ONLINE APPENDIX for Descriptive Legitimacy and International Organizations: Evidence from United Nations High Commissioner for Refugees

Wilfred M. Chow
Department of Politics and Public Administration
The University of Hong Kong
wilfred.chow@hku.hk

Enze Han
Department of Politics and Public Administration
The University of Hong Kong
enzehan@hku.hk

TABLE OF CONTENTS

A1: UNHCR PICTURE SELECTION PRETEST 3

FIGURE A1: PICTURE POSITION OF UNHCR PANEL 3

FIGURE A2: PANEL PICTURE RATINGS 4

A2: REPLICATION SURVEY 6

FIGURE A3: UNHCR PANEL WITH COUNTRY LABELS, BY RACIAL AND GENDER DISTRIBUTION 6

A3: SAMPLE TREATMENT BALANCE 7

TABLE A1: DEMOGRAPHIC SUMMARY STATISTICS COMPARISON 7

TABLE A2: SAMPLE TREATMENT BALANCE (MAIN SURVEY) 8

TABLE A3: SAMPLE TREATMENT BALANCE (FOLLOW-UP REPLICATION WITHOUT COUNTRY LABELS)..... 9

TABLE A4: SAMPLE TREATMENT BALANCE (FOLLOW-UP REPLICATION WITH COUNTRY LABELS) 10

A4: FULL PROBIT REGRESSION TABLES WITH ALL CONTROLS..... 11

FIGURE A3: FACTORIAL VS. ITEM RESPONSE THEORY OF WHITE NATIONALISM 13

TABLE A5: TRUST AND FAIRNESS OF THE UNHCR DECISION-MAKING PROCESS..... 14

TABLE A6: UNHCR REPORT OUTCOME FOR REFUGEES AND AMERICANS..... 15

A5: ALTERNATIVE DEPENDENT VARIABLE MEASUREMENTS 16

TABLE A7: FACTOR ANALYSIS VARIABLES OF SUBSTANTIVE AND PROCEDURAL LEGITIMACY..... 16

TABLE A8: DESCRIPTIVE REPRESENTATION & SUBSTANTIVE AND PROCEDURAL LEGITIMACY 17

A6: PRETREATMENT EFFECTS AND ATTENTION CHECK 19

FIGURE A5: UNHCR PANEL DIVERSITY AND LEGITIMACY, BY MEDIA CONSUMPTION..... 19

FIGURE A6: UNHCR PANEL DIVERSITY AND LEGITIMACY, BY RESPONDENT ATTENTION 20

FIGURE A7: UNHCR DIVERSITY AND PROCEDURAL LEGITIMACY REPLICATION COMPARISON..... 21

FIGURE A8: UNHCR DIVERSITY AND SUBSTANTIVE LEGITIMACY REPLICATION COMPARISON 22

A7: HETEROGENUOUS EFFECTS 23

FIGURE A9: UNHCR PANEL DIVERSITY AND PROCEDURAL LEGITIMACY, BY POLITICAL ATTITUDE..... 23

FIGURE A10: UNHCR PANEL DIVERSITY AND SUBSTANTIVE LEGITIMACY, BY POLITICAL ATTITUDE .. 25

FIGURE A11: CONDITIONAL MARGINAL TREATMENT EFFECTS OF DIVERSITY, BY PARTY (2021)..... 27

A1: UNHCR Panel Picture Selection Criteria

For our descriptive representation treatments, we displayed photos of an eight-person UNHCR panel. To select our photos, we conducted a pre-test survey of fifty photos of generic, middle-aged professionals from visually different racial groups purchased from Shutterstock. To ensure comparability, we asked respondents to rate each photo on a hundred-point scale of the age, attractiveness, likeability, and competence. From the data, we selected eight white-male photos closest to the global mean on an equal weighted scale of

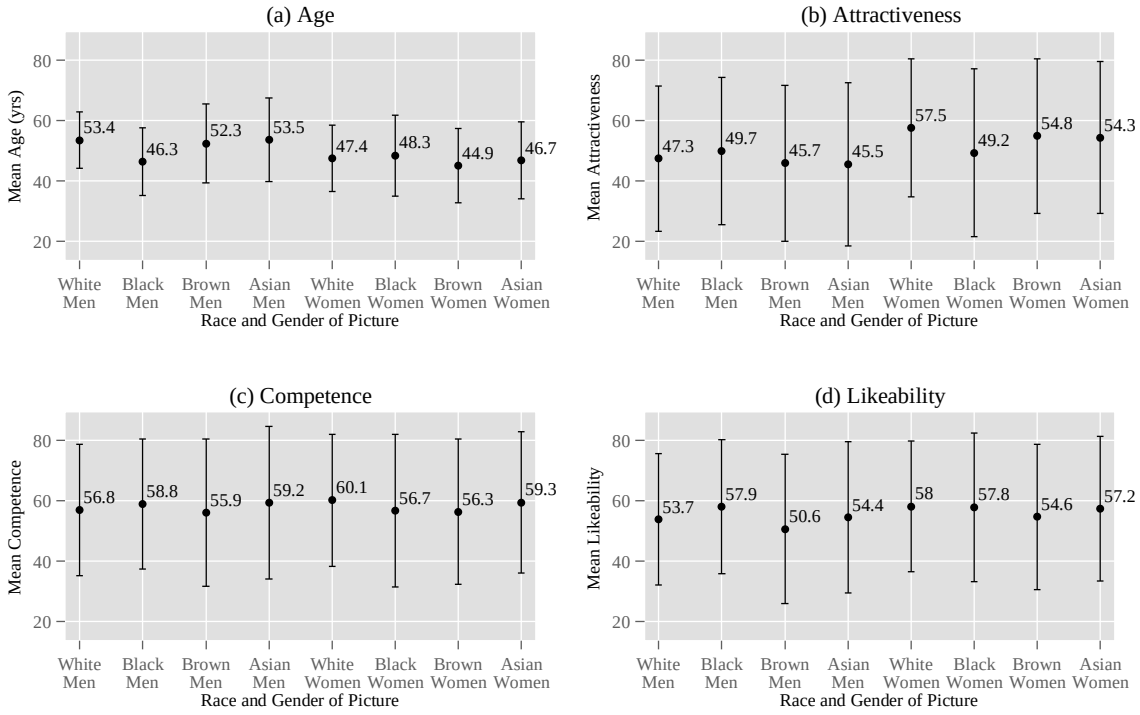
Figure A1: Picture Position for UNHCR Panel

Position 1	**Position 2**	**Position 3**	**Position 4**
Position 5	**Position 6**	**Position 7**	**Position 8**

age, attractiveness, likeability, and competence for the all-white, all male panel. For the all-white, mixed-gender panel and the mixed-race, all-male panel, the same white-male photos were in photo positions 1, 3, 6, and 8 while the remaining positions were occupied by either all-white, all-female or mixed-race, all-male photos (see Figure A1). For the mixed-race, mixed-gender panel, only photo position 1 retains a white-male photo (same photo throughout all four treatments). This white-male photo was selected because it had the smallest difference to all other photo groups. Positions 1, 3, 6, and 8 were male photos of varying racial groups. Position 1 held the white-male photo; position 3 held the black-male

photo; position 6 held the Hispanic/Arab-male photo; and position 8 held the Asian-male photo. For the female members, positions 2, 4, 5, and 7 were female photos where position 4 was the white-female photo (same position as the all-white, all-female panel). Position 2 had the Asian-female photo; position 5 had the black-female photo; and position 8 had the Hispanic/Arab-female.

