From the lab to the poll

The use of survey experiments in political research

Sergio Martini
sergio.martini@unisi.it

Francesco Olmastroni
olmastroni3@unisi.it

Department of Social, Political and Cognitive Sciences
University of Siena

Contents

1 Appendix A: Samples characteristics

1.1 Criteria for the content analysis of APSR and EJPR articles

In the case of APSR, considered articles include manuscripts published as “articles”, “controversies”, “forum”, “letters”, “research articles” and “research notes”. “Book reviews”, “review essays”, “review symposia”, “symposia”, “communications” and “errata” are excluded. As for EJPR, “articles” “fora” and “research notes” are used for computation, whereas “introductions” to the annual review or to special issues, “nominations and reflections”, “prefaces”, “errata”, “editorials” and all articles published in the ‘Political Data Yearbook’ are excluded. Year of publication refers to the time at which the article appeared in a volume of the two journals. ‘Early view’ and ‘online first’ articles are excluded from this analysis.

	APSR			EJPR		
	N	%	% of articles	N	%	% of articles
Field	28	29.2	30.1	2	11.1	12.5
Natural	3	3.1	3.2	0	0.0	0.0
Laboratory	30	31.3	32.3	1	5.6	6.3
Survey	35	36.5	37.6	15	83.3	93.8
Total	96	100.0	103.2	18	100.0	112.5

Table 1: Typology of experiments in ‘APSR’ (N=93) and ‘EJPR’ (N=16) articles, 2000-2019.

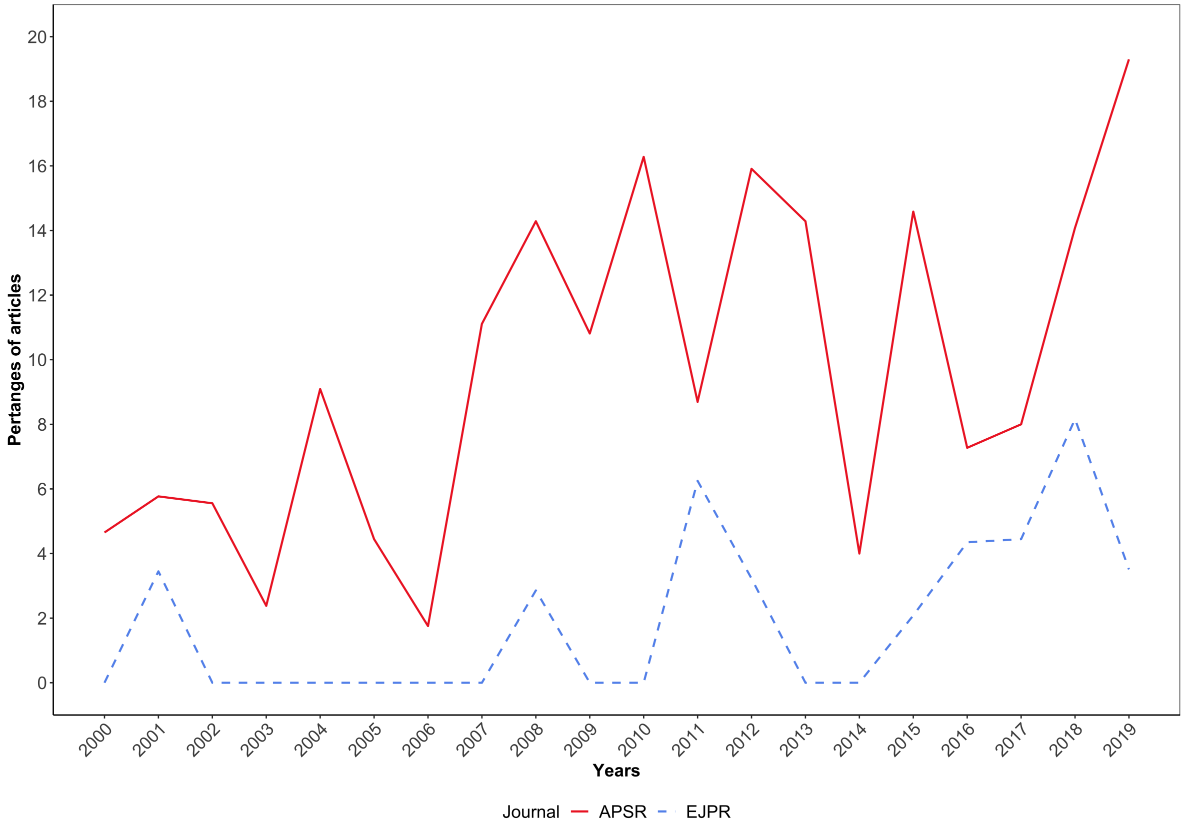

Figure 1: Percentage of experimental articles in APSR and EJPR, 2000-2019

1.2 The EUENGAGE and DISPOC-GfK surveys

Table 2 shows a comparison of the distribution of some basic socio-demographic characteristics (i.e., gender, age, educational attainment) of the EUENGAGE sample, the DISPOC-GfK sample and the Italian adult population, respectively (sources: ISTAT 2017a, 2017b, 2019a, 2019b). As it can be seen, the distribution of the considered variables in the EUENGAGE sample approximates quite well that obtained through Census data. The only exception is educational attainment, with more educated people over-represented and less educated people under-represented. During the fieldwork, 2,735 invitations were sent, while the overall number of completed interviews was 1,278, with a response rate of 47.7%. A small number of cases was excluded because of device failure or screened out (screening questions were country (1) and age (2), asked in that order). Panellists declaring to reside in the country and to be 18 or older were screened in, whereas those who did not accept the confidentiality statement at the beginning of the questionnaire were not allowed to take the survey. Finally, some respondents were excluded because of quota restrictions (gender*age-group, region), while 134 abandoned before completing the interview. Respondents were recruited from the pool of panellists who had completed the first wave of the EUENGAGE citizen survey (June 14 – July 20, 2016).

	ISTAT 2017	EUENGAGE (2017) Italy - Wave 2	ISTAT 2019	DISPOC-GfK (2019) Wave 2
Gender				
Male	52.0	48.2	51.7	41.8
Female	48.0	51.8	48.3	58.2
Age				
14-17			4.3	2.1
18-24	8.2	4.6	7.9	3.7
25-39	21.0	25.8	19.4	22.4
40-54	28.5	30.7	26.8	38.1
55-64	15.6	17.7	15.6	17.0
65+	26.7	21.6	26.0	16.7
Education				
Primary or none	18.7	2.07	16.2	4.3
Secondary	66.3	59.0	68.8	74.6
Tertiary	15.0	38.9	15.0	21.1

Table 2: Samples' characteristics. Descriptive statistics for education refer to population aged 20 years or older in the case of ISTAT 2017 and to population aged 15 years or older in the case of ISTAT 2019.

In the DISPOC-GfK survey subjects were recruited offline by the survey company, regardless of their access to the Internet. If they accepted, they were, then, provided with a tablet to participate in incentivised online surveys. In the first stage, households were selected via a stratified random sampling using region, urban area and number of household members as main strata. In the second stage, a sample of individuals is selected via stratified proportional random sampling according to gender, age-group, region and demographic size of the municipality of residence. To obtain the final sample for the survey, individuals were randomly selected according to quotas of gender, age-group, geographical area and demographic size of the municipality of residence. As in the previous case, the DISPOC-GfK sample is skewed with respect to education, though deviations are smaller.

We applied post-stratification weights (capped between 0.33 and 3.0), based on gender and age-group (interlocked), geographical area and size of the municipality of residence (interlocked), and educational attainment to reflect the actual demographic composition of Italy's adult population. In DISPOC-GfK wave 1, a total of 4,244 invitations were sent to panellists. Overall, 3,523 interviews were completed, with a response rate of 83%. In DISPOC-GfK wave 2, the number of invitations was 4,244, whereas the number of completed interviews was 3,224, with a response rate of 76%.

2 Appendix B: Experimental stimuli

2.1 Factorial experiment

Let's talk now about an individual who is interested in migrating to Italy.
Please, take a minute and read about his background.

He is
[CHOOSE 1 AT RANDOM: Syrian/Ukrainian man],
[CHOOSE 1 AT RANDOM: skilled/not skilled],
[CHOOSE 1 AT RANDOM: fleeing from war/looking for a job]

Given what you know about this potential immigrant do you think his application for asylum should be approved or rejected?

