.0622	-0.0625
		(-3.787)***		(-1.809)*	(-1.820)*
Right-wing respondent			-0.00406		-0.00911
			(-0.180)		(-0.414)
Unknown respondent			0.00259		0.00413
			(0.102)		(0.169)
Respondent's gender			-0.00619		0.00111
			(-0.324)		(0.0603)
Observations	2,983	2,983	2,983	2,983	2,983
Pseudo R-squared	0.0224	0.00715	4.63e-05	0.0245	0.0246
Log likelihood	-2018	-2050	-2064	-2014	-2014
Chi-squared	59.69	18.92	0.172	70.24	70.44

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5a: Dependent variable: Share of right-wing classifications in Survey 1. Male deputies.

	(1)	(2)	(3)
Red tie	-0.118		-0.117
	(-6.720)***		(-6.783)***
Other tie	-0.0702		-0.0676
	(-4.883)***		(-4.713)***
Open collar	-0.250		-0.248
	(-3.990)***		(-4.274)***
Beard	-0.312		-0.305
	(-14.77)***		(-14.27)***
Moustache	-0.244		-0.230
	(-7.515)***		(-7.011)***
Non-Caucasian	-0.297		-0.305
	(-5.672)***		(-6.204)***
Smile	0.0251		0.0236
	(1.910)*		(1.829)*
Glasses	-0.0469		-0.0425
	(-3.489)***		(-3.146)***
Attractiveness		0.00574	0.0115
		(0.285)	(0.716)
Competence		0.110	0.0660
		(4.103)***	(2.992)***
Trustworthiness		-0.120	-0.0690
		(-4.473)***	(-3.052)***
Constant	0.725	0.618	0.696
	(48.62)***	(12.46)***	(18.08)***
Observations	392	392	392
R-squared	0.496	0.050	0.513
Adj. R-squared	0.486	0.0431	0.499
F-test	57.98	7.648	45.71

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5b: Dependent variable: Guessed orientation in Survey 2. Male deputies.

	(1)	(2)	(3)	(4)	(5)
Red tie	-0.128			-0.128	-0.126
	(-4.290)***			(-4.405)***	(-4.440)***
Other tie	-0.0485			-0.0463	-0.0452
	(-1.922)*			(-1.872)*	(-1.858)*
Open collar	-0.351			-0.337	-0.330
	(-4.939)***			(-5.323)***	(-5.325)***
Beard	-0.183			-0.177	-0.175
	(-5.198)***			(-5.147)***	(-5.175)***
Moustache	-0.236			-0.226	-0.223
	(-4.468)***			(-4.552)***	(-4.573)***
Non Caucasian	-0.00234			-0.00226	0.00213
	(-0.0245)			(-0.0255)	(0.0244)
Smile	0.0229			0.0215	0.0214
	(0.989)			(0.956)	(0.962)
Glasses	-0.0971			-0.0897	-0.0893
	(-4.388)***			(-4.084)***	(-4.082)***
Attractiveness		-0.0141		-0.0195	-0.0191
		(-0.519)		(-0.789)	(-0.784)
Competence		0.0714		0.0409	0.0402
		(1.957)*		(1.214)	(1.210)
Trustworthiness		-0.130		-0.0755	-0.0752
		(-3.533)***		(-2.144)**	(-2.157)**
Right-wing respondent			-0.0139		-0.0182
			(-0.550)		(-0.789)
Unknown respondent			0.0176		0.0128
			(0.619)		(0.487)
Respondent's gender			0.00484		0.0177
			(0.233)		(0.953)
Observations	2,380	2,380	2,380	2,380	2,380
Pseudo R-squared	2,380	2,380	2,380	2,380	2,380
Log likelihood	0.0383	0.00698	0.000329	0.0416	0.0422
Chi-squared	-1574	-1626	-1636	-1569	-1568

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6a: Dependent variable: Share of right-wing classifications in Survey 1. Female deputies.