Figure A2: Panel Picture Ratings of Competence, Likeability, Attractiveness, and Age



Notes: All subfigures contain the mean ratings for each of the four categories (age, attractiveness, competence, and likeability) with their accompanying standard deviation.

To describe the comparability measures in greater detail, consider Figure A2, which displays the mean respondent scoring in the pretest on age, attractiveness, competence, and likeability. The selected photos were chosen to be as similar as possible to the all-white, all-male panel on all four indices. Figures A2c and A2d show that we were able to obtain balanced scores

(no statistical difference) on competence and likeability with a slight deviation for Hispanic/Arab-males, which may largely be attributed to certain feelings of national security threat. As for attractiveness, females were rated more attractive than males with exception of black-females (Figure A2b). The largest challenge was achieving similarity in age since females were rated younger than their male counterparts except for black-males (see Figure A2a). However, it should be noted that the largest difference in age is around 8 years (between white-males and Asian-females), but this age difference did not lead to substantial differences in perceived competence ratings or likeability across groups. It should be noted that attractiveness remains gender-specific (women are rated more attractive than men) rather than race-specific, but this did not lead to systematic differences in likability and competence. These results give us confidence that the main factors in the photos are not driven by any systematic desirability bias generated by the linking photos with specific personality characteristics (Larson 2019).

A2: Replication

For the follow-up survey in 2021, we introduced country labels to the UNHCR panel members' pictures in the original experimental design as well as a replication of the original experimental design. As discussed in the main text, we kept every part of the experiment the same from the position of the photos to the treatment outcomes. For the country labels, we chose four European countries and four Latin American countries, which were selected based on racial and ethnic diversity. For Latin American countries, we selected the four most ethnically diverse countries to ensure that photos do not appear out of place. Figure A3 displays the potential four panels that respondents may have viewed with the corresponding country labels.

Figure A3: UNHCR Panel with Country Labels, by Racial and Gender Distribution

(a) All-White, All-Male

(b) Mixed-Race, All-Male

(c) All-White, Mixed-Gender

(d) Mixed-Race, Mixed-Gender

A3: Summary Statistics Between Our Sample and National Demographics

Table A1: Sample Statistics Characteristics

Gender	**2019 Sample (%)**	**2021 Sample (%)**
Female	50.06	48.66
Male	49.94	51.34
Race	**2019 Sample (%)**	**2021 Sample (%)**
White/Caucasian	68.77	66.19
Non-white Hispanic	10.80	15.52
African-American	13.09	12.89
Asian/Indigenous/Other	7.35	5.40
Age (years)	**2019 Sample (%)**	**2021 Sample (%)**
18-24	15.88	13.83
25-34	18.93	19.44
35-44	19.12	16.55
45-54	14.67	16.28
55-64	14.63	16.00
>65	16.78	17.89
Party Identification	**2019 Sample (%)**	**2021 Sample (%)**
Independent	32.97	33.55
Democrat	38.08	35.20
Republican	28.95	31.25
Education	**2019 Sample (%)**	**2021 Sample (%)**
High School or Less	30.21	30.35
Some College	31.89	37.41
College/University	23.40	20.85
Postgraduate/Professional	14.50	11.39

Our online surveys (for both the main survey and the follow-up survey) generally match key demographics on gender, race, age, education, and party affiliation with respect to nationally representative samples. The gender ratio matches national percentages while our sample is slightly younger than the national sample. For racial demographics, our sample had fewer non-white Hispanics and more whites/Caucasians. Our sample is also more partisan than the national average (8 percent more Democrats and 3 percent more Republicans). The 2021 follow-up survey had similar demographics. Sample treatment balance is displayed in Table A2 (main survey), Tables A3 (replication), Table A4 (replication with country labels).

Table A2: Sample Treatment Balance (Main Survey)

	UNHCR Report Not Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	52.13	50.78	46.89	54.14
Male	47.87	49.22	53.11	45.86
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	70.72	65.29	68.15	69.81
Non-white Hispanic	9.97	11.15	8.60	8.12
African-American	9.97	13.69	16.56	14.29
Asian/Indigenous/Other	9.35	9.87	6.69	7.79
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	14.63	15.89	14.91	19.43
25-34	21.65	15.89	16.77	20.70
35-44	20.73	21.50	18.94	16.56
45-54	13.11	15.58	13.04	11.46
55-64	14.63	12.46	13.04	17.20
>65	15.24	18.69	23.29	14.65
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	33.44	34.08	35.58	33.11
Democrat	36.56	38.26	36.54	38.3
Republican	30.00	27.65	27.88	28.52
	UNHCR Report Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	48.73	48.89	50.15	47.80
Male	51.27	51.11	49.85	52.20
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	66.56	71.43	69.43	68.71
Non-white Hispanic	12.25	11.69	13.38	11.29
African-American	14.90	11.36	10.83	13.23
Asian/Indigenous/Other	6.29	5.52	6.37	6.77
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	16.56	15.87	12.92	16.98
25-34	19.43	20.32	18.77	17.92
35-44	16.56	16.51	22.15	19.81
45-54	15.29	17.46	16.00	15.41
55-64	15.29	15.56	13.85	15.09
>65	16.88	14.29	16.31	14.78
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	30.64	33.33	31.73	31.70
Democrat	41.41	34.98	39.74	38.89
Republican	27.95	31.68	28.53	29.41

Table A3: Sample Treatment Balance (Follow-up Replication without Country Labels)

	UNHCR Report Not Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	55.68	48.67	50.52	53.59
Male	44.32	51.33	49.48	46.41
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	63.24	63.76	64.77	66.48
Non-white Hispanic	16.22	16.11	18.65	16.76
African-American	11.89	17.45	12.44	12.85
Asian/Indigenous/Other	8.65	2.68	4.15	3.91
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	18.38	16.00	16.49	12.71
25-34	12.43	16.67	18.56	22.10
35-44	17.84	18.67	16.49	19.34
45-54	16.22	12.67	14.95	14.92
55-64	16.76	18.67	12.89	14.36
>65	18.38	17.33	20.62	16.57
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	32.97	32.67	38.14	27.62
Democrat	40.00	34.67	30.93	39.23
Republican	27.03	32.67	30.93	33.15
	UNHCR Report Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	55.95	53.76	50.82	52.49
Male	44.05	46.24	49.18	47.51
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	72.29	61.08	67.96	63.33
Non-white Hispanic	11.45	16.22	13.81	17.22
African-American	12.05	17.84	12.15	13.33
Asian/Indigenous/Other	4.22	4.86	6.08	6.11
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	11.31	14.52	10.93	16.02
25-34	14.29	18.28	21.86	19.34
35-44	19.05	19.35	14.21	17.13
45-54	22.02	17.20	15.85	12.15
55-64	15.48	12.90	16.39	13.81
>65	17.86	17.74	20.77	21.55
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	27.98	31.72	38.80	29.83
Democrat	34.52	36.02	37.70	34.81
Republican	37.50	32.26	23.50	35.36

