sociological science

Supplement to:

Quillian, Lincoln, Anthony Heath, Devah Pager, Arnfinn H. Midtbøen, Fenella Fleischmann, and Ole Hexel. 2019. "Do Some Countries Discriminate More than Others? Evidence from 97 Field Experiments of Racial Discrimination in Hiring." Sociological Science 6: 467-496.

Online Supplement for "Do Some Countries Discriminate More than Others? Evidence from 97 Field Experiments of Racial Discrimination in Hiring"

Includes: Appendix 1 and Appendix 2
Table S1 to Table S4

Some parts of the description of procedure in the appendices below reprise parts of Quillian et al. (2017).

Appendix 1: Publication Bias Analysis

A potential problem for any meta-analysis is publication bias - that studies that find no differences or effects are less likely to be published and included in a meta-analysis (Sutton 2009). If this is the case, publication bias would produce an upwards bias in observed discrimination ratios, because studies finding little or no discrimination are less likely to be published (or even written up).

To reduce the potential problem we made efforts to include all available unpublished studies in our analysis. Our search included repositories of working papers and included an e-mail survey to authors to attempt to locate studies that are unpublished (see "Study Search Methods" in "Procedures" in text). This resulted in many studies that were not published as academic journal articles: of 97 studies, only 47 were journal publications at the time that they were initially coded for our meta-analysis. The others are combinations of reports, working papers, and unpublished theses.

As a further check on the problem, we employed statistical tests indicative of potential publication bias from the meta-analysis literature. These tests consider if the distribution of effect sizes (discrimination estimates in our study) are, in a graph of effect size versus standard error of effect, shaped like a funnel. Funnel asymmetry can result if some studies are "missing" systematically on one side of the graph due to publication bias. As the literature emphasizes, however, there can be causes other than publication bias of asymmetric funnel plots. These tests only provide a potential indicator rather than a clear diagnostic of publication bias. Nevertheless, they provide a sensitivity check assuming that any asymmetry is the result of publication bias (Sutton 2009).

Most tests of funnel plot asymmetry cannot include covariates. Because our focus is on differences between countries, we performed publication bias diagnostics for each of our nine countries. Our analysis found that patterns appear significantly different for white from non-white immigrant groups. For this reason, we only include effect estimates of discrimination

against non-white target groups in our publication bias analyses, excluding European immigrant target groups.

We use the trim-and-fill test from the meta-analysis literature (Duval and Tweedie 2000). This has the advantage over other meta-analysis tests that it both provides a formal test for funnel plot asymmetry and also estimates the change in results if the funnel plot were more symmetric. That is, it both provides a significance test and an estimate of how much this would change results (assuming all of the asymmetry is the result of publication bias).[1]

Results based on the trim-and-fill test are shown in appendix table S4. For comparison, column 1 shows results without adjustment for publication bias. That is, column 1 of table S4 shows estimates of the discrimination ratio for non-whites by country from country-specific random-effects meta-analyses without controls or the trim-and-fill procedure. Column 2 shows the number of studies that need to be added to produce a symmetric funnel-plot, as estimated from the trim-and-fill procedures. Finally column 3 shows the estimated country average discrimination ratio adding in the "missing" studies.

Results of the trim-and-fill test show evidence of funnel plot asymmetry in some countries: all countries except for France, Norway, and Sweden are shown as having significant funnel plot asymmetry.[2] However, when effects of adding missing studies are included, this only produces very small changes in estimated mean discrimination ratio for most countries. The relative ordering in discrimination among countries is almost unchanged. The original ordering, from highest to lowest discrimination ratios, are: France, Sweden, Britain, Belgium, Canada, Norway, USA, Netherlands, and Germany. Adding in missing studies with the trim-and-fill procedure the ordering is: France, Sweden, Belgium, Britain, Canada, Norway, Netherlands, USA, and Germany. The significantly higher position of France and Sweden overall (compared to other countries) is unchanged. The biggest change in the point estimate is for Belgium, a country with only three studies and a large standard error. Overall, our core results are unchanged adjusting for funnel-plot asymmetry. We conclude that tests for funnel plot asymmetry find some evidence of possible publication bias, but that the extent of this potential bias is too small to be of any consequence for our results.

Appendix 2: Adjustment to Discrimination Ratios in Some Multi-Stage Studies

A few studies in our sample follow a multi-stage design in measuring discrimination. This was a study design used by some studies commissioned by the International Labor Organization. In these studies the applicants first called employers by phone to inquire if a job was still available. We would like to incorporate these responses into our measures of discrimination, to get total discrimination from initial application to the callback. For situations where either both

[1] The Eggers' test showed statistically significant publication bias in three countries at $p<.05$.

[2] Interestingly, for Belgium we estimate funnel plot asymmetry based on asymmetry on the right side, suggesting potential publication bias against studies that find high levels of discrimination against non-whites relative to those that do not.

applicants were told the job is still available or both were told it is not available, this is straightforward: we include these counts in calculating the rates of callbacks.