The application should be approved
The application should be rejected

Table 3: Vignette with manipulated conditions

2.2 Conjoint experiment

"There is some talk about the characteristics a candidate should have to enter politics at the European level. We will provide you with several pieces of information about people who might have run for the European elections. For each pair of candidates, please indicate which of the two you would personally have preferred to win a seat in the European Parliament. This exercise is purely hypothetical. Even if you aren't entirely sure, please indicate which of the two you prefer."

[RANDOM ORDER]	**Candidate 1**	**Candidate 2**
Gender	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Male	Male
	Female	Female
Job experience	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Was a manual worker outside of politics	Was a manual worker outside of politics
	Was a farmer outside of politics	Was a farmer outside of politics
	Was a manager outside of politics	Was a manager outside of politics
	Was a university professor outside of politics	Was a university professor outside of politics
	Was an engineer outside of politics	Was an engineer outside of politics
	Was a professional politician	Was a professional politician
Communication	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Uses proper and refined language to convey messages	Uses proper and refined language to convey messages
	Uses coarse and rude type of language to convey messages	Uses coarse and rude type of language to convey messages
Social skills	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Tends to be emotionally involved in problems and really enjoy caring for other people	Tends to be emotionally involved in problems and really enjoy caring for other people
	Tends to be distant without involving emotionally in problems of other people	Tends to be distant without involving emotionally in problems of other people
Integrity	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Has a clean criminal record	Has a clean criminal record
	Is under investigation for using public reimbursements for personal expenses	Is under investigation for using public reimbursements for personal expenses
Competence	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Has no skills in specific policy areas and does not speak English	Has no skills in specific policy areas and does not speak English
	Has skills in specific policy areas but does not speak English	Has skills in specific policy areas but does not speak English
	Has no skills in specific policy areas but speaks English fluently	Has no skills in specific policy areas but speaks English fluently
	Has skills in specific policy areas and speaks English fluently	Has skills in specific policy areas and speaks English fluently
View of role	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Is focused on needs of the wide public even at the expenses of the interests of voters he/she represents and promises he/she made	Is focused on needs of the wide public even at the expenses of the interests of voters he/she represents and promises he/she made
	Is focused on the interests of voters he/she represents and promises he/she made, at expenses of the general public	Is focused on the interests of voters he/she represents and promises he/she made, at expenses of the general public
Leadership	[CHOOSE 1 AT RANDOM]	[CHOOSE 1 AT RANDOM]
	Does not provide strong and charismatic leadership, but s/he is able to listen at different views	Does not provides strong and charismatic leadership, but s/he is able to listen at different views
	Provides strong and charismatic leadership, but s/he falls short of listening at different views	Provides strong and charismatic leadership, but s/he falls short of listening at different views

Table 4: Vignette on candidate preference for the EU Parliament

[Do not show the same profile for the 2 candidates.]

If you had to choose between them, which of these two candidates should be given priority to win a seat in the European Parliament?

[Rotate items]

Candidate 1
Candidate 2

Rotate following questions

On a scale from 1 to 7, where 1 indicates that you do not favour at all the candidate and 7 indicates that you favour completely the candidate, how would you rate Candidate 1?

1 Do not favour at all
2
3
4
5
6
7 Favour completely

On a scale from 1 to 7, where 1 indicates that you do not favour at all the candidate and 7 indicates that you favour completely the candidate, how would you rate Candidate 2?

1 Do not favour at all
2
3
4
5
6
7 Favour completely

We now ask you to repeat the exercise. You will be shown another pair of candidates, please indicate again which of the two candidates would have preferred to win a seat in the EU Parliament.

[Repeat the previous exercise and related questions. Order of categories must be the same of that shown previously. Do not show the same profile for the 2 candidates and do not show the same profile of the 2 candidates presented in the previous experiment.]

3 Appendix C: Diagnostics

3.1 Balance tests: factorial experiment

	G2	G3	G4	G5	G6	G7	G8
Gender (ref: male)							
Female	-0.065	0.106	-0.195	-0.082	0.133	0.068	-0.286
	(0.226)	(0.230)	(0.229)	(0.227)	(0.228)	(0.227)	(0.237)
Age	0.001	0.007	-0.009	-0.001	-0.003	-0.002	-0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Education (ref. Primary)							
Secondary not completed	2.167	0.499	0.602	0.045	0.580	2.048	0.377
	(1.214)	(0.823)	(0.914)	(0.895)	(0.856)	(1.219)	(0.829)
Secondary completed	1.290	-0.359	0.010	0.077	-0.218	1.305	-0.593
	(1.138)	(0.705)	(0.795)	(0.737)	(0.741)	(1.139)	(0.709)
Degree	1.559	0.058	0.178	0.236	-0.140	1.550	-0.319
	(1.153)	(0.728)	(0.818)	(0.760)	(0.766)	(1.154)	(0.733)
Post-graduate	1.273	-0.889	-0.046	-0.525	-0.479	1.229	-0.902
	(1.164)	(0.768)	(0.831)	(0.790)	(0.786)	(1.166)	(0.762)
Ideology (ref: Left)							
Centre	-0.047	-0.279	0.039	0.289	-0.450	-0.280	0.257
	(0.303)	(0.306)	(0.314)	(0.307)	(0.308)	(0.304)	(0.323)
Right	-0.095	-0.320	-0.114	0.037	-0.074	-0.378	0.037
	(0.340)	(0.342)	(0.346)	(0.343)	(0.334)	(0.339)	(0.360)
Party Id (ref: Don't know)							
Democratic Party	0.047	-0.124	0.244	0.003	0.093	-0.243	-0.095
	(0.367)	(0.391)	(0.396)	(0.376)	(0.377)	(0.379)	(0.397)
Go Italy	-0.138	-0.580	0.396	0.119	-0.314	-0.207	-0.088
	(0.491)	(0.561)	(0.496)	(0.479)	(0.520)	(0.506)	(0.509)
Northern League	-0.127	0.126	0.769	0.017	0.117	-0.081	-0.113
	(0.471)	(0.469)	(0.453)	(0.464)	(0.470)	(0.471)	(0.480)
Five Star Movement	-0.140	0.372	-0.037	-0.304	0.334	-0.086	-0.254
	(0.344)	(0.342)	(0.374)	(0.355)	(0.342)	(0.343)	(0.361)
Other	-0.428	-0.306	0.293	0.115	-0.561	-0.152	-0.044
	(0.399)	(0.409)	(0.399)	(0.378)	(0.419)	(0.384)	(0.400)
No party Id	-0.138	0.458	0.620	-0.206	-0.060	-0.246	-0.107
	(0.520)	(0.492)	(0.497)	(0.533)	(0.538)	(0.533)	(0.536)
Attitudes towards immigration (scale)	-0.071	-0.216	-0.145	-0.090	0.001	-0.143	-0.169
	(0.144)	(0.145)	(0.145)	(0.143)	(0.145)	(0.144)	(0.150)
Constant	-1.262	-0.059	0.134	-0.068	0.347	-1.054	1.017
	(1.253)	(0.885)	(0.956)	(0.906)	(0.911)	(1.253)	(0.892)
N	1,276	1,276	1,276	1,276	1,276	1,276	1,276
Pseudo R-squared	0.017	0.017	0.017	0.017	0.017	0.017	0.017
Log-likelyhood empty model	-2.655.821	-2.655.821	-2.655.821	-2.655.821	-2.655.821	-2.655.821	-2.655.821
Log-likelyhood	-2.605.686	-2.605.686	-2.605.686	-2.605.686	-2.605.686	-2.605.686	-2.605.686
Log-likelyhood ratio	91.760	91.760	91.760	91.760	91.760	91.760	91.760
P-value	0.818	0.818	0.818	0.818	0.818	0.818	0.818

Table 5: Balance tests for the factorial experiment, multinomial logistic regression. Baseline: G1 (He is a Syrian Man, skilled, fleeing from war). Standard errors in parentheses: * p<0.05; ** p<0.01. Results confirm that the random procedure was correct: none of the variables are statistically significant, pseudo-R squared terms are always very small and the likelihood ratio chi-squared tests are not significant.