Table A4: Sample Treatment Balance (Follow-up Replication with Country Labels)

	UNHCR Report Not Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	48.77	50.50	51.40	48.94
Male	51.23	49.50	48.60	51.06
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	65.52	67.82	66.48	66.13
Non-white Hispanic	17.24	15.35	16.48	13.98
African-American	10.84	12.38	10.23	13.44
Asian/Indigenous/Other	6.40	4.46	6.82	6.45
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	14.29	12.38	13.41	11.17
25-34	21.67	18.32	19.55	23.94
35-44	19.21	15.35	16.76	13.83
45-54	14.29	19.31	17.88	16.49
55-64	14.29	18.32	15.64	17.55
>65	16.26	16.34	16.76	17.02
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	33.50	35.64	29.05	38.30
Democrat	33.99	35.64	35.75	42.02
Republican	32.51	28.71	35.20	19.68
	UNHCR Report Critical of U.S.			
Gender	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Female	53.33	51.74	50.26	44.79
Male	46.67	48.26	49.74	55.21
Race	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
White/Caucasian	64.25	68.02	67.55	71.17
Non-white Hispanic	16.76	15.12	13.30	12.88
African-American	13.41	13.37	13.30	9.82
Asian/Indigenous/Other	5.59	3.49	5.85	6.13
Age (years)	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
18-24	15.56	15.12	13.61	9.20
25-34	18.33	18.02	23.56	23.31
35-44	12.22	18.60	9.95	17.79
45-54	15.00	18.02	18.85	14.11
55-64	16.11	15.12	14.66	24.54
>65	22.78	15.12	19.37	11.04
Party Identification	**White-Male**	**Mixed-Male**	**White-Male**	**Mixed-Mixed**
Independent	37.22	31.40	34.55	36.20
Democrat	26.11	38.37	31.94	31.29
Republican	36.67	30.23	33.51	32.52

A4: Analysis with Control Variables

We recorded demographic characteristics that could potentially confound our results. We asked questions about the respondents' gender, political party affiliation, income, age, self-identified racial identity, patriotism, and our measurement of anti-immigration sentiment and white nationalism. We included these demographic controls because Republicans, men, white Americans, older individuals, and those with greater antipathy towards immigrants or strong attachment to white nationalism, or exhibit extreme patriotism are potentially associated with attitudes toward foreign policy.

For these demographic variables, we measure patriotism using a conventional, additive index of three commonly used questions on patriotism (Mansfield and Mutz 2009). As for our white nationalism measure, we used two sets of questions that focused on contemporary perceptions of racism both in the broader U.S. society and regarding individual racial groups (Hooghe and Dassonneville 2018). By doing so, we sought to capture two elements of white identity based on racial resentment and perceived white victimization (Sides, Tesler, and Vavreck 2018). Using responses to these questions, we constructed a normalized measure of white nationalism using factor analysis.[1]

For contemporary perception of racism, we asked respondents "to what degree do you agree that racism is still a problem in the US?" For the perception towards different groups, we borrow from the ANES and ask respondents "how much discrimination is there against the following groups?" The groups mentioned range from racial groups (e.g., whites, blacks) to gendered groups (e.g., women, gays and lesbians). Using responses to these questions, we

[1] We narrowed the discrimination groups by selecting Blacks, non-white Hispanics, Whites, men, women, and Christians. While Muslims and Asian are included in the broader questionnaire, using Cronbach's alpha (estimated accuracy of .89) and a Bayesian Information Criteria, we found that the aforementioned listed variables are optimal for generating our group of responses (factors) into a single dimension that we believe to be expressed as white nationalism.

constructed an index of white nationalism using factor analysis.[2] From this, we were able to construct a single dimension of white nationalism. We also used three sets of questions to construct the anti-immigration measure. Specifically, we asked respondents three sets of questions: (1) whether the number of foreign immigrants admitted to the United States should be increased, decreased or stay the same; (2) to what extent do respondents agree or disagree that immigration is good for the American economy or culture in the U.S.; and (3) to what extent do respondents agree or disagree that the U.S. should accept refugees fleeing violence and persecution.[3]

To construct the unidimensional measure, we preferred to use factor analysis (henceforth referred to as FA) compared to Item Response Theory (IRT) in our analysis since we are using a method to reduce the number of dimensions (questions) into a single dimension, which is explicitly assumed in IRT but not with FA. Moreover, FA handles missing data much better than IRT. Regardless of the approach taken, the final variable is quite similar.

[2] We specifically narrowed the discrimination groups by selecting blacks, non-white Hispanics, whites, men, women, and Christians. While Muslim and Asian are included in the broader questionnaire, using Cronbach's alpha (estimated accuracy of .89) and a Bayesian Information Criteria, we found that our selected variables are optimal for generating our group of responses (factors) into a single dimension of white nationalism.

[3] Cronbach's alpha for the three sets of immigration questions is an estimated accuracy of 0.81.

Figure A4: FA and IRT Construction of White Nationalism and Anti-Immigration

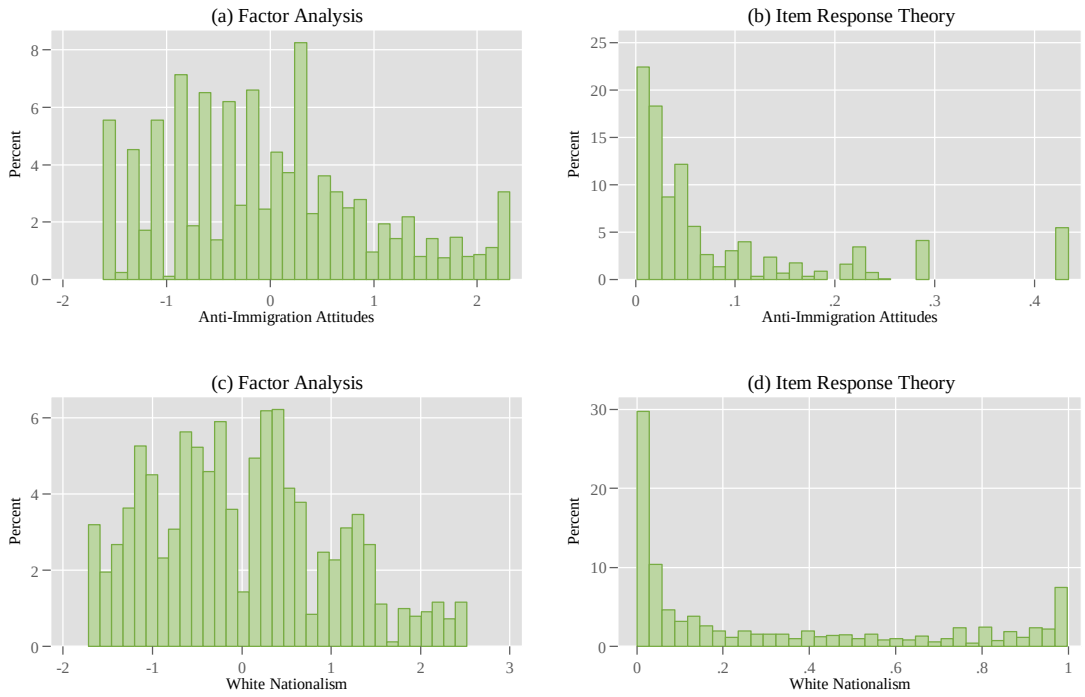


Figure A4 shows the distribution of the white nationalism (bottom-half) and anti-immigration (top-half) variable using factor analysis (A4a and A4c) and IRT (A4b and A4d). While the distributions differ based on their different assumptions of the underlying distribution, a bivariate linear regression reveals a tight linear correlation with the correlation of .87 and .76 between the two variables for white nationalism (bottom-figures) and anti-immigration (top-figures). Given their similarity, we prefer to use FA given its standardization and easier interpretation.

