PNAS

Supporting Information for
Trends in Racial and Ethnic Discrimination in Hiring in Six Western Countries

Lincoln Quillian* and John J. Lee

Lincoln Quillian
E-mail: l-quillian@northwestern.edu

This PDF file includes:

Supporting text: Appendices A to D
Figures S1 to S3
Tables S1 to S11
SI References

Supporting Information Text

Some parts of the procedures described in the appendices below reprise text from Quillian et al. (1) and Quillian et al. (2)

Appendix A. Study Search Methods and Coding

Study Search: Our search for studies included the methods of bibliographic search, citation search, and an e-mail request of authors of field experiments.

Our bibliographic search covered the following bibliographic databases and working paper repositories: Thomson's Web of Science (Social Science Citation Index), ProQuest Sociological Abstracts, ProQuest Dissertations and Theses, Lexis Nexis, Google Scholar, and NBER working papers. We searched for some combination of "field experiment" or "audit study" or "correspondence study" and sometimes included the term "discrimination," with some variation depending on the search functions of the database. To improve our coverage of non-English publications, we also searched two French-language indexes, Cairn.info and Persée; two international sources: IZA discussion papers, a German working paper archive; and ILO International Migration Papers. Finally, we conducted a search with Italian, Spanish, Portuguese and Dutch translations of the search terms and other terms frequently used in these languages to describe field experiments in hiring discrimination in Google Scholar. The search was first performed in March 2014 and repeated in August and September 2014 and in November 2015. We conducted searches in Italian, Spanish, Portuguese, and Dutch in November 2015 and February 2016. These searches were conducted by colleagues or research assistants fluent in the language of the search.

Our second technique for identifying relevant studies relied on citation search. Working from the initial set of studies located through bibliographic search, we examined the bibliographies of all review articles and eligible audit studies to find additional field experiments of hiring discrimination.

The last technique employed was an e-mail request to authors of existing field experiments of discrimination. From our list of audit studies identified by bibliographic and citation searches, we compiled a list of e-mail addresses of authors of existing field experiments of discrimination. We added the e-mail addresses of several well-known experts on field experiments, notably authors of literature review articles on field experiments. Our e-mail request asked for citations or copies of experimental field studies of discrimination that were published, unpublished, or ongoing. We also asked that the authors refer us to other researchers who may have recent or ongoing field experiments.

We conducted the e-mail requests in two phases. In the initial wave 131 apparently valid e-mail addresses were contacted. We received 56 responses. We also sent out a second wave of 68 e-mails which consisted of additional authors identified from the initial wave of surveys and some corrected e-mail addresses. We received 19 responses to this second wave of e-mail surveys.

Overall, our search located more than 100 studies that included contrasts between white and non-white groups who were on-average equivalent in their labor-market relevant characteristics (e.g. education, experience level in the labor market, etc.) and who otherwise met our inclusion criterion.[1] Some of these studies included contrasts between more than one target group and whites (e.g. blacks and Hispanics), producing multiple estimates of discrimination against non-whites.

Finally, since this procedure was originally undertaken, we have added new studies to our sample based on a refreshed bibliographic and citation search in 2021 and reference searches of new studies located through bibliographic search.

[1] We excluded some studies where it was unclear if employers were making decisions producing discrepant outcomes because applications were made through an employment agency. We excluded a few other studies because they lacked basic information on counts of outcomes by target group and the authors could not be located or declined to provide these data when contacted.

Coding: We coded key characteristics of the studies into a database for our analysis. Coding was based on a coding rubric we developed, which listed the characteristics and included coding instructions. To ensure reliability, most studies in our analysis were coded independently by two raters. Studies were coded by readers fluent in the language of the study report, which included English, French, German, and Dutch. The coders were the authors of this study, plus colleagues (co-authors of previous publications using this data), and research assistants. We reconciled the results of the two codings, performing further investigations to find the correct answer on coding decisions in cases of disagreement. Further details of the coding procedures are discussed in SI ref. 1.

Appendix B. 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, applicants applied for advertised jobs in pairs, and the applicants first called employers by phone to inquire if a job was still available.

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. The 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 callback rates for the non-white racial-ethnic applicants and white applicants from the point of initial application. We know that respondents who were told "no job is available" did not receive a callback. For situations where both 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. However, when one pair member was told the job is available and the other was not, we do not know how often the pair member 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.

