

Online Appendix for *Portraits of Power: Facial Appearance and the Tacit Domain of Political Selection in China*

Table of Contents

A	Summary Statistics	A-1
B	Details on Automated Facial Rating Procedures and Output	A-2
B.1	Information about Human Rating Process	A-3
B.2	Validation of Rater Familiarity with Officials	A-6
B.3	Model Performance	A-8
B.4	Supplementary Validation on a Curated Dataset	A-9
B.5	Output Validation	A-10
C	Additional Information on Robustness Checks	A-14
D	Additional Information on Conjoint Experiment	A-23
D.1	Experiment Interface	A-23
D.2	Validation and Numerical Results	A-24

A Summary Statistics

Table A.1: Summary Statistics

Variable	N	Mean	S.D.	Min	Max
<i>Panel A: leaders' characteristics</i>					
Highest political rank	4,174	1.41	0.83	1	5
Demoted, arrested or received party disciplinary sanctions (1 = yes)	4,174	0.09	0.29	0	1
Female	4,174	0.06	0.24	0	1
Education level	4,174	1.74	0.47	0	2
Han ethnicity	4,174	0.88	0.33	0	1
City leaders: average GDP growth rate	2,277	-0.11	0.07	-0.30	0.68
City leaders: average fiscal income growth rate	2,016	0.15	0.12	-0.31	0.67
City leaders: Connections with Politburo members	2,279	1.13	2.38	0.00	16.01
<i>Panel B: facial traits</i>					
Perceived attractiveness	4,157	2.78	0.22	1.70	3.57
Perceived competence	4,157	3.05	0.22	1.74	3.91
Perceived trustworthiness	4,157	3.12	0.19	2.04	4.28
Perceived aggressiveness	4,157	2.78	0.16	1.82	3.62
Forehead ratio	4,154	0.25	0.01	0.20	0.28
Eyebrow position	4,154	0.13	0.01	0.08	0.17
Eye size	4,154	0.10	0.01	0.08	0.12
Facial width-to-height ratio	4,154	2.33	0.12	1.88	2.84
Nose length	4,154	0.27	0.01	0.24	0.31
Nose width	4,154	0.16	0.01	0.13	0.20
Nose width-to-height ratio	4,154	0.59	0.03	0.47	0.74
Mouth width	4,154	0.29	0.02	0.22	0.39
Lip thickness	4,154	0.08	0.01	0.07	0.11
Chin length	4,154	0.19	0.01	0.12	0.23
Chin width	4,154	0.38	0.02	0.29	0.43
<i>Panel C: respondents' characteristics in conjoint</i>					
Female	159	0.30	0.46	0	1
Education level	159	1.28	0.50	0	2
Age group (1 = 20-30, 2 = 30-40, 3 = 40-50, 4 = 50-60)	159	1.78	0.46	1	4
Employment type (1 = SOEs, 2 = Party and government organs)	159	1.50	0.50	1	2
Administrative rank (1 = none, 2 = deputy division and below, 3 = division level)	159	1.52	0.51	1	3

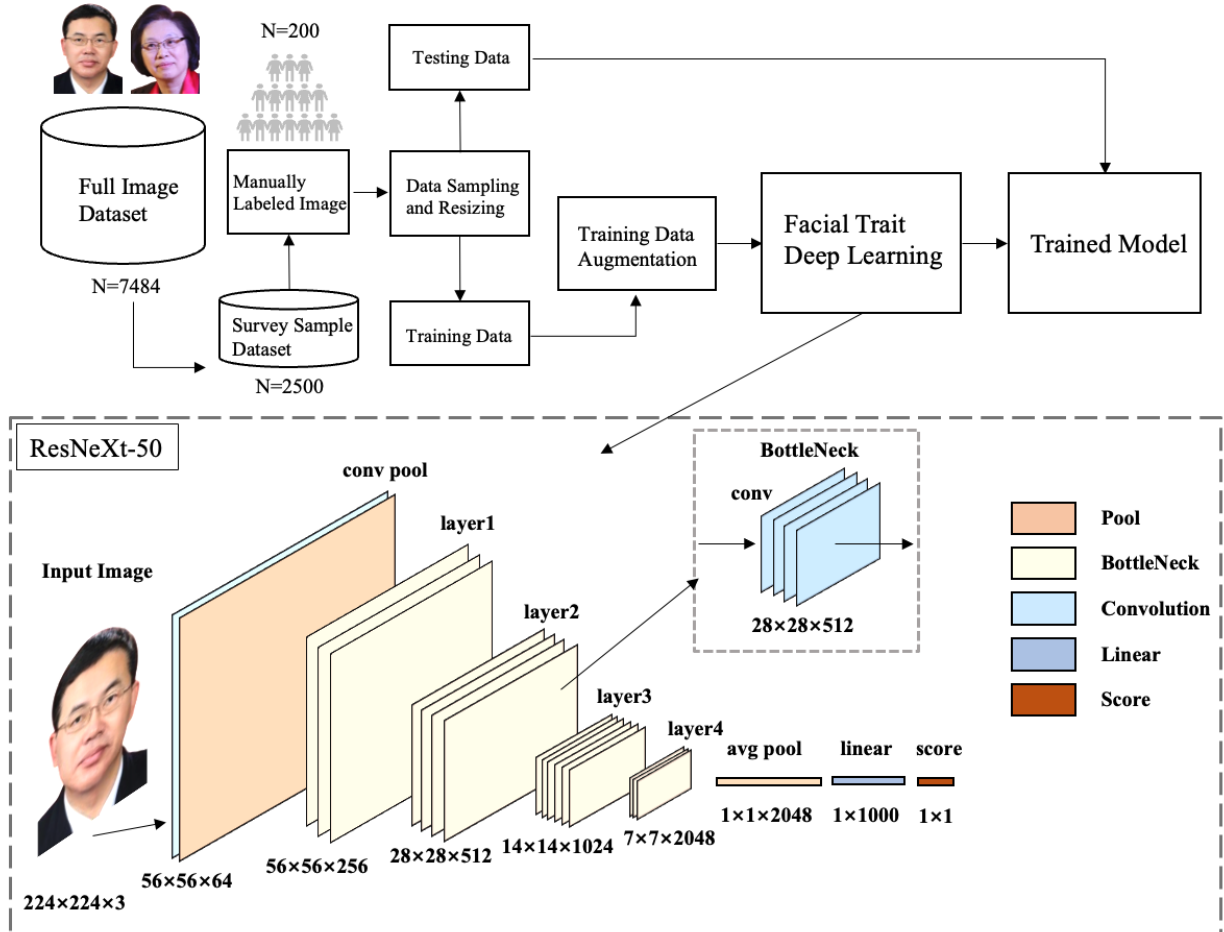
Table A.2: Image Availability Rates by Rank

Rank	N	Head Shot Available	Meeting Available	Either Available
Prefecture	4203	85.2%	88.1%	91.7%
Deputy provincial	1639	91.6%	94.5%	97.8%
Full provincial	744	93.4%	91.8%	97.3%

Note: This table summarizes the availability rates of officials' photos across three subnational ranks. "Head Shot Available" refers to officials with a valid head-shot image; "Meeting Available" refers to the person having meeting-scene images; "Either Available" indicates cases where at least one image of either type is available. National-level officials are not shown because their images are fully available.

B Details on Automated Facial Rating Procedures and Output

Figure B.1: Schematic Representation of the Training Process.



B.1 Information about Human Rating Process

We first present raters with a brief prompt explaining the task and then give them 125 headshots to rate across four dimensions. Of these 125 photos, 25 are fixed and rated by all raters, while the remaining 100 are randomly selected from the pool of the 2,500 headshots excluding the 25 fixed ones. The 125 photos are displayed to raters, who evaluate them in one dimension at a time before moving on to the next dimension to rate the same set of photos again. The order of the photos is randomized for each dimension. The wording of the prompt and an example of a photo being rated are as follows.

Prompt “请观察以下照片中人物，根据您对他/她相貌的第一印象，从1至5进行打分” (“Please rate the person in the following photo on a scale from 1 to 5 based on your first impression of their appearance.”)



