

WZB

Wissenschaftszentrum Berlin
für Sozialforschung



Ruud Koopmans
Susanne Veit
Ruta Yemane

**Ethnische Hierarchien in der
Bewerbersauswahl:
Ein Feldexperiment zu den Ursachen von
Arbeitsmarktdiskriminierung**

Discussion Paper

SP VI 2018-104

Mai 2018

Forschungsschwerpunkt

Migration und Diversität

Forschungsabteilung

Migration, Integration, Transnationalisierung

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
www.wzb.eu

Das Urheberrecht liegt bei den Autoren.

Discussion Papers des WZB dienen der Verbreitung von Forschungsergebnissen aus laufenden Arbeiten im Vorfeld einer späteren Publikation. Sie sollen den Ideenaustausch und die akademische Debatte befördern. Die Zugänglichmachung von Forschungsergebnissen in einem WZB Discussion Paper ist nicht gleichzusetzen mit deren endgültiger Veröffentlichung und steht der Publikation an anderem Ort und in anderer Form ausdrücklich nicht entgegen. Discussion Papers, die vom WZB herausgegeben werden, geben die Ansichten des jeweiligen Autors wieder und nicht die der gesamten Institution WZB.

Ruud Koopmans, Susanne Veit, Ruta Yemane

Ethnische Hierarchien in der Bewerberauswahl: Ein Feldexperiment zu den Ursachen von Arbeitsmarktdiskriminierung

Discussion Paper SP VI 2018-104

Wissenschaftszentrum Berlin für Sozialforschung (2018)

Affiliation der Autoren

Ruud Koopmans

WZB Wissenschaftszentrum Berlin für Sozialforschung
E-mail: ruud.koopmans@wzb.eu

Susanne Veit

WZB Wissenschaftszentrum Berlin für Sozialforschung
E-mail: susanne.veil@wzb.eu

Ruta Yemane

WZB Wissenschaftszentrum Berlin für Sozialforschung
E-mail: ruta.yemane@wzb.eu

Ethnische Hierarchien in der Bewerberauswahl: Ein Feldexperiment zu den Ursachen von Arbeitsmarktdis- kriminierung

Ruud Koopmans, Susanne Veit, Ruta Yemane

Zusammenfassung

In einem großen Feldexperiment haben wir die Ursachen von Diskriminierung gegenüber Bewerbern mit Migrationshintergrund untersucht. Dazu versendeten wir tausende Bewerbungen von fiktiven Personen an reale Stellenausschreibungen in acht Berufen im gesamten Bundesgebiet. Neben der Ethnizität der Bewerber (deutschstämmig oder Migrationshintergrund in einem von 34 Herkunftsländern), ihrem phänotypischen Erscheinungsbild (Schwarz, Weiß oder Asiatisch) und ihrer Religionszugehörigkeit (keine, Christlich, Muslimisch oder Buddhistisch/Hinduistisch) variierten wir weitere Merkmale der Bewerbungen, wie das Geschlecht des Bewerbers, den Notendurchschnitt, ob der Bewerbung ein Referenzschreiben beilag und ob Informationen über die derzeitige Vertragssituation bereitgestellt wurden. Unsere Ergebnisse bestätigen, dass Bewerber mit Migrationshintergrund gegenüber Bewerbern ohne Migrationshintergrund diskriminiert werden. Allerdings variiert das Ausmaß der Diskriminierung deutlich zwischen Herkunftsgruppen: Bewerber mit Migrationshintergrund in West- und Südeuropa sowie Ostasien werden nicht signifikant diskriminiert, während andere Herkunftsgruppen erhebliche Nachteile erfahren. Auch Bewerber mit schwarzem Phänotyp und mit muslimischer Religion erfahren signifikante Diskriminierung. Mit Blick auf die klassischen Erklärungsansätze für Diskriminierung auf dem Arbeitsmarkt - d.h. präferenzbasierte und statistische Diskriminierung - zeigen unsere Analysen, dass die kulturelle Distanz zwischen Herkunftsländern und Deutschland die Diskriminierung gegenüber verschiedenen Gruppen deutlich besser erklärt als leistungsbezogene Gruppenmerkmale wie der durchschnittliche Bildungsstand. Somit sprechen unsere Befunde stärker für präferenzbasierte Diskriminierung als für statistische Diskriminierung.

Abstract

We conducted a large-scale field experiment to investigate the drivers of discrimination against second generation immigrant job applicants. To these ends, we sent thousands of applications from fictitious persons to real job openings in eight professions all over Germany. Next to job applicants' ethnicity (German or migration background in one out of 34 origin countries), phenotype (Asian, Black, White), and religious affiliation (none, Buddhist or Hindu, Christian, or Muslim), we varied several other characteristics of the applications, such as applicants' gender, final grades, whether or not a reference letter was included, as well as information about applicants' current contract. Our results confirm that employers discriminate against immigrant job applicants. The magnitude of discrimination, however, varies strongly between origin groups. Whereas employers do not discriminate against Western and Southern European and East Asian immigrants, other origin groups experience significant disadvantages. In addition, we observe substantial disadvantages for Black and Muslim job applicants. With respect to classic theories about the drivers of discrimination on the labor market, that is, taste-based and statistical discrimination, we find that the cultural distance between origin countries and Germany explains discrimination against different groups much better than productivity-related group characteristics, such as average levels of education. Consequently, our empirical findings are more supportive of taste-based discrimination than they are of statistical discrimination theories.

Inhalt

Tabellenverzeichnis	1
Abbildungsverzeichnis	2
Hintergrund der Studie	3
Präferenzbasierte oder statistische Diskriminierung?.....	5
Forschungsfragen und Hypothesen.....	7
Methoden	8
Studiendesign.....	8
Positive Rückmeldung.....	10
Experimentell variierte Merkmale.....	11
Geschlecht.....	11
Ethnizität.....	11
Religion.....	12
Phänotyp.....	13
Noten, Referenzschreiben und Vertragslaufzeit.....	14
Gruppenmerkmale.....	15
Produktivität.....	17
Kulturelle Distanz.....	18
Ergebnisse	20
Deskriptive Ergebnisse.....	20
Welche Rolle spielt das Herkunftsland?.....	21
Welche Rolle spielt die Religion?.....	23
Welche Rolle spielt der Phänotyp?.....	25
Weitere Merkmale: Geschlecht, Referenzschreiben, Noten und Vertragsende... 26	
Multivariate Analysen.....	29
Prüfungen zur Robustheit der Ergebnisse.....	34

Zusammenfassung und Diskussion.....	35
Literaturverzeichnis.....	37
Anhang	40
Tabelle A1. Herkunftsländer und Namen.....	40
Tabelle A2. Zuweisungshäufigkeit von Phänotyp und Religion nach Ethnizität....	41
Tabelle A3. Mittleres Bildungsniveau, Erwerbslosigkeit, ISEI und Wertedistanz zu Deutschland nach Ethnizität.....	42
Tabelle A4. Bewerbungen nach Berufsgruppe und Geschlecht	43

Tabellenverzeichnis

Tabelle 1: Berufe.....	10
Tabelle 2: Multivariate Regression.....	32

Abbildungsverzeichnis

Abbildung 1: Herkunftsländer der fiktiven Bewerber	12
Abbildung 2: Rückmeldung nach Berufen.....	21
Abbildung 3: Rückmeldung je Herkunftsgruppe	22
Abbildung 4: Religion und Phänotyp.....	24
Abbildung 5: Produktivitätsmerkmale und Geschlecht.....	27

Hintergrund der Studie

Laut dem Datenreport des Statistischen Bundesamtes war der Erwerbslosenanteil in der Bevölkerung mit Migrationshintergrund im Jahr 2014 mit sieben Prozent fast doppelt so hoch wie in der Bevölkerung ohne Migrationshintergrund. Der relative hohe Anteil an Erwerbslosen mit Migrationshintergrund lässt sich zum Teil durch fehlende formale Qualifikationen oder deren Anerkennung sowie durch soziokulturelle Faktoren wie fehlende interethnische Kontakte erklären. Auch unter Migranten¹ der zweiten Generation, die in Deutschland aufgewachsen sind, ist der Anteil der Erwerbslosen mit sechs Prozent deutlich erhöht. Empirische Studien bestätigen, dass Migranten in Deutschland ein überdurchschnittlich hohes Risiko haben, erwerbslos zu werden, und dies selbst nach Kontrolle für Humankapitalmerkmale (Uhlendorff und Zimmermann 2006; Kogan 2004; Burkert und Seibert 2007). Unterschiede zwischen Menschen mit und ohne Migrationshintergrund, die auch nach Kontrolle für wichtige Humankapitalmerkmale wie Bildungsabschlüsse oder Sprachkenntnisse bestehen bleiben, werden häufig als „ethnic penalty“ bezeichnet - also als ein Nachteil aufgrund der ethnischen Zugehörigkeit. Auch wird häufig angenommen, dass solche Nachteile die Folge von Diskriminierung sein könnten. Dies zuverlässig zu belegen ist jedoch schwierig, weil Diskriminierung unter Verwendung herkömmlicher Forschungsformate nicht direkt gemessen werden kann.

Um Diskriminierung aufzuzeigen, bedarf es experimenteller Studien. Ein Beispiel dafür sind Korrespondenztests. Hierbei handelt es sich um Feldexperimente, in denen Bewerbungen fiktiver Personen an reale Arbeitgeber verschickt werden. Dabei sind die versendeten Bewerbungen in allen Aspekten miteinander vergleichbar, mit Ausnahme desjenigen Merkmals, für das überprüft werden soll, ob es Grundlage für Diskriminierung ist. So wird beispielsweise die ethnische Gruppenzugehörigkeit, das Geschlecht oder die sexuelle Orientierung der Bewerber variiert, während andere Merkmale wie Bildungsstand und Berufserfahrung konstant gehalten werden. Ergeben sich dann unterschiedliche Rücklaufquoten für die getesteten Gruppen, so ist dies ein eindeutiger Beleg für Diskriminierung.

¹ Im Folgenden wird aus Gründen der Lesbarkeit das generische Maskulinum verwendet. Dennoch sind ausdrücklich alle Geschlechter gemeint.