Appendix C. Variances Estimation for Effect Sizes

For studies that are unpaired or do not report paired outcomes, the variance of the logged discrimination ratio for the mth non-white racial-ethnic group in the ith study for callbacks is estimated by:

$$\sigma_{ij}^2 = Var\left(\ln\left(y_{ij}\right)\right) = \frac{1}{c_{ij}^w} - \frac{1}{n_{ij}^w} + \frac{1}{c_{ij}^m} - \frac{1}{n_{ij}^m}$$

This is Bornstein, Hedges, Higgins, and Rothstein's (3) formula 5.3. For studies that use a paired design – with one non-white and one white applicant applying for each job – and report paired outcomes, we use an alternative formula to account for the pairing from Zhou (2007). If p^a is the number of pairs in which both white and non-white testers receive a callback, p^b is the number of pairs in which the white tester received a callback but not the non-white, p^c is the number of pairs in which the non-white tester received a callback but not the white, and p^d is the number of pairs in which neither tester received a callback, then the variance of the logged discrimination ratio for the jth non-white group in the ith study with paired data is:

$$\sigma_{im}^2 = Var\left(\ln\left(y_{ij}\right)\right) = \frac{p_{ij}^b + p_{ij}^c}{\left(p_{ij}^a + p_{ij}^b\right)\left(p_{ij}^a + p_{ij}^c\right)}$$

Appendix D. Publication Bias

Publication bias results when studies that fail to find statistically significant effects are less likely to be published. If studies that find no statistically significant discrimination are less likely to be published, this will lead to overestimating discrimination in meta-analysis. See Borenstein et al. (3), chapter 30.

Our primary interest is change over time. The key question then becomes if the extent of publication bias has changed over time, which would confound the time trend. A constant level of publication bias would bias estimates of the level of discrimination upwards evenly across time but not change trend estimates.

We tried two methods to examine publication bias. First, we examined publication bias based on funnel plots and the trim-and-fill method (4) with the pooled data broken into three periods: before 2000, 2001-2010, and 2011 and later. Funnel plots are shown in Supporting Information Figure S3. The trim-and-fill analysis found no evidence of publication bias in any period – funnel plots are fairly symmetric and no points were found to be "missing" in any of the three periods based on the trim-and-fill analysis (estimated with “metafor” with default options).

Second, we used the PEESE procedure. PEESE readily accommodates moderators such as our year variables and country and group controls (5). This method uses a weighted regression with the variance of the effect as a predictor to capture publication bias.

Estimates of PEESE models are shown in Supporting Information Table S11. Models 3-4 allow for interaction of study year and variance, as suggested by Stanley and Doucouliagos (6). We find evidence of publication bias in some (but not all) models. Importantly, we never find that the slope of year is significantly changed by adjusting for publication bias, which is of primary significance for our analysis. Not do we find a significant interaction of variability and the slope of year in models including this interaction.

We note, however, that some simulation studies have found that the PEESE method works well when there is little residual variability but does not work well under “realistic” conditions of heterogeneity in population effects (7, 8).

In sum, neither the trim-and-fill by decade nor the PEESE finds evidence that publication bias is likely to affect the trend over time.