Table A5: Trust and Fairness of the UNHCR Decision-Making Process

	Refugees Treated Fairly		UNHCR Can Be Trusted	
	Positive Report	Negative Report	Positive Report	Negative Report
Mixed-Mixed	**0.55*** **	**0.28** **	**0.39*** **	**0.33** **
	(0.11)	**(0.11)**	**(0.11)**	**(0.11)**
Mixed-Males	0.13	0.06	0.32**	0.11
	(0.11)	(0.11)	(0.11)	(0.11)
White-Mixed	0.08	0.18	0.08	0.23*
	(0.11)	(0.11)	(0.11)	(0.11)
Household Income	0.03*	-0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Age (yrs.)	0.01	0.02	-0.05	0.01
	(0.03)	(0.02)	(0.03)	(0.02)
Democrats	-0.06	0.33***	0.02	0.31**
	(0.10)	(0.09)	(0.10)	(0.09)
Republicans	0.45***	0.20*	0.44***	0.18
	(0.10)	(0.10)	(0.10)	(0.10)
White	0.23*	0.13	0.31***	-0.08
	(0.09)	(0.09)	(0.09)	(0.09)
Female	-0.23**	-0.00	-0.46***	-0.08
	(0.08)	(0.08)	(0.08)	(0.08)
Education	0.05	-0.01	0.12**	-0.04
	(0.05)	(0.04)	(0.05)	(0.04)
Anti-immigration	0.14**	-0.29***	0.18***	-0.28***
	(0.05)	(0.05)	(0.05)	(0.05)
White Nationalism	0.20***	-0.16***	0.21***	-0.17***
	(0.05)	(0.05)	(0.05)	(0.05)
Patriotism	0.06***	0.01	0.08***	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-1.46***	-0.58**	-1.42***	-0.57**
	(0.22)	(0.21)	(0.22)	(0.21)
Pseudo R^2	0.160	0.079	0.177	0.076
Observations	1213	1188	1213	1188

Robust standard errors in parentheses *p<0.05 **p<0.01 ***p<0.001

Table A5 replicates the analysis in Figure 4 of the main text with the inclusion of the control variables. The results are qualitatively identical: mixed-mixed panels are positively and significantly viewed as more procedurally legitimate compared to white-male panels.

Table A6: UNHCR Report Outcome for Refugees and U.S. Citizens

	Outcome Good for Refugees		Outcome Good for Americans	
	Non-Critical Report	Critical Report	Non-Critical Report	Critical Report
Mixed-Mixed	**0.30****	**0.11**	**0.46*****	**0.10**
	(0.11)	**(0.12)**	**(0.11)**	**(0.11)**
Mixed-Males	0.14	-0.10	0.25*	-0.03
	(0.11)	(0.11)	(0.12)	(0.11)
White-Mixed	0.02	-0.00	0.19	0.06
	(0.11)	(0.11)	(0.11)	(0.11)
Household Income	0.05***	0.02	0.05***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Age (yrs.)	-0.08**	-0.03	-0.06*	-0.08**
	(0.03)	(0.03)	(0.03)	(0.03)
Democrats	-0.04	0.26**	-0.02	0.40***
	(0.10)	(0.10)	(0.10)	(0.10)
Republicans	0.33**	0.21	0.50***	0.08
	(0.10)	(0.11)	(0.11)	(0.11)
White	0.25**	0.12	0.29**	0.10
	(0.09)	(0.10)	(0.09)	(0.10)
Female	-0.36***	0.13	-0.39***	0.08
	(0.08)	(0.08)	(0.08)	(0.08)
Education	0.01	-0.02	0.04	0.01
	(0.05)	(0.05)	(0.05)	(0.05)
Anti-immigration	0.23***	-0.47***	0.25***	-0.39***
	(0.05)	(0.05)	(0.05)	(0.05)
White Nationalism	0.25***	-0.21***	0.25***	-0.25***
	(0.05)	(0.05)	(0.05)	(0.05)
Patriotism	0.11***	-0.00	0.09***	0.00
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-1.46***	0.18	-1.40***	0.11
	(0.22)	(0.23)	(0.22)	(0.23)
Pseudo R^2	0.211	0.161	0.228	0.169
Observations	1213	1188	1213	1188

Robust standard errors in parentheses *p<0.05 **p<0.01 ***p<0.001

Table A6 replicates the analysis in Figure 5 of the main text with the inclusion of the control variables. The results are qualitatively identical: mixed-mixed panels are positively and significantly perceived as more substantively legitimate compared to white-male panels when the UNHCR issues a non-critical report.

A5: Alternative Dependent Variable Measurement Analysis

Table A7: Alternative Dependent Variable (Factor Analysis Construction)

	Procedural Legitimacy		Substantive Legitimacy	
	Non-Critical Report	Critical Report	Non-Critical Report	Critical Report
Mixed-Mixed	**0.28*** **	**0.15** **	**0.34*** **	**0.12***
	(0.06)	**(0.06)**	**(0.07)**	**(0.06)**
Mixed-Males	0.06	0.09	0.21**	0.00
	(0.05)	(0.06)	(0.07)	(0.06)
White-Mixed	0.07	0.21***	0.11	0.03
	(0.06)	(0.06)	(0.07)	(0.06)
Household Income	-0.01*	-0.01	0.03***	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Age (yrs.)	0.05***	0.04**	-0.03*	-0.01
	(0.01)	(0.01)	(0.02)	(0.01)
Democrats	0.02	0.04	-0.12	0.19***
	(0.05)	(0.05)	(0.06)	(0.05)
Republicans	0.11*	0.07	0.29***	0.05
	(0.05)	(0.06)	(0.06)	(0.06)
White	0.01	-0.07	0.17**	0.05
	(0.05)	(0.05)	(0.06)	(0.05)
Female	0.06	-0.07	-0.27***	0.07
	(0.04)	(0.04)	(0.05)	(0.04)
Education	0.03	0.02	0.03	-0.02
	(0.02)	(0.02)	(0.03)	(0.02)
Anti-immigration	-0.05*	0.08**	0.21***	-0.31***
	(0.02)	(0.03)	(0.03)	(0.03)
White Nationalism	0.03	0.03	0.21***	-0.17***
	(0.02)	(0.03)	(0.03)	(0.03)
Patriotism	-0.02*	0.00	0.07***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Constant	-0.04	-0.30*	-1.16***	0.02
	(0.11)	(0.12)	(0.14)	(0.12)
R^2	0.047	0.056	0.372	0.265
Observations	1213	1188	1213	1188

Robust standard errors in parentheses *p<0.05 **p<0.01 ***p<0.001

Table A7 replicates the analysis in the main text using factor analysis to construct the procedural and substantive legitimacy analysis. Results are identical to the main text results, confirming H1c and H2a and H2b.