In den vergangenen Jahren wurde eine Vielzahl von Korrespondenzstudien auf dem Arbeitsmarkt durchgeführt, die nachweisen konnten, dass Bewerber mit Migrationshintergrund geringere Chancen auf eine positive Rückmeldung haben. Die Studien zeigten robuste Diskriminierungseffekte, und zwar über verschiedene Länder, Minderheiten und Branchen hinweg (Agerström et al. 2012; Bertrand und Mullanaithan 2004; Oreopoulos 2011, Edo et al. 2013; Blommaert et al. 2013; Gaddis 2015; Kaas und Manger 2012; Schneider, Yemane und Weinmann 2014; Weichselbaumer 2016). In Deutschland konzentrierte sich die Forschung bisher auf den Nachweis von Diskriminierung gegenüber männlichen Bewerbern mit türkischem Migrationshintergrund. So zeigten Schneider, Yemane und Weinmann (2014) sowie Kaas und Manger (2012), dass Bewerber mit türkischem Namen bei der Suche nach einem Ausbildungs- oder Praktikumsplatz in Deutschland diskriminiert werden. Auf Grundlage dieser Studien kann jedoch nicht geklärt werden, warum Arbeitgeber diskriminieren, da mehrere Merkmale untrennbar miteinander verbunden sind, wie Migrationshintergrund, türkische Herkunft und eine vermutete Religionszugehörigkeit zum Islam. Zudem stammt ein Großteil der Menschen mit Migrationshintergrund in Deutschland aus anderen Ländern, insbesondere aus dem europäischen Ausland. Ob diese Gruppen ebenfalls von Diskriminierung betroffen sind, und welche Rolle Religion oder das phänotypische Erscheinungsbild dabei spielen, wurde im Rahmen dieser Studien bisher noch nicht untersucht.

Fragen nach den zugrundeliegenden Mechanismen ethnischer Ungleichbehandlung auf dem Arbeitsmarkt sowie nach Gründen von Unterschieden in der Diskriminierung verschiedener Geschlechter, ethnischer oder religiöser Gruppen und dem Zusammenspiel dieser Faktoren blieben somit weitgehend unbeantwortet. Um effektive Maßnahmen zur Sicherung von Chancengleichheit und zur Bekämpfung von Diskriminierung entwickeln zu können, muss man die Gründe für Diskriminierung verstehen. Wichtige Hinweise kann der Vergleich zwischen verschiedenen Gruppen geben. Bevor wir uns dieser Fragen im Detail widmen, fassen wir im Folgenden jedoch zunächst die zentralen Theorien der Diskriminierungsforschung zusammen.

Präferenzbasierte oder statistische Diskriminierung?

In der Forschung zu Arbeitsmarktdiskriminierung existieren zwei theoretische Ansätze, die sich mit den Ursachen für diskriminierendes Verhalten von Arbeitgebern befassen: Theorien der präferenzbasierten und der statistischen Diskriminierung.

Der amerikanische Ökonom und Soziologe Gary Becker hat erstmals den Begriff der „taste-based discrimination“ eingeführt um zu beschreiben, dass sich Arbeitgeber bei ihren Einstellungsentscheidungen von persönlichen Präferenzen und Abneigungen gegenüber bestimmten Gruppen leiten lassen. Der Theorie der präferenzbasierten Diskriminierung zufolge diskriminieren Arbeitgeber Angehörige bestimmter Gruppen (z.B. Migranten, Afro-Amerikaner, etc.) aufgrund eigener Abneigung oder aufgrund von antizipierten Vorbehalten von Mitarbeitern oder Kunden. Um den Kontakt mit Angehörigen dieser Gruppe zu vermeiden, sind Arbeitgeber sogar bereit, ökonomische Nachteile (wie weniger Profit oder geringere Produktivität) in Kauf zu nehmen. Laut Beckers Theorie führt dies zu einem Marktnachteil für diskriminierende Firmen, der wiederum dazu führen sollte, dass diese langfristig vom Markt verschwinden (Becker 1957).

Demgegenüber sehen Vertreter der statistischen Diskriminierung (Aigner und Cain 1977; Arrow 1973, 1998; Phelps 1972) Zweifel an der Leitungsfähigkeit und Produktivität von Angehörigen bestimmter Gruppen als zentrale Ursache für Diskriminierung. Arbeitgeber haben immer nur unvollständige Informationen über die tatsächliche Produktivität von Bewerbern. Daher greifen Arbeitgeber auf ihr Wissen über die durchschnittliche Produktivität von Gruppen zurück. Wenn Arbeitgeber das Wissen über Gruppenmerkmale in ihre Auswahlentscheidung einfließen lassen, wirkt sich dies jedoch ungünstig auf die Arbeitsmarktchancen individueller Bewerber aus, die Gruppen angehören, die im Durchschnitt häufiger arbeitslos sind oder ein geringeres Bildungsniveau haben. Ein klassisches Beispiel für statistische Diskriminierung ist die Diskriminierung von jungen Frauen aufgrund des im Durchschnitt erhöhten Risikos von Arbeitsausfällen infolge von Schwangerschaft und Elternzeit.

Im Kern unterscheiden sich die beiden theoretischen Ansätze vor allem darin, aufgrund welcher Motivation diskriminiert wird: Unbehagen und Antipathie

gegenüber bestimmten Gruppen (präferenzbasiert) oder Zweifel an der Leistungsfähigkeit und Produktivität von Angehörigen bestimmter Gruppen (statistisch). Beide Ansätze erklären jedoch Diskriminierung, d.h. die Benachteiligung von Angehörigen bestimmter Gruppen im Vergleich zu gleich qualifizierten Angehörigen anderer Gruppen aufgrund der Gruppenzugehörigkeit. Diese Ungleichbehandlung ist auch dann noch diskriminierend, wenn die Gründe für die Ungleichbehandlung gut nachvollziehbar sind (unabhängig von ethischen und moralischen Aspekten) – wie im eben erwähnten Beispiel der Diskriminierung von jungen Frauen. Denn nicht jede junge Frau bekommt Kinder und nicht jede Mutter fällt langfristig aus.

Obwohl zahlreiche Studien Diskriminierung auf dem Arbeitsmarkt belegen, können die wenigsten Studien zwischen präferenzbasierter und statistischer Diskriminierung unterscheiden². Inwiefern es sich bei der in früheren Studien aufgedeckten Diskriminierung um präferenzbasierte oder um statistische Diskriminierung handelt ist jedoch eine zentrale Frage, deren Relevanz über den akademischen Kontext hinausreicht. Denn hieraus ergeben sich wichtige Implikationen mit Blick auf Maßnahmen zur Bekämpfung von Diskriminierung auf dem Arbeitsmarkt. Wenn Becker mit seiner Theorie der präferenzbasierten Diskriminierung richtig liegt, wären antirassistische Kampagnen, Kontaktinterventionen und affirmative Maßnahmen vielversprechende Ansätze, um der Benachteiligung von Minderheiten entgegenzuwirken. Wenn Diskriminierung hingegen aufgrund „statistischer“ Unterschiede in der durchschnittlichen Produktivität von Gruppen erfolgt, dann muss vor allem daran gearbeitet werden, diese Gruppen besser zu qualifizieren – durch bessere Angebote, Beratung und höhere Anreize, in Bildung zu investieren und sich am Erwerbsleben zu beteiligen.

² Bartoš et al. (2016) leisteten hier mit ihrer Studie zu „Attention Discrimination“ einen innovativen Beitrag. Auch die Studien von Agerström et al. 2012; Ahmed, Andersson, und Hammarstedt 2013; Auspurg, Hinz, und Schmid 2017; Drydakis 2014; Edo, Jacquemet, und Yannelis 2013; Gaddis 2015 und Kaas und Manger 2012 haben in ihren Studien untersucht, ob sich das Ausmaß der Diskriminierung durch mehr Information über den Bewerber reduziert. Studien, die explizit Hypothesen der präferenzbasierten Diskriminierung testen, sind jedoch selten.

Forschungsfragen und Hypothesen

Ob ethnische Diskriminierung gegenüber Migranten der zweiten Generation in Deutschland eher an Antipathie oder an Wissen über durchschnittliche Produktivitätsunterschiede liegt, ist Kern unserer Studie und soll anhand folgender Fragen beantwortet werden: Richtet sich Diskriminierung gegen Menschen mit Migrationshintergrund im Allgemeinen oder gegen Migranten aus bestimmten Herkunftsländern im Speziellen? Welche Migrantengruppen werden am stärksten benachteiligt, und warum? Welche Rolle spielen Religion, Geschlecht und *race*³, also phänotypische Unterschiede? Welche Gründe gibt es für Diskriminierung und für welche Theorie der Arbeitsmarktdiskriminierung, präferenzbasiert oder statistisch, finden wir Belege?

Um diese Fragen zu beantworten, führten wir eine Korrespondenzstudie durch, in der wir Diskriminierung anhand folgender Merkmale untersuchten: Ethnizität, also Zugehörigkeit zu einer bestimmten nationalen Herkunftsgruppe (z.B. „Türken“, „Italiener“ oder „Nigerianer“), *race*, also Zugehörigkeit zu einer Gruppe, die über ihr phänotypisches Äußeres definiert wird (z.B. schwarz, weiß, asiatisch) und Religionszugehörigkeit (z.B. muslimisch, christlich, buddhistisch/hinduistisch).

Ausgehend von der Theorie statistischer Diskriminierung explorieren wir den Einfluss produktivitäts-relevanter Informationen über einzelne Bewerber (d.h. Noten, Referenzen und derzeitige Arbeitssituation) auf Diskriminierung und überprüfen, inwieweit Diskriminierung mit produktivitäts-bezogenen Merkmalen von Herkunftsgruppen (d.h. durchschnittlicher Bildungsgrad, beruflicher Status und Erwerbslosigkeit) variiert. Mit Blick auf produktivitätsbezogene Gruppenmerkmale erwarten wir, dass die Diskriminierung gegenüber unterschiedlichen Herkunftsgruppen mit eben diesen Informationen systematisch variiert: je 'produktiver' eine Gruppe im Durchschnitt ist, umso positiver sollten Arbeitgeber auf Bewerbungen von Angehörigen dieser Gruppe reagieren (Hypothese 1).

³ Im Folgenden werden wir den englischen Begriff „*race*“ nicht übersetzen, da es keine geeignete deutsche Entsprechung gibt. Ersatzweise werden wir auf den Begriff „Phänotyp“ zurückgreifen.

Für die Untersuchung von präferenzbasierter Diskriminierung greifen wir auf das Konzept der „Homophilie“ (McPherson, Smith-Lovin und Cook 2001) zurück, welches besagt, dass Menschen eine Präferenz für andere haben, die ihnen sozial und kulturell ähnlich sind, da sie beispielsweise die gleiche Religion, gleiche Interessen und gleiche Wertvorstellungen haben. Diese Tendenz zeigt sich am stärksten im privaten Bereich, lässt sie jedoch auch auf den Arbeitskontext übertragen. Wir erwarten entsprechend, dass die Diskriminierung gegenüber Bewerbern mit Migrationshintergrund insbesondere dann stark ausfällt, wenn die jeweilige Herkunftsgruppe sich mit Blick auf Werte und Normen stark von Deutschland unterscheidet (Hypothese 2). Für Herkunftsgruppen, die den Deutschen kulturell sehr ähnlich sind, erwarteten wir hingegen keine Diskriminierung.