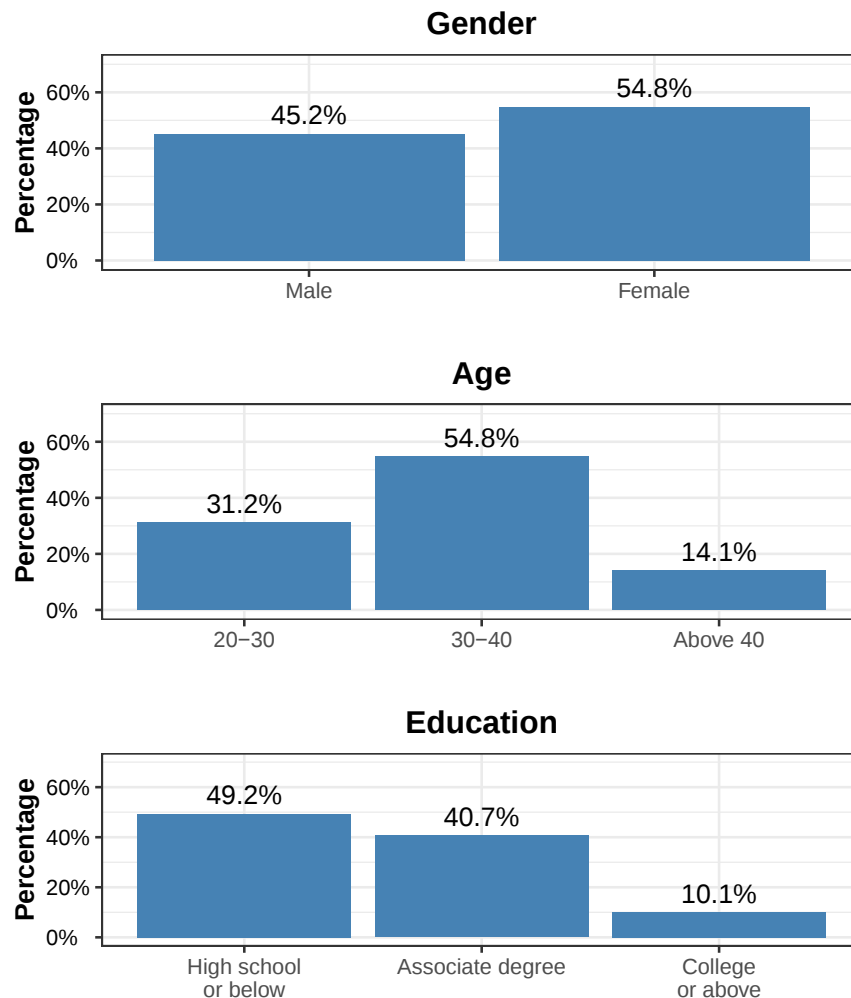
Attractiveness question 您认为图中人物有多好看？请您为他（她）的颜值打分（5代表非常好着，1代表非常不好看） (Does the person in the photo look attractive to you? Please rate his/her appearance on a scale from 1 to 5 (5 represents very attractive, 1 represents not attractive at all.))

Competence question 图中人物是否给您聪明精干的感觉？请从1到5打分（5代表非常聪明精干，1代表非常不聪明精干） (Does the person in the photo look competent to you? Please rate his/her appearance on a scale from 1 to 5 (5 represents very competent, 1 represents not competent at all.))

Trustworthiness question 图中人物是否给您忠厚可靠的感觉？请从1到5打分（5代表非常忠厚可靠，1代表非常不忠厚可靠） (Does the person in the photo look trustworthy to you? Please rate his/her appearance on a scale from 1 to 5 (5 represents very trustworthy, 1 represents not trustworthy at all.))

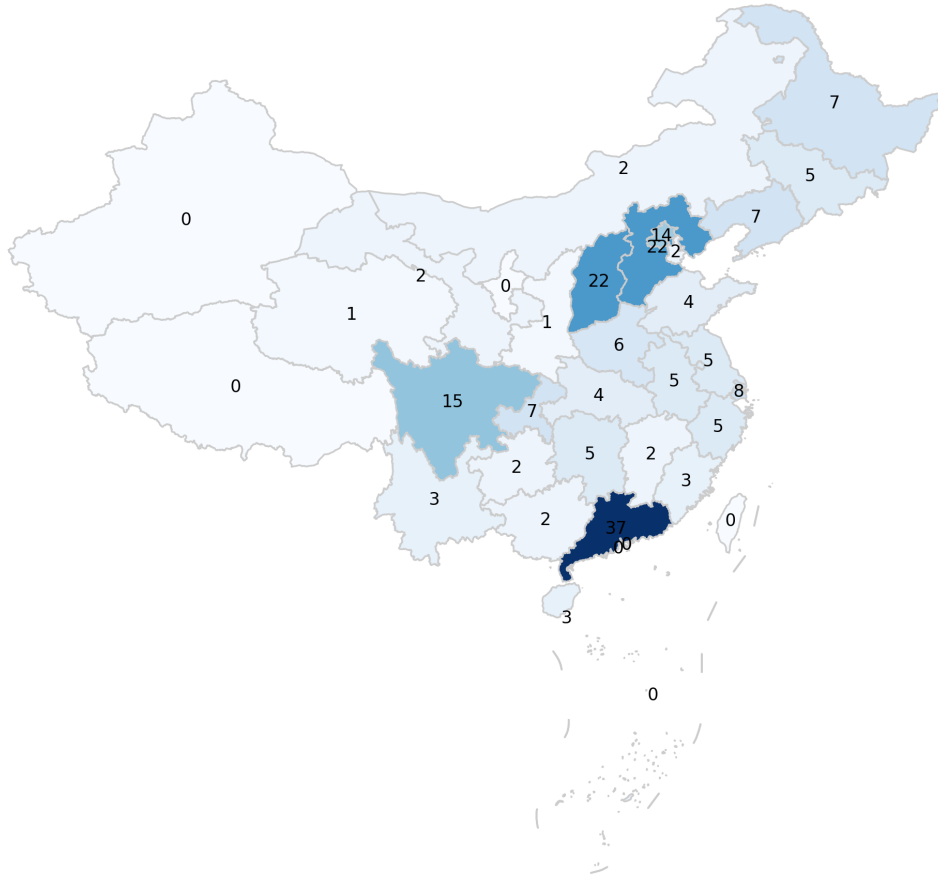
Aggressiveness question 图中人物是否给您凶狠强悍的感觉？请从1到5打分（5代表非常凶狠强悍，1代表非常不凶狠强悍）（Does the person in the photo look aggressive to you? Please rate his/her appearance on a scale from 1 to 5 (5 represents very aggressive, 1 represents not aggressive at all.))

Figure B.2: Demographic Background of Raters for the Training Dataset



Note: This figure displays the distribution of key demographic attributes for human raters (n = 199).

Figure B.3: Distribution of Raters by Province



Note: This figure displays the distribution of raters by their birth province ($n = 199$), with the number in each province representing the count of raters. The map shows a relatively balanced distribution of raters across China's northern, southern, eastern, and western regions.

Table B.1: Inter-rater Reliability

Facial Dimension	Cronbach's α	Guttman's λ_6
Perceived attractiveness	0.91	0.92
Perceived trustworthiness	0.69	0.73
Perceived competence	0.71	0.74
Perceived aggressiveness	0.84	0.86

Note: This table provides information about the inter-rater reliability for each facial dimension. We use two measures: Cronbach's α and Guttman λ_6 . Both range from 0 to 1, with larger values indicating greater agreement among raters. The values are calculated based on the 25 photos rated by all raters.

B.2 Validation of Rater Familiarity with Officials

As noted in the main text, a key assumption of our human-rating task is that officials in our annotation sample are largely unrecognizable to the public. To test this, we conducted a supplementary recognition survey, summarized below. The survey used a pool of 227 officials' images randomly drawn from all ranks. We recruited 186 respondents online and applied post-stratification weights to match the provincial distribution of the original rater pool (see Table B.2). Each respondent was asked to rate a random subset of 20 photos and answered a question of whether they could give the name or the position of the individual in the photo. We reviewed all the answers ourselves and coded an answer as correct if the respondent is right about either the name or the position.