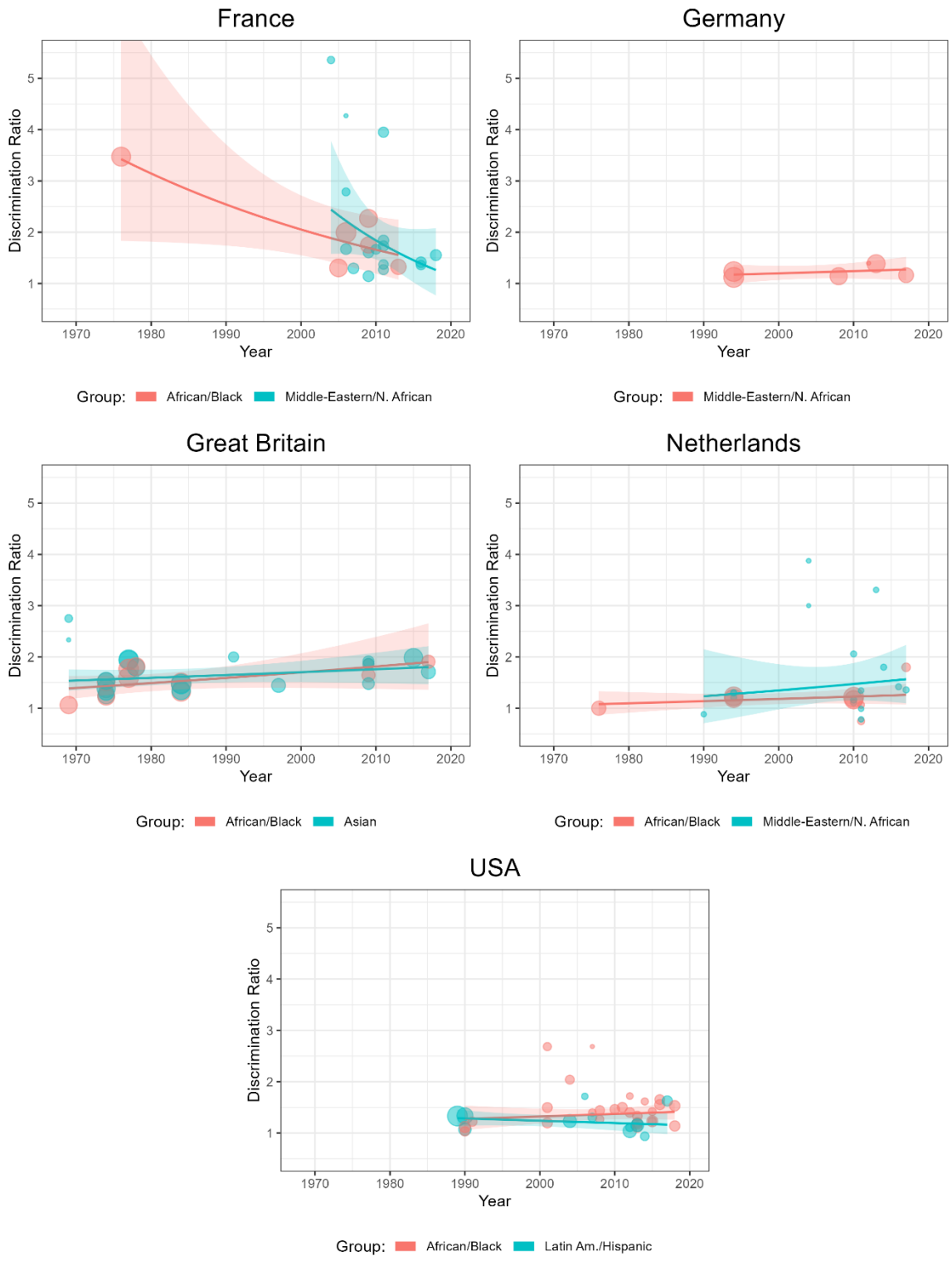


Fig. S1. Discrimination Ratios over Time by Country and Group

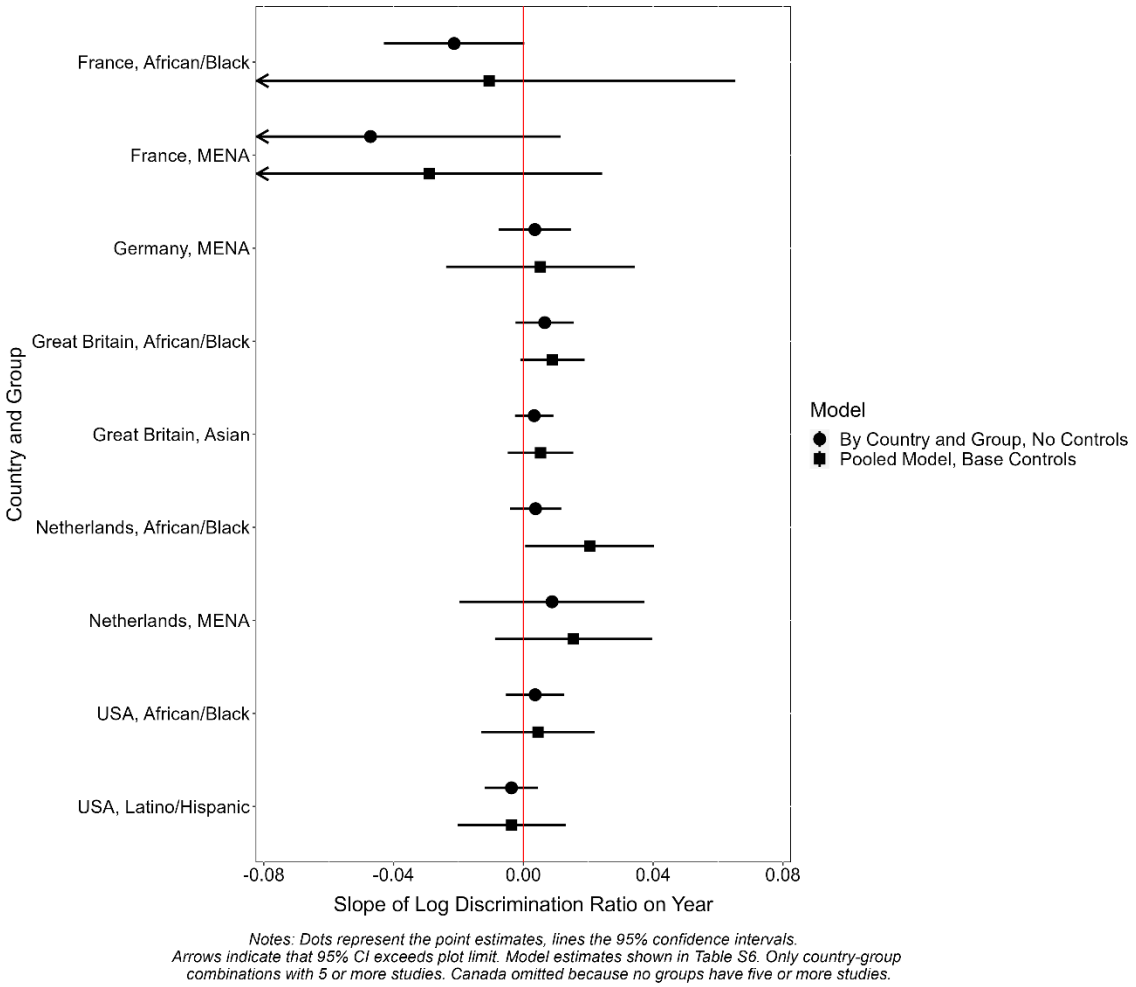
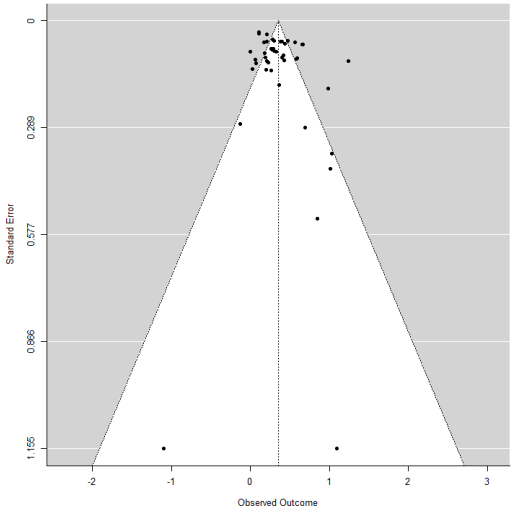
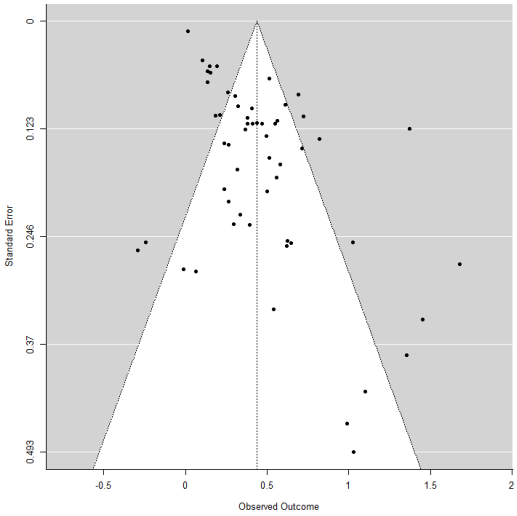