Methoden

Studiendesign

Um den Ursachen für ethnische Diskriminierung nachzugehen, haben wir eine Korrespondenzstudie durchgeführt. Wie bereits erwähnt, werden bei dieser Methode Bewerbungen fiktiver Personen, die sich nur in den für die Fragestellung relevanten Merkmalen (wie z.B. in deren ethnischer Zugehörigkeit) unterscheiden, an reale Unternehmen versendet. Unterschiede im Rückmeldeverhalten der Unternehmen sind Belege für Diskriminierung.

Unsere Studie weicht vom traditionellen Design solcher Studien insofern ab, als dass wir, statt wie üblich mehrere Bewerbungen an eine Stellenausschreibung zu senden, nur *eine* Bewerbung pro Ausschreibung versendeten. Dies war nötig, da wir eine Reihe von ethnischen Gruppen untersuchen und eine Vielzahl an experimentellen Bedingungen testen wollten. Dieses Vorgehen wurde bereits von Weichselbaumer (2015, 2016) und Ahmed, Andersson und Hammarstedt (2013) in Studien zur Diskriminierung aufgrund von sexueller Orientierung gewählt.

Die Stichprobe unserer Studie besteht aus insgesamt 5.819⁴ Bewerbungen, mit denen wir auf Stellenausschreibungen in acht Ausbildungsberufen antworteten, die deutschlandweit zwischen Oktober 2014 und April 2016 auf der Webseite der Bundesagentur für Arbeit ausgeschrieben waren. Wie in Deutschland üblich, beinhalteten die Bewerbungen neben dem Motivationsschreiben und dem Lebenslauf mit Foto auch Kopien der Ausbildungs- und Schulabschlusszeugnisse. Alle Bewerber waren laut ihrem Lebenslauf 1992 in Deutschland geboren worden, deutsche Staatsbürger, und hatten ihren gesamten Bildungsweg in Deutschland durchlaufen. In ihren Bewerbungsunterlagen gaben sie an, dass sie nach erfolgreichem Abschluss ihrer Ausbildung noch einige Jahre Berufserfahrung in ihrem Ausbildungsbetrieb gesammelt hatten und nun nach einer neuen Stelle suchten.

Um die ausschlaggebenden Gründe für Diskriminierung genauer zu betrachten, variierten wir verschiedene Merkmale der Bewerber. Die meisten Merkmale (z.B. Noten, Referenzschreiben, Vertragsende) wurden randomisiert einer Bewerbung zugewiesen. Für andere Merkmale (Phänotyp, Religion) war die Zufallszuweisung eingeschränkt, um nur plausible Merkmalskombinationen zuzulassen. So gab es beispielsweise zwar in jeder Herkunftsgruppe Bewerber mit und ohne Hinweis auf eine Religionszugehörigkeit, aber die jeweilige Religion variierte je nach Herkunftsland. So war das Religionssignal bei deutschstämmigen Bewerbern immer „christlich“, bei irakischstämmigen Bewerbern immer „muslimisch“ und bei Bewerbern mit Wurzeln in Bosnien und Herzegowina entweder „christlich“ oder „muslimisch“. Der Projektbericht (Veit & Yemane, 2018) gibt detaillierte Hinweise zum Studiendesign und eine Übersicht über die konkreten Zuweisungswahrscheinlichkeiten.

Insgesamt haben wir uns auf Ausschreibungen in acht Ausbildungsberufen beworben, darunter vier Berufe, die sowohl von Männern als auch Frauen häufig ausgeübt werden (Hotelfachmann/frau, Koch/Köchin, Industriekaufmann/frau und Verkäufer/in), sowie zwei typische Männerberufe (Anlagenmechaniker-SHK und Mechatroniker) und zwei typische Frauenberufe (medizinische und

⁴ Insgesamt haben wir rund 7.500 Bewerbungen versandt, allerdings schlossen wir für die vorliegenden Analysen alle Bewerbungen ohne Foto und diejenigen von fiktiven Bewerberinnen mit Kopftuch aus.

zahnmedizinische Fachangestellte). Die Berufe wurden auf Grundlage verschiedener Überlegungen ausgewählt. So mussten während der Feldphase der Studie für jeden Beruf ausreichend viele freie Stellen ausgeschrieben sein. Außerdem wählten wir Berufe mit unterschiedlichen Zugangsqualifikationen (Haupt- oder Realschulabschluss) und verschiedenem Ausmaß an Kundenkontakt (siehe Tabelle 1).

Tabelle 1: Berufe

Berufsfeld	Geschlecht	Beruf	Kundenkontakt	Qualifikation
Technisch	männlich	Mechatroniker	gering	Realschule
		Anlagenmechaniker-SHK	gering	Hauptschule
Gastronomie	männlich & weiblich	Hotelfachmann/frau	hoch	Realschule
		Koch/Köchin	gering	Hauptschule
Handel	weiblich	Verkäufer/in	hoch	Hauptschule
		Industriekaufmann/frau	gering	Realschule
Gesundheitswesen	weiblich	Medizinische Fachangestellte	hoch	Realschule
		Zahnmedizinische Fachangestellte	hoch	Hauptschule

Positive Rückmeldung

Die zentrale abhängige Variable unserer Studie ist die Reaktion von Arbeitgebern auf unsere Bewerbungen, d.h. ob sie Interesse signalisierten oder nicht. Dafür registrierten wir alle möglichen Reaktionen von Arbeitgebern auf unsere Bewerbungen, sowohl per Post, per E-Mail als auch per Telefon. Sobald ein Arbeitgeber uns eine Nachricht zukommen ließ (z.B. eine Einladung zum Vorstellungsgespräch, eine Bitte um Rückmeldung oder eine Nachfrage zur Bewerbung), zogen wir unsere Bewerbung freundlich zurück, um den Aufwand für

die getesteten Firmen und den möglichen Nachteil für andere Bewerber minimal zu halten.

Auf Grundlage der Reaktion der Arbeitgeber erstellten wir eine dichotome Variable, die zwischen positiven und negativen Rückmeldungen unterscheidet. Als positive Rückmeldung galten dabei Einladungen zu einem Bewerbungsgespräch oder zum Probearbeiten, Nachfragen zur Bewerbung sowie die Bitte um einen Rückruf. Als negative Rückmeldungen werteten wir, wenn wir eine explizite Absage, nicht mehr als eine Eingangsbestätigung oder gar keine Antwort erhielten.

Experimentell variierte Merkmale

Geschlecht

Wir versendeten Bewerbungen von sowohl männlichen als auch weiblichen Stellensuchenden. Das Geschlecht der Bewerber haben wir dabei auf verschiedene Weisen verdeutlicht: durch den Namen der Bewerber, durch die geschlechtsspezifische Berufsbezeichnung in der Überschrift der Bewerbungen (z.B. Industriekaufmann oder Industriekauffrau) sowie durch das ggf. beigefügte Foto. Wie oben beschrieben, bewarben wir uns in vier Berufen mit weiblichen und männlichen Profilen und in jeweils zwei Berufen mit jeweils nur männlichen oder nur weiblichen Profilen.

Ethnizität

Insgesamt erstellten wir 35 Profile mit verschiedenen ethnischen Zugehörigkeiten: 25 Prozent der Bewerber waren Deutsche ohne Migrationshintergrund, 25 Prozent hatten einen türkischen Migrationshintergrund, und die restlichen 50 Prozent verteilten sich auf 33 weitere Herkunftsländer (siehe Abbildung 1). Bewerber mit türkischem Migrationshintergrund waren in der Stichprobe überrepräsentiert, um die Ergebnisse unserer Studie mit früheren Korrespondenzstudien in Deutschland (Kaas und Manger 2012; Schneider, Yemane und Weinmann 2014) vergleichen zu können. Um die ethnische Herkunft zu signalisieren, wählten wir ländertypische Vor- und Nachnamen (siehe Tabelle A1 im Anhang). Zusätzlich war im

entsprechenden Abschnitt des Lebenslaufs neben der Muttersprache Deutsch eine zweite Muttersprache (z.B. Rumänisch) erwähnt.



Abbildung 1: Herkunftsländer der fiktiven Bewerber

Die Auswahl der Herkunftsländer erfolgte aufgrund von mehreren zentralen Kriterien. Zunächst wählten wir Herkunftsländer aus, die in Deutschland besonders stark vertreten sind (z.B. Türkei, Bosnien und Herzegowina, Polen, Russland und Italien). Ein weiteres wichtiges Kriterium war die Verteilung von Religionszugehörigkeiten. Um die Effekte von ethnischer Zugehörigkeit und Religion entzerren zu können, haben wir gezielt Länder ausgewählt, in denen zwei dominante Religionsgruppen vertreten sind (z.B. Christen und Muslime in Nigeria). Um die Effekte von ethnischer Zugehörigkeit und Phänotyp unterscheiden zu können, wählten wir zusätzlich Länder aus, in denen es eine große phänotypische Heterogenität gibt (z.B. Ägypter mit weißem und schwarzem Phänotyp). Ein weiteres Kriterium war, dass wir Herkunftsländer auf allen Kontinenten berücksichtigen wollten.

Religion

Nach dem Vorbild anderer Korrespondenztests (Wallace, Wright und Hyde 2014) signalisierten wir die Religionszugehörigkeit durch ehrenamtliches Engagement in

einem Verein. Alle Kandidaten engagierten sich neben der Arbeit ehrenamtlich in einem sozialen Nachbarschaftsverein „Sozialverein AKTIV e.V.“. In rund der Hälfte der Fälle war dieser Verein zusätzlich mit einer Konfession verbunden, z.B. „Christlicher Sozialverein AKTIV e.V.“. Welche Religion signalisiert wurde, hing vom Anteil der Gläubigen im jeweiligen Herkunftsland ab (siehe Tabelle A2 im Anhang). Pro Herkunftsland wurden maximal zwei Religionen berücksichtigt, wenn im jeweiligen Herkunftsland ein Bevölkerungsanteil von mindestens 5 % den jeweiligen Religionen angehörte.

Phänotyp

Um herauszufinden, welchen Einfluss das phänotypische Erscheinungsbild der Bewerber auf die Rückmelderaten hat, variierten wir die der Bewerbung beigefügten Fotos. Insgesamt gab es zwei Fotoreihen mit jeweils sieben Fotos, die nach drei stereotypischen Herkunftsregionen, und zwar Nord- und Südeuropa (weiß), Ost- und Westafrika (schwarz) und Südost- und/Ostasien (asiatisch) gruppiert wurden. Die Hälfte der Fotos wurde, ausgehend von einem einzelnen Originalfoto, mit Hilfe von Photoshop erstellt. Die andere Hälfte waren „echte“ Bilder, die als plausibel bzw. stereotypisch für die jeweilige Herkunftsregion gelten konnten. Da wir sowohl weibliche als auch männliche Profile erstellten, gab es insgesamt 28 Fotos, jeweils 14 für Männer und 14 für Frauen. Die verschiedenen Fotos wurden den Bewerbungen getrennt nach Geschlecht zufällig zugewiesen, wobei wir unplausible oder untypische Kombinationen von Ethnizität und Phänotyp ausschlossen (z.B. Personen mit chinesischem Namen und schwarzem Phänotyp).