As shown in Figure B.4, only a very small percentage of respondents could identify officials they rated. For Ranks 1 to 3 (prefecture, deputy provincial, and full provincial), the correct identification rate was **1.4%**, **0.7%**, and **1.5%**, respectively, with overlapping 95% confidence intervals. Even for national-level officials (Rank 4) and PSC members (Rank 5), recognition rates were merely **2.3%** and **3.6%**.

This survey confirms that officials from Ranks 1-3 are virtually anonymous to the public. The facial trait scores in our main analysis therefore reflect first impressions of the images rather than prior knowledge or familiarity with the officials.

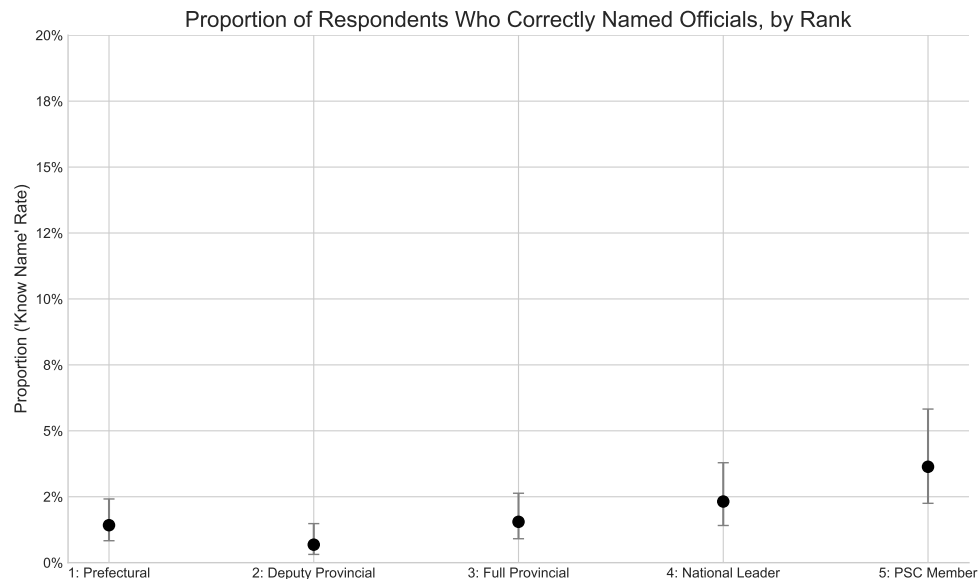


Figure B.4: Proportion of Respondents Correctly Identifying Officials, by Rank

Note: The figure displays the proportion of respondents who correctly provided the name or position of the official. Dots represent the point estimate, and the vertical bars represent 95% confidence intervals.

Table B.2: Summary Statistics of Recognition Survey Respondents

Variable	Unweighted	Weighted
<i>Panel A: Recognition Performance</i>		
Correctly Identified an Official (%)	2.2%	2.5%
<i>Panel B: Demographics</i>		
Mean Education Level	3.60	3.72
Male (%)	65.6%	60.6%
<i>Panel C: Media Habits</i>		
Primary Media: Social Media (%)	55.9%	62.3%
Primary Media: Television (%)	43.5%	37.6%
<i>Panel D: News Preference</i>		
Preferred News: Political (%)	21.0%	22.1%
Preferred News: Social/Lifestyle (%)	9.7%	6.0%
Total Number of Respondents	186	

Note: This table displays the summary statistics for the 186 unique respondents in the recognition survey. The 'Unweighted' column shows the simple averages from the raw survey data. The 'Weighted' column shows statistics after applying post-stratification weights to match the provincial distribution of the original human raters from the main study. Mean Education Level is measured on a 1-6 scale, where 1 = Middle School or less, 2 = High School, 3 = Junior College, 4 = Bachelor, 5 = Master, and 6 = PhD.

B.3 Model Performance

Table B.3: Predictive Accuracy of Facial Traits Across Models

Model	MSE	MAE	RMSE
Perceived Attractiveness			
AlexNet	0.69	0.79	0.83
DenseNet121	0.46	0.57	0.68
DenseNet201	0.65	0.72	0.81
ResNet18	0.68	0.79	0.82
ResNet152	0.55	0.67	0.74
ResNeXt50	0.47	0.60	0.69
ResNeXt101_32x8d	0.67	0.75	0.82
ViT	0.78	0.84	0.88
Perceived Competence			
AlexNet	0.25	0.41	0.50
DenseNet121	0.21	0.38	0.46
DenseNet201	0.20	0.36	0.45
ResNet18	0.24	0.41	0.49
ResNet152	0.25	0.41	0.50
ResNeXt50	0.25	0.42	0.50
ResNeXt101_32x8d	0.23	0.39	0.48
ViT	0.29	0.44	0.53
Perceived Trustworthiness			
AlexNet	0.20	0.37	0.45
DenseNet121	0.24	0.40	0.49
DenseNet201	0.28	0.41	0.53
ResNet18	0.20	0.36	0.45
ResNet152	0.21	0.38	0.46
ResNeXt50	0.22	0.37	0.47
ResNeXt101_32x8d	0.20	0.36	0.45
ViT	0.24	0.40	0.49
Perceived Aggressiveness			
AlexNet	0.17	0.33	0.41
DenseNet121	0.21	0.36	0.46
DenseNet201	0.22	0.38	0.47
ResNet18	0.17	0.32	0.41
ResNet152	0.24	0.38	0.49
ResNeXt50	0.20	0.36	0.45
ResNeXt101_32x8d	0.19	0.35	0.44
ViT	0.17	0.33	0.41

Note: This table provides predictive accuracy metrics (MSE, MAE, RMSE) for different models (AlexNet, DenseNet121, ResNet18, ResNeXt50, DenseNet201, ResNet152, ResNeXt101_32x8d, ViT) across four facial traits: Perceived attractiveness, Perceived competence, Perceived trustworthiness, and Perceived aggressiveness.

B.4 Supplementary Validation on a Curated Dataset

We conducted an additional validation test by sampling another 160 images. This set was constructed through a stratified random sampling method that slightly oversamples photos rated at the top and bottom 10%. This is to ensure that there is sufficient variation in human rating in a relative small sample. We then recruited a new group of 122 human raters to provide scores for the four traits on these 160 images.

Figure B.5 presents the scatter plots of the new human ratings against the model's original predictions for this curated dataset. The results show a marked improvement in the correlation between human and machine ratings. The Pearson's r coefficients are substantially higher than those from the cross-validation, ranging from $r = 0.397$ (Trustworthiness) to $r = 0.502$ (Aggressiveness), and all are highly statistically significant ($p < 0.001$).