	(1)	(2)	(3)
Suit	0.0555		0.0691
	(1.333)		(1.712)*
Jewelry	0.0683		0.0615
	(1.735)*		(1.628)
Blond hair	0.146		0.134
	(4.349)***		(3.810)***
Short hair	-0.0360		-0.0261
	(-0.856)		(-0.634)
Long hair	0.00696		-0.0151
	(0.0992)		(-0.222)
Tied hair	-0.00342		0.0305
	(-0.0414)		(0.366)
Non Caucasian	-0.156		-0.184
	(-2.800)***		(-3.335)***
Smile	0.0128		0.0158
	(0.387)		(0.508)
Glasses	-0.150		-0.122
	(-4.170)***		(-3.079)***
Attractiveness		0.0658	0.0200
		(1.962)*	(0.509)
Competence		-0.131	-0.0238
		(-2.218)**	(-0.431)
Trustworthiness		-0.0647	-0.0994
		(-1.203)	(-2.016)**
Constant	0.396	0.830	0.670
	(6.989)***	(7.645)***	(5.426)***
Observations	99	99	99
R-squared	0.359	0.150	0.417
Adj. R-squared	0.295	0.123	0.336
F-test	9.126	5.522	8.690

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6b: Dependent variable: Guessed orientation in Survey 2. Female MPs.

	(1)	(2)	(3)	(4)	(5)
Suit	0.122			0.108	0.0976
	(1.963)**			(1.746)*	(1.585)
Jewelry	0.0158			0.0157	0.0178
	(0.267)			(0.268)	(0.309)
Blond hair	0.0542			0.0413	0.0448
	(0.976)			(0.730)	(0.798)
Short hair	0.0249			0.0266	0.0227
	(0.354)			(0.402)	(0.347)
Long hair	0.0538			0.0707	0.0704
	(0.588)			(0.761)	(0.760)
Tied hair	0.230			0.217	0.223
	(2.113)**			(1.845)*	(1.962)**
Non Caucasian	-0.423			-0.392	-0.386
	(-2.327)**			(-2.065)**	(-1.970)**
Smile	-0.000108			-0.00433	-0.00284
	(-0.00186)			(-0.0776)	(-0.0510)
Glasses	-0.186			-0.196	-0.196
	(-3.130)***			(-2.869)***	(-2.759)***
Attractiveness		0.0338		-0.00863	-0.00713
		(0.764)		(-0.172)	(-0.143)
Competence		-0.153		-0.0913	-0.0955
		(-1.885)*		(-1.094)	(-1.170)
Trustworthiness		0.127		0.120	0.127
		(1.769)*		(1.541)	(1.655)*
Right-wing respondent			0.0230		0.0277
			(0.486)		(0.558)
Unknown respondent			-0.0664		-0.0794
			(-1.282)		(-1.385)
Respondent's gender			-0.0533		-0.0422
			(-1.238)		(-0.937)
Observations	603	603	603	603	603
Pseudo R-squared	0.0463	0.00776	0.00390	0.0497	0.0534
Log likelihood	-389.2	-404.9	-406.5	-387.8	-386.3
Chi-squared	36.26	5.089	3.619	37.62	46.58

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7a: Dependent variable: Share of right-wing classification in Survey 1. Whole sample. Leftwing vs. rightwing respondents.

	(1) LW	(2) RW	(3) LW	(4) RW	(5) LW	(6) RW
Male deputy	0.155	0.147			0.122	0.131
	(7.674)***	(8.315)***			(5.467)***	(6.727)***
Non Caucasian	-0.314	-0.262			-0.322	-0.275
	(-8.624)***	(-8.090)***			(-9.444)***	(-8.605)***
Smile	0.0328	0.0357			0.0306	0.0315
	(2.077)**	(2.382)**			(1.970)**	(2.136)**
Glasses	-0.0711	-0.0658			-0.0612	-0.0554
	(-4.315)***	(-4.208)***			(-3.624)***	(-3.459)***
Attractiveness			-0.0119	0.0324	-0.000708	0.0450
			(-0.629)	(1.800)*	(-0.0396)	(2.639)***
Competence			0.103	0.112	0.0635	0.0689
			(4.232)***	(4.713)***	(2.534)**	(2.852)***
Trustworthiness			-0.158	-0.142	-0.105	-0.0862
			(-6.543)***	(-5.919)***	(-4.170)***	(-3.498)***
Constant	0.493	0.452	0.758	0.556	0.619	0.404
	(22.28)***	(23.28)***	(15.93)***	(11.96)***	(11.79)***	(8.126)***
Observations	491	491	491	491	491	491
R-squared	0.213	0.198	0.085	0.069	0.244	0.230
Adj. R-squared	0.207	0.191	0.0792	0.0631	0.233	0.219
F-test	44.85	44.95	17.48	12.96	34.30	29.43

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7b: Dependent variable: Guessed orientation in Survey 2. Left-wing vs. right-wing and undeclared respondents. All deputies.