Fig. S2. Trends in Hiring Discrimination by Country and Group

Funnel plot, year 2000 and before:



Funnel Plot 2001 to 2010:



Funnel Plot 2011 and later:

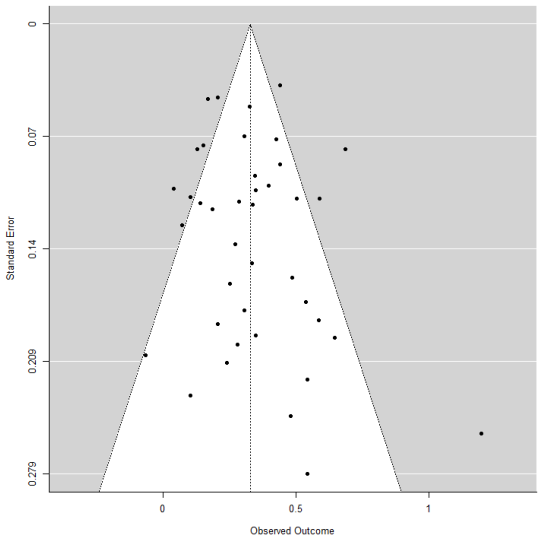


Fig. S3. Funnel Plots by Year Categories

Table S1: Pooled Meta-Regression Estimates of Log Discrimination Ratio on Country, Target Group, and Controls

Variable	Linear Trend Only (1)	Linear Trend Only, 1985-2019 (2)	Country + Group (3)	Base Controls (4)	Add Foreign Ed, Nationality (5)	Add Local UE Rate + % Foreign-Born (6)	Excluding single-occupation studies, base controls (7)	Base Controls, Resume Audits Only (8)
Year of Fieldwork	0.0001	0.0032	0.0024	0.0037	0.0051	0.0087	0.0008	0.0061
(Four Digit Year)	(0.0018)	(0.0021)	(0.0019)	(0.0043)	(0.0048)	(0.0062)	(0.0059)	(0.0047)
Dummy variable for countries (6 countries)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Dummy variables for racial-ethnic groups (4 groups)	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Basic Study Controls (see Table 2) (12 variables)	No	No	No	Yes	Yes	Yes	Yes	Yes
Controls for Foreign Educ. and Nationality (4 variables)	No	No	No	No	Yes	Yes	No	No
Controls for Unemployment and % Foreign Born in Region	No	No	No	No	No	Yes	No	No
N effects / N studies	140/90	119 / 81	140/90	140/90	140/90	113/72	119/72	118 / 75
Tau-squared, USA	0.014	0.011	0.017	0.009	0.009	0.011	0.010	0.009
Tau-squared, Canada	0.031	0.036	0.042	0.140	0.153	0.199	0.126	0.066
Tau-squared, France	0.224	0.185	0.159	0.157	0.129	0.157	0.085	0.172
Tau-squared, Germany	0.019	0.012	0.004	0.020	0.024	0.002	0.017	0.030
Tau-squared, Great Britain	0.036	0.058	0.013	0.000	0.000	0.002	0.001	0.000
Tau-squared, Netherlands	0.072	0.059	0.080	0.145	0.143	0.058	0.178	0.165

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. All models use the "subgroup correlated effects" model of Pustejovsky and Tipton (2022). Robust standard errors clustered at the study level.
Base controls are variables shown in table 2: study method dummy, tester gender (2 dummies), applicant education (3 dummies), applicant occupation (3 dummies), immigrant status, job source online (2 dummies).