3.2 Balance tests: conjoint experiment

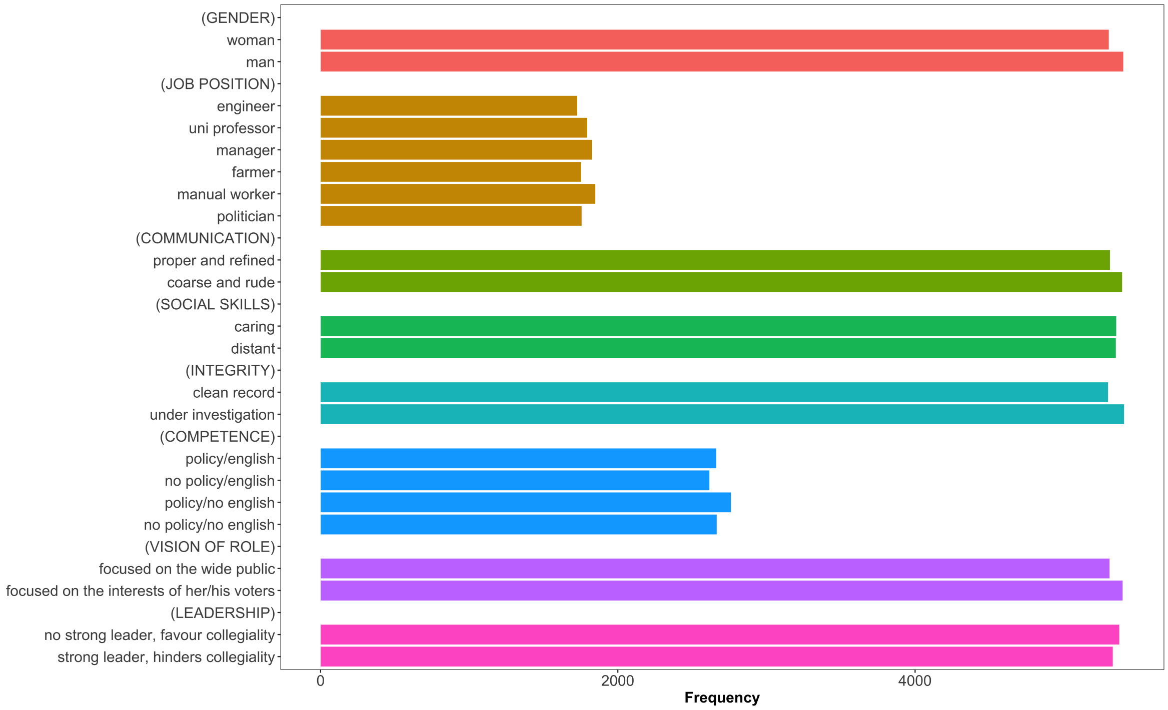

Figure 2: Frequencies of attributes by experimental trait.

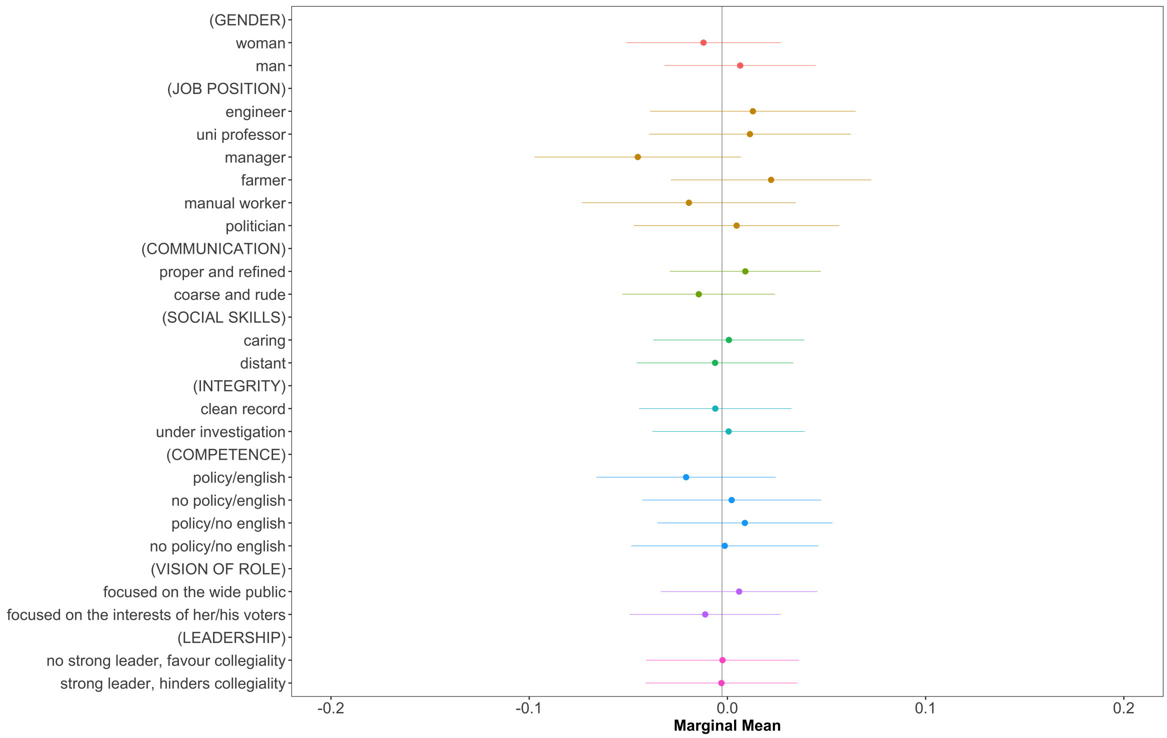

Figure 3: Comparing levels of populism across attributes and experimental traits. Confidence intervals for each feature set around the grand mean; this should indicate that imbalance is not a problem.

4 Appendix D: Models and robustness checks

4.1 Logistic models: factorial experiment

Our dependent variable is dichotomous in format, we have first estimated a logistic regression model to identify the effects of our treatments (Long 1997), expressed via the following general equation:

$$P(y_i = 1) = \text{Logit}^{-1}(X_i\beta) \tag{1}$$

where we have modelled the probability of success (y=1) in each observation i, namely whether the respondent approved the asylum application. This is linked to the linear predictor $(X_i\beta)$ via the inverse of the logit function. The linear predictor $(X_i\beta)$ is a combination of an intercept, β_0, and other covariates; in our case, these are the dummy variables capturing the effects of the different treatments, as well as their interactions. Thus, our final solution has been a three-way interaction model represented via the following notation:

$$\begin{aligned} P(y_i = 1) = \text{Logit}^{-1}(\beta_0 - \beta_1 \text{qualification}_i^{skilled} + \beta_2 \text{reason}_i^{war} \\ + \beta_3 \text{origin}_i^{Syrian} + \beta_4 \text{qualification}_i^{skilled} * \text{reason}_i^{war} \\ + \beta_5 \text{qualification}_i^{skilled} * \text{origin}_i^{Syrian} \\ + \beta_6 \text{reason}_i^{war} * \text{origin}_i^{Syrian} \\ + \beta_7 \text{qualification}_i^{skilled} * \text{reason}_i^{war} * \text{origin}_i^{Syrian}) \end{aligned} \tag{2}$$

This model allowed us to test H1 through H4 (model 1 in Table 6). When considering, instead, the last expectation (H5a, H5b, H5c) (model 2 in Table 7), we have employed a two-way interaction model via the following equation:

$$\begin{aligned} P(y_i = 1) = \text{Logit}^{-1}(\beta_0 - \beta_1 \text{qualification}_i^{skilled} + \beta_2 \text{reason}_i^{war} \\ + \beta_3 \text{origin}_i^{Syrian} + \beta_4 \text{ideology}_i^{left} + \beta_5 \text{ideology}_i^{right} \\ + \beta_6 \text{qualification}_i^{skilled} * \text{ideology}_i^{left} \\ + \beta_7 \text{qualification}_i^{skilled} * \text{ideology}_i^{right} \\ + \beta_8 \text{reason}_i^{war} * \text{ideology}_i^{left} \\ + \beta_9 \text{reason}_i^{war} * \text{ideology}_i^{right} \\ + \beta_{10} \text{origin}_i^{Syrian} * \text{ideology}_i^{left} \\ + \beta_{11} \text{origin}_i^{Syrian} * \text{ideology}_i^{right}) \end{aligned} \tag{3}$$

For the sake of simplicity, we display results showing predicted probabilities of approval of asylum applications as a function of our covariates. To get the average treatment effects, we computed average marginal effects and performed Wald tests.