Table A8: Descriptive Representation Effect on Substantive and Procedural Legitimacy

	Good for Refugees		Good for Americans		UNHCR is Fair	UNHCR is Trusted
	Non-critical Report	Critical Report	Non-critical Report	Critical Report	Combined Reports	Combined Reports
Mixed-mixed	**0.32*** **	**0.10**	**0.41*** **	**0.10**	**0.40*** **	**0.40*** **
	(0.10)	**(0.09)**	**(0.10)**	**(0.09)**	**(0.07)**	**(0.07)**
Mixed-Males	0.26**	-0.04	0.25*	-0.01	0.11	0.22**
	(0.10)	(0.09)	(0.10)	(0.09)	(0.07)	(0.07)
White-Mixed	0.10	-0.06	0.18	-0.01	0.13	0.16*
	(0.10)	(0.09)	(0.10)	(0.09)	(0.07)	(0.07)
Household Income	0.04***	0.01	0.04***	-0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age (yrs.)	-0.06**	-0.03	-0.05*	-0.07***	0.03*	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Democrats	-0.17	0.23**	-0.17	0.31***	0.04	0.10
	(0.09)	(0.08)	(0.09)	(0.08)	(0.06)	(0.06)
Republicans	0.33***	0.04	0.40***	-0.10	0.26***	0.22**
	(0.09)	(0.09)	(0.09)	(0.10)	(0.07)	(0.07)
White	0.22**	0.10	0.22**	0.07	0.12*	0.06
	(0.08)	(0.07)	(0.08)	(0.08)	(0.06)	(0.06)
Female	-0.38***	0.13*	-0.39***	0.12	-0.04	-0.21***
	(0.07)	(0.06)	(0.07)	(0.07)	(0.05)	(0.05)
Education	0.03	-0.04	0.05	-0.03	0.02	0.04
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Anti-immigration	0.31***	-0.45***	0.29***	-0.44***	-0.06	-0.03
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)
White Nationalism	0.27***	-0.23***	0.28***	-0.28***	0.09**	0.05
	(0.04)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)
Patriotism	0.10***	0.01	0.09***	0.01	0.02*	0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Constant	1.96***	3.65***	2.07***	3.68***	2.50***	2.39***
	(0.19)	(0.17)	(0.19)	(0.18)	(0.14)	(0.14)
R-squared	0.341	0.261	0.343	0.300	0.048	0.057
Observations	1213	1188	1213	1188	2401	2401

Robust Standard errors in parentheses * p<0.05 ** p<0.01 *** p<0.001

Table A8 replicates our main analysis with an alternative measurement of the dependent variable (ordered scale). Since the normative effects were the same for the procedural measurements, for presentational ease, we combined the fairness and trust measurements of the dependent variable in this table. The results are qualitatively the same: mixed-mixed panel compared to the white-male panels increased perceptions of substantive legitimacy

conditional on whether the report is non-critical of the U.S. (columns 1 and 3) and increased procedural legitimacy (columns 5 and 6).

A6: Pre-treatment Effects

Figure A5: UNHCR Panel Diversity and Legitimacy, by Media Attention

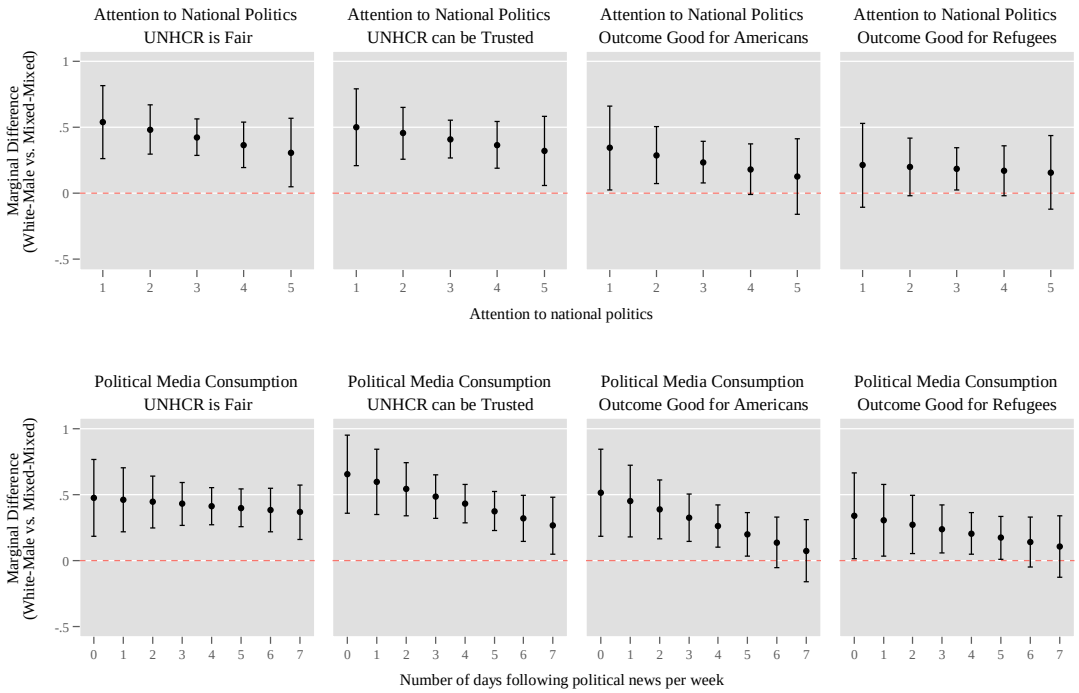


Figure A5 displays the marginal treatment effect on legitimacy between all-white, all-male and mixed-race, mixed-gender panels conditional on respondent media attention (top-half) and weekly media consumption (bottom-half). Generally, media attention and frequency of media consumption does not qualitatively change the results in the main text. There are some effects at the margins for frequency of media consumption with respect to the perceived outcome measures, but overall, the figures suggest that pre-treatment effects are minimal. However, this evidence is not enough to completely dismiss pre-treatment concerns as we do not know exactly what type of media and corresponding information that respondents consumed in 2019. Thus, we decided to run another round of surveys to replicate our results in 2021 to ensure that our results were not affected by the 2019 context, which was discussed in the main text (see Figure 8) and further discussed later in Appendix A7.