The complication we run into is the following: In five studies, if one applicant was told the job was available and the other was not, no application was submitted by *either* tester. This last aspect of this design – that when one applicant received a positive response and the other did not, the applicant who could have then submitted a resume did not – requires some adjustment. We want to capture rates of receiving a callback for all minority and majority applicants from the point of initial application. We know that respondents who were told "no job is available" did not receive a callback. But when one member of a pair was told the job is available, and the other was not, we do not know how often the member of the pair who was told the job was available would have received a callback if they had applied. We need to estimate this to get complete callback outcomes from the point of application.

To estimate callback rates in these studies, we assume that the member of the pair who received the invitation to interview but did not submit a resume (because their partner was told the job was no longer available) was as likely to get a callback if they had submitted a resume as applicants of the same race/ethnic group in the same study for which an application was submitted.

More formally, we adjust the discrimination ratios in these five studies in the following way. Define:

n_1^w is the number of applicants from the native majority (white) group who initially call the employer to inquire if jobs are available. b_1^m is the number of applicants from the minority group who initially call the employer to inquire if jobs are available.

f_1^w is the number of applicants from the native majority (white) group who are told the job is still available. f_2^m is the number of applicants from the minority group who are told the job is still available.

n_2^w is the number of applicants from the native majority (white) group who submit application materials. n_2^m is the number of applicants from the minority group who submit application materials.

c_2^w is the number of applicants from the native majority (white) group who actually receive a callback. c_2^m is the number of applicants from the minority group who actually receive a callback.

We calculate the estimated discrimination ratio for minority group j in study i from the point of initial application with:

$$Y_{ij}^* = \left(\frac{f_1^w / n_1^w}{f_1^m / n_1^m} \right) \left(\frac{c_2^w / n_2^w}{c_2^m / n_2^m} \right)$$

This just multiplied the discrimination ratio at the stage of asking if job is still available with discrimination ratio at the stage of receiving a callback. We use this estimated discrimination ratio for these five studies.

We calculate the estimated variance of the log adjusted discrimination ratio with:

$$\text{Var}(\ln(Y_{ij}^*)) = \frac{1}{n_1^w\left(f_1^w/n_1^w\right)\left(c_2^w/n_2^w\right)} - \frac{1}{n_1^w} + \frac{1}{n_1^m\left(f_1^m/n_1^m\right)\left(c_2^m/n_2^m\right)} - \frac{1}{n_1^m}$$

This is the standard formula for the variance of a risk ratio with unpaired groups (Bornstein, Hedges, Higgins, and Rothstein 2009, formula 5.3), substituting the implied count of successes based on our estimation. Using the unpaired formula for these paired studies will slightly overstate the variance of the ratio, while treating the counts as actual rather than estimated rates somewhat understates it.

References (Online Supplement)

Duval, Sue and Richard Tweedie. 2000. "A NonParametric 'Trim and Fill' Method of Assessing Publication Bias in Meta-Analysis." *Journal of the American Statistical Association* 95(449): 89-98.

Quillian, Lincoln, Devah Pager, Ole Hexel, Arnfinn Midtbøen. 2017. "Meta-Analysis of Field Experiments shows No Change in Racial Discrimination in Hiring Over Time." *The Proceedings of the National Academy of Sciences of the United States of America* 114(41): 10870-10875.

Sutton, A.J., 2009. Publication bias, in: Cooper, H.M., Hedges, L.V., Valentine, J.C. (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation, New York.

Table S1: Meta-Regression of Discrimination Ratios, Studies Since 1989 Only

Variable	Coef (SE)
Country (reference=US)	
Belgium	-0.02
	(0.13)
Canada	0.05
	(0.14)
France	0.28 *
	(0.11)
Germany	-0.09
	(0.13)
Great Britain	0.27 *
	(0.12)
Netherlands	0.03
	(0.11)
Norway	0.01
	(0.14)
Sweden	0.21
	(0.16)
Minority Group (reference=African/black)	
European Immigrant (1=yes)	-0.07
	(0.11)
Middle-Eastern	0.06
/N. African (1=yes)	(0.08)
Hispanic (1=yes)	-0.07
	(0.09)
Asian (1=yes)	0.03
	(0.10)
Applicant Gender (reference=both)	
Testers Male Only (1=yes)	-0.02
	(0.08)
Testers Female Only (1=yes)	0.00
	(0.08)
Most Common Level of Applicant Education (reference = high school or less)	
Some College or Post-HS Vocational Degree (1=yes	0.12
	(0.10)
College or More (1=yes)	-0.15
	(0.12)
Education information missing (1=yes)	-0.17
	(0.11)

Table S1 Continued:

	Coef (SE)
Study Attributes	
Audit Study (1=yes)	0.10
	(0.07)
Year of Fieldwork (Four Digit Year)	0.01
	(0.00)
Occupation Controls	
Includes Blue Collar Jobs (1=yes)	-0.07
	(0.08)
Includes Jobs with Customer Contact (1=yes)	0.12
	(0.11)
Includes Jobs with an Office Focus (1=yes)	0.05
	(0.10)
Intercept	0.23
	(0.16)
Tau-squared (between-study var.)	0.0338
I-squared (% between-study var.)	80.1
N effects / N studies	118 / 80

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses.