Für die Auswahl der Fotos und die Zuweisung von Fotos zu Herkunftsgruppen führten wir zwei Vorstudien durch: eine zur Bewertung der Attraktivität und Kompetenz der Kandidaten und eine zur Plausibilität der Kombination von Phänotyp und Ethnizität. Die erste Vorstudie führten wir zunächst an einem *convenience sample* von Studierenden der Sozialwissenschaften der Humboldt Universität in Berlin durch. Jedoch erhielten die Fotos von Angehörigen von Minderheiten besonders hohe Bewertungen für Attraktivität und Kompetenz, was wir als sozial erwünschte Antworttendenz unter Studierenden interpretierten. Daher führten wir eine weitere, repräsentativere Umfrage in Fußgängerzonen und

Stationen des öffentlichen Nahverkehrs von Berlin und Potsdam durch. Für jedes Foto fragten wir ca. 25 Personen, wie attraktiv sie die abgebildete Person auf einer zehnstufigen Skala (sehr unattraktiv bis sehr attraktiv) beurteilten und wie kompetent sie die Person in Bezug auf den Arbeitsmarkt auf einer zehnstufigen Skala von sehr inkompetent bis sehr kompetent einschätzten. Auf Grundlage dieser Testergebnisse wurden Bilder ausgewählt, die vergleichbare Bewertungen für die Dimensionen Kompetenz und Attraktivität erhielten.

Die zweite Vorstudie wurde unter Mitarbeitern unseres Forschungsinstituts durchgeführt. Insgesamt befragten wir 38 Personen, welche Phänotypen sie als sehr unwahrscheinlich für bestimmte Herkunftsländer einschätzten. Die Befragten beurteilten auf einer vierstufigen Skala (von „sehr unwahrscheinlich“ bis „sehr wahrscheinlich“) Kombinationen von Fotos und den 35 Herkunftsländern unserer Studie. Ausgehend von den Ergebnissen schlossen wir unplausible Kombinationen zwischen Phänotypen und Herkunftsländern aus. Als unwahrscheinlich galten dabei Kombinationen, bei denen die Zustimmung unter 20 Prozent fiel; d.h. wenn mehr als 80 Prozent der Befragten ein Foto als sehr unwahrscheinlich für eine bestimmte ethnische Gruppe beurteilte. Aus der Tabelle A2 im Anhang geht hervor, welche Kombinationen von Phänotyp und ethnischer Gruppe für die Studie genutzt wurden.

Noten, Referenzschreiben und Vertragslaufzeit

Noten: Um herauszufinden, ob gute Noten unterschiedliche Effekte für Bewerber mit und ohne Migrationshintergrund haben, variierten wir die Noten auf fiktiven Schul- und Ausbildungszeugnissen. Die Hälfte aller Zeugnisse beinhaltete Informationen über gute Noten (Schulnote 2 im deutschen Notensystem), die andere Hälfte der fiktiven Kandidaten hatte befriedigende Noten (Schulnote 3 im deutschen Notensystem). Für den Beruf des Kochs bzw. der Köchin senkten wir das Notenniveau auf die Noten 3 und 4, da Auszubildende in diesem Beruf im Schnitt etwas schlechtere Abschlussnoten haben als Absolventen in den restlichen untersuchten Berufen.

Referenzschreiben: Um die Ergebnisse von Kaas und Manger (2012) replizieren zu können, enthielten etwa die Hälfte der Bewerbungsunterlagen ein

Referenzschreiben, in dem der aktuelle Arbeitgeber den jeweiligen Bewerber positiv bewertete. Innerhalb der Bedingung „Referenzschreiben“ gab es noch zwei weitere, unabhängige Bedingungen: Die Hälfte der Referenzschreiben betonten die *sozialen Fähigkeiten* des Bewerbers, während die andere Hälfte der Schreiben sich dazu nicht explizit äußerte. Zweitens betonten wir in 50 Prozent der Referenzschreiben die *sprachlichen Fähigkeiten* des Bewerbers, während 50 Prozent der Schreiben keine Aussage zu sprachlichen Fähigkeiten enthielten. Aus diesen Bedingungen ergaben sich vier verschiedene Kombinationen von Referenzschreiben mit gleicher Zuweisungswahrscheinlichkeit.

Vertragslaufzeit: Am Ende des Motivationsschreibens erwähnten alle fiktiven Bewerber, dass sie ihren aktuell laufenden Arbeitsvertrag innerhalb einer Kündigungsfrist von vier Wochen beenden könnten. In 50 Prozent der Fälle schrieben die Kandidaten, dass der aktuelle Vertrag in wenigen Monaten auslaufen würde, also ein befristeter Vertrag vorliege. In den anderen 50 Prozent wurde keine Angabe diesbezüglich gemacht, was suggerieren sollte, dass der Bewerber in einem unbefristeten Arbeitsverhältnis angestellt ist.

Gruppenmerkmale

Um die Ursachen für Diskriminierung genauer analysieren zu können, berücksichtigten wir zudem sowohl produktivitätsbezogene als auch kulturelle Merkmale auf Gruppenebene. Alle Informationen über produktivitätsbezogene Gruppenmerkmale stammen aus dem Mikrozensus und betreffen Migrantengruppen (d.h. Ausländer und Deutsche mit Migrationshintergrund). Zur Messung kultureller Merkmale nutzten wir die Values Surveys (World Values Survey und European Values Survey), auf deren Grundlage wir für jedes Herkunftsland Mittelwerte berechneten.

Wir nutzten Informationen über folgende drei Arten von Gruppen: ethnische Gruppen (i.S. der 35 Herkunftsländer), phänotypische Gruppen (i.S. von *race*) und schließlich Religionsgruppen, da wir eben diese Merkmale auf individueller Ebene experimentell variierten. Die Idee dabei war, dass Arbeitgeber möglicherweise annehmen, dass weiße, christliche Ägypter und schwarze, muslimische Ägypter einerseits gewisse Gemeinsamkeiten aufgrund ihrer ethnischen

Gruppenzugehörigkeit haben; sich gleichzeitig aber aufgrund der verschiedenen Religionen und Phänotypen voneinander unterscheiden.

Um Merkmale auf den drei Gruppenebene bzw. unterschiedliche Merkmalskombinationen in unseren Analysen berücksichtigen zu können, berechneten wir zunächst Mittelwerte des jeweiligen Merkmals für alle drei Gruppen:

- Für ethnische Gruppen errechneten wir Mittelwerte für Migrantengruppen bzw. Herkunftsländer.
- Für phänotypische Gruppen berechneten wir Mittelwerte auf Grundlage aller Länder der Welt, in denen ein bestimmter Phänotyp eindeutig überwiegt, gewichtet nach dem jeweiligen Bevölkerungsanteil in Deutschland. Die Angaben für Bewerber mit schwarzem Phänotyp sind folglich dominiert durch Merkmalsausprägungen von Migranten aus afrikanischen Ländern.
- Für Religionsgruppen berechneten wir Mittelwerte über alle Länder hinweg, in denen das Christentum, der Islam oder der Hinduismus bzw. Buddhismus die Mehrheitsreligion darstellt, gewichtet nach dem jeweiligen Bevölkerungsanteil in Deutschland. Die Angaben für Muslime sind folglich überdurchschnittlich geprägt durch die Merkmalsausprägung unter türkischen Migranten bzw. Türken.

Letztendlich nutzten wir in unseren Analysen gemittelte Werte für die Kombination aus ethnischer, phänotypischer und religiöser Gruppenzugehörigkeit. Das heißt, jedem fiktiven Bewerber wurden jeweils drei Werte auf Gruppenebene zugeordnet, einer für Ethnizität, einer für Phänotyp und einer für Religion. Diese drei Werte haben wir aufsummiert und durch drei geteilt, sodass der resultierende Wert für jeden Bewerber die entsprechende Kombination aus Ethnizität, Phänotyp und Religion widerspiegelte. Weiße, christliche Ägypter erhielten beispielsweise einen Wert, der auf den Mittelwerten für Weiße, Christen und Ägyptern beruhte und sich von dem Wert für schwarze muslimische Ägypter unterschied, da dieser

wiederum aus den durchschnittlichen Werten für Schwarze, Muslime und Ägypter errechnet wurde.⁵

Produktivität

Als Indikatoren der mittleren Produktivität einer Gruppe nutzten wir jeweils das durchschnittliche Bildungsniveau, den beruflichen Status sowie die Erwerbslosigkeitsrate der jeweiligen Gruppen. Das durchschnittliche Bildungsniveau ist dabei der wichtigste der drei Indikatoren, da beruflicher Status und Erwerbslosigkeit nicht nur Ursache sondern auch Folge von Diskriminierung sein können.

Die Daten bezogen wir aus dem Mikrozensus von 2012. Hierbei handelt es sich um die Befragung einer repräsentativen Stichprobe, die ein Prozent der Bevölkerung in Deutschland darstellt. Um den durchschnittlichen Bildungsgrad einer Migrantengruppe abzubilden, erfassten wir die Bildung von Personen mit ausländischer Staatsangehörigkeit in Deutschland sowie Menschen mit erweitertem Migrationsstatus, also Migranten vor einer Einbürgerung, sowie Migranten in zweiter Generation:⁶

⁵ Für Bewerber, die keine bestimmte Religion signalisierten, wurde auf den Wert der im jeweiligen Herkunftsland dominanten Religion zurückgegriffen. Diese Vorgehensweise lässt sich durch Vorabanalysen rechtfertigen, die bestätigten, dass für die jeweiligen Religionsgruppen kein signifikanter Unterschied in positiven Rückmeldungen zwischen den Bewerbungen ohne Religionssignal und den Bewerbungen, für die eine Zugehörigkeit zur Mehrheitsreligion signalisiert wurde, besteht. Signifikante Unterschiede gibt es nur zwischen Bewerbern, die keine Religionszugehörigkeit signalisieren und solchen, die einer Minderheitsreligion zugehörig sind.

⁶ Für einige Gruppen waren keine eigenen Einträge im Mikrozensus verfügbar, weswegen wir die Werte für diese Gruppen aufgrund von größeren Herkunftskategorien schätzen mussten. Zum Beispiel stammen die Werte für Mexikaner und Trinidadier aus der Kategorie „Zentralamerika und Karibik“. Auch wenn diese Kategorien teilweise sehr grob sind, handelt es sich hierbei dennoch um die genauesten Daten, die in Deutschland erhältlich sind. Außerdem ist davon auszugehen, dass Arbeitgeber kein fundiertes Wissen über den durchschnittlichen Bildungsstand von Gruppen haben, für die es nicht einmal detaillierte Werte in offiziellen Statistiken gibt.