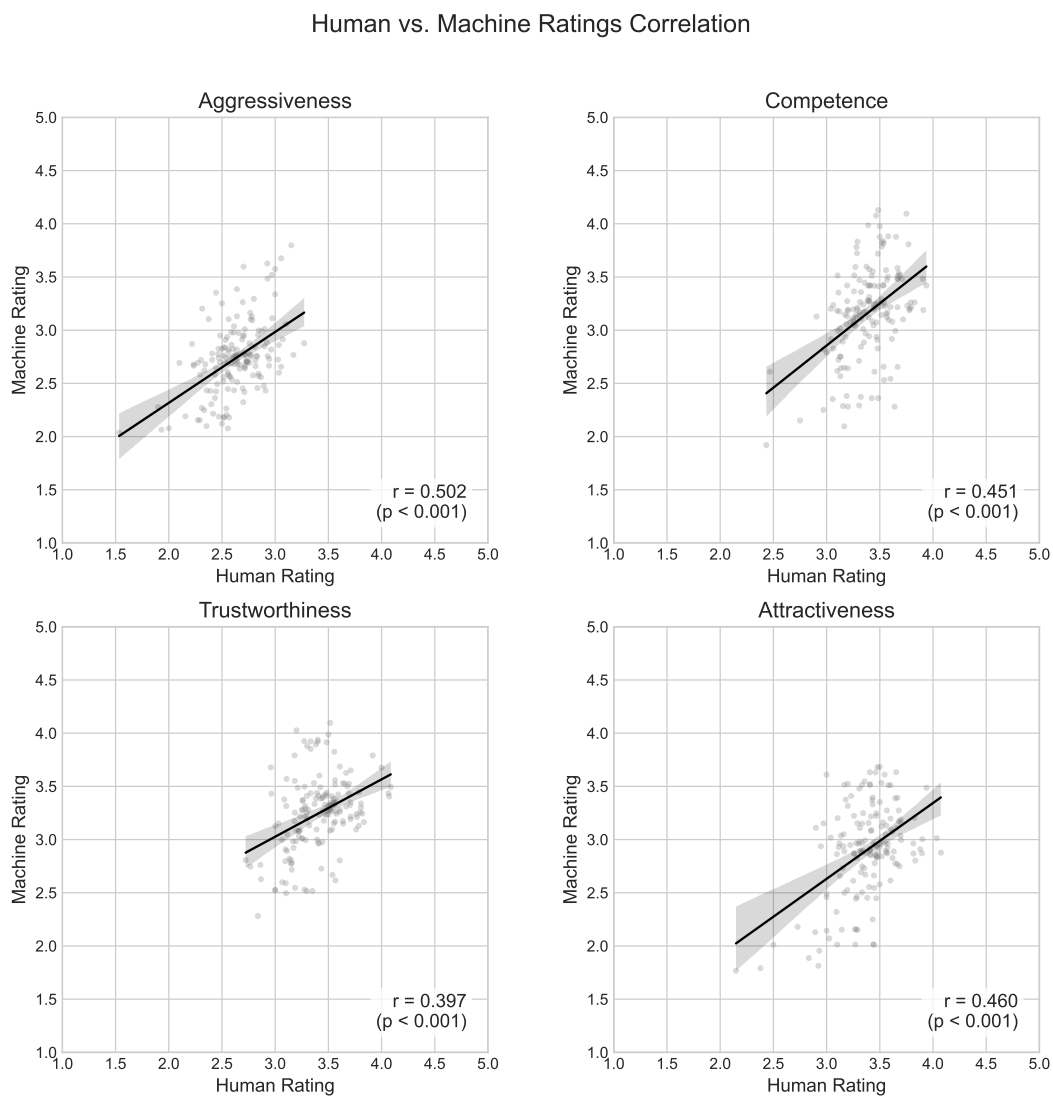
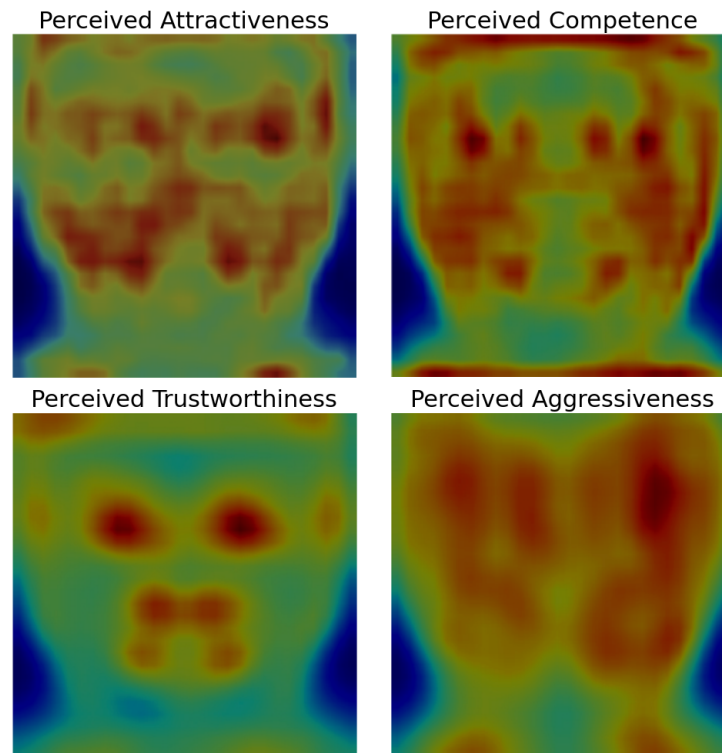


Figure B.5: Human vs. Machine Ratings on a Curated 160-Image Set

B.5 Output Validation

Figure B.6: Contributions of Facial Regions to Trait Perceptions

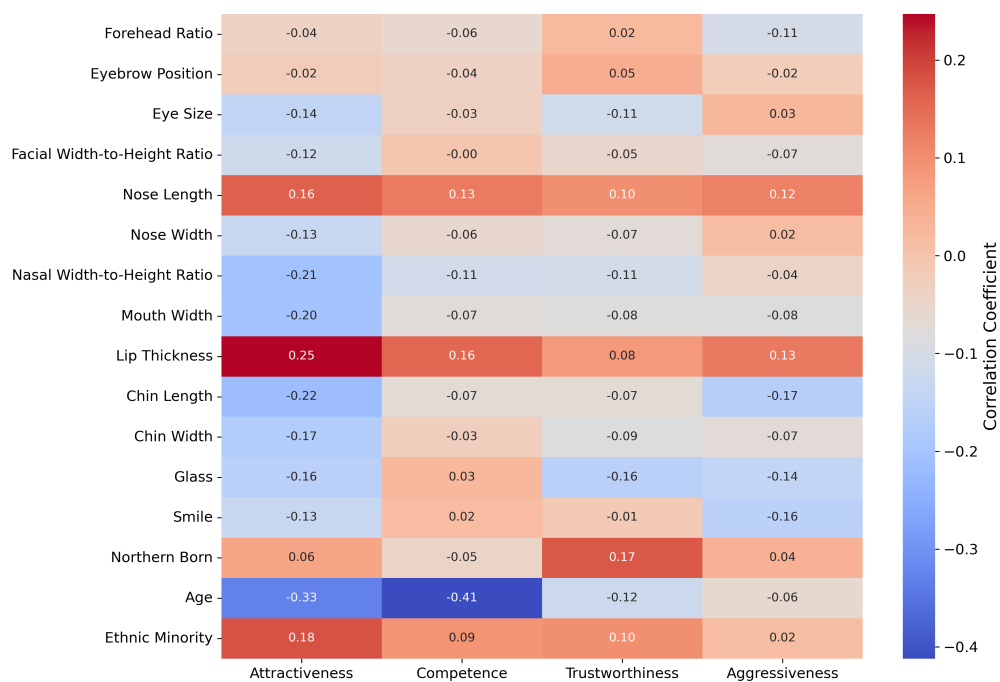


Note: This figure illustrates the contributions of different facial regions to the machine-based ratings of the four facial traits. The facial heat maps are produced by Gradient-weighted Class Activation Mapping (Grad-CAM), a visualization technique for interpreting deep-learning models (Selvaraju et al. 2017). Regions highlighted in red indicate significant contributions to the prediction of the trait, while blue regions indicate minimal contributions. We see that inferences of attractiveness are primarily drawn from features in the central facial area, including the eyes, nose, and cheeks; perceptions of competence are most heavily influenced by the eyes, eyebrows, and jawline; evaluations of trustworthiness focus on the eye and mouth; finally, perceptions of aggressiveness rely on diffuse cues from the entire face, including the forehead, eyebrows, and broader lower facial regions. These patterns are broadly consistent with existing research on how humans infer traits from facial cues (e.g., Oosterhof and Todorov 2008; Vernon et al. 2014), suggesting that the machine-based algorithm replicates the processes underlying human perceptions of facial traits with reasonable accuracy and realism.