Table S2: Predicted Discrimination Ratios by Country and Minority Group

Country	Minority Group	Estimate	Lower CI	Upper CI	N (Studies)
Belgium	African/Black	1.41	1.03	1.92	1
Belgium	European/White	1.15	0.84	1.57	1
Belgium	Middle-Eastern/N. African	1.40	1.03	1.89	4
Canada	African/Black	1.50	1.14	1.97	4
Canada	European/White	1.22	0.94	1.59	4
Canada	Middle-Eastern/N. African	1.48	1.08	2.04	2
Canada	Latin Am./Hispanic	1.35	0.97	1.87	1
Canada	Asian	1.56	1.17	2.07	7
France	African/Black	1.94	1.48	2.55	9
France	Middle-Eastern/N. African	1.92	1.45	2.55	17
France	Asian	2.02	1.49	2.74	2
Germany	Middle-Eastern/N. African	1.23	0.88	1.73	5
Great Britain	African/Black	1.50	1.24	1.81	10
Great Britain	European/White	1.22	0.95	1.56	6
Great Britain	Asian	1.56	1.25	1.95	16
Netherlands	African/Black	1.39	1.04	1.86	7
Netherlands	European/White	1.13	0.82	1.56	1
Netherlands	Middle-Eastern/N. African	1.38	1.00	1.89	10
Netherlands	Asian	1.45	1.03	2.03	1
Norway	Asian	1.41	1.00	2.00	4
Sweden	Middle-Eastern/N. African	1.75	1.24	2.47	8
USA	African/Black	1.35	1.04	1.76	27
USA	European/White	1.10	0.83	1.47	1
USA	Middle-Eastern/N. African	1.34	0.97	1.85	1
USA	Latin Am./Hispanic	1.22	0.89	1.65	9
USA	Asian	1.41	1.04	1.92	1

Note: predictions from table 3 model 2, controls at reference values except year = 2000.

Table S3: Meta-Regression of Log Odds of Discrimination, Native White to Minority

Variable	Coef (SE)
Country (reference=US)	
Belgium	0.07
	(0.16)
Canada	0.33
	(0.17)
France	0.46 **
	(0.14)
Germany	-0.22
	(0.20)
Great Britain	0.36
	(0.23)
Netherlands	0.09
	(0.16)
Norway	0.04
	(0.18)
Sweden	0.45 *
	(0.20)
Minority Group (reference=African/black)	
European (1=yes)	-0.42 *
	(0.16)
Middle-Eastern	-0.03
/N. African (1=yes)	(0.11)
Hispanic (1=yes)	-0.18 +
	(0.11)
Asian (1=yes)	0.07
	(0.10)
Applicant Gender (reference=both)	
Testers Male Only (1=yes)	-0.07
	(0.10)
Testers Female Only (1=yes)	-0.08
	(0.13)
Most Common Level of Applicant Education (reference = high school or less)	
Some College or Post-HS Vocational Degree (1=yes	0.00
	(0.17)
College or More (1=yes)	-0.32 *
	(0.15)
Education information missing (1=yes)	-0.23 +
	(0.13)

Table S3 Continued:

	Coef (SE)	
Study Attributes		
Audit Study (1=yes)	0.16	
	(0.10)	
Year of Fieldwork (Four Digit Year)	-0.01	+
	(0.00)	
Occupation Controls		
Includes Blue Collar Jobs (1=yes)	-0.09	
	(0.10)	
Includes Jobs with Customer Contact (1=yes)	-0.05	
	(0.14)	
Includes Jobs with an Office Focus (1=yes)	0.19	
	(0.12)	
Intercept	0.50	*
	(0.20)	
Tau-squared (between-study var.)	0.037655	
I-squared (% between-study var.)	84.98475	
N effects / N studies	159 / 97	

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses.

Table S4: Funnel-plot Asymmetry Analysis of Discrimination Against Non-Whites

	(1)	(2)	(3)
Country	Average Log Discrimination Ratio, White Natives vs. Nonwhites	Number of Effects Dropped by Trim-and-Fill Adjustment	Trim-and-Fill Adjusted Average Log Discrimination Ratio, White Natives vs. Nonwhites
	average (se)	Number Dropped	average (se)
Belgium	0.35	2	0.45
	(0.10)		(0.10)
Canada	0.35	2	0.34
	(0.06)		(0.06)
France	0.62	0	0.62
	(0.07)		(0.07)
Germany	0.20	1	0.19
	(0.04)		(0.04)
Britain	0.43	6	0.39
	(0.03)		(0.04)
Netherlands	0.25	2	0.29
	(0.06)		(0.06)
Norway	0.29	0	0.29
	(0.06)		(0.06)
Sweden	0.50	0	0.50
	(0.07)		(0.07)
US	0.28	6	0.23
	(0.04)		(0.04)

Notes: Averages and standard errors are from random-effects meta-analysis by country.
Effects sizes from contrasts of native whites and nonwhites (excluding European immigrant targets)