- *Bildungsniveau:* Das durchschnittliche Bildungsniveau wurde auf einer vierstufigen Skala erfasst (keine Bildung, niedrige, mittlere und hohe Bildung⁷).
- *Beruflicher Status:* Außerdem bezogen wir aus dem Mikrozensus den mittleren ISEI (*International Socio-Economic Index of Occupational Status* bzw. internationales sozioökonomisches Maß des beruflichen Status) für alle berufstätigen Personen der jeweiligen Herkunftsgruppe als weiteren Indikator für Produktivität.
- *Erwerbslosigkeit:* Zudem wurde, ebenfalls aus dem Mikrozensus, die Erwerbslosigkeitsrate für jede Herkunftsgruppe berücksichtigt. Diese wurde definiert als Prozentsatz der Personen zwischen 15 und 64 Jahren, die arbeitssuchend und Teil der Erwerbsbevölkerung waren. Ausgeschlossen waren demnach Hausmänner/-frauen, Frühberentete und Studierende etc.

Kulturelle Distanz

Um „Homophilie“ als mögliche Ursache für Diskriminierung überprüfen zu können, haben wir kulturelle Ähnlichkeit bzw. Distanz mittels Werteunterschieden zwischen Herkunftsgruppen und Deutschen erfasst. Da keine verlässlichen Statistiken über Werte in unterschiedlichen Migrantengruppen in Deutschland verfügbar waren, berechneten wir die kulturelle Distanz auf Grundlage durchschnittlicher Werteausprägungen in den jeweiligen Herkunftsländern. Hierfür nutzten wir für jedes Herkunftsland die jeweils aktuellste verfügbare Welle des World Values Survey (WVS) oder des European Values Survey (EVS). Wir berechneten für jedes Herkunftsland die Distanz in der durchschnittliche Ausprägung emanzipativer Werte (vs. Überlebenswerte) und säkularer Werte (vs. traditioneller Werte) von den durchschnittlichen Werten in Deutschland:

- *Emanzipative Werte:* Emanzipative Werte umfassen vier Bereiche emanzipativer Orientierung: postmaterielle Werte, Toleranz gegenüber

⁷ Keine Bildung: kein Schulabschluss, maximal siebenjährige Schulzeit; niedrige Bildung: Hauptschulabschluss, maximal neunjährige Schulzeit; mittlere Bildung: Realschulabschluss, maximal zehnjährige Schulzeit; hohe Bildung: Fachabitur, Abitur.

Scheidung, Abtreibung und Homosexualität, Autonomie in der Kindeserziehung, sowie Gleichberechtigung der Geschlechter.

- *Säkulare Werte*: Säkulare Werte beziehen sich auf die Loslösung von Konformitätsnormen und von religiösen, staatlichen sowie anderen nichtreligiösen Autoritäten, wie zum Beispiel Familie.

Die beiden Dimensionen (emanzipativ und säkular) sind stark korreliert ($r = .83$) und konnten daher nicht gleichzeitig in multivariaten Modellen getestet werden. Da sich Ländern insbesondere hinsichtlich ihrer emanzipativen Werte unterscheiden⁸, konzentrieren wir uns im Ergebnisteil dieser Studie auf die emanzipative Wertedimension, auch wenn eine kombinierte Skala zu ähnlichen Ergebnissen führt. Da ausgehend von der Theorie der sozialen Homophilie die Ähnlichkeit das zentrale Maß ist, maßen wir den relativen Abstand zum durchschnittlichen deutschen Wertemuster bzw. zu den Werten in christlich geprägten Ländern und Ländern mit überwiegend phänotypisch weißer Bevölkerung.

⁸ In unseren Analysen haben wir sowohl den Einfluss der beiden Einzeldimensionen als auch der kombinierten Skala untersucht. Dabei stellte sich heraus, dass emanzipative Werte der bessere Schätzer sind. Dieser Befund deckt sich mit der Erkenntnis von Inglehart und Norris (2003), dass sich die Wertevorstellung von Bürgern muslimischer und westlicher Länder vor allem hinsichtlich ihrer Ansichten zu Sexualität und Geschlecht unterscheiden, also Aspekten, die den Kern der emanzipativen Dimension ausmachen, und weniger in Vorstellungen über Demokratie, die eher die säkulare Dimension abbilden.

Ergebnisse

Deskriptive Ergebnisse

Im Schnitt gaben rund 54 Prozent aller angeschriebenen Arbeitgeber den Bewerbern eine positive Rückmeldung, 39 Prozent luden die Bewerber bei der ersten Kontaktaufnahme sogar direkt zu einem Vorstellungsgespräch oder zum Probearbeiten ein.

Allerdings zeigen sich sehr große Unterschiede zwischen den acht Berufsgruppen. Während die positive Rückmeldequote für Köche, Anlagenmechaniker, Mechatroniker, medizinische und zahnmedizinische Fachangestellte zwischen 56 und 70 Prozent liegt, fällt die Quote für Verkäufer/innen mit 41 Prozent und Industriekaufmännern/-frauen mit 21 Prozent deutlich niedriger aus. Im Vergleich zu Köchen und Köchinnen sind die positiven Rückmeldequoten für Industriekaufmänner und -frauen, Verkäufer/innen und medizinische Fachangestellte signifikant geringer und für zahnmedizinische Fachangestellte signifikant höher (siehe Abbildung 2). Betrachtet man die Analysen getrennt nach Bewerbungen mit deutschen und ausländischen Namen, wird deutlich, dass Unterschiede zwischen Berufen für beide Gruppen in die gleiche Richtung gehen. Zudem übertrifft die positive Rückmeldequote von Bewerbern ohne Migrationshintergrund in allen Berufen diejenige von Bewerbern mit Migrationshintergrund. Eine einzige Abweichung vom Muster ergibt sich für Hotelfachmänner/-frauen. In dieser Berufsgruppe liegt die Quote für Bewerber mit und ohne Migrationshintergrund jeweils bei 61 Prozent. Dies könnte unter anderem auf die zusätzlichen Sprachkenntnisse zurückzuführen sein, die jeder Bewerber mit ausländischen Namen in seinem Lebenslauf angab.

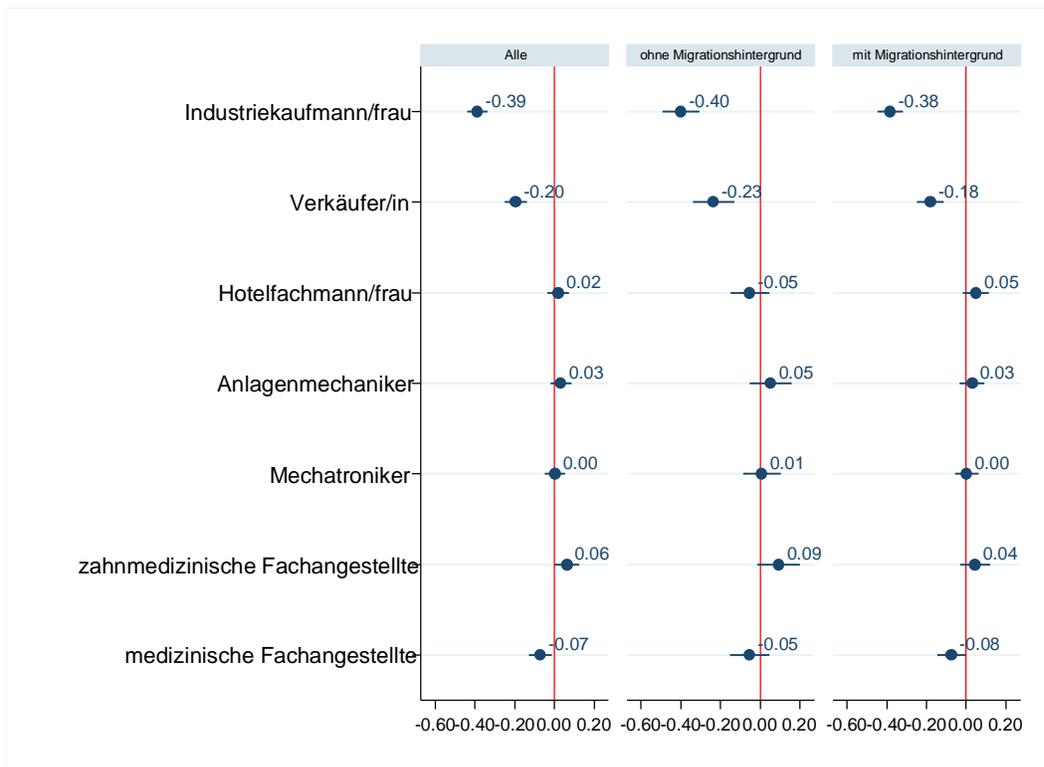


Abbildung 2: Rückmeldung nach Berufen

Die Abbildung zeigt, wie sich die positive Rückmeldungsquote je nach Beruf von der Quote für Köche und Köchinnen unterscheidet. Dargestellt sind die jeweiligen Koeffizienten mit Vertrauensintervall (.95), basierend auf einer linearen Regression mit folgenden zusätzlichen Variablen: Geschlecht, Migrationshintergrund, Religion, Phänotyp, Monat der Veröffentlichung, Noten, Vertragsinformation und Referenzschreiben (im ersten und dritten Modell wurden Mehrebenenmodelle mit Herkunftsklustern berechnet).

Welche Rolle spielt das Herkunftsland?

Die Ergebnisse zeigen eine deutliche Diskriminierung gegenüber Bewerbern mit Migrationshintergrund. Während Bewerber mit deutschem Namen in 60 Prozent aller Fälle eine positive Rückmeldung erhalten, ist dies nur bei 51 Prozent der Bewerber mit Migrationshintergrund der Fall. Dieser absolute Unterschied von neun Prozentpunkten reduziert sich zwar auf sechs Prozentpunkte Unterschied, wenn man für Religion und Phänotyp sowie für weitere Merkmale der Bewerbungen (z.B. Geschlecht und Beruf) kontrolliert, bleibt aber dennoch statistisch hochsignifikant ($p < .001$).