	(1)	(2)
Skilled	0.866**	0.780**
	(0.230)	(0.250)
Fleeing from war	0.848**	0.989**
	(0.232)	(0.253)
Ukrainian	-0.343	-0.469
	(0.236)	(0.249)
Skilled*Fleeing from war	-0.228	
	(0.341)	
Skilled*Ukrainian	0.178	
	(0.332)	
Fleeing from war*Ukrainian	0.184	
	(0.334)	
Skilled*Fleeing from war*Ukrainian	-0.234	
	(0.483)	
Centre		-0.961**
		(0.298)
Right		-1.310**
		(0.306)
Skilled*Centre		0.051
		(0.323)
Skilled*Right		0.112
		(0.322)
Fleeing from war*Centre		-0.194
		(0.324)
Fleeing from war*Right		-0.257
		(0.324)
Ukrainian*Centre		0.498
		(0.321)
Ukrainian*Right		0.214
		(0.320)
Constant	-0.164	0.624**
	(0.159)	(0.222)
N	1,273	1,273

Table 6: Results of the factorial experiment: logit model (unweighted models). Standard errors in parenthesis: * $p<0.05$; ** $p<0.01$.

Last, some robustness checks were performed. To improve the quality of the data (Baker et al. 2010) and control for respondent's attention, we removed from the analyses those respondents who completed the interview in less than 50% of the median response time (N=68). Then, we weighted the data to adjust our sample to known population distributions of selected socioeconomic and demographic variables. In this respect, we applied a capped weight (between 0.2 and 5.0) based on gender, age-group, region, educational attainment and Internet usage at country level to reflect the actual demographic composition of each country's adult population with access to the Internet. Finally, we ran our analysis controlling for a possible order effect. In all the three cases, results do not change substantially, confirming that these three potential sources of bias are negligible (see Appendix D for further details).

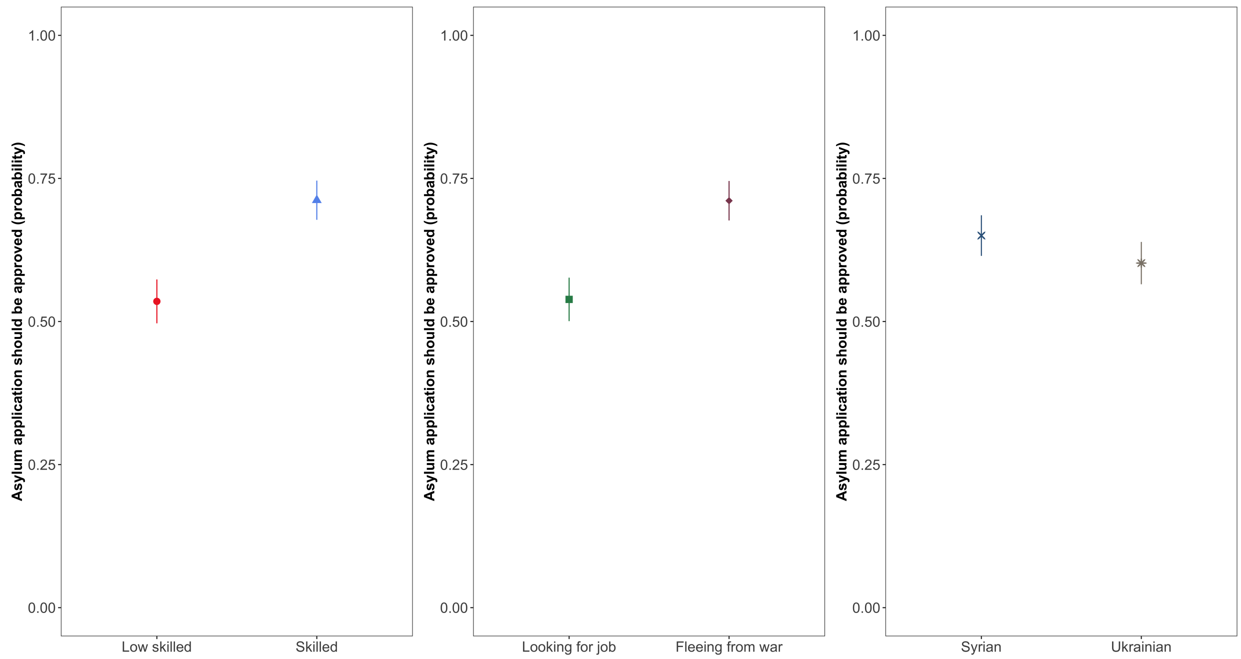

Figure 4: Main effects of skills and reasons for leaving the country and ethnic group on the probability to accept asylum applications (Model 1). Lines on both sides of the points represent 95% confidence intervals.

4.2 Logistic models (spill-over test): factorial experiment

The factorial experiment we have presented is not the only one included in the survey. In fact, other two on the topics of the economy and globalisation were assigned to our respondents, implying potential spill-overs among the three. In other words, responses to one experiment may be contaminated by those already provided in a previous one. All the three experiments, however, were presented in a randomised order, which can alleviate this type of bias.

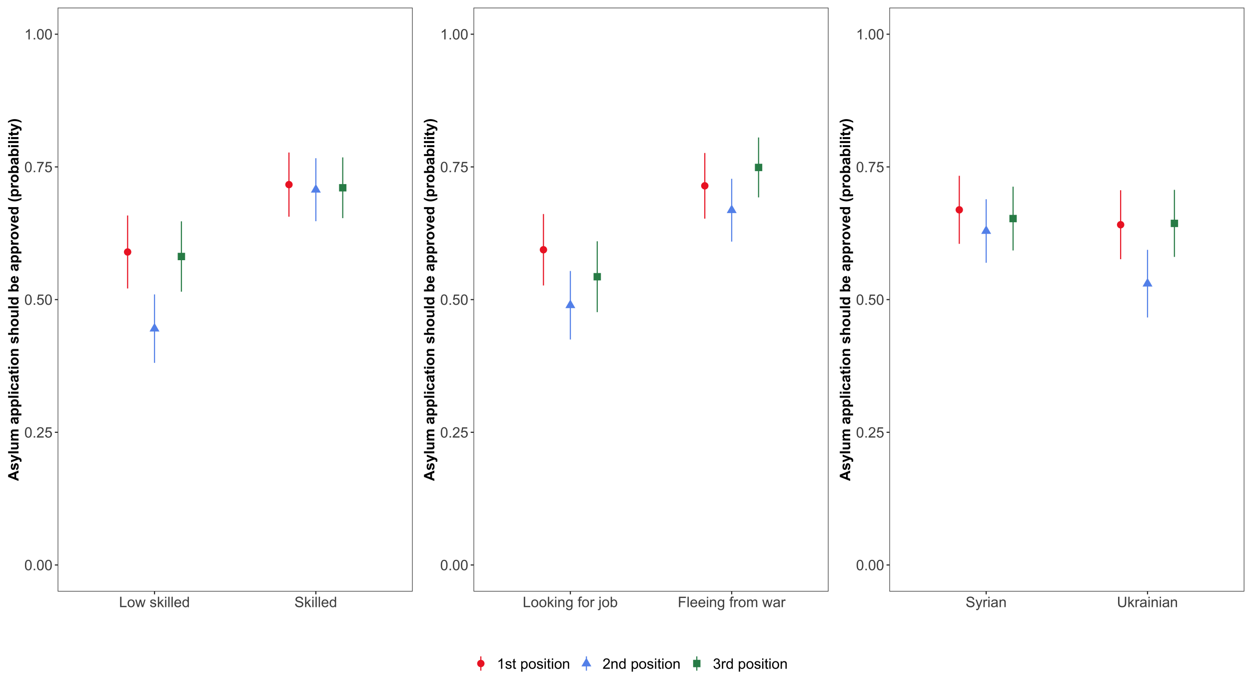

Figure 5: Main effects of skills, reasons for leaving the country and ethnic group on the probability to accept asylum applications by the order in which the experiment was presented. Lines on both sides of the points represent 95% confidence intervals.

To check this possibility, we ran a logistic model with 2-way interactions between each of our treatments and the order in which the experiment on immigration was presented. As it can be observed from Figure 5, the effect of the level qualification – the approval rate gap between low-skilled and skilled – seems to be larger when our experiment was presented in the second position (13 per cent higher than when it was either in the first position (Chi-squared=4.34; $p<0.05$) or in the third one (Chi-squared=4.37; $p<0.05$). Nevertheless, if we run a three-way interaction logistic model controlling for the order of the experiment and calculate the predicted probabilities for each of our treatments, we get similar results, suggesting that bias due to potential spill-over effects among experiments is minimal.