Table S2: Pooled Meta-Regression Estimates of Log Discrimination Ratio on Country, Target Group, and Controls, Post-1984 Studies Only

Variable	Linear Trend Only (1)	Country + Group (2)	Base Controls (3)	Add Foreign Ed, Nationality (4)	Add Local UE Rate + % Foreign-Born (5)	Excluding single-occupation studies, base controls (6)	Base Controls, Resume Audits Only (7)
Year of Fieldwork	0.0032	0.0019	-0.0008	-0.0009	0.0099	0.0027	-0.0032
(Four Digit Year)	(0.0021)	(0.0025)	(0.0083)	(0.0086)	(0.0095)	(0.0073)	(0.0086)
Dummy variable for countries (6 countries)	No	Yes	Yes	Yes	Yes	Yes	Yes
Dummy variables for racial-ethnic groups (4 groups)	No	Yes	Yes	Yes	Yes	Yes	Yes
Basic Study Controls (see Table 2) (12 variables)	No	No	Yes	Yes	Yes	Yes	Yes
Controls for Foreign Educ. and Nationality (4 variables)	No	No	No	Yes	Yes	No	No
Controls for Unemployment and % Foreign Born in Region	No	No	No	No	Yes	No	No
N effects / N studies	119 / 81	119 / 81	119 / 81	119 / 81	94 / 65	100 / 64	100 / 67
Tau-squared, USA	0.011	0.017	0.009	0.019	0.014	0.016	0.006
Tau-squared, Canada	0.036	0.042	0.141	0.129	0.191	0.068	0.086
Tau-squared, France	0.185	0.131	0.139	0.132	0.160	0.000	0.140
Tau-squared, Germany	0.012	0.005	0.015	0.003	0.000	0.008	0.027
Tau-squared, Great Britain	0.058	0.000	0.000	0.000	0.000	0.000	0.000
Tau-squared, Netherlands	0.059	0.084	0.162	0.107	0.081	0.204	0.181

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. All models use the "subgroup correlated effects" model of Pustejovsky and Tipton (2022). Robust standard errors clustered at the study level.
Basic study controls are variables shown in table 2: study method dummy, tester gender (2 dummies), applicant education (3 dummies), applicant occupation (3 dummies), immigrant status, job source online (2 dummies).

Model 1 is the same as model 2 in table S1.

Table S3: Meta-Regression Estimates of the Trend in the Log Discrimination Ratio by Racial-Ethnic Group

	Models by Group			Pooled Model with Group-Specific Slopes	
Linear Trends by Group	Linear Trend Only (1)	Add Country Controls (2)	Add Base Controls (3)	Base Controls (4)	N effects / N studies
African/Black, Year of Fieldwork	-0.0012	0.0023	-0.0034	0.0082 +	57 / 51
(Four Digit Year)	(0.0029)	(0.0044)	(0.0048)	(0.0047)	
MENA, Year of Fieldwork	0.0052	-0.0037	0.0136	0.0095	41 / 39
(Four Digit Year)	(0.0073)	(0.0143)	(0.0136)	(0.0054)	
Latin/Hispanic, Year of Fieldwork	-0.0012	-0.0040	Insufficient	0.0000	12 / 12
(Four Digit Year)	(0.0043)	(0.0041)	Data	(0.0061)	
Asian, Year of Fieldwork	-0.0015	0.0008	Insufficient	0.0032	30 / 21
(Four Digit Year)	(0.0024)	(0.0027)	Data	(0.0043)	
Trend Post-1984					
African/Black, Year post -1984	0.0020	0.0102 *	0.0067	0.0050	47 / 42
	(0.0035)	(0.0041)	(0.0058)	(0.0090)	
Asian, Year post -1984	-0.0007	-0.0027	Insufficient	-0.0029	19 / 14
	(0.0026)	(0.0029)	Data	(0.0089)	
Controls					
Dummy variables for Country	No	Yes	Yes	Yes	
Dummy for In-Person vs. Resume	No	Yes	Yes	Yes	
Basic Study Controls	No	No	Yes	Yes	

Notes: Columns 1-3 based on a separate random-effects meta-regression estimated for each racial-ethnic group. Column 4 based on a single model with country-specific year trends, base model controls, and country-specific Tau parameters. Tau parameters not shown. The base controls are twelve dummy variables shown in Table 2.

+=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Robust standard errors clustered at the study level.

With the Bonjamini-Hochberg adjustment for multiple comparisons with four tests (see Materials & Methods) no tests are significant.