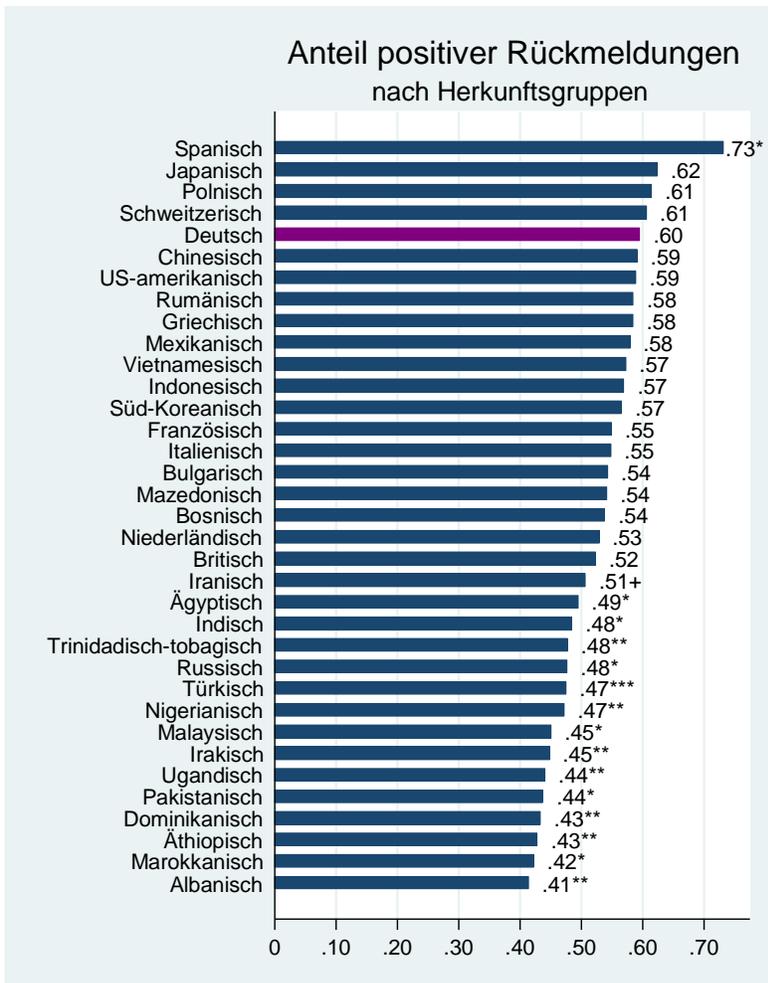


Abbildung 3: Rückmeldung je Herkunftsgruppe

Die abgebildeten Balken zeigen die Rücklaufquoten je Gruppe ohne weitere Kontrollvariablen. Bewerbungen von Personen ohne Migrationshintergrund sind farblich hervorgehoben. Die Pluszeichen und Sternchen hinter den jeweiligen Quoten geben an, ob der dargestellte Unterschied zur Rücklaufquote von Bewerbungen ohne Migrationshintergrund statistisch signifikant ist (mit $p < .001^{***}$, $p < .01^{**}$, $p < .05^*$ und $p < .10^+$, zweiseitige Testung), wenn in linearen Regressionen eine Reihe weiterer Merkmale der Bewerbungen (wie Geschlecht, Noten, Referenzschreiben, Vertragsinformation, Berufsgruppe, und Bewerbungszeitpunkt) berücksichtigt werden.

Abbildung 3 gibt einen Überblick über den Anteil an positiven Rückmeldungen je nach Herkunftsgruppe. Während die meisten Herkunftsgruppen deutlich seltener eine positive Rückmeldung erhalten als deutschstämmige Bewerber, gibt es auch eine Vielzahl an Herkunftsgruppen, für die sich die Rücklaufquoten kaum von derjenigen für Bewerber ohne Migrationshintergrund unterscheiden. Dazu zählen Bewerber aus westlichen Demokratien, aus ost- sowie südosteuropäischen Ländern, aber auch aus ostasiatischen Ländern wie Japan und China. Die höchste Rückmelderate haben Bewerber mit spanischen Namen, während insbesondere

Bewerber mit einem albanischen, marokkanischen, äthiopischen oder dominikanischen Migrationshintergrund signifikant seltener eine positive Rückmeldung erhalten (17-19 Prozentpunkte Unterschied). Auch die Rückmelderaten für Bewerber mit türkischem Migrationshintergrund bewegen sich im unteren Drittel - und fallen mit einem Nachteil von 13 Prozentpunkten signifikant geringer aus als die Rückmeldequoten für deutschstämmige Bewerber ($p < .001$).

Welche Rolle spielt die Religion?

Insgesamt liegt die positive Rückmeldequote für Bewerber ohne Religionssignal bei 54%, für Christen bei 57%, für Muslime bei 46% und für Buddhisten und Hindus bei 53%. Allerdings waren die Religionssignale aufgrund des Designs der Studie über die Herkunftsgruppen ungleich verteilt. Bewerber ohne Migrationshintergrund waren beispielsweise niemals Muslime oder Buddhisten/Hindus.

Vergleicht man deutschstämmige Bewerber mit unterschiedlichen Religionssignalen (siehe oberer Abschnitt in Abbildung 4), so zeigt sich, dass sich die Häufigkeit positiver Rückmeldungen für Bewerber ohne Religionssignal von christlichen Bewerbern nicht unterscheidet⁹. Betrachtet man die Ergebnisse für Bewerber mit Migrationshintergrund, so zeigt sich ein anderes Bild: sie profitieren leicht davon, wenn sie Nähe zum Christentum signalisieren ($p < .05$; im Vergleich zu keinem Religionssignal). Gleichzeitig fällt die Wahrscheinlichkeit einer positiven Rückmeldung für muslimische Bewerber mit Migrationshintergrund im Vergleich zu christlichen Bewerbern deutlich geringer aus: sie sinkt um neun Prozentpunkte ($p < .001$). Im Gegensatz dazu hat es keine Auswirkungen, wenn sich Bewerber mit Migrationshintergrund in einem hinduistischen oder buddhistischen Verein engagieren.

Der Nachteil für muslimische Bewerber ist erheblich. Allerdings wurde das muslimische Religionssignal nur für Bewerber aus Herkunftsländern mit einem substantiellen Anteil von Muslimen in der Bevölkerung vergeben. Gleiches gilt für

⁹ Für Bewerber ohne Migrationshintergrund wurde entweder kein Religionssignal gegeben, oder ein Hinweis auf Christentum. Sie erhielten jedoch niemals ein Signal, das auf den Islam oder Hinduismus/Buddhismus verwies.

das christliche Religionssignal. Daher reduzierten wir in einem zweiten Schritt die Analysen auf die wenigen Herkunftsländer mit substantieller christlicher *und* muslimischer Bevölkerung und betrachteten nur Bewerber mit weißem Phänotyp (d.h. Migranten aus Albanien, Bosnien und Herzegowina, Bulgarien, Mazedonien, Russland, und Ägypten). Auch unter diesen starken Einschränkungen bleibt ein statistisch signifikanter Nachteil für muslimische Bewerber von zehn Prozentpunkten bestehen (Rückmelderate von 52% für Christen versus 42% für Muslime; $p < .05$). Für Bewerber mit Migrationshintergrund in diesen Ländern und asiatischem oder schwarzem Phänotyp (d.h. Migranten aus Albanien, Indonesien, Mazedonien, Nigeria, Uganda, Äthiopien und Ägypten) sind die Unterschiede etwas kleiner (53% versus 45%) und aufgrund der kleineren Fallzahl ($N = 460$) nicht mehr statistisch signifikant. Insgesamt spiegeln die Ergebnisse dieser Binnenanalysen die Ergebnisse der Gesamtstichprobe wieder und deuten darauf hin, dass nicht gegen „Nicht-Christen“ per se oder gegenüber Migranten aus muslimisch geprägten Ländern, sondern speziell gegenüber Muslimen diskriminiert wird.

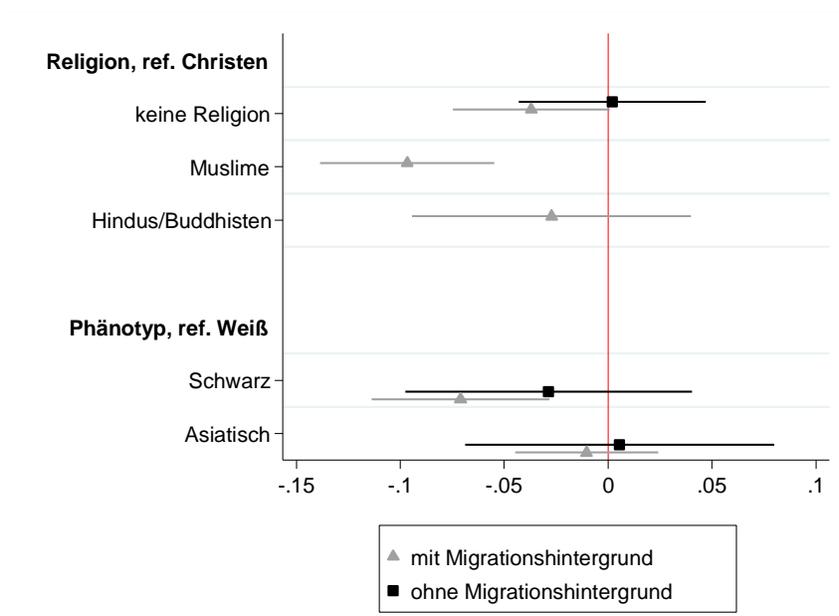


Abbildung 4: Religion und Phänotyp

Die Abbildung zeigt, wie sich die positive Rückmeldungsquote für Bewerber mit und ohne Migrationshintergrund je nach Religionssignal und Phänotyp unterscheidet. Dargestellt sind die jeweiligen Koeffizienten mit Vertrauensintervall (.95), basierend auf einer linearen Regression (für Bewerber mit Migrationshintergrund wurden Mehrebenenmodelle mit Herkunftsklustern berechnet), kontrolliert für Berufsgruppe, Geschlecht, Monat der Veröffentlichung, Noten, Vertragsinformation, Referenzschreiben und Religion bzw. Phänotyp.

Welche Rolle spielt der Phänotyp?

Die positiven Rückmeldequoten unterscheiden sich zwischen den drei phänotypischen Gruppen¹⁰ (Asiatisch: 53%, Schwarz: 49%, Weiß: 55%), wobei im Vergleich zu weißen Phänotypen insbesondere ein schwarzer Phänotyp zu signifikant niedrigeren Rückmelderaten führt ($p < .01$), nicht aber ein asiatischer Phänotyp.

In getrennten Analysen für Bewerber mit und ohne Migrationshintergrund (siehe unterer Abschnitt in Abbildung 4) wird deutlich, dass sich für Bewerber mit ausländischem Namen die Quote an positiven Rückmeldungen um sieben Prozentpunkte verschlechtert ($p < .01$), wenn der Bewerbung ein Bewerbungsfoto einer phänotypisch schwarzen Person beiliegt. Für die Bewerber mit deutschem Namen geht der Trend in die gleiche Richtung, auch wenn die Effektstärke mit drei Prozentpunkten Unterschied niedriger ausfällt und statistisch nicht signifikant ist. Für asiatische Bewerbungsfotos bleiben die Effekte in den getrennten Analysen statistisch unbedeutend.