Figure A6: UNHCR Panel Diversity and Legitimacy, by Respondent Attentiveness

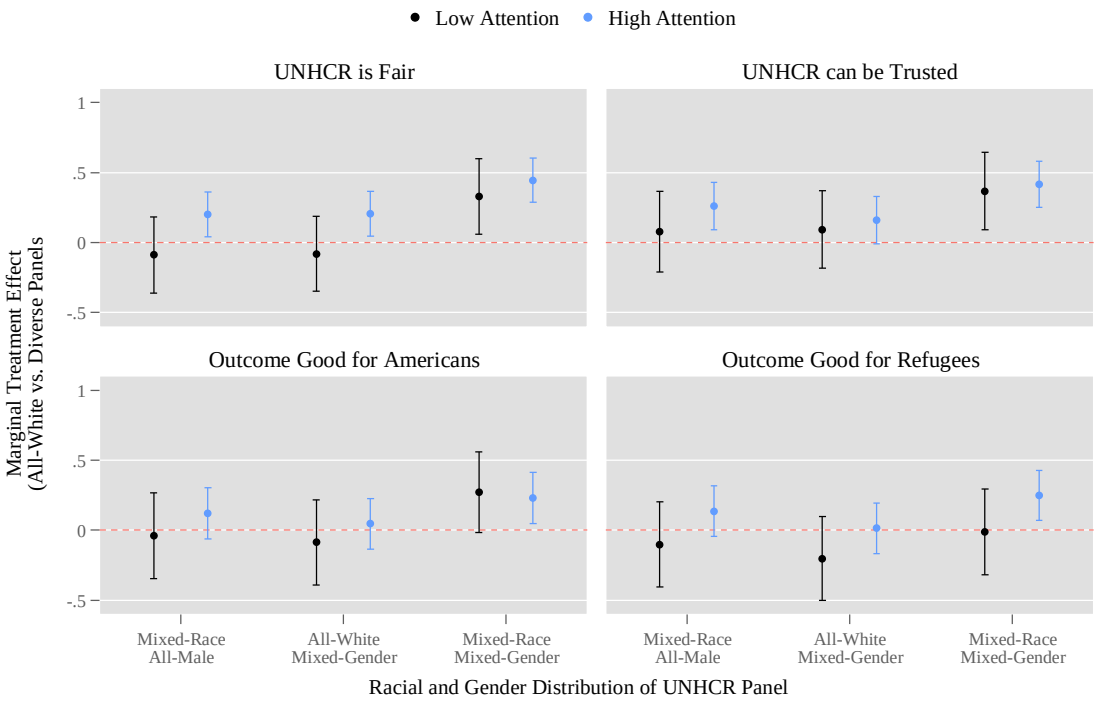


Figure A6 displays the marginal treatment effect between the all-white, all-male panels and the diversity panels on UNHCR legitimacy. The vertical axis displays the marginal difference between the treatment effects while the horizontal axis displays the individual diversity panels with the all-white, all-male panel as the reference group. Each subfigure is broken into a low-attention (black) and high-attention group (blue). There are concerns that Lucid respondents may exhibit less attentiveness, leading to underestimation effects (Aronow et al. 2020). While Figure A6 suggests that low attentive respondents have attention bias and exhibit attention bias, it should be noted that roughly 76 percent of our respondents passed the simple attention check of following a simple knowledge instruction question. So, it could simply be a smaller sample effect (lack of statistical power) or an inattentiveness problem for a smaller subset of the sample.

Figure A7: UNHCR Diversity and Procedural Legitimacy Replication Comparison

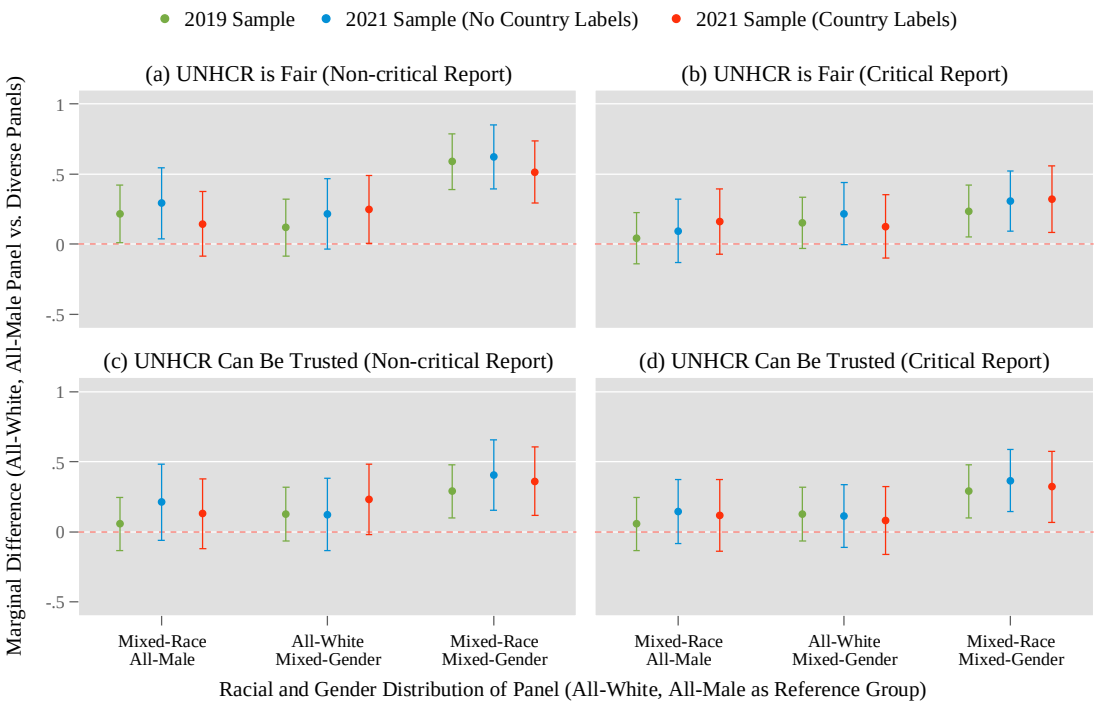


Figure A7 compares the diversity panels and procedural legitimacy results found in Figure 4 in the main text (green) with the follow-up survey replication (blue) and survey replication with country labels (red). The results are identical across all three samples.