Table S4: Meta-Regression Estimates of Decade Changes in Log Discrimination Ratio, MENA and Muslim-Origin Minority Groups

	Subsample: MENA			Subsample: Muslim-Majority Origin Country		
Variable	Period Dummies Only (1)	+ Country (2)	+ Base Controls (3)	Period Dummies Only (4)	Country + Group (5)	+ Base Controls (6)
Period Dummies (Reference = 2000 or earlier)						
Year 2001 to 2011	0.3351 *	0.0510	0.9672 **	0.1927	0.0399	0.7783 **
	(0.1072)	(0.1113)	(0.2345)	(0.1125)	(0.0919)	(0.1395)
Year After 2011	0.2417 *	0.1146	0.8936 **	0.1333	0.1306 +	1.0677 ***
	(0.0773)	(0.0761)	(0.2361)	(0.0767)	(0.0579)	(0.1690)
Dummy variable for country (6 countries)	No	Yes	Yes	No	Yes	Yes
Dummy vars. for racial-ethnic groups (4 groups)	No	No	No	No	Yes	Yes
Basic Study Controls	No	No	Yes	No	No	Yes
N effects / N studies	41 / 39	41 / 39	41 / 39	57 / 51	57 / 51	57 / 51
Tau-squared, USA	0.053	0.022	0.011	0.000	0.000	0.000
Tau-squared, Canada	0.126	0.107	0.628	0.076	0.064	0.328
Tau-squared, France	0.145	0.161	0.000	0.134	0.155	0.014
Tau-squared, Germany	0.040	0.005	0.033	0.042	0.003	0.010
Tau-squared, Great Britain	0.000	0.022	0.011	0.045	0.020	0.000
Tau-squared, Netherlands	0.085	0.091	0.083	0.082	0.091	0.138

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Robust standard errors clustered at the study level.
Model 1 to 3 include only Middle-East North Africa (MENA) ethnic groups. Model 4 to 6 include all ethnic groups with origins in Muslim-majority countries.
Basic study controls are shown in Table 2.

Table S5: Meta-Regression Estimates of Linear Trend in Log Discrimination Ratio by Country

Variable	Linear Trend Only, Models by Country (1)		Group Controls, Models by Country (2)		Pooled Model, Country-Specific Slopes, Base Controls (3)		N effects / studies
Linear Trend by Country (Slope of Year of Fieldwork Variable)							
Canada	-0.0030		0.0007		-0.0095		14 / 7
	(0.0050)		(0.0046)		(0.0170)		
France	-0.0228	***, †	-0.0260	**, †	-0.0195	+, ns	23 / 20
	(0.0044)		(0.0069)		(0.0055)		
Germany	0.0023		0.0035		0.0072		8 /6
	(0.0032)		(0.0052)		(0.0119)		
Great Britain	0.0046	*, †	0.0044	+, ns	0.0042		30 / 12
	(0.0018)		(0.0021)		(0.0042)		
Netherlands	0.0081	+, ns	0.0063		0.0188	*, ns	25 / 15
	(0.0045)		(0.0059)		(0.0067)		
USA	0.0028		0.0016		0.0017		40 / 30
	(0.0031)		(0.0039)		(0.0075)		
Trend Post-1984							
Great Britain post-1984	0.0082	+	0.0090		-0.0065		11 / 5
	(0.0027)		(0.0054)		(0.0122)		
France post-1984	-0.0344		-0.0401		-0.0289		22 / 19
	(0.0217)		(0.0237)		(0.0255)		
Netherlands post-1984	0.0074		0.0064		0.0205		24 / 14
	(0.0060)		(0.0081)		(0.0122)		
Model Controls							
Dummy variables for race-ethnic group	No		Yes		Yes		
Base Controls	No		No		Yes		

Notes: Models 1 and 2 based on separate random-effects meta-regression for each country. Model 1 only includes a linear year predictor, model 2 adds controls for group. Model 3 is a single pooled model with country-specific slopes, base controls, and country-specific tau parameters.

+=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests, no multiple-comparison adjustment. Standard error in parentheses. Robust standard errors clustered at the study level.

† = Benjamini-Hochberg multiple comparison adjusted test significant with 6 tests, ns = not significant after adjustment.

Table S6: Meta-Regression Estimates of Linear Trend in Log Discrimination Ratio by Country and Group