Da der Phänotyp aus Plausibilitätsgründen nicht gänzlich zufällig zugewiesen werden konnte, sind die Bewerbungsfotos über Herkunftsgruppen ungleich verteilt. So gab es beispielsweise nie einen Bewerber mit chinesischem Namen und schwarzem Phänotyp. Um jedoch eindeutig Rückschlüsse darauf ziehen zu können, ob es sich um phänotypische Diskriminierung handelt, müssen Bewerbergruppen miteinander verglichen werden, für die die gleiche Auswahl an Fotos genutzt wurde. Insgesamt gibt es sieben Herkunftsländer, für die wir sowohl weiße als auch schwarze Bewerber ins Rennen geschickt haben: Deutschland, Großbritannien, Frankreich, Vereinigte Staaten von Amerika, Dominikanische Republik, Ägypten und Marokko. Um die Effekte des Phänotyps von den Effekten der Religion sauber trennen zu können, schlossen wir für diese Analyse alle Bewerbungen mit muslimischem Religionssignal aus. Übrig bleiben also nur Bewerbungen mit

¹⁰ Insgesamt gab es in den beiden Fotoreihen (Photoshop und „echte“ Bilder) jeweils sieben Phänotypen, die nach drei stereotypischen Herkunftsregionen Nord- und Südeuropa (Weiß), Ost- und Westafrika (Schwarz) und Südost- und Ostasien (Asiatisch) gruppiert wurden. Da wir keine signifikanten Unterschiede in den Ergebnissen zwischen den Fotoreihen feststellen konnten, werden im Folgenden die Ergebnisse nach Phänotyp für beide Fotoreihen (Photoshop und „echte“ Bilder) gemeinsam dargestellt.

weißen oder schwarzen Bewerbungsfotos, bei denen entweder ein christliches oder neutrales Ehrenamt signalisiert wurde. Auch bei dieser reduzierten Stichprobe bleibt der negative Effekt für schwarze Phänotypen bestehen und ist vergleichbar mit der Effektstärke in der Gesamtstichprobe (60% positive Rückmeldungen für Bewerber mit weißem Phänotyp vs. 53% positive Rückmeldungen für Bewerber mit schwarzem Phänotyp; $p < .05$).

Das gleiche Verfahren kann auch für Bewerber mit asiatischem Phänotyp angewendet werden. In der Gesamtstichprobe gibt es 14 Herkunftsländer, für die Bewerbungen von sowohl weißen als auch asiatischen-aussehenden Bewerbern versendet wurden (Deutschland, Großbritannien, Frankreich, Albanien, Indien, Irak, Iran, Mazedonien, Mexico, Pakistan, Rumänien, Spanien, Türkei und Vereinigte Staaten von Amerika). Auch hier wurden nur die Bewerbungen berücksichtigt, bei der der Sozialverein entweder christlich oder ohne religiösen Hinweis war. Für weiße Bewerber liegt die Quote an positiven Rückmeldungen bei 57 Prozent, für asiatisch aussehende Bewerber bei 52 Prozent. Der Unterschied von fünf Prozentpunkten ist größer als in der Gesamtstichprobe und in diesem Fall auch statistisch bedeutsam ($p < .05$). Da asiatische Phänotypen überdurchschnittlich häufig für bestimmte Herkunftsländer (Japan oder China) genutzt wurden und diese Herkunftsländer wiederum mit bestimmten Merkmalen verknüpft sind (relativ geringere kulturelle Distanz und hohes durchschnittliches Bildungsniveau), ist es möglich, dass die negativen Effekte eines asiatischen Phänotyps in der Gesamtstichprobe nicht zutage treten.

Weitere Merkmale: Geschlecht, Referenzschreiben, Noten und Vertragsende

Abbildung 5 gibt einen Überblick über die Auswirkungen aller weiteren experimentell variierten Merkmale der Bewerber auf die Wahrscheinlichkeit, eine positive Rückmeldung zu erhalten, sowohl für die Gesamtstichprobe als auch getrennt für Bewerber mit und ohne Migrationshintergrund.

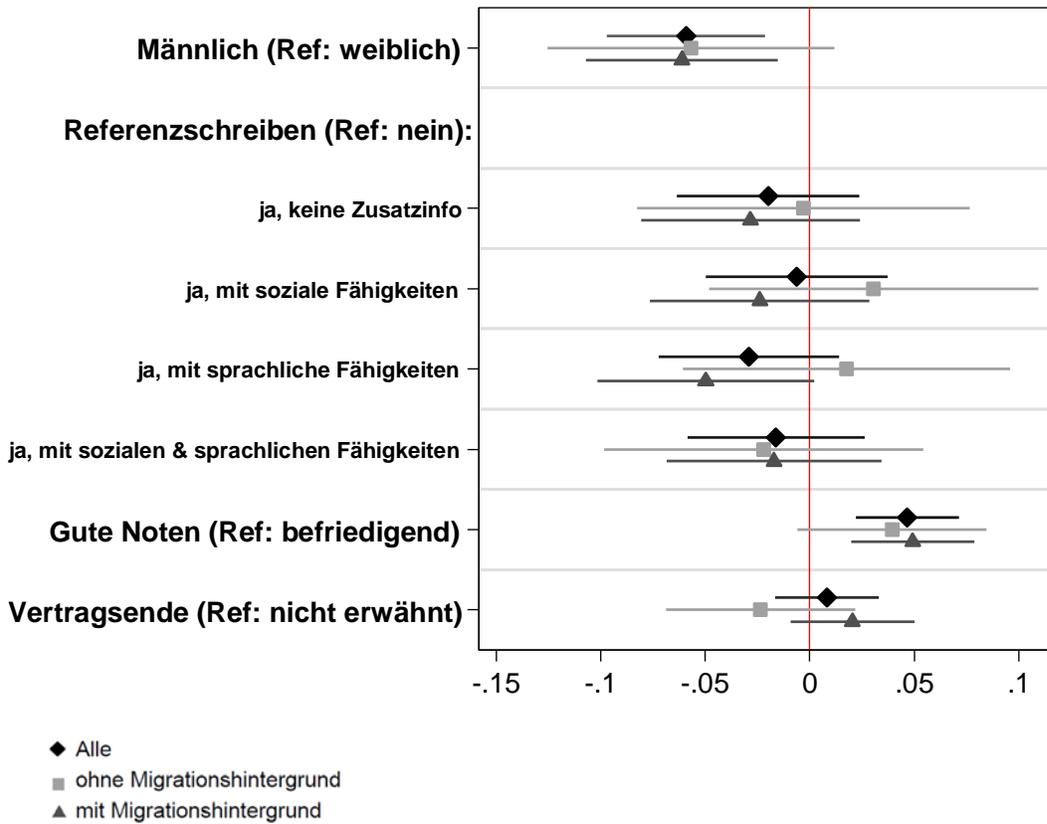


Abbildung 5: Produktivitätsmerkmale und Geschlecht

Die Abbildung zeigt sowohl für alle Bewerber als auch getrennt für Bewerber mit und ohne Migrationshintergrund, wie sich die positive Rückmeldungsquote mit dem Geschlecht, Merkmalen des Referenzschreibens, Noten in Abschlusszeugnissen und Informationen über ein baldiges Vertragsende verändert. Dargestellt sind die jeweiligen Koeffizienten mit Vertrauensintervall (.95), basierend auf einer linearen Regression, kontrolliert für Berufsgruppe, Monat der Veröffentlichung, Migrationshintergrund, Religion und Phänotyp (für die Gesamtstichprobe und Bewerber mit Migrationshintergrund wurden Mehrebenenmodelle mit Herkunftsklustern berechnet).

Die Zahl der positiven Rückmeldungen unterscheidet sich signifikant zwischen männlichen und weiblichen Bewerbern, zum Vorteil von Frauen. Während 57 Prozent der Bewerberinnen eine positive Rückmeldung erhalten, gilt dies für nur 51 Prozent der männlichen Bewerber ($p < .01$). Auch wenn man nur die Rückmeldungen in den vier Berufen vergleicht, in denen wir uns mit weiblichen und männlichen Profilen beworben haben, bleibt ein Unterschied von acht Prozentpunkten bestehen.

Insgesamt hat das Beifügen eines Referenzschreibens keine signifikanten Auswirkungen auf die positive Rücklaufquote (siehe Abbildung 5). Betrachtet man die Effekte getrennt für Bewerber mit und ohne Migrationshintergrund, so zeigt sich für deutschstämmige Bewerber ein positiver Trend für Referenzschreiben, die soziale oder sprachliche Fähigkeiten hervorheben, während Bewerber mit Migrationshintergrund die höchste positive Rücklaufquote haben, wenn der Bewerbung überhaupt kein Referenzschreiben beiliegt. Interessant ist auch, dass es für Bewerber mit ausländischem Namen am nachteiligsten ist, wenn im Referenzschreiben gute sprachliche Fähigkeiten hervorgehoben werden. Keiner dieser Unterschiede ist jedoch statistisch signifikant. Es zeigt sich auch kein signifikanter Interaktionseffekt: Bewerber mit ausländischen Namen haben keine besseren Chancen auf eine positive Rückmeldung durch die Beigabe eines Referenzschreibens. Dieses Ergebnis widerspricht deutlich den Befunden von Kaas und Manger (2012), die zu dem tendenziellen Schluss kamen, dass durch die Beigabe eines Referenzschreibens Diskriminierung verringert werden kann.¹¹

Als weiteres Produktivitätssignal haben wir das Notenniveau der Bewerber sowohl für die Schulzeugnisse als auch für die Abschlussnoten der Ausbildung variiert. Es zeigt sich, dass sich die Rate positiver Rückmeldungen für Bewerber mit guten Noten (im Vgl. zu befriedigenden Noten) um fünf Prozentpunkte ($p < .001$) verbessert. Für Bewerber mit und ohne Migrationshintergrund geht der Trend in die gleiche Richtung (siehe Abbildung 5). Der Unterschied ist für deutschstämmige Bewerber aufgrund der geringeren Stichprobengröße nicht signifikant. Zusatzanalysen bestätigten, dass es keinen signifikanten Interaktionseffekt gibt.

¹¹ Diese Schlussfolgerung übersieht jedoch ein wichtiges Merkmal des Forschungsdesigns von Kaas und Manger: beim Versenden der Bewerbungspaare wurden Bewerbungstyp (Typ A: ohne Referenzschreiben, Typ B: mit Referenzschreiben) und Name (deutsch oder türkisch) zufällig der ersten Bewerbung zugewiesen und für die zweite Bewerbung dann das jeweils andere Merkmal ausgewählt. Das bedeutet, dass ein türkischstämmiger Deutscher mit Referenzschreiben immer mit einem Deutschen ohne Referenzschreiben konkurrierte und umgekehrt. Das bedeutet auch, dass vor allem die Ergebnisse für Typ B nur bedingt vergleichbar sind: die Arbeitgeber, die sich bei den türkischstämmigen Deutschen (mit Referenzschreiben) gemeldet haben, hatten (zumindest im Rahmen der Studie) keine gleichwertige Bewerbung mit Referenzschreiben von Deutschen erhalten. Es kann also nicht mit abschließender Sicherheit geklärt werden, ob Bewerber mit ausländischen Namen tatsächlich von einem Referenzschreiben profitieren. Im Rahmen unserer Studie konnte dies zumindest nicht nachgewiesen werden.