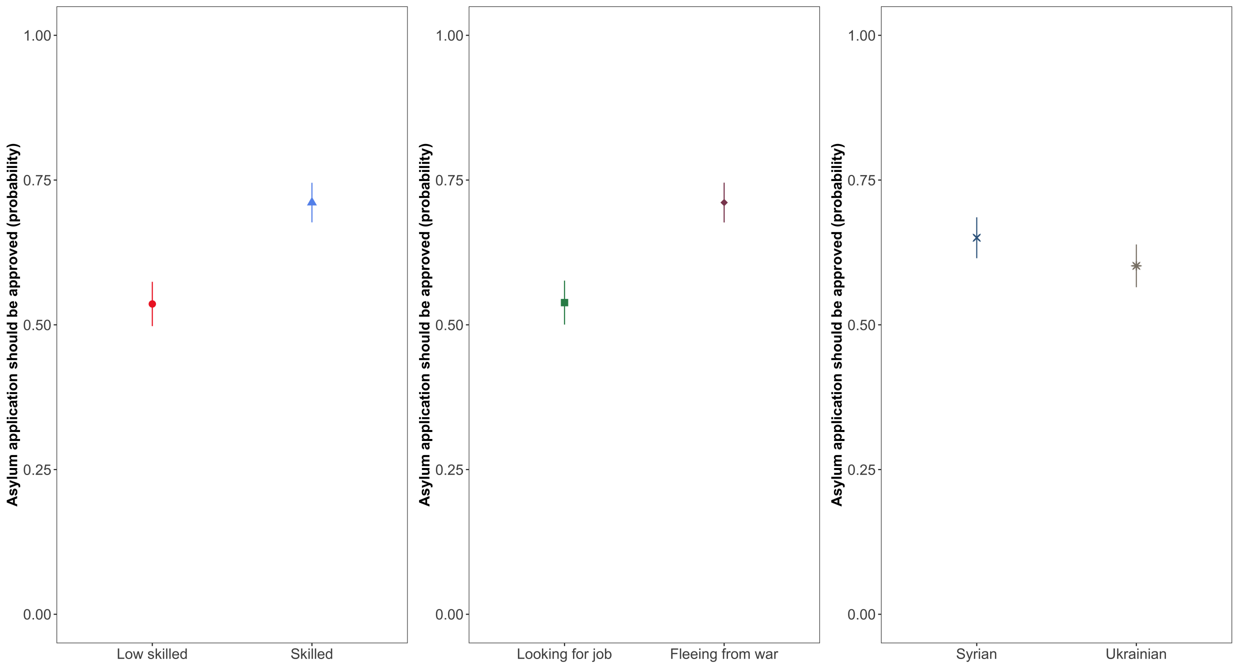

Figure 6: Main effects of skills, reasons for leaving the country and ethnic group on the probability to accept asylum applications conditioned on the order in which the experiment was presented. Lines on both sides of the points represent 95% confidence intervals.

4.3 Logistic models (weighted): factorial experiment

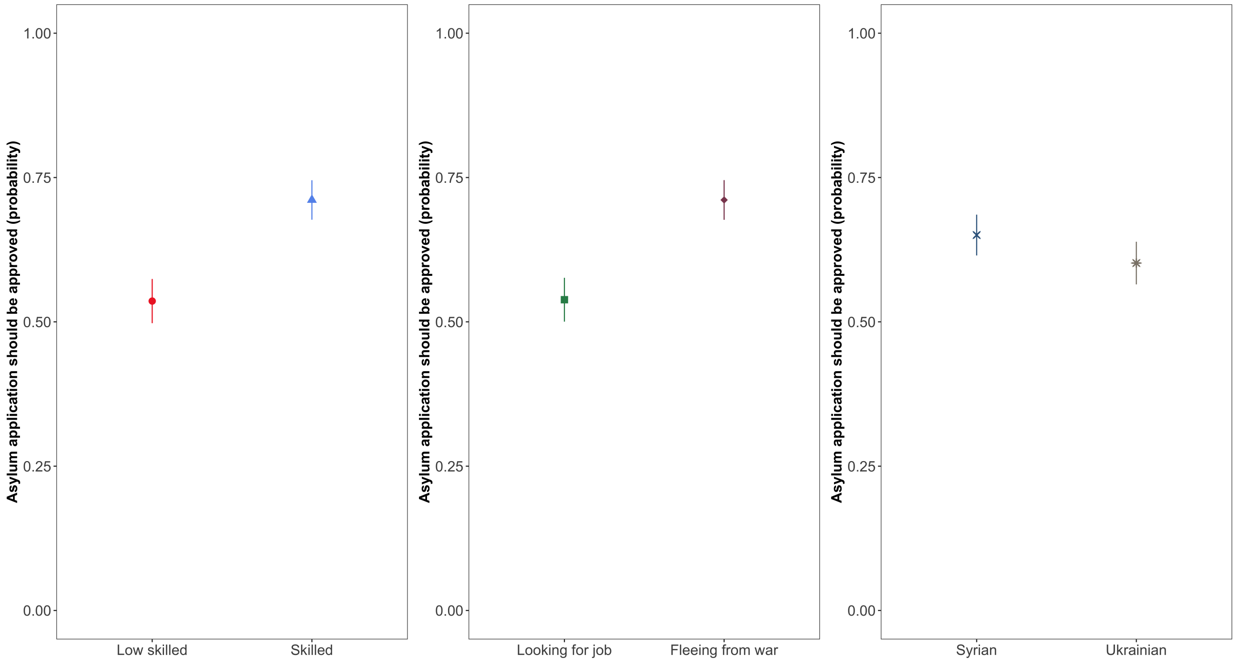

Figure 7: Main effects of skills, reasons for leaving the country and ethnic group on the probability to accept asylum applications (weights applied). Lines on both sides of the points represent 95% confidence intervals.

4.4 Logistic models (without speeders): factorial experiment

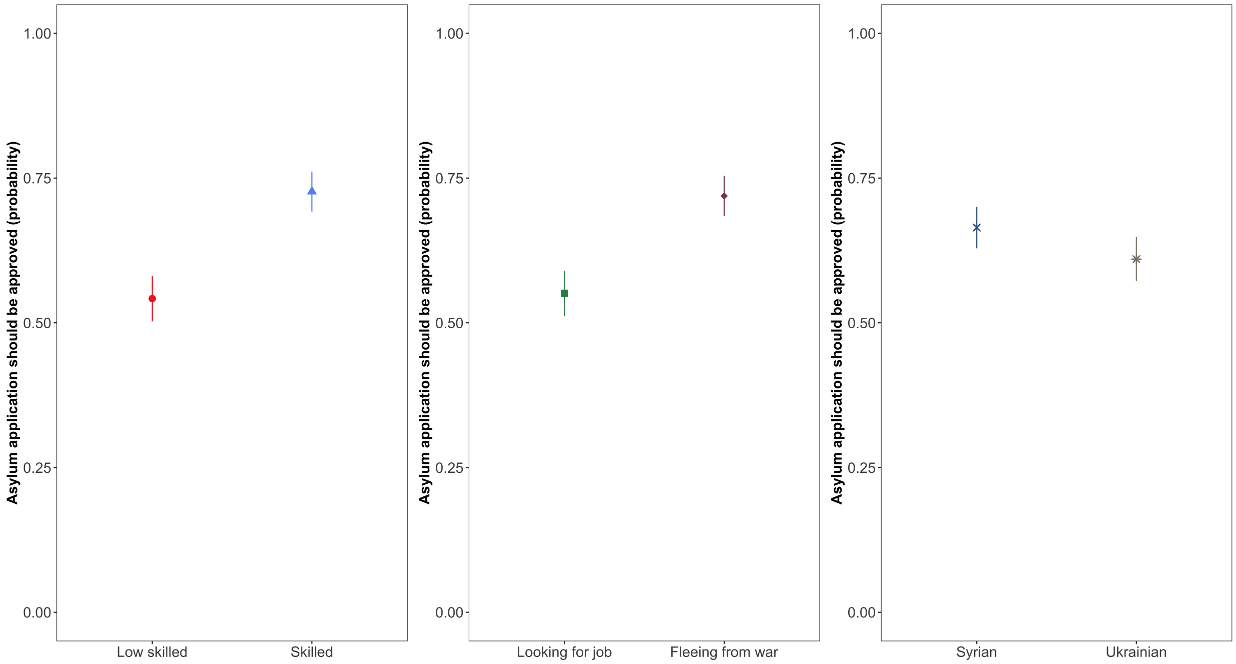

Figure 8: Main effects of skills, reasons for leaving the country and ethnic group on the probability to accept asylum applications (filtering speeders). Lines on both sides of the points represent 95% confidence intervals.