	(1) LW	(2) RW	(3) U	(4) LW	(5) RW	(6) U	(7) LW	(8) RW	(9) U
Male deputy	0.148	0.121	0.212				0.113	0.140	0.164
	(4.663)***	(2.537)**	(4.341)***				(3.311)***	(2.747)***	(2.900)***
Non Caucasian	-0.150	-0.143	-0.404				-0.145	-0.158	-0.454
	(-1.704)*	(-1.944)*	(-1.991)**				(-1.656)*	(-2.212)**	(-2.400)**
Smile	0.0231	0.0128	0.0715				0.0278	0.00971	0.0681
	(0.797)	(0.351)	(1.730)*				(0.980)	(0.267)	(1.598)
Glasses	-0.135	-0.115	-0.0967				-0.138	-0.107	-0.104
	(-4.886)***	(-3.089)***	(-2.315)**				(-4.980)***	(-2.818)***	(-2.369)**
Attractiveness				-0.0507	0.0632	-0.0581	-0.0643	0.0703	-0.0447
				(-1.680)*	(1.571)	(-1.187)	(-2.198)**	(1.743)*	(-0.936)
Competence				0.0646	-0.000971	0.178	0.0302	-0.0372	0.152
				(1.422)	(-0.0167)	(2.693)***	(0.664)	(-0.633)	(2.366)**
Trustworthiness				-0.138	-0.0645	-0.184	-0.0702	-0.00641	-0.116
				(-3.218)***	(-1.169)	(-3.005)***	(-1.640)	(-0.113)	(-1.831)*
Observations	1,527	976	480	1,527	976	480	1,527	976	480
Pseudo R-squared	0.0246	0.0161	0.0373	0.0133	0.00296	0.0192	0.0322	0.0189	0.0464
Log likelihood	-1031	-664.9	-319.6	-1043	-673.8	-325.6	-1023	-663.0	-316.6
Chi-squared	40.70	18.99	23.90	20.72	3.252	12.63	56.66	21.79	29.50

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8a: Dependent variable: Share of right-wing classifications in Survey 1. Male MPs.

	(1) LW	(2) RW	(3) LW	(4) RW	(5) LW	(6) RW
Red tie	-0.116	-0.124			-0.115	-0.122
	(-6.447)***	(-6.915)***			(-6.540)***	(-6.990)***
Other tie	-0.0650	-0.0843			-0.0623	-0.0819
	(-4.438)***	(-5.725)***			(-4.267)***	(-5.645)***
Open collar	-0.255	-0.235			-0.249	-0.245
	(-3.933)***	(-4.075)***			(-4.162)***	(-4.429)***
Beard	-0.321	-0.285			-0.314	-0.280
	(-14.65)***	(-13.87)***			(-14.21)***	(-13.47)***
Moustache	-0.248	-0.233			-0.234	-0.222
	(-7.110)***	(-8.314)***			(-6.637)***	(-7.837)***
Non Caucasian	-0.311	-0.256			-0.318	-0.271
	(-5.710)***	(-5.425)***			(-6.248)***	(-5.939)***
Smile	0.0252	0.0248			0.0237	0.0230
	(1.875)*	(1.846)*			(1.802)*	(1.769)*
Glasses	-0.0488	-0.0420			-0.0438	-0.0392
	(-3.553)***	(-3.067)***			(-3.175)***	(-2.882)***
Attractiveness			-0.00515	0.0361	0.000872	0.0410
			(-0.252)	(1.793)*	(0.0536)	(2.513)**
Competence			0.112	0.104	0.0674	0.0619
			(4.112)***	(3.872)***	(3.009)***	(2.743)***
Trustworthiness			-0.127	-0.101	-0.0751	-0.0516
			(-4.685)***	(-3.707)***	(-3.310)***	(-2.198)**
Constant	0.736	0.694	0.665	0.489	0.741	0.572
	(48.62)***	(45.63)***	(13.25)***	(9.730)***	(19.19)***	(13.92)***
Observations	392	392	392	392	392	392
R-squared	0.498	0.453	0.054	0.052	0.516	0.481
Adj. R-squared	0.488	0.441	0.0467	0.0442	0.502	0.466
F-test	56.83	54.11	8.268	7.740	45.54	43.45

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8b: Dependent variable: Guessed orientation in Survey 2. Left-wing vs. right-wing respondents and undeclared. Male deputies.