Country and Group	(1) Trend from Meta-Regressions by Country and Group		(2) Country-Group Specific Trend, Pooled Model, Base Controls		N effects / N studies
USA, African/Black	0.0036		0.0045		27 / 27
	(0.0044)		(0.0084)		
USA, Latin Am./ Hispanic	-0.0037		-0.0036		11 / 11
	(0.0036)		(0.0079)		
France, African/Black	-0.0213	+	-0.0105		6 / 6
	(0.0078)		(0.0116)		
France, MENA	-0.0471		-0.0290		16 / 16
	(0.0273)		(0.0232)		
Germany, MENA	0.0035		0.0053		6 / 6
	(0.0040)		(0.0125)		
Great Britain, African/Black	0.0066		0.0090	+	11 / 11
	(0.0040)		(0.0045)		
Great Britain, Asian	0.0033		0.0053		18 / 18
	(0.0028)		(0.0048)		
Netherlands, African/Black	0.0038		0.0204	*	8 / 8
	(0.0032)		(0.0076)		
Netherlands, MENA	0.0088		0.0155		15 / 15
	(0.0132)		(0.0109)		
<u>Trend Dropping Early Study for France</u>					
France, African/Black, post-1984	-0.0075		0.0017		5 / 5
	(0.0485)		(0.0596)		
Great Britain, African/Black post-1984	0.0129		-0.0014		3 / 3
	(0.0117)		(0.0145)		
Great Britain, Asian post-1984	0.0088		-0.0092		7 / 7
	(0.0061)		(0.0137)		
Netherlands, African/Black post-1984	0.0001		0.0263	*	7 / 7
	(0.0044)		(0.0113)		

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Robust standard errors clustered at the study level. In model 1, each year coefficient is estimated in a separate meta-regression with year of fieldwork as the only predictor. Model 2 is subgroup correlated effect model including base controls. With the Bonjamini-Hochberg adjustment for multiple comparisons with 10 tests, (see Materials & Methods) no tests are significant.

Table S7: Descriptive Statistics, Predictor Variables

Base Model Variables

Annual Trend	Mean	Std. Dev.	N (effects / studies)
Year of Fieldwork	2002.6	14.0	140 / 90

Study Method	Effects	Studies
Resume Audit / Correspondence	118	75
In-person Audit	22	15

Tester Gender	Effects	Studies
Testers Male Only	43	37
Testers Female Only	9	9
Testers Both Male and Female	88	46

Applicant Education (most common level)	Effects	Studies
High School or Less	55	39
Some College or Post-HS Vocational Degree	34	23
College or More	30	20
Education information missing	21	8

Occupation Controls (all that apply)	Effects	Studies
Includes Blue Collar Jobs (1=yes)	69	41
Includes Jobs with Customer Contact (1=yes)	100	64
Includes Jobs with an Office Focus (1=yes)	104	64

Job Source	Effects	Studies
Online	72	46
Offline	50	32
Both Online and Offline	18	12

Minority Group Includes Foreign-Born Persons?	Effects	Studies
Native-Born Only	97	70
Includes Foreign-Born	43	22

Additional Controls (used in some models in Table S1 and Figure 2)

Nationality	Effects	Studies
Non-white Applicants Citizens of Host Nation	134	86
Non-white Applicants Not Citizens of Host Nation	4	3
Non-white Applicants Mix of Citizens/Noncitizens	2	2

Non-White Applicant Highest Ed. Credential Foreign?	Effects	Studies
Domestic Highest Education Credential	123	84
Foreign Highest Education Credential	1	1
Highest Credential Mix of Foreign/Domestic	16	7

Contextual Controls	Mean	Std. Dev.	N (effects / studies)
Unemployment Rate of Local City/Region	7.0%	2.4%	115 / 74
Percentage Immigrants in Local City/Region	13.4%	10.9%	115 / 73

Notes: Effects are distinct estimates of discrimination against minority groups. Some studies include discrimination estimates against multiple minority groups. Studies do not always sum to 90 because an effect can have different values for the same study.

Table S8: Specific Minority Groups in Field Experiments by Country

Country	Non-White Groups with Effect Sizes (Study Term)
Canada	African, Arab, Black, Chinese, Indian, Indo-Pakistani, Latino,Middle Eastern, West Indian
France	African, Antillean, Asian, Franco-North African, Moroccan, North African, Senegalese, Subsaharan African, Vietnamese
Germany	MENA, Turkish, Southeast Asian, Sub-Saharan African
Great Britain	African, Asian (South Asian), Black African, Black Caribbean, Chinese, Indian, Pakistani, Pakistani/Bangladeshi, West Indian
Netherlands	Antillean, Arab, Black Surinamer, Hindustani, Moroccan, Spanish, Surinamese, Turkish
US	African American, Arab American, Asian, Black, Hispanic, Latino, Somali

Table S9: Pooled Meta-Regression of Log Discrimination Ratio on Year, Non-Linearity Analysis