4.5 Populist scale used in the conjoint experiment

Variable	Factor	Uniqueness
pop1	0.638	0.593
pop2	0.619	0.617
pop3	0.627	0.607
pop4	0.611	0.627
pop5	0.594	0.647
pop6	0.538	0.711

Table 7: Results of exploratory factor analysis.

How much do you agree or disagree with each one of the following statements...

pop1 The politicians in the [country] parliament need to follow the will of the people.
pop2 The people, and not politicians, should make our most important policy decisions.
pop3 The political differences between the elite and the people are larger than the differences among the people.
pop4 I would rather be represented by a citizen than by a specialised politician.
pop5 Elected officials talk too much and take too little action.
pop6 What people call "compromise" in politics is really just selling out on one's principles.

1. Strongly disagree
2. Fairly disagree
3. Neither... nor...
4. Fairly agree
5. Strongly agree
98. NA
99. DK

4.6 Linear probability models: conjoint experiment

Following Hainmueller, Hopkins and Yamamoto (2013), we estimated a linear probability model (Long 1997) to assess the role of the different profile traits and the relative assigned attributes on people's candidate choice, summarised by the following notation:

$$y_i = X_i \beta + \varepsilon_i \tag{4}$$

where y_i is whether the respondent chose a certain candidate (y_i=1) or not (y_i=0), while $X_i \beta$ is a linear predictor of explanatory variables with their respective β coefficients and ε_i is a residual error term.

We ran a series of robustness tests to check reliability of our results. First, we ran models using the 7-point candidate rating scale recoding it as 1 if the rating was above the midpoint and 0 otherwise. Results concerning general AMCEs are overall robust. Then, we ran models removing subjects who did not pass any of the attention checks included in the survey and again AMCEs do not change substantially.[1] Finally, we ran sub-group analysis using the populist scale included in the previous wave to avoid potential priming and the impact of corruption persists while dislike of professional politician disappears. It is worth mentioning that in wave 2 some cases were re-introduced so we lose information for these cases when running these tests.

Lastly, it is worth noting that a problem with randomisation occurred while running the conjoint experiment. In fact, in the original form, job position also included the attribute "Was an artisan outside politics". However, due to a problem in the randomisation procedure, it was shown only to 420 subjects. To overcome this problem, we dropped this attribute and the subjects assigned to it from the analysis with no harm. In fact, for the rest of the sample randomisation worked successfully and we have completed data. We also ran the analysis on the full sample and results do not change substantially.

To estimate models and report results, we used the Cregg R-package by Leeper (2018).

[1] Checks read as follows: (1) "Now, before we go any further, we ask you to select the value four on a scale from 1 to 7, where 1 means "strongly disagree" and 7 "strongly agree" (answers recorded on a 7-point scale, 1 completely disagree - 7 completely agree); (2) "Some people are interested in politics but don't read carefully the questions they are asked. Other people are not interested in politics but pay attention to the questions they are asked. To prove that you have read this question, please select the "fairly disagree" option. Please respond by following our instructions" (answers recorded on a 4-point scale: 1 completely agree; 2 fairly disagree; 3 fairly agree; 4 completely agree); (3) "Many modern decision-making theories recognise that decisions are made with attention to their possible implications. To prove that you have read our instructions, please select the third and fourth options from those listed below" (answers: 1 Terrorism; 2 Immigration; 3 Economic growth; 4 Unemployment; 5 Taxes; 6 Precarious job; 7 Climate change; 8 European Union crisis; 9 Populism; 10 Racism; 11 Crime organisations; 12 Political corruption,; 96 None of the above; 97 Don't know / Prefer not to answer).

	All	Populists	Non-populists
Woman	0.006	0.008	0.013
	(0.012)	(0.018)	(0.017)
Manual worker	0.055**	0.096**	-0.006
	(0.020)	(0.028)	(0.030)
Farmer	0.031	0.091**	-0.024
	(0.019)	(0.029)	(0.029)
Manager	0.038	0.060*	0.006
	(0.020)	(0.029)	(0.029)
Uni professor	0.041*	0.082**	0.009
	(0.020)	(0.029)	(0.030)
Engineer	0.054**	0.073*	0.044
	(0.020)	(0.029)	(0.030)
Proper and refined	0.076**	0.054**	0.096**
	(0.012)	(0.016)	(0.017)
Caring	0.041**	0.045*	0.027
	(0.012)	(0.019)	(0.018)
Clean record	0.223**	0.268**	0.199**
	(0.012)	(0.017)	(0.017)
Policy/no English	0.060**	0.075**	0.063*
	(0.016)	(0.024)	(0.024)
No policy/no English	0.053**	0.013	0.071**
	(0.016)	(0.024)	(0.025)
Policy/English	0.086**	0.069**	0.102**
	(0.016)	(0.022)	(0.026)
Focused on the wide public	0.033**	0.032	0.027
	(0.012)	(0.017)	(0.018)
No strong leader, favour collegiality	0.035**	0.040*	0.032
	(0.012)	(0.017)	(0.018)
Constant	0.209**	0.169**	0.241**
	(0.020)	(0.029)	(0.031)
N	10,704	4,884	4,844

Table 8: Average Marginal Conditional Effects (AMCE): conjoint experiment, all subjects, populists, non-populists. Standard errors in parentheses: * p<0.05; ** p<0.01.

4.7 Marginal Means: conjoint experiment

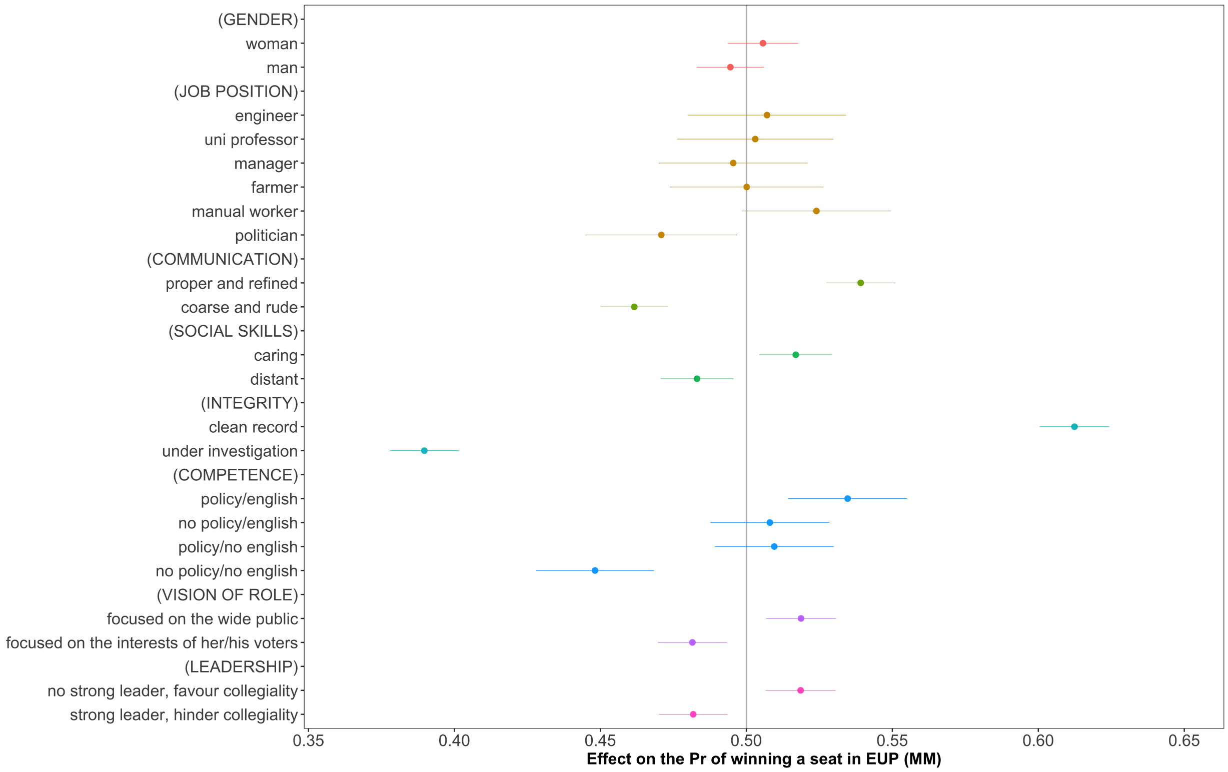

Figure 9: Marginal means (MMs). Lines on both sides of the points represent 95% confidence intervals.