	(1) LW	(2) RW	(3) U	(4) LW	(5) RW	(6) U	(7) LW	(8) RW	(9) U
Red tie	-0.105	-0.161	-0.115				-0.108	-0.159	-0.113
	(-2.963)***	(-3.231)***	(-1.456)				(-3.213)***	(-3.284)***	(-1.456)
Other tie	-0.0437	-0.0648	-0.0267				-0.0419	-0.0650	-0.0248
	(-1.274)	(-1.735)*	(-0.493)				(-1.255)	(-1.820)*	(-0.454)
Open collar	-0.262	-0.431					-0.225	-0.438	
	(-2.357)**	(-3.020)***					(-2.328)**	(-3.222)***	
Beard	-0.187	-0.178	-0.169				-0.180	-0.178	-0.164
	(-3.706)***	(-2.385)**	(-2.524)**				(-3.635)***	(-2.395)**	(-2.523)**
Moustache	-0.208	-0.322	-0.0298				-0.199	-0.326	0.00535
	(-3.031)***	(-3.595)***	(-0.187)				(-2.965)***	(-3.814)***	(0.0326)
Non Caucasian	0.0110	0.0594					0.0219	0.0392	
	(0.101)	(0.566)					(0.222)	(0.384)	
Smile	0.0300	-0.00439	0.0606				0.0332	-0.00556	0.0526
	(1.014)	(-0.121)	(1.193)				(1.180)	(-0.157)	(1.038)
Glasses	-0.0958	-0.103	-0.0733				-0.0933	-0.0947	-0.0662
	(-3.266)***	(-2.927)***	(-1.534)				(-3.310)***	(-2.663)***	(-1.345)
Attractiveness				-0.0594	0.0881	-0.0555	-0.0651	0.0806	-0.0663
				(-1.655)*	(1.804)*	(-1.017)	(-2.088)**	(1.850)*	(-1.168)
Competence				0.0695	-0.00478	0.183	0.0293	-0.0274	0.187
				(1.353)	(-0.0734)	(2.504)**	(0.650)	(-0.478)	(2.489)**
Trustworthiness				-0.136	-0.0885	-0.178	-0.0722	-0.0334	-0.154
				(-2.806)***	(-1.433)	(-2.548)**	(-1.706)*	(-0.599)	(-2.085)**
Observations	1,207	791	375	1,207	791	375	1,207	791	375
Pseudo R-squared	0.0342	0.0522	0.0224	0.0127	0.00575	0.0165	0.0438	0.0588	0.0342
Log likelihood	-801.4	-517.3	-249.2	-819.2	-542.6	-250.7	-793.4	-513.7	-801.4
Chi-squared	53.56	45.72	12.50	15.02	4.958	8.145	68.90	51.60	53.56

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9a: Dependent variable: Share of right-wing classifications in Survey 1. Leftwing vs. rightwing respondents. Female MPs.

	(1) LW	(2) RW	(3) LW	(4) RW	(5) LW	(6) RW
Suit	0.0546	0.0581			0.0698	0.0677
	(1.261)	(1.496)			(1.665)*	(1.743)*
Jewelry	0.0725	0.0558			0.0660	0.0542
	(1.772)*	(1.564)			(1.721)*	(1.549)
Blond hair	0.162	0.101			0.149	0.0958
	(4.477)***	(3.407)***			(4.022)***	(2.803)***
Short hair	-0.0352	-0.0359			-0.0268	-0.0411
	(-0.883)	(-0.988)			(-0.678)	(-1.141)
Long hair	0.00403	0.0168			-0.0256	-0.00387
	(0.0564)	(0.289)			(-0.370)	(-0.0653)
Non Caucasian	-0.167	-0.129			-0.179	-0.157
	(-2.734)***	(-2.522)**			(-2.935)***	(-2.963)***
Smile	0.00903	0.0230			0.0180	0.0173
	(0.261)	(0.752)			(0.565)	(0.586)
Glasses	-0.148	-0.154			-0.133	-0.115
	(-3.935)***	(-4.993)***			(-3.367)***	(-3.327)***
Attractiveness			0.0508	0.108	-0.00212	0.0622
			(1.455)	(3.446)***	(-0.0545)	(1.825)*
Competence			-0.147	-0.0842	-0.0318	0.00640
			(-2.423)**	(-1.488)	(-0.571)	(0.116)
Trustworthiness			-0.0608	-0.0757	-0.0956	-0.0969
			(-1.089)	(-1.486)	(-1.897)*	(-1.992)**
Constant	0.401	0.381	0.908	0.614	0.738	0.479
	(6.924)***	(7.252)***	(7.998)***	(6.161)***	(5.816)***	(4.050)***
Observations	99	99	99	99	99	99
R-squared	0.356	0.336	0.164	0.139	0.429	0.377
Adj. R-squared	0.299	0.277	0.137	0.111	0.357	0.299
F-test	9.788	10.48	6.141	5.425	9.450	8.307