Figure B.7: Relationship between Objective Facial Features and Subjective Perceptions

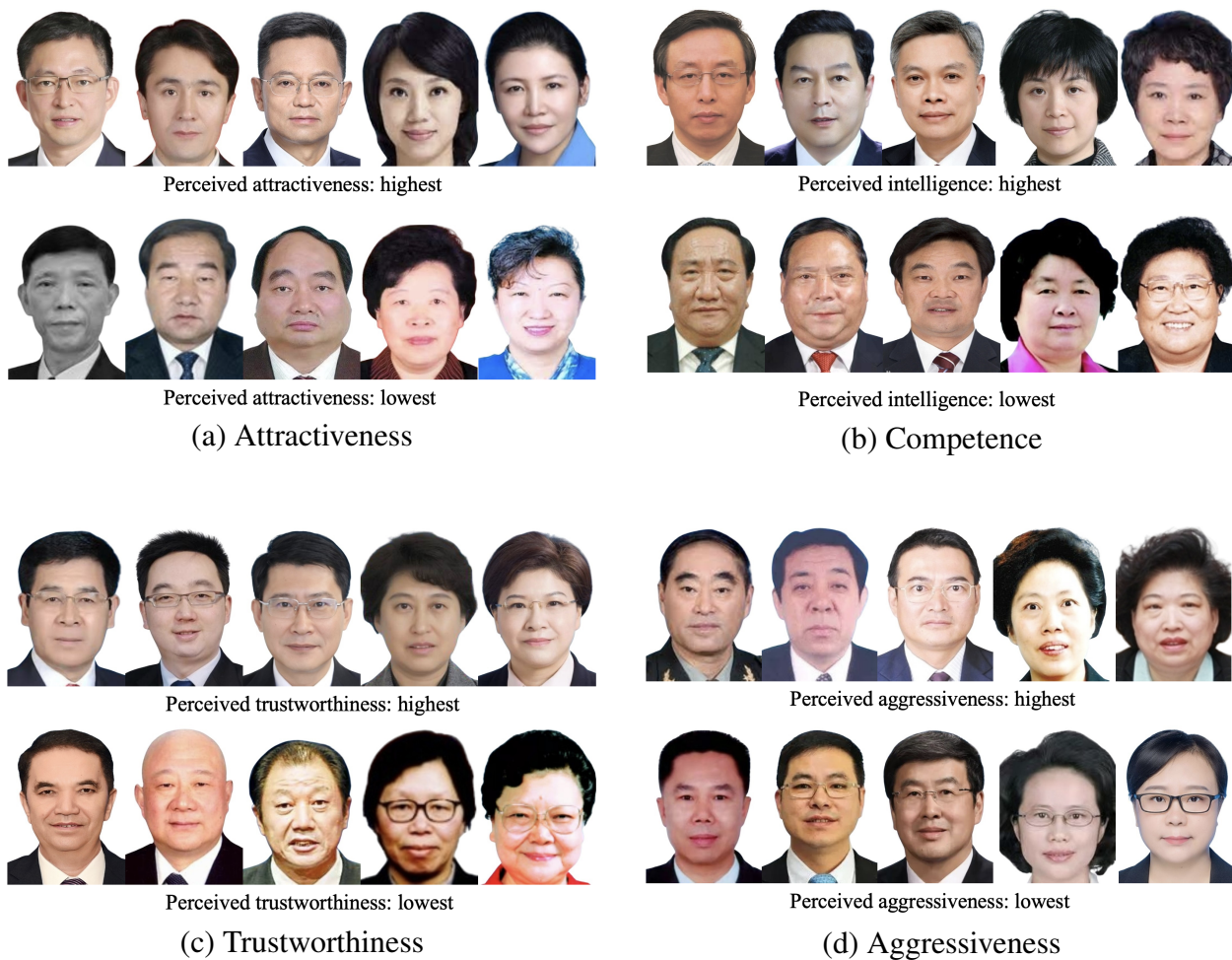


Male



Female

Figure B.8: Examples of Photos Receiving Extreme Values in Automated Rating



Note: This figure shows the actual photos that receive the highest and lowest ratings for each trait. The rating is assigned by the ResNext-50 model with no human intervention.

Table B.4: Correlation Matrix of Perceived Facial Traits (Human Rating)

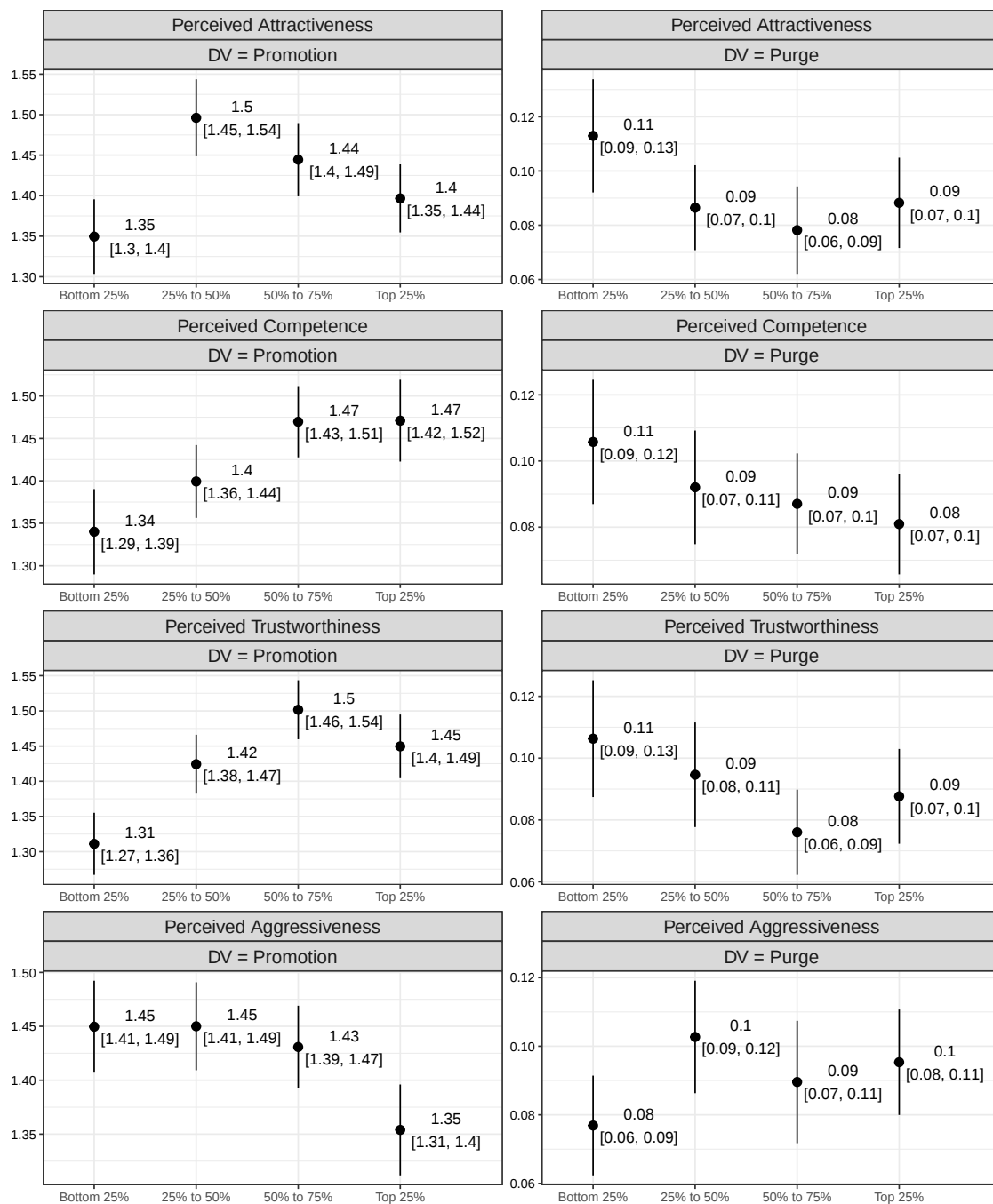
	Perceived Attractiveness	Perceived Competence	Perceived Trustworthiness	Perceived Aggressiveness
Perceived Attractiveness	1.00	0.42	0.33	0.08
Perceived Competence	0.41	1.00	0.32	0.02
Perceived Trustworthiness	0.33	0.32	1.00	-0.002
Perceived Aggressiveness	0.08	0.02	-0.002	1.00

Table B.5: Correlation Matrix of Perceived Facial Traits (Automated Rating)

	Perceived Attractiveness	Perceived Competence	Perceived Trustworthiness	Perceived Aggressiveness
Perceived Attractiveness	1.00	0.43	0.33	-0.03
Perceived Competence	0.43	1.00	0.38	-0.10
Perceived Trustworthiness	0.33	0.38	1.00	-0.11
Perceived Aggressiveness	-0.03	-0.10	-0.11	1.00

C Additional Information on Robustness Checks

Figure C.1: Estimated Rank and Purge Probabilities Using Rating Quartiles



Note: This figures shows the predicted rank and probability of purge for officials at different quartiles of facial ratings. The left and right columns are based on results from two separate regressions with different dependent variables (maximum rank and purge). In both regressions, we use the quartiles of facial ratings as the key independent variables. The rest of the specifications are identical to Models 5 and 10 of Table 1, respectively.