Figure A8: UNHCR Diversity and Substantive Legitimacy Replication Comparison

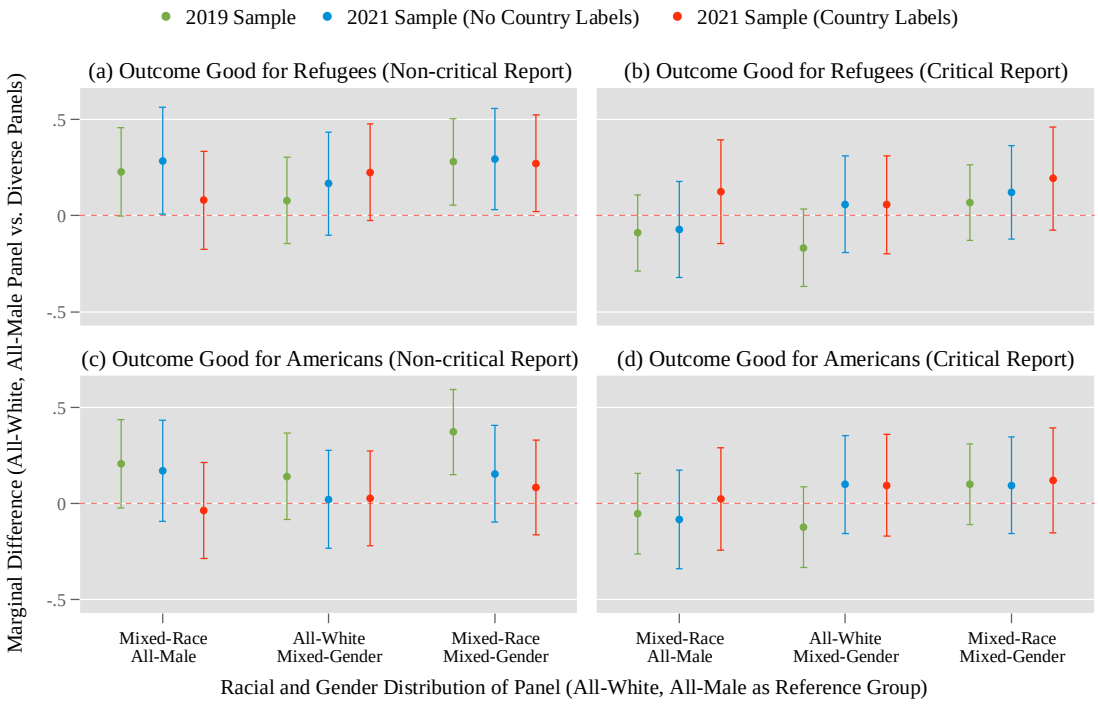


Figure A8 compares the diversity panels and procedural legitimacy results found in Figure 5 in the main text (green) with the follow-up survey replication (blue) and survey replication with country labels (red). The results are identical across all three samples for the refugee outcome measurement but not the outcome measurement for Americans, which we discussed in the main text (see page 30).

A7: Heterogeneous Effects

To explore heterogeneous effects, we test for conditional effects by respondents' prior political attitudes on their perception of UNHCR legitimacy. Consistent with prior literature, we expect such perception should be correlated with respondents' immigration and racial attitudes (Bansak, Hainmueller, and Hangartner 2016; Sides, Tesler, and Vavreck 2018). We discussed the construction of these measures earlier in Appendix A4 (see pp. 11-13).

Figure A9: UNHCR Panel Diversity and Procedural Legitimacy, by Political Attitudes

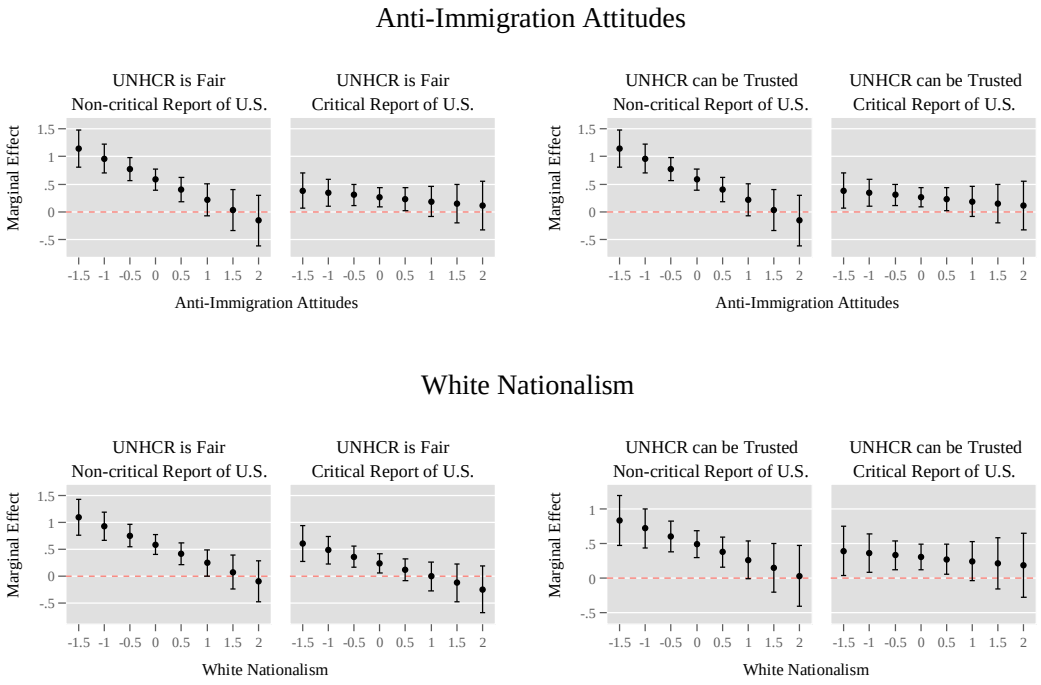


Figure A9 displays how immigration attitudes and white nationalist tendencies influence the legitimacy-enhancing effect of diversity on procedural dimension of IO legitimacy. Each set of subfigures is split between the UNHCR report that is critical of the U.S. and non-critical of the U.S. The horizontal axis for the top-half depicts a respondent's normalized anti-immigration sentiment where higher values correspond to greater anti-immigration attitudes. The bottom-half displays our normalized measure of white nationalism where higher values

indicate greater sentiments of white nationalism. The vertical axis depicts the marginal treatment effect between the mixed-race, mixed-gender panel compared to the all-white, all-male panel conditioned by political attitudes with accompanying 95 percent confidence intervals.

All four subfigures depict the same relationship: respondents who exhibited lower values of anti-immigration or white nationalist sentiment rated the mixed-mixed UNHCR panel as procedurally more legitimate than the white-male panel. As anti-immigration and white nationalist sentiment rises, the marginal effect on procedural legitimacy on the difference between mixed-mixed and white-male panels becomes statistically indistinguishable for anti-immigration sentiments values of 0.5 (A9b and A9d) or 1 (A9a and A9c) and for the white nationalist values of 0.5 (A9e, A9g, A9h) and 0 (A9f). These results put a scope condition on which types of respondents respond best to diversity: those with less anti-immigration and white nationalist sentiment. These results further support H2a as the type of report issued by the UNHCR does not influence respondents one way or another.

Figure A10: UNHCR Panel Diversity and Substantive Legitimacy, by Political Attitude