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9b: Dependent variable: Guessed orientation in Survey 2. Left-wing vs. right-wing and undeclared respondents. Female deputies.

	(1) LW	(2) RW	(3) U	(4) LW	(5) RW	(6) U	(7) LW	(8) RW	(9) U
Suit	0.0613	0.312	0.0477				0.0388	0.284	0.0102
	(0.768)	(3.085)***	(0.434)				(0.508)	(2.682)***	(0.0890)
Jewelry	0.0348	-0.0107	-0.0931				0.0394	-0.00645	-0.0618
	(0.489)	(-0.0939)	(-0.845)				(0.558)	(-0.0581)	(-0.527)
Blond hair	0.0758	0.00260	0.0775				0.0679	-0.0642	0.114
	(1.144)	(0.0236)	(0.709)				(1.005)	(-0.501)	(0.912)
Short hair	0.0167	-0.0190	0.0853				0.0261	-0.00866	0.102
	(0.227)	(-0.152)	(0.627)				(0.371)	(-0.0759)	(0.686)
Long hair	0.0579	-0.0841	0.296				0.0738	-0.0612	0.336
	(0.436)	(-0.490)	(1.330)				(0.560)	(-0.382)	(1.385)
Tied hair	0.201	0.422	0.189				0.146	0.446	0.328
	(1.571)	(1.714)*	(0.850)				(0.953)	(1.793)*	(1.142)
Non Caucasian	-0.410	-0.471	-0.390				-0.338	-0.491	-0.556
	(-1.650)*	(-1.966)**	(-1.605)				(-1.300)	(-1.956)*	(-1.872)*
Smile	-0.0590	0.0628	0.125				-0.0525	0.0121	0.127
	(-0.941)	(0.622)	(0.951)				(-0.836)	(0.119)	(0.930)
Glasses	-0.244	-0.0511	-0.138				-0.303	-0.0341	-0.144
	(-3.577)***	(-0.436)	(-1.031)				(-3.473)***	(-0.268)	(-0.924)
Attractiveness				0.0317	0.0151	0.0436	-0.0724	0.0130	0.102
				(0.623)	(0.193)	(0.410)	(-1.026)	(0.157)	(0.769)
Competence				-0.160	-0.240	-0.0457	-0.106	-0.248	0.0832
				(-1.566)	(-1.586)	(-0.371)	(-0.925)	(-1.606)	(0.545)
Trustworthiness				0.0739	0.368	0.0221	0.147	0.382	-0.0826
				(0.741)	(2.400)**	(0.183)	(1.394)	(2.104)**	(-0.641)
Observations	320	185	98	320	185	98	320	185	98
Pseudo R-squared	0.0608	0.0701	0.0597	0.00783	0.0355	0.00236	0.0677	0.114	0.0692
Log likelihood	-203.7	-117.7	-60.06	-215.2	-122.0	-63.72	-202.2	-112.1	-59.46
Chi-squared	34.24	17.62	10.65	2.911	6.543	0.319	30.99	23.34	11.78

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10a: Dependent variable: Share of categorization errors in Survey 1 and categorization errors in Survey 2. Whole sample.