Table C.1: Robustness: Using Median Ratings Instead of Mean Ratings

	DV: Maximum Rank					DV: Purge				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Perceived attractiveness (median)	0.202** (0.053)				0.084 (0.056)	-0.059** (0.022)				-0.024 (0.026)
Perceived competence (median)		0.315** (0.060)			0.227** (0.064)		-0.080** (0.022)			-0.057* (0.025)
Perceived trustworthiness (median)			0.255** (0.059)		0.146* (0.065)			-0.078** (0.024)		-0.051* (0.026)
Perceived aggressiveness (median)				-0.195** (0.064)	-0.165** (0.063)				-0.005 (0.024)	-0.013 (0.025)
Female	0.210** (0.051)	0.227** (0.051)	0.221** (0.051)	0.215** (0.052)	0.214** (0.051)	-0.033* (0.015)	-0.038* (0.016)	-0.036* (0.016)	-0.037* (0.016)	-0.036* (0.015)
Han ethnicity	-0.043 (0.055)	-0.044 (0.054)	-0.041 (0.054)	-0.043 (0.054)	-0.042 (0.054)	-0.008 (0.020)	-0.008 (0.021)	-0.009 (0.020)	-0.008 (0.021)	-0.009 (0.020)
College degree	0.453* (0.191)	0.462* (0.186)	0.486* (0.189)	0.478* (0.187)	0.480* (0.187)	0.136** (0.027)	0.133** (0.029)	0.126** (0.028)	0.132** (0.028)	0.131** (0.028)
Graduate degree	0.599** (0.197)	0.609** (0.192)	0.637** (0.196)	0.629** (0.194)	0.626** (0.193)	0.153** (0.027)	0.150** (0.028)	0.142** (0.028)	0.149** (0.028)	0.149** (0.028)
Face quality (median)	0.005** (0.001)	0.005** (0.001)	0.005** (0.001)	0.005** (0.001)	0.004** (0.001)	-0.001* (0.000)	-0.001+ (0.000)	-0.001+ (0.000)	-0.001* (0.000)	-0.001 (0.000)
Blurriness (median)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Birth year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Birth city FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.27	0.27	0.27	0.26	0.27	0.13	0.13	0.13	0.13	0.13
Observations	4071	4071	4071	4071	4071	4071	4071	4071	4071	4071

Note: This table presents the estimated effects of facial features on two sets of career outcomes (as of 2022): maximum political rank achieved and purge. For each official, we use the median rating of all his/her collected photos instead of the mean. The specification is otherwise identical to Table C.3. FE = fixed effects. Standard errors clustered at the birth city level are reported in parentheses. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table C.2: Baseline Results with Facial Similarity Controls

	DV: Maximum Rank					DV: Purge				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Perceived attractiveness	0.268** (0.055)				0.061 (0.060)	-0.074** (0.024)				-0.027 (0.028)
Perceived competence		0.431** (0.064)			0.287** (0.069)		-0.096** (0.021)			-0.063* (0.025)
Perceived trustworthiness			0.414** (0.065)		0.262** (0.071)			-0.096** (0.026)		-0.060* (0.028)
Perceived aggressiveness				-0.287** (0.069)	-0.230** (0.068)				0.011 (0.026)	-0.002 (0.027)
Similarity to Xi Jinping	-0.002 (0.004)	-0.002 (0.003)	-0.003 (0.004)	-0.002 (0.004)	-0.002 (0.004)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Similarity to Hu Jintao	0.008** (0.003)	0.008** (0.003)	0.008** (0.003)	0.008* (0.003)	0.008** (0.003)	0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)
Similarity to Jiang Zemin	0.002 (0.003)	0.001 (0.003)	0.001 (0.003)	0.001 (0.003)	0.001 (0.003)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)
Female	0.287** (0.058)	0.299** (0.057)	0.292** (0.057)	0.280** (0.058)	0.281** (0.057)	-0.042* (0.018)	-0.046* (0.018)	-0.044* (0.018)	-0.044* (0.018)	-0.044* (0.018)
Han ethnicity	-0.059 (0.057)	-0.052 (0.056)	-0.057 (0.055)	-0.058 (0.055)	-0.054 (0.055)	-0.005 (0.021)	-0.007 (0.021)	-0.006 (0.021)	-0.006 (0.021)	-0.007 (0.021)
College degree	0.488* (0.194)	0.487* (0.190)	0.522** (0.190)	0.517** (0.190)	0.516** (0.189)	0.126** (0.028)	0.126** (0.029)	0.118** (0.027)	0.122** (0.028)	0.123** (0.028)
Graduate degree	0.635** (0.200)	0.634** (0.196)	0.674** (0.198)	0.670** (0.196)	0.662** (0.196)	0.143** (0.027)	0.143** (0.029)	0.134** (0.027)	0.138** (0.028)	0.140** (0.028)
Birth year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Birth city FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.25	0.26	0.26	0.25	0.27	0.13	0.13	0.13	0.12	0.13
Observations	4071	4071	4071	4071	4071	4071	4071	4071	4071	4071

Note: This table presents the estimated effects of facial features on two sets of career outcomes (as of 2022): maximum political rank achieved and purge. *Similarity to Xi Jinping*, *Similarity to Hu Jintao*, and *Similarity to Jiang Zemin* are variables that are constructed to measure the facial similarity between an official and the three consecutive top leaders following the approach of Schroff, Kalenichenko, and Philbin (2015). FE = fixed effects. Standard errors clustered at the birth city level are reported in parentheses. * $p < 0.05$, ** $p < 0.01$

Table C.3: Baseline Results with Image Quality Controls: Promotion and Purge

	DV: Maximum Rank					DV: Purge				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Perceived attractiveness	0.230** (0.057)				0.058 (0.061)	-0.062** (0.024)				-0.026 (0.028)
Perceived competence		0.375** (0.065)			0.259** (0.069)		-0.082** (0.022)			-0.057* (0.025)
Perceived trustworthiness			0.346** (0.065)		0.213** (0.072)			-0.073** (0.026)		-0.041 (0.029)
Perceived aggressiveness				-0.260** (0.069)	-0.217** (0.068)				0.001 (0.026)	-0.008 (0.027)
Female	0.212** (0.051)	0.226** (0.051)	0.221** (0.051)	0.213** (0.052)	0.213** (0.051)	-0.032* (0.015)	-0.036* (0.015)	-0.035* (0.015)	-0.035* (0.015)	-0.035* (0.015)
Han ethnicity	-0.043 (0.056)	-0.038 (0.055)	-0.043 (0.055)	-0.043 (0.054)	-0.041 (0.054)	-0.009 (0.020)	-0.010 (0.021)	-0.010 (0.020)	-0.010 (0.021)	-0.010 (0.020)
College degree	0.477* (0.190)	0.479* (0.187)	0.509** (0.189)	0.502** (0.187)	0.505** (0.187)	0.135** (0.028)	0.134** (0.029)	0.127** (0.028)	0.132** (0.029)	0.132** (0.029)
Graduate degree	0.624** (0.196)	0.625** (0.193)	0.659** (0.196)	0.653** (0.193)	0.651** (0.193)	0.153** (0.028)	0.152** (0.029)	0.144** (0.028)	0.149** (0.029)	0.150** (0.028)
Face quality	0.004** (0.001)	0.004** (0.001)	0.003** (0.001)	0.004** (0.001)	0.003** (0.001)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.002** (0.000)	-0.001** (0.000)
Blurriness	-0.001 (0.001)	-0.001 (0.001)	-0.002 (0.001)	-0.001 (0.001)	-0.002 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Birth year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Birth city FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.26	0.26	0.26	0.26	0.27	0.13	0.13	0.13	0.13	0.13
Observations	4071	4071	4071	4071	4071	4071	4071	4071	4071	4071