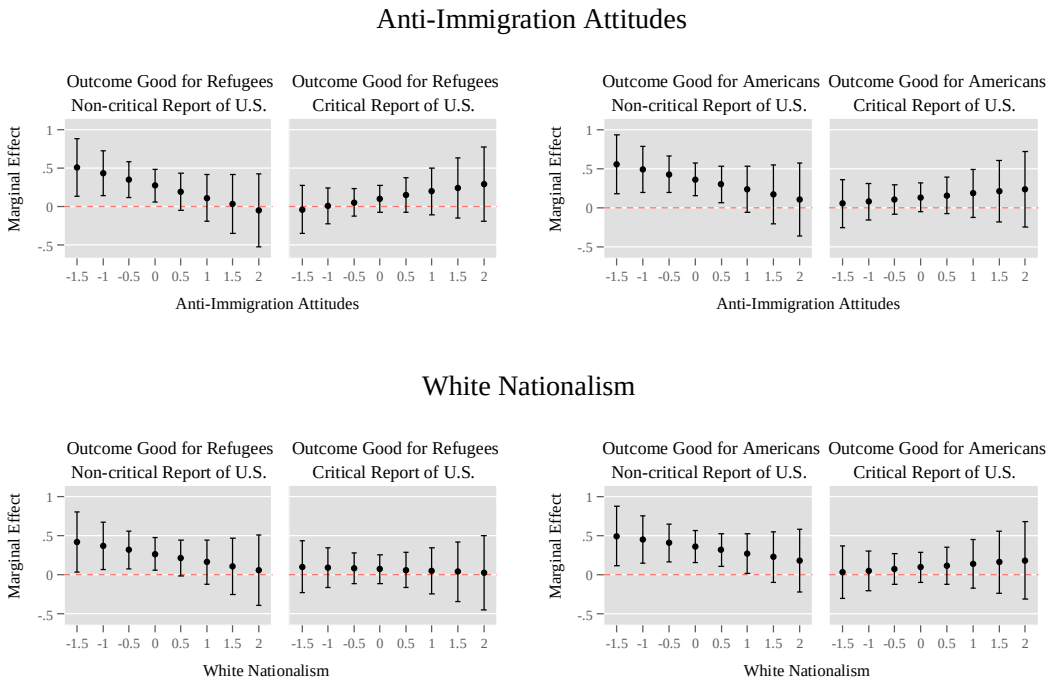


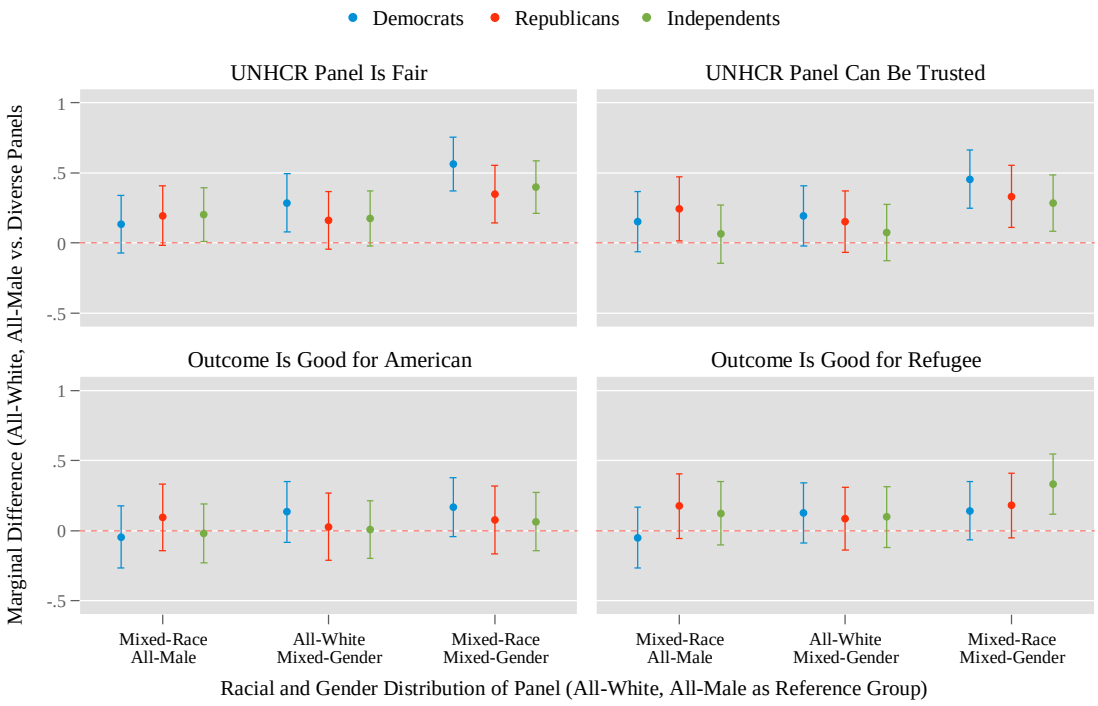
Figure A10 repeats the results in Figure A7 for our substantive legitimacy measures. As before, all subfigures are split between the UNHCR report that is critical of the U.S. and non-critical of the U.S. The horizontal axis for the top-half depicts a respondent's normalized anti-immigration sentiment while and the bottom-half displays our normalized measure of white nationalism. The vertical axis depicts the marginal treatment effect between the mixed-race, mixed-gender panel compared to the all-white, all-male panel conditioned by political attitudes with accompanying 95 percent confidence intervals.

The four subfigures for non-critical UNHCR reports (A10a, A10c, A10e, and A10g) show that respondents with lower values of anti-immigration or white nationalist sentiment rated the mixed-mixed UNHCR panel as substantively more legitimate than the white-male panel. As respondent anti-immigration and white nationalist sentiment rises, the marginal effect on the difference between mixed-mixed and white-male panels dissipates and becomes insignificant. For the case when the UNHCR issues a critical report of the U.S., the marginal

effect of mixed-mixed panels for these four subfigures is broadly insignificant for all values of anti-immigration and white nationalist sentiment.

These results suggest that respondents with prior views on immigration and race influence how they view the outcome as more favorable for refugees and Americans (roughly a quarter of a standard deviation) if the decision was made by a mixed-mixed panel as opposed to a white-male panel. These results are partly consistent with findings from Clayton et al. (2019) on the crystallization mechanism, and how prior political beliefs identified by Hayes and Hibbing (2017) might shape respondent views on the legitimacy enhancing effects of diversity.

Figure A11: Conditional Marginal Treatment Effects of Diversity, by Partisanship (2021)



The difference in media coverage of two events may account for some of the partisan heterogeneous effects found in the 2019 survey. In Figure A11, we repeated the analysis from Figure 7 in the main text for the 2021 sample and found that regardless of whether respondents were Republican, Democrat, or Independent, they perceived that the mixed-race, mixed-gender UNHCR panel to be fairer and more trustworthy compared to the all-white, all-male panel. Republicans but not Democrats also perceived that the outcome was better for refugees when comparing a mixed-race, mixed-gender panel and all-white, all-male panel. These differences suggest that the media environment on refugee policy may have influenced how Republicans and Democrats reacted to our treatments. That is, respondent partisanship may have been influenced by media reporting on refugees, which affected the impact of our diversity treatments on perceived UNHCR legitimacy.

Appendix References

Aronow, P. M., Joshua Kalla, Lilla Orr, and John Ternovski. 2020. "Evidence of Rising Rates of Inattentiveness on Lucid in 2020." SocArXiv. https://doi.org/10.31235/osf.io/8sbe4.

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner. 2016. "How Economic, Humanitarian, and Religious Concerns Shape European Attitudes toward Asylum Seekers." *Science* 354 (6309): 217–22.

Hooghe, Marc, and Ruth Dassonneville. 2018. "Explaining the Trump Vote: The Effect of Racist Resentment and Anti-Immigrant Sentiments." *PS: Political Science & Politics* 51 (3): 528–34.

Larson, Ronald B. 2019. "Controlling Social Desirability Bias." *International Journal of Market Research* 61 (5): 534–47.

Mansfield, Edward D., and Diana C. Mutz. 2009. "Support for Free Trade: Self-Interest, Sociotropic Politics, and Out-Group Anxiety." *International Organization* 63 (3): 425–57.

Sides, John, Michael Tesler, and Lynn Vavreck. 2018. *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America*. Princeton: Princeton University Press.