	(1) Share All	(2) Share All	(3) Share All	(4) Share LW	(5) Share RW	(6) Error All	(7) Error All	(8) Error All	(9) Error All	(10) Error LW	(11) Error RW	(12) Error U
Gender	-0.0392		-0.0390	-0.0413	-0.0327	0.0221			0.0230	0.00468	0.0242	0.0754
	(-1.919)*		(-1.750)*	(-1.761)*	(-1.582)	(0.734)			(0.702)	(0.123)	(0.452)	(1.140)
Non Caucasian	-0.103		-0.100	-0.0990	-0.104	-0.0685			-0.0723	-0.0704	-0.0184	-0.227
	(-1.723)*		(-1.655)*	(-1.584)	(-1.862)*	(-0.804)			(-0.851)	(-0.708)	(-0.210)	(-1.258)
Smile	-0.0189		-0.0180	-0.0202	-0.0118	-0.0149			-0.0165	-0.0437	0.00624	0.0319
	(-1.092)		(-1.033)	(-1.101)	(-0.753)	(-0.624)			(-0.700)	(-1.454)	(0.165)	(0.668)
Glasses	0.0117		0.0100	0.0119	0.00494	-0.0310			-0.0250	-0.0420	-0.0183	0.00138
	(0.661)		(0.547)	(0.616)	(0.294)	(-1.276)			(-1.021)	(-1.324)	(-0.456)	(0.0274)
Attractiveness		-0.0133	-0.0128	-0.0114	-0.0167		0.0191		0.0193	0.0199	-0.0525	0.139
		(-0.717)	(-0.656)	(-0.553)	(-0.941)		(0.798)		(0.796)	(0.631)	(-1.266)	(2.540)**
Competence		-0.0237	-0.0121	-0.0119	-0.0133		-0.0126		-0.0183	-0.0694	0.0966	-0.0807
		(-0.855)	(-0.424)	(-0.391)	(-0.526)		(-0.364)		(-0.511)	(-1.444)	(1.606)	(-1.204)
Trustworthiness		0.0283	0.0114	0.0132	0.00660		-0.0225		-0.0104	0.0418	-0.0715	-0.0459
		(1.066)	(0.406)	(0.450)	(0.254)		(-0.650)		(-0.282)	(0.892)	(-1.192)	(-0.693)
Right-wing respondent								0.0229	0.0209			
								(1.042)	(0.934)			
Unknown respondent								0.0250	0.0238			
								(1.013)	(0.960)			
Respondent's gender								-0.0143	-0.0134	-0.00514	-0.0706	0.0492
								(-0.763)	(-0.719)	(-0.199)	(-2.081)**	(1.010)
Constant	0.452	0.431	0.482	0.475	0.503							
	(20.76)	(8.013)	(8.029)	(7.404)	(9.323)							
	***	***	***	***	***							
Observations	491	491	491	491	491	2,983	2,983	2,983	2,983	1,527	976	480
Pseudo R-squared						0.00139	0.000559	0.000562	0.00230	0.00442	0.00853	0.0193
Log likelihood						-2044	-2046	-2046	-2042	-1039	-666.2	-324.5
Chi-squared						3.277	1.494	2.096	6.372	7.039	9.816	13.74
R-squared	0.015	0.003	0.016	0.016	0.018							
Adj. R-squared	0.00641	-0.00282	0.00194	0.00130	0.00391							
F-test	1.503	0.552	0.976	0.960	1.033							

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10b: Dependent variable: Share of categorization errors in Survey 1 and categorization errors in Survey 2. Male deputies.