Note: This table presents the estimated effects of facial features on two sets of career outcomes (as of 2022): maximum political rank achieved and purge. *Face quality* and *Blurriness* are two measures of image quality generated by the Face++ API. FE = fixed effects. Standard errors clustered at the birth city level are reported in parentheses. * $p < 0.05$, ** $p < 0.01$

Table C.4: Official Portraits vs. Meeting and Event Photos

	Maximum Rank		Purge	
	(1) Head shot	(2) Meeting	(3) Head shot	(4) Meeting
Perceived attractiveness	0.083 ⁺ (0.043)	0.066 (0.058)	-0.020 (0.019)	-0.029 (0.024)
Perceived competence	0.193** (0.045)	0.173** (0.062)	-0.022 (0.019)	-0.041 ⁺ (0.024)
Perceived trustworthiness	0.138** (0.051)	0.163** (0.058)	-0.004 (0.020)	-0.034 (0.025)
Perceived aggressiveness	-0.158** (0.046)	-0.128 ⁺ (0.069)	-0.008 (0.018)	0.025 (0.025)
Female	0.229** (0.054)	0.199** (0.053)	-0.038* (0.016)	-0.036* (0.016)
Han ethnicity	-0.044 (0.056)	-0.043 (0.054)	-0.009 (0.022)	-0.011 (0.021)
College degree	0.435 (0.339)	0.545** (0.205)	0.143** (0.039)	0.120** (0.031)
Graduate degree	0.580 ⁺ (0.347)	0.686** (0.212)	0.163** (0.038)	0.138** (0.031)
Birth year FE	✓	✓	✓	✓
Birth city FE	✓	✓	✓	✓
R-squared	0.29	0.27	0.14	0.13
Observations	3771	3920	3771	3920

Note: This table presents the estimated effects of facial features on promotions and purges. We use facial feature ratings generated solely from official portraits (Models 1 and 3) and meeting photos (Models 2 and 4). Demographic covariates include *Gender* and *Han ethnicity*, and *Education level*. FE = fixed effects.

Standard errors clustered at the birth city level are reported in parentheses. ⁺ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Figure C.2: Temporal Stability in Facial Rating for Selected Figures

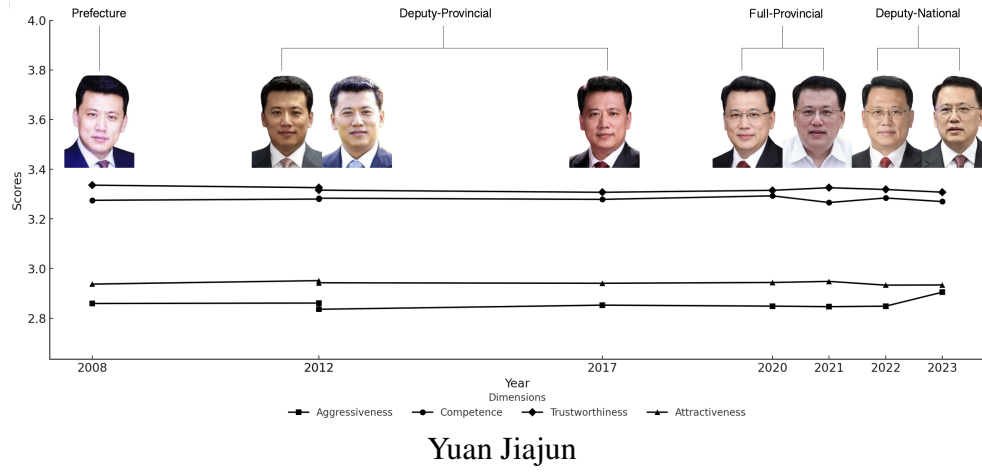


Table C.5: Effects of Promotion on Facial Features

	DV: Facial Traits Rating			
	(1) Perceived attractiveness	(2) Perceived competence	(3) Perceived trustworthiness	(4) Perceived ag- gressiveness
Photos taken after Politburo	0.076 (0.067)	0.024 (0.066)	-0.019 (0.051)	-0.008 (0.032)
Age	-0.007 (0.005)	-0.014** (0.005)	0.002 (0.004)	0.000 (0.003)
Blurriness	0.008* (0.003)	0.008** (0.002)	0.007** (0.002)	0.001 (0.003)
Face quality	0.004* (0.002)	0.004** (0.001)	0.003* (0.001)	-0.001 (0.001)
Individual FE	✓	✓	✓	✓
Observations	352	352	352	352
Adj. R-squared	0.204	0.276	0.209	0.261

Note: This table presents the estimated effects of how promotion and age affect facial features of Politburo members. The sample includes 352 photos for 37 non-military members in the 19th and 20th Central Committee. *Photos taken after Politburo* is a binary indicator that takes the value of 1 if the photo was taken after the person had been promoted to the Politburo, and 0 otherwise. *Age* is a person's actual age at that time the photo was taken. We also include two Image quality controls derived from the Face++ API: *Face quality*, which assesses general image quality, and *Blurriness*, which evaluates the level of blur in the image. FE = fixed effects. Standard errors are clustered at the individual level and reported in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table C.6: Effects of Facial Features on City Leaders' Economic and Fiscal Performance

	DV: GDP Growth at $t + 2$					DV: Fiscal Revenue Growth at $t + 2$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Perceived attractiveness	-0.022 (0.023)				-0.015 (0.026)	-0.011 (0.050)				0.004 (0.053)
Perceived competence		-0.037 (0.027)			-0.038 (0.031)		-0.056 (0.051)			-0.065 (0.055)
Perceived trustworthiness			0.004 (0.024)		0.024 (0.027)			0.008 (0.047)		0.033 (0.048)
Perceived aggressiveness				0.006 (0.030)	0.002 (0.031)				0.056 (0.059)	0.048 (0.061)
Female	0.044 (0.023)	0.042 (0.022)	0.042 (0.023)	0.042 (0.023)	0.043 (0.023)	0.054 (0.033)	0.054 (0.033)	0.052 (0.033)	0.055 (0.033)	0.055 (0.034)
Han ethnicity	0.027 (0.030)	0.028 (0.030)	0.027 (0.030)	0.027 (0.030)	0.027 (0.030)	0.038 (0.044)	0.037 (0.045)	0.037 (0.044)	0.037 (0.044)	0.035 (0.045)
College degree	-0.022 (0.018)	-0.022 (0.018)	-0.022 (0.018)	-0.023 (0.018)	-0.022 (0.018)	-0.020 (0.043)	-0.020 (0.044)	-0.020 (0.043)	-0.021 (0.043)	-0.022 (0.044)
Graduate degree	-0.029 (0.017)	-0.028 (0.017)	-0.029 (0.017)	-0.029 (0.017)	-0.028 (0.017)	-0.030 (0.039)	-0.030 (0.040)	-0.031 (0.039)	-0.031 (0.039)	-0.031 (0.040)
Birth year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Birth city FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.52	0.52	0.52	0.52	0.52	0.53	0.53	0.53	0.53	0.53
Observations	2016	2016	2016	2016	2016	1815	1815	1815	1815	1815

Note: This table presents the estimated effects of facial features on city leaders' economic and fiscal performance at $t + 2$. The $t + 2$ GDP and fiscal revenue growth rates are calculated as the percentage change relative to the first year of the city leader's tenure. FE = fixed effects. Standard errors clustered at the birth city level are reported in parentheses. * $p < 0.05$, ** $p < 0.01$

Table C.7: Pooled Model with Interaction Terms (By Step)