4.8 Marginal Means for populists and non-populists: conjoint experiment

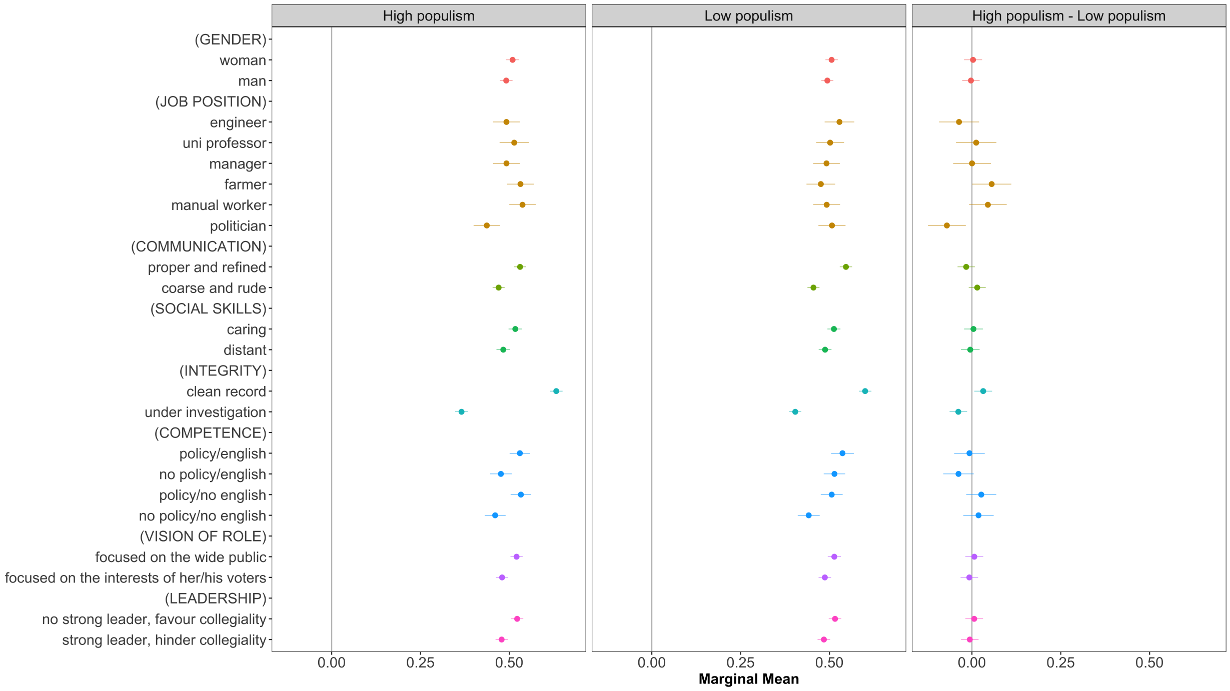

Figure 10: Marginal means (MMs) for populists and non-populists. Lines on both sides of the points represent 95% confidence intervals.

4.9 Average Marginal Conditional Effects (AMCE): conjoint experiment (full sample)

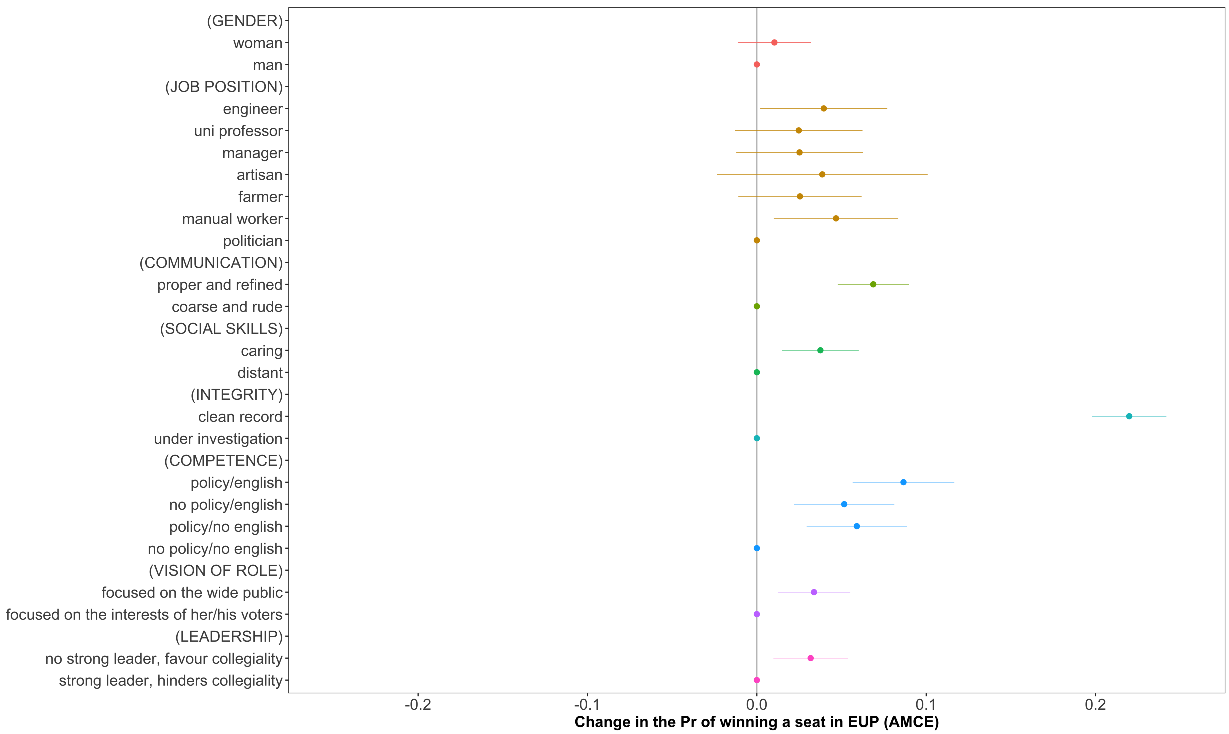

Figure 11: Average Marginal Conditional Effects (AMCE): model based on the full sample (including attribute "artisan" for job position). Lines on both sides of the points represent 95% confidence intervals.

4.10 Average Marginal Conditional Effects (AMCE): conjoint experiment (rating scale)

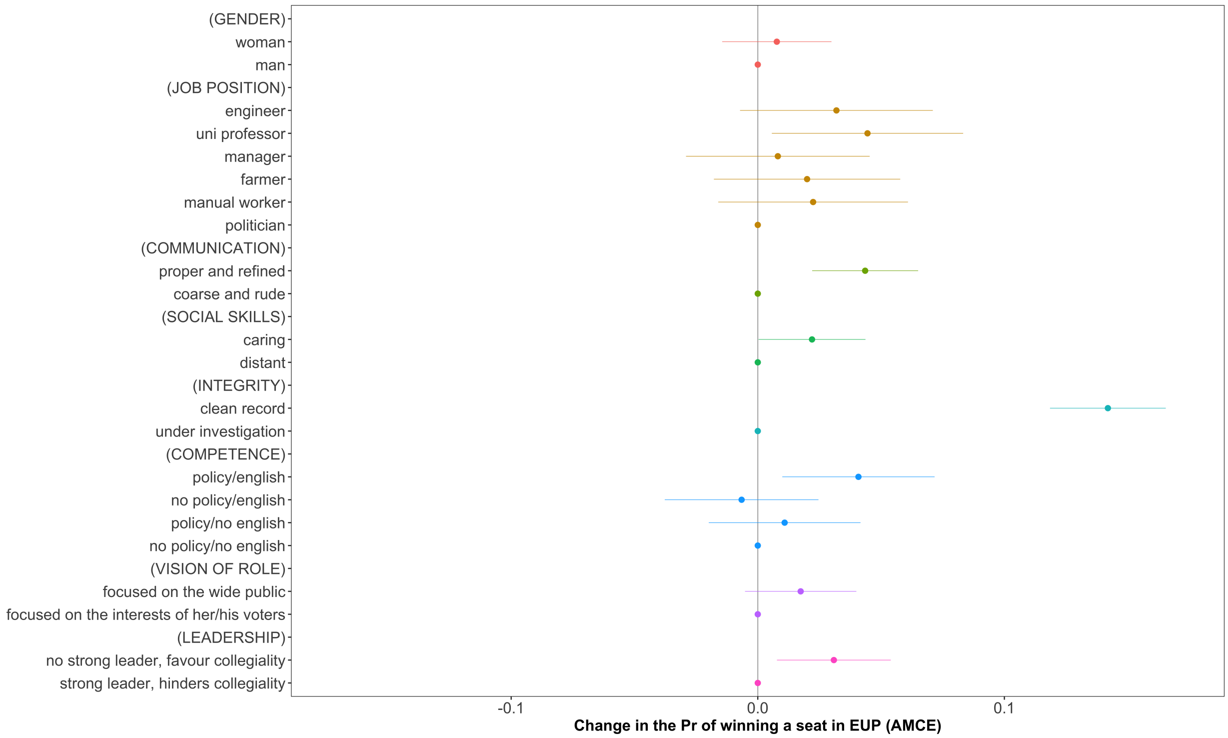

Figure 12: Average Marginal Conditional Effects (AMCE): model based on the 7-point rating scale dichotomised (1 if the rating is above the midpoint and 0 otherwise). Lines on both sides of the points represent 95% confidence intervals.