	(1) Share All	(2) Share All	(3) Share All	(4) Share LW	(5) Share RW	(6) Error All	(7) Error All	(8) Error All	(9) Error All	(10) Error LW	(11) Error RW	(12) Error U
Red tie	0.0692		0.0680	0.0711	0.0599	0.0776			0.0810	0.113	0.0893	-0.0303
	(2.625)***		(2.553)**	(2.530)**	(2.506)**	(2.218)**			(2.392)**	(2.997)***	(1.559)	(-0.371)
Other tie	0.0320		0.0317	0.0319	0.0312	0.0548			0.0534	0.0481	0.0996	-0.0151
	(1.414)		(1.396)	(1.337)	(1.533)	(1.890)*			(1.892)*	(1.302)	(2.400)**	(-0.262)
Open collar	0.0373		0.0432	0.0441	0.0412	0.0613			0.0584	0.179	0.0306	
	(0.461)		(0.528)	(0.524)	(0.536)	(0.583)			(0.567)	(1.618)	(0.182)	
Beard	0.0209		0.0205	0.0267	0.00382	0.0113			0.00786	0.0668	-0.0372	-0.00808
	(0.512)		(0.496)	(0.639)	(0.0924)	(0.246)			(0.176)	(1.204)	(-0.479)	(-0.113)
Moustache	0.0433		0.0435	0.0437	0.0432	-0.0207			-0.0188	-0.0274	-0.00487	0.0694
	(0.810)		(0.803)	(0.788)	(0.818)	(-0.282)			(-0.264)	(-0.341)	(-0.0415)	(0.480)
Non Caucasian	-0.0263		-0.0203	-0.0144	-0.0369	0.144			0.135	0.132	0.127	
	(-0.256)		(-0.197)	(-0.136)	(-0.380)	(1.766)*			(1.651)*	(1.349)	(1.207)	
Smile	-0.0199		-0.0191	-0.0213	-0.0130	-0.00779			-0.00943	-0.0358	0.0158	0.0184
	(-1.033)		(-0.985)	(-1.041)	(-0.754)	(-0.304)			(-0.379)	(-1.155)	(0.406)	(0.351)
Glasses	0.0329		0.0314	0.0327	0.0279	-0.00558			-0.000479	-0.00796	-1.42e-05	0.0328
	(1.676)*		(1.558)	(1.539)	(1.532)	(-0.214)			(-0.0185)	(-0.240)	(-0.00034)	(0.599)
Attractiveness		-0.0292	-0.0218	-0.0182	-0.0317		0.0230		0.0196	0.0218	-0.0463	0.110
		(-1.310)	(-0.966)	(-0.758)	(-1.597)		(0.814)		(0.720)	(0.623)	(-0.990)	(1.786)*
Competence		-0.00136	0.00276	0.00365	-0.000196		-0.0326		-0.0271	-0.0600	0.0810	-0.116
		(-0.0413)	(0.0857)	(0.106)	(-0.00706)		(-0.824)		(-0.716)	(-1.216)	(1.284)	(-1.489)
Trustworthiness		0.0205	0.00521	0.00675	0.00120		-0.00816		-0.00840	0.0330	-0.0686	-0.0280
		(0.652)	(0.165)	(0.203)	(0.0419)		(-0.204)		(-0.214)	(0.690)	(-1.125)	(-0.362)

Table continued on the next page.

Table9b continued from previous page.

Right-wing respondent								0.0360	0.0347			
								(1.458)	(1.438)			
Unknown respondent								0.0390	0.0405			
								(1.408)	(1.487)			
Respondent's gender								-0.0215	-0.0195	-0.0180	-0.0781	0.0862
								(-1.054)	(-0.990)	(-0.673)	(-2.225)**	(1.538)
Constant	0.371	0.420	0.397	0.381	0.442							
	(16.82)***	(6.765)***	(6.208)***	(5.568)***	(7.909)***							
Observations	392	392	392	392	392	2,380	2,380	2,380	2,380	1,207	791	375
Pseudo R-squared						0.00378	0.000760	0.00135	0.00586	0.0119	0.0149	0.0175
Log likelihood						-1629	-1634	-1633	-1626	-814.1	-538.0	-254.9
Chi-squared						9.036	1.694	3.997	15.59	14.41	15.15	9.553
R-squared	0.029	0.005	0.032	0.031	0.035							
Adj. R-squared	0.00914	-0.00313	0.00373	0.00284	0.00743							
F-test	1.326	0.637	1.114	1.079	1.251							

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10c: Dependent variable: Share of categorization errors in Survey 1 and categorization errors in Survey 2. Female MPs.