	DV: Promoted (1 = yes)
Promotion step = 2 (deputy provincial to full provincial)	0.522 (0.361)
Promotion step = 3 (full provincial to national)	-1.119** (0.362)
Perceived attractiveness	0.033 (0.032)
Perceived competence	0.108** (0.034)
Perceived trustworthiness	0.010 (0.034)
Perceived aggressiveness	-0.051 (0.038)
Attractiveness $\times$ Step 2	-0.106 (0.071)
Attractiveness $\times$ Step 3	-0.060 (0.079)
Competence $\times$ Step 2	-0.038 (0.079)
Competence $\times$ Step 3	0.162+ (0.096)
Trustworthiness $\times$ Step 2	0.179* (0.072)
Trustworthiness $\times$ Step 3	0.146 (0.092)
Aggressiveness $\times$ Step 2	-0.241** (0.081)
Aggressiveness $\times$ Step 3	0.063 (0.087)
Female	0.073** (0.020)
Han ethnicity	-0.028 (0.021)
College degree	0.158** (0.027)
Graduate degree	0.244** (0.026)
Birth year FE	✓
Birth city FE	✓
Observations	5,996
R-squared	0.124

Note: This table reports the results of a pooled regression model where each observation corresponds to an official at a given promotion step. The model includes interactions between each facial trait and the indicator for the promotion step. Standard errors clustered at the birth city level are shown in parentheses. FE = fixed effects. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table C.8: Pooled Model with Interaction Terms
(By Gender)

	Maximum Rank (1)	Purge (2)
Female (=1)	-2.044 (1.714)	0.832 (0.543)
Perceived attractiveness	0.072 (0.064)	-0.020 (0.029)
Perceived competence	0.289** (0.070)	-0.063* (0.025)
Perceived trustworthiness	0.237** (0.073)	-0.054+ (0.030)
Perceived aggressiveness	-0.265** (0.071)	-0.003 (0.028)
Attractiveness $\times$ Female	-0.242 (0.313)	-0.101 (0.087)
Competence $\times$ Female	-0.018 (0.354)	-0.080 (0.093)
Trustworthiness $\times$ Female	0.454 (0.329)	-0.110 (0.087)
Aggressiveness $\times$ Female	0.575 (0.464)	0.003 (0.117)
Han ethnicity	-0.056 (0.054)	-0.010 (0.020)
College degree	0.514** (0.189)	0.126** (0.028)
Graduate degree	0.662** (0.196)	0.144** (0.028)
Constant	-0.248 (0.386)	0.386** (0.127)
Birth year FE	✓	✓
Birth city FE	✓	✓
Observations	4010	4010

Note: This table reports the results of a pooled regression model that includes interaction terms between each facial trait and an indicator for female candidates. Unlike the sub-sample analysis in the main text, this model estimates gender heterogeneity by allowing the effects of facial traits to vary systematically between male and female officials within a single unified specification. Standard errors clustered at the birth-city level are reported in parentheses. FE = fixed effects. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

D Additional Information on Conjoint Experiment

D.1 Experiment Interface

We first present respondents with a brief prompt about the task and then show them eight successive pairs of candidate profiles. The wording of the prompt and the examples of the profiles are as follows. The two photos are matched on attractiveness, competence, and aggressiveness, but the one on the right has higher trustworthiness rating than the one on the left. The colored texts highlight the randomized components of the conjoint experiment (The subjects only saw black texts in the actual experiments).

Prompt “假设您是某政府机关单位的主要负责同志，即将对以下两位干部进行考察，并选择提拔其中一位，以下是这两名干部的一些基本信息和简历照片。综合他们的背景，以及您看到他们相貌的第一印象，您觉得哪位更有可能被提拔？” (“Suppose you are the main decision maker in a government agency and are about to evaluate the following two cadres, choosing one for promotion. Below is the photo and background information of these two candidates. Based on your reading of their profiles and impression of their appearance, which one do you think is more likely to be promoted?”)



干部1：1977年出生，本科学历，曾在基层政府任职

Candidate 1 was born in 1977, has a bachelor's degree, and has worked in grassroots government



干部2：1973年出生，博士研究生，曾担任上级领导秘书

Candidate 2 was born in 1973, has a Ph.D degree, and has worked as a personal aide to a senior leader

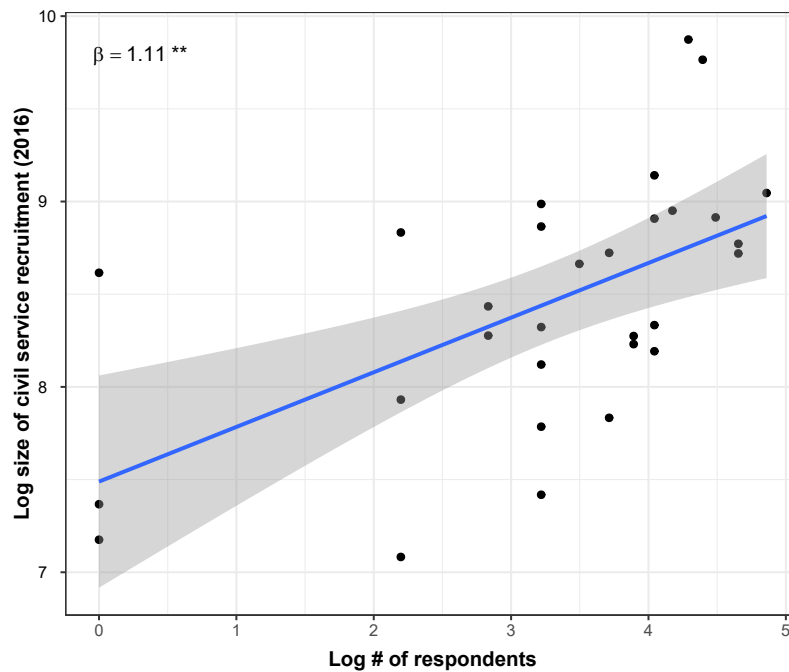
D.2 Validation and Numerical Results

Table D.1: Numerical Results for Conjoint Experiment

	DV: Profile Selected for Promotion			
	(1) Perceived attractive- ness	(2) Perceived compe- tence	(3) Perceived trustwor- thiness	(4) Perceived aggres- siveness
Face with higher rating	-0.017 (0.055)	0.249** (0.053)	0.312** (0.051)	-0.334** (0.056)
Birth year (baseline = 1973)				
- 1975	-0.005 (0.107)	0.076 (0.096)	0.235* (0.099)	-0.122 (0.101)
- 1977	0.072 (0.111)	0.063 (0.087)	0.113 (0.092)	0.154 (0.104)
- 1979	0.018 (0.117)	0.100 (0.103)	0.213* (0.095)	0.047 (0.103)
Education (baseline = bachelor's degree)				
- Master's degree	0.124 (0.094)	0.323** (0.077)	0.348** (0.078)	0.156+ (0.085)
- Ph.D degree	0.375** (0.102)	0.564** (0.080)	0.309** (0.084)	0.406** (0.092)
Work experience (baseline = SOE)				
- Grassroots government	0.137 (0.110)	0.189* (0.094)	0.034 (0.101)	0.096 (0.102)
- Central government	0.429** (0.123)	0.531** (0.100)	0.437** (0.099)	0.524** (0.110)
- Personal aide to senior leader	0.454** (0.117)	0.480** (0.101)	0.365** (0.091)	0.324** (0.108)
Respondent-pair FE	✓	✓	✓	✓
R-squared	0.12	0.28	0.25	0.24
# of subjects	159	159	159	159
Observations	636	636	636	636

Note: This table presents the numerical estimates from the conjoint experiment on civil servants. The units of observation are individual official profiles (photo + short biography). The dependent variable is whether a given profile is selected by a respondent for promotion in a pair. Respondent-pair fixed effects are included in all models and standard errors clustered at respondent level are reported in parentheses. SOE = state-owned enterprises. * $p < 0.05$, ** $p < 0.01$

Figure D.1: Representativeness of Conjoint Experiment Respondents



Note: This table presents the correlation between the regional origin of respondents in our conjoint sample and the actual regional distribution of government size in China. Each circle represents a province. The x-axis is the logged number of conjoint experiment respondents from that province, and the y-axis is the logged number of new civil servants recruited in 2016. The coefficient printed at the top left is from the bivariate regression between the two variables.