Variable	Period Dummies Only 1	Country + Group 2	Base Controls 3	Squared Year Predictor 4
Period Dummy Variables (Reference=Before 1991)				
Year 1991 to 2000 (1=yes)	0.0218	-0.0089	-0.0436	
	(0.0656)	(0.0810)	(0.0797)	
Year 2001 to 2010 (1=yes)	0.1765 *	0.1316	0.2233	
	(0.0677)	(0.0823)	(0.1273)	
Year After 2010 (1=yes)	0.1014 +	0.0731	0.2232	
	(0.0530)	(0.0731)	(0.1402)	
Year as Continuous Variable with Squared Term				
Year (Year 2000=0)				0.0013
				(0.0057)
Year Squared (Year 2000=0)				-0.0002
				(0.0004)
Dummy variable for country (6 countries)	No	Yes	Yes	Yes
Dummy variables for minority group (4 minority groups)	No	Yes	Yes	Yes
Basic Study Controls (12 variables)	No	No	Yes	Yes
N effects / N studies	140/90	140/90	140/90	140/90
Tau-squared, USA	0.0077	0.0158	0.0097	0.0079
Tau-squared, Canada	0.0700	0.0596	0.1387	0.1380
Tau-squared, France	0.1624	0.1588	0.1541	0.1674
Tau-squared, Germany	0.0100	0.0096	0.0160	0.0228
Tau-squared, Great Britain	0.0122	0.0131	0.0000	0.0000
Tau-squared, Netherlands	0.0611	0.0727	0.1219	0.1488

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Standard errors clustered at the study level.
Basic study controls are shown in Table 2.

Table S10: Random-Effects Meta-Regression Models with Odds-Ratio Outcome, Pooled Models with Base Controls

Variables in Model	Single Trend	Country-Specific Trends		Group-Specific Trends		MENA Trends	
Linear Trend Coef(s)	-0.0089	Canada	-0.0281	African/Black	-0.0034	Year 2000 or Earlier	(ref.)
(Slope of Year)	(0.0077)		(0.0170)		(0.0082)		
		France	-0.0268 +	MENA	-0.0012	Year 2001 to 2011	1.1066 **
			(0.0074)		(0.0101)		(0.3047)
		Germany	0.0116	Latin/Hispanic	-0.0150	Year 2012 or Later	1.0605 **
			(0.0151)		(0.0093)		(0.2976)
		Great Britain	-0.0118	Asian	-0.0108		
			(0.0081)		(0.0083)		
		Netherlands	0.0287 *				
			(0.0096)				
		USA	0.0003				
			(0.0120)				
Control Variables							
Dummy variables for country (6 countries)	Yes		Yes		Yes		Yes
Dummy variables for minority group (4 minority groups)	Yes		Yes		Yes		Yes
Basic Study Controls (12 variables, see Table 2)	Yes		Yes		Yes		Yes
Variance Structure							
Country-Specific Tau's	Yes		Yes		Yes		Yes
N effects / N studies	140 / 90		140 / 90		140 / 90		41 / 39

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Standard errors clustered at the study level.
The models are pooled subgroup correlated effects models.
Basic study controls are shown in Table 2.

Table S11: Publication Bias (PEESE) Estimates of Pooled Models

Variable	Linear Trend Only (1)	+ Base Controls (2)	Linear Trend Only (3)	+ Base Controls (4)
Year of Fieldwork (Four Digit Year)	-0.0047	0.0070	-0.0066	0.0055
	(0.0041)	(0.0040)	(0.0047)	(0.0043)
Estimated Variance of Effect	4.3055	2.2296 **	5.7208	2.9894 *
	(1.8687)	(0.6456)	(2.4717)	(0.9439)
Year * Variance			0.2704	0.1107
			(0.1480)	(0.0618)
Dummy variable for country (6 countries)	No	Yes	No	Yes
Dummy variables for minority group (4 minority groups)	No	Yes	No	Yes
Basic Study Controls (9 variables)	No	Yes	No	Yes
N effects / N studies	140 / 90	140 / 90	140 / 90	140 / 90

Notes: +=p<.1; * = p<.05; ** = p<.01; *** = p<.001. Two-tailed tests. Standard error in parentheses. Robust standard errors clustered at the study level.

Basic study controls are dummies for year and minority group plus twelve variables shown in Table 2.

SI References

1. L. Quillian, *et al.*, Do Some Countries Discriminate More than Others? Evidence from 97 Field Experiments of Racial Discrimination in Hiring. *Sociol. Sci.* **6**, 467–496 (2019).
2. L. Quillian, D. Pager, O. Hexel, A. H. Midtbøen, Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proc. Natl. Acad. Sci. USA* **114**, 10870–10875 (2017).
3. M. Borenstein, L. V. Hedges, J. P. T. Higgins, H. R. Rothstein, *Introduction to Meta-Analysis* (John Wiley & Sons, 2009).
4. G. Y. Zou, One relative risk versus two odds ratios: implications for meta-analyses involving paired and unpaired binary data. *Clin. Trials* **4**, 25–31 (2007).
5. S. Duval, R. Tweedie, A Nonparametric “Trim and Fill” Method of Accounting for Publication Bias in Meta-Analysis. *J. Am. Stat. Assoc.* **95**, 89–98 (2000).
6. T. D. Stanley, H. Doucouliagos, *Meta-Regression Analysis in Economics and Business*, 1st Ed. (Routledge, 2012).
7. N. Alinaghi, W. R. Reed, Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work? *Res. Synth. Methods* **9**, 285–311 (2018).
8. S. Hong, Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work? A replication study of Alinaghi & Reed (Research Synthesis Methods, 2018). *International Journal for Re-Views in Empirical Economics (IREE)* **3**, 1–22 (2019).