	(1) Share All	(2) Share All	(3) Share All	(4) Share LW	(5) Share RW	(6) Error All	(7) Error All	(8) Error All	(9) Error All	(10) Error LW	(11) Error RW	(12) Error U
Suit	-0.00129		-0.000356	0.00456	-0.0139	-0.0300			-0.0241	-0.00905	-0.0193	-0.0183
	(-0.0242)		(-0.00646)	(0.0800)	(-0.258)	(-0.428)			(-0.335)	(-0.105)	(-0.175)	(-0.140)
Jewelry	0.0241		0.0226	0.0221	0.0245	0.0168			0.0151	0.0143	-0.0429	0.0997
	(0.578)		(0.535)	(0.498)	(0.620)	(0.281)			(0.258)	(0.202)	(-0.438)	(0.781)
Blond hair	-0.00602		-0.0109	-0.00941	-0.0148	-0.0325			-0.0251	0.0408	-0.0469	-0.110
	(-0.147)		(-0.263)	(-0.208)	(-0.410)	(-0.561)			(-0.434)	(0.540)	(-0.443)	(-0.912)
Short hair	0.0244		0.0280	0.0301	0.0227	0.0390			0.0385	-0.00395	0.104	0.0791
	(0.589)		(0.658)	(0.665)	(0.587)	(0.542)			(0.547)	(-0.0525)	(0.892)	(0.502)
Long hair	0.0376		0.0343	0.0268	0.0548	0.122			0.120	-0.00936	0.414	0.0622
	(0.477)		(0.412)	(0.306)	(0.736)	(1.236)			(1.223)	(-0.0642)	(2.686)***	(0.267)
Tied hair	0.0522		0.0483	0.0473	0.0506	0.175			0.184	0.00242	0.891	0.259
	(0.503)		(0.436)	(0.415)	(0.478)	(1.425)			(1.344)	(0.0131)	(3.101)***	(0.877)
Non Caucasian	-0.176		-0.170	-0.174	-0.158	-0.415			-0.429	-0.366	-0.547	-0.774
	(-2.96)***		(-2.67)***	(-2.60)**	(-2.74)***	(-2.260)**			(-2.334)**	(-1.413)	(-2.113)**	(-2.543)**
Smile	-0.00638		-0.00318	-0.00294	-0.00340	-0.0419			-0.0424	-0.0203	-0.112	0.0378
	(-0.168)		(-0.0820)	(-0.0713)	(-0.0973)	(-0.716)			(-0.756)	(-0.303)	(-1.234)	(0.296)
Glasses	-0.0879		-0.0922	-0.0910	-0.0960	-0.0958			-0.0966	-0.193	0.0630	-0.158
	(-1.976)*		(-1.812)*	(-1.688)*	(-2.025)**	(-1.497)			(-1.407)	(-2.247)**	(0.643)	(-1.031)
Attractiveness		0.0176	-0.0125	-0.0172	0.000384		0.0211		0.0108	-0.0609	-0.00120	0.348
		(0.475)	(-0.272)	(-0.363)	(0.00844)		(0.435)		(0.187)	(-0.810)	(-0.0151)	(2.200)**
Competence		-0.0485	-0.0235	-0.0236	-0.0235		0.0258		0.0547	-0.0102	0.103	0.114
		(-0.832)	(-0.430)	(-0.426)	(-0.396)		(0.301)		(0.646)	(-0.0858)	(0.725)	(0.669)
Trustworthiness		-0.00321	0.00518	0.00358	0.00993		-0.0413		-0.0419	0.0586	-0.0642	-0.237
		(-0.0576)	(0.0873)	(0.0599)	(0.153)		(-0.480)		(-0.464)	(0.501)	(-0.401)	(-1.764)*

Table continued on the next page.

Table9c continued from previous page.

Right-wing respondent								-0.0334	-0.0386			
								(-0.691)	(-0.814)			
Unknown respondent								-0.0270	-0.0269			
								(-0.491)	(-0.469)			
Respondent's gender								0.0135	0.0128	0.0711	-0.0110	-0.170
								(0.297)	(0.282)	(1.199)	(-0.138)	(-1.423)
Constant	0.437	0.529	0.512	0.524	0.478							
	(6.461)***	(4.777)***	(4.273)***	(4.098)***	(4.236)***							
Observations	99	99	99	99	99	603	603	603	603	320	185	98
Pseudo R-squared						0.0302	0.000712	0.000849	0.0321	0.0368	0.0942	0.127
Log likelihood						-398.4	-410.5	-410.4	-397.6	-211.2	-113.1	-57.86
Chi-squared						19.06	0.336	0.663	18.73	13.64	20.70	16.81
R-squared	0.127	0.013	0.133	0.120	0.156							
Adj. R-squared	0.0384	-0.0177	0.0117	-0.00234	0.0388							
F-test	3.110	0.388	2.460	2.275	2.919							

z-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A1: Descriptive statistics

	Mean	Std. dev.
Guessed orientation	0.53	0.5
Share of right-wing classifications	0.6	0.19
Male MP	0.81	0.4
Non Caucasian	0.02	0.15
Smile	0.52	0.5
Glasses	0.33	0.47
Red tie	0.19	0.39
Other tie	0.46	0.5
Open collar	0.03	0.17
Beard	0.06	0.24
Moustache	0.04	0.21
Suit	0.79	0.41
Jewelry	0.72	0.45
Blond hair	0.28	0.45
Short hair	0.58	0.49
Long hair	0.09	0.28
Tied hair	0.06	0.25
Attractiveness	2.15	1.00
Competence	2.6	0.98
Trustworthiness	2.53	0.99
Right-wing respondent	0.33	0.47
Unknown respondent	0.16	0.37
Respondent's gender	0.63	0.48
National office	0.11	0.31
Right-wing deputy	0.6	0.49