4.11 Average Marginal Conditional Effects (AMCE) for populists and non-populists: conjoint experiment

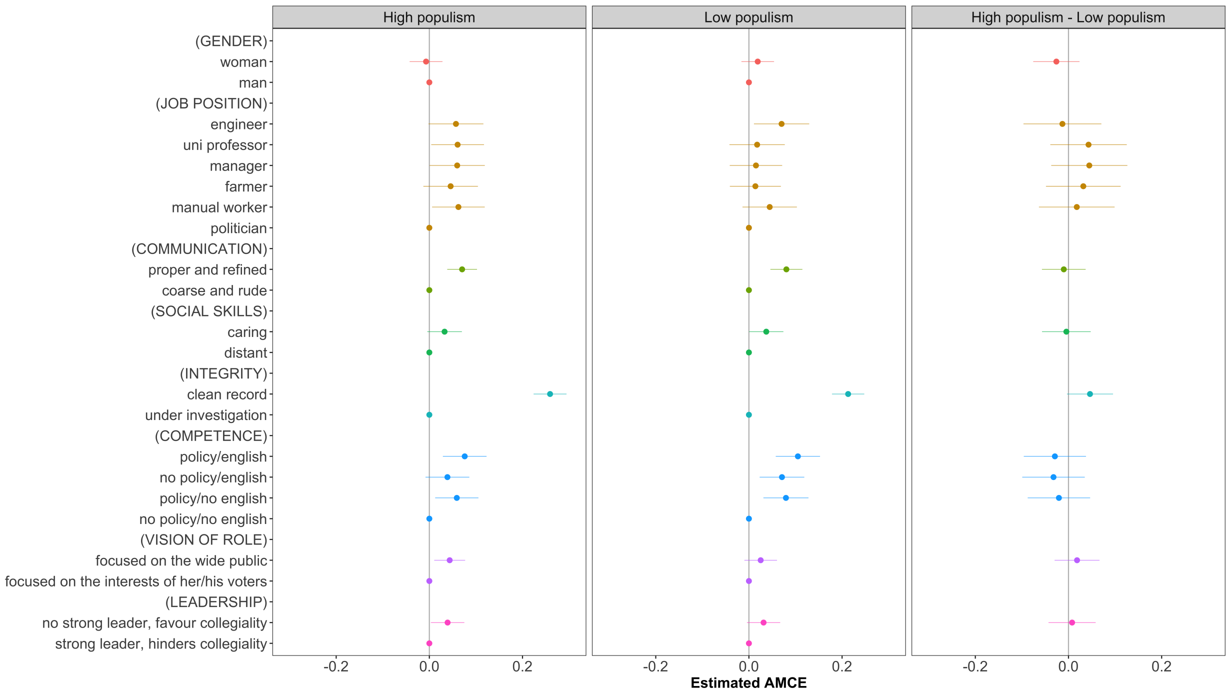

Figure 13: Average Marginal Conditional Effects (AMCE) for populists and non-populists using the populist scale from the first survey wave. Lines on both sides of the points represent 95% confidence intervals.

4.12 Average Marginal Conditional Effects (AMCE): conjoint experiment (excluding those failing attention checks)

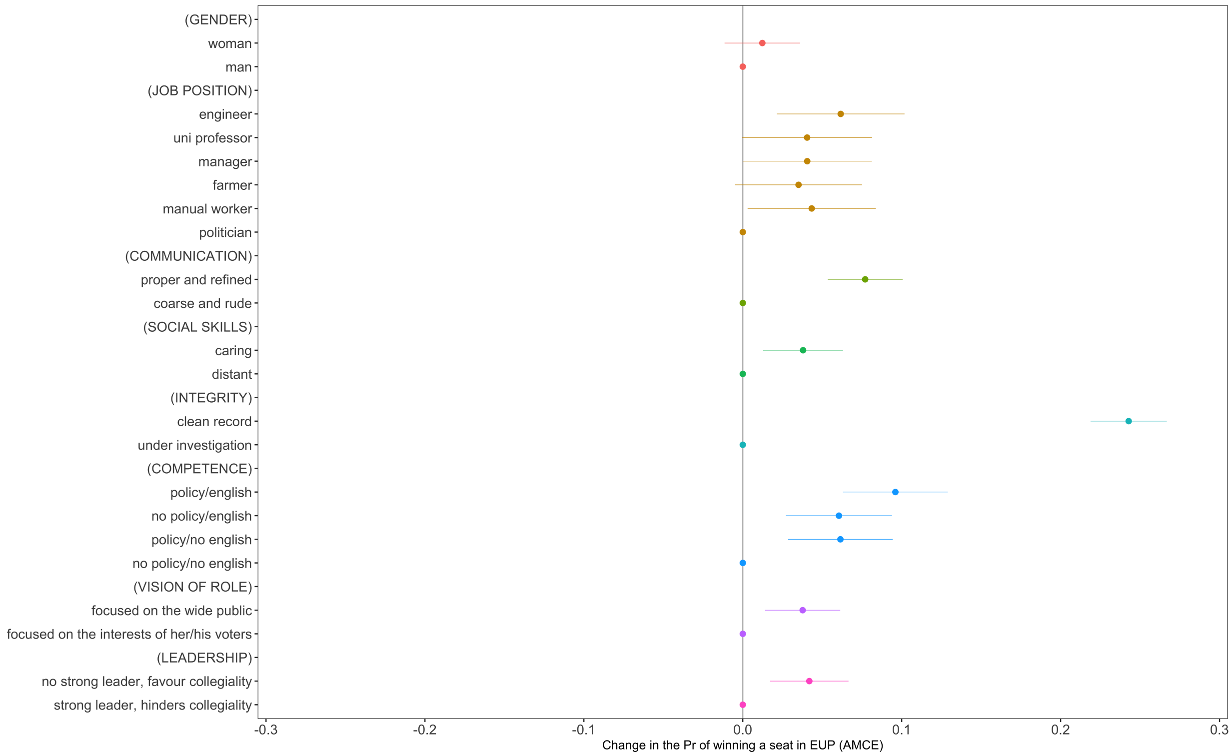

Figure 14: Average Marginal Conditional Effects (AMCE) excluding respondents failing all attention checks. Lines on both sides of the points represent 95% confidence intervals.

5 References

Baker R, Blumberg SJ, Brick JM, Couper MP, Courtright M, Dennis JM, Dillman D, Frankel MR, Garland P, Groves RM, Kennedy C, Krosnick J and Lavrakas PJ (2010) Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly* **74**, 711–781. https://doi.org/10.1093/poq/nfq048

Hainmueller J, Hopkins D and Yamamoto T (2013) Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis* **22**, 1–30. https://doi.org/10.1177/0002716203254763

ISTAT (2017a) Popolazione Residente. http://demo.istat.it/

ISTAT (2017b) Rilevazione Sulle Forze di Lavoro. http://siqual.istat.it/SIQual/visualizza.do?id=0019700&refresh=true&language=IT

ISTAT (2019a) Popolazione Residente. http://demo.istat.it/

ISTAT (2019b) Rilevazione Sulle Forze di Lavoro. http://siqual.istat.it/SIQual/visualizza.do?id=0019700&refresh=true&language=IT

Leeper TJ (2018) Cregg: Simple Conjoint Analyses and Visualization. R package version 0.2.1.

Long SJ (1997) *Regression Models for Categorical and Limited Dependent variables.* London: Sage.