Im Anschreiben erwähnte die Hälfte der Bewerber, dass ihr Vertrag in den nächsten Monaten auslaufe und sie u.a. deshalb nach einer neuen Stelle suchten. Die andere Hälfte gab keinerlei Information zur Vertragslaufzeit. Mit dieser Variation wollten wir die Konsequenzen negativer Information testen, da in Deutschland eine unbefristete Anstellung als Normalarbeitsverhältnis gilt und die Befristung bedeuten könnte, dass der jetzige Arbeitgeber den Vertrag des Bewerbers nicht verlängern möchte. Allerdings könnten Arbeitgeber einen befristeten Vertrag auch als Signal dafür interpretieren, dass Bewerber eine ihm angebotene Stelle mit hoher Wahrscheinlichkeit antreten würden. Insgesamt hat die Vertragslaufzeit keinen Effekt auf Rückmeldungen (siehe Abbildung 5). Unsere Ergebnisse zeigen, dass das Vertragssignal für Bewerber mit deutschen Namen eher negativ wahrgenommen wird und zu einer Reduzierung der Rückmelderaten führt. Für Bewerber mit Migrationshintergrund wird das Signal eher positiv wahrgenommen und führt zu einer leicht höheren Rückmelderate im Vergleich zu Bewerbungen ohne Information zum Vertragsende. In beiden Fällen sind die Unterschiede jedoch klein und statistisch nicht signifikant.

Multivariate Analysen

Wie können die oben beschriebenen Unterschiede zwischen ethnischen Gruppen, verschiedenen Phänotypen und der Religion erklärt werden? Der statistischen Diskriminierungstheorie nach greifen Arbeitgeber auf Wissen über die durchschnittliche Produktivität von Angehörigen bestimmter Gruppen zurück, wenn sie Auswahlentscheidungen treffen. Laut der Theorie der präferenzbasierten Diskriminierung fließen andere, nicht produktivitätsrelevante Merkmale in die Entscheidungsfindung ein, wie beispielsweise kulturelle Distanz.

Die deskriptiven Ergebnisse unserer Studie zeigen, dass Bewerber mit Migrationshintergrund, Bewerber mit schwarzem Phänotyp und muslimische Bewerber seltener positive Rückmeldungen erhalten als deutschstämmige Bewerber, Bewerber mit weißem Phänotyp und christliche Bewerber. Wie Abbildung 3 zeigt, unterscheiden sich die positiven Rücklaufquoten stark zwischen unterschiedlichen Herkunftsgruppen. In Übereinstimmung mit Theorien der statistischen Diskriminierung deuten die Ergebnisse an, dass diejenigen

Herkunftsgruppen eine hohen Rückmelderate haben, die im Durchschnitt über ein hohes Bildungsniveau verfügen (z.B. Japaner, Südkoreaner, Chinesen, US-Amerikaner und Schweizer); während Gruppen mit niedrigen Rückmeldequoten im Durchschnitt geringere Bildungsabschlüsse haben (Türken, Iraker und Marokkaner). Allerdings sind die deskriptiven Ergebnisse unserer Studie auch mit präferenzbasierte Erklärungsansätzen vereinbar. Die Herkunftsländer der Migranten mit den höchsten Rücklaufquoten (d.h. westeuropäische und ostasiatische Migranten) sind überwiegend Länder mit ähnlichen Werten wie Deutschland, während die Herkunftsländer der Migranten mit den niedrigsten Rücklaufquoten (z.B. die mit Wurzeln in Pakistan, Nigeria, Türkei, Uganda und Marokko) sich kulturell stark von Deutschland unterscheiden.

Im folgenden Abschnitt soll geklärt werden, inwieweit die zwei Erklärungsansätze der statistischen und der präferenzbasierten Diskriminierung geeignet sind, um die festgestellten Nachteile für Muslime, Schwarze und Migranten im Allgemeinen zu erklären. In den Modellen der Tabelle 2 zeigen wir die Ergebnisse von Panelregressionen, um die hierarchische Datenstruktur unserer Studie zu berücksichtigen (einzelne experimentelle Bedingungen, die innerhalb der 35 ethnischen Gruppen eingebettet sind). Wir nehmen als erklärende Variablen Ethnizität (Referenz = kein Migrationshintergrund, 1 = türkischer Migrationshintergrund, 2 = nicht-türkischer Migrationshintergrund), Religion (Referenz = Christlich, 1 = keine Religion, 2 = Muslimisch, 3 = Buddhistisch /Hinduistisch) und Phänotyp (Referenz = weiß, 1 = schwarz, 2 = asiatisch) auf und kontrollieren für Geschlecht, Noten, Referenzschreiben, Vertragslaufzeit, Berufsgruppe und den Bewerbungsmonat. Somit berücksichtigen wir die Tatsache, dass wir mit einigen Berufen erst später ins Feld gegangen sind (für Details siehe Veit & Yemane, 2018), sowie etwaige Schwankungen auf dem Arbeitsmarkt.

In Spalte 1 von Tabelle 2 wird deutlich, dass die positive Rückmelderate für Bewerber mit ausländischem Namen um fünf Prozentpunkte geringer ausfällt als für Bewerber mit deutschen Namen. Für Bewerber mit türkischem Namen liegt der Unterschied sogar bei neun Prozentpunkten. Schwarze Bewerber haben eine im Vergleich zu weißen Bewerbern um sieben Prozentpunkte reduzierte Rückmelderate; und die Mitgliedschaft in einem muslimischen Sozialverein (im Vergleich zu einem christlichen Sozialverein) reduziert die positive Rückmelderate

ebenfalls um sieben Prozentpunkte. All diese Effekte sind statistisch signifikant, während die einzelnen Interaktionen zwischen Herkunftsland, Phänotyp und Religion statistisch unbedeutend sind. Das bedeutet, dass die berichteten Effekte additiv sind, mit den größten Nachteilen für schwarze muslimische Bewerber mit Migrationshintergrund (-.05 Prozent für Migranten, -.07 für schwarze Phänotypen und -.07 für Muslime).

In den Spalten 2-4 prüften wir die Argumente der *statistischen Diskriminierung* und versuchten aufzuklären, inwiefern durchschnittliche Gruppenunterschiede in der Produktivität die Unterschiede zwischen ethnischen Gruppen, verschiedenen Phänotypen und Religionszugehörigkeiten aufklären können. In Spalte 2 wird als Variable das durchschnittliche Bildungsniveau hinzugenommen, das einen signifikanten und positiven Einfluss auf die Rückmelderate hat. Der Koeffizient von .17 impliziert, dass ein Anstieg um eine Standardabweichung im durchschnittlichen Bildungsniveau (0.21 Punkte auf der Bildungsskala) zu einer 3,5 Prozent höheren Rücklaufquote für Angehörige dieser Gruppe führt. Dieser Befund bestätigt unsere erste Hypothese. Interessant hierbei ist, dass sich unter Berücksichtigung von Bildung der negative Effekt für Muslime von -.07 auf -.04 reduziert und nur noch marginal signifikant ist. Der nachteilige Effekt für Bewerber mit türkischem Namen reduziert sich ebenfalls von -.09 auf -.03 und verliert an Signifikanz. Allerdings bleiben die Nachteile für Bewerber mit ausländischem Namen unverändert; der Nachteil für schwarze Bewerber steigt sogar auf -.08. Auch wenn ein Großteil der Diskriminierung gegen Bewerber mit türkischem Namen durch die Berücksichtigung von durchschnittlichem Bildungsniveau aufgeklärt werden kann, bleiben die Nachteile für Muslime, Schwarze und nicht-türkische Migranten bestehen und können durch die Berücksichtigung von Bildung nicht aufgeklärt werden.

Tabelle 2: Multivariate Regression

	Modell 1: Baseline	Modell 2: Bildung	Modell 3: Berufl. Status	Modell 4: Erwerbs- losigkeit	Modell 5: Werte- distanz	Modell 6: Bildung + Werte- distanz
Ethnizität (Ref: deutschstämmig)						
Türkisch	-.09 (.02) ***	-.04 (.03)	-.01 (.03)	-.03 (.03)	-.00 (.03)	.01 (.03)
Andere (nicht-türkisch)	-.05 (.02) **	-.05 (.02) ***	-.02 (.02)	-.01 (.02)	.01 (.02)	-.00 (.03)
Phänotyp (Ref: Weiß)						
Schwarz	-.07 (.02) **	-.08 (.02) ***	-.03 (.02)	-.02 (.03)	-.03 (.02)	-.04 (.03)
Asiatisch	-.01 (.02)	-.03 (.02)	-.00 (.02)	-.01 (.02)	.01 (.02)	-.00 (.02)
Religion (Ref: Christlich)						
Keine Religion	-.02 (.02)	-.01 (.02)	-.01 (.02)	-.00 (.02)	-.00 (.02)	-.00 (.02)
Muslimisch	-.07 (.02) ***	-.04 (.02) +	-.03 (.02)	-.03 (.02)	-.03 (.02)	-.03 (.02)
Buddhistisch/Hinduistisch	-.02 (.03)	-.05 (.04)	-.03 (.03)	-.03 (.03)	.00 (.03)	-.01 (.04)
Gute Noten						
(Ref.: befriedigend)	.05 (.01) ***	.05 (.01) ***	.05 (.01) ***	.05(.01) ***	.05 (.01) ***	.05 (.01) ***
Referenzschreiben						
(Ref.: ohne)	-.02 (.02)	-.02 (.02)	-.02 (.02)	-.02 (.02)	-.02 (.02)	-.02 (.02)
Männlich						
(Ref.: weiblich)	-.06 (.02) **	-.06 (.02) **	-.06 (.02) **	-.06 (.02) **	-.06 (.02) **	-.06 (.02) **
GRUPPENMERKMALE						
Bildungsniveau	-	.17 (.06) *	-	-	-	.08 (.07)
Beruflicher Status (ISEI)	-	-	.02 (.00)***	-	-	-
Erwerbslosigkeit	-	-	-	-.02 (.01)***	-	-
Wertedistanz	-	-	-	-	-.05 (.01)***	-.04 (.02) *
Berufsgruppe	Ja	Ja	Ja	Ja	Ja	Ja
Monat	Ja	Ja	Ja	Ja	Ja	Ja
N (Gruppen)	35	35	35	35	35	35
N (Bewerbungen)	5,819	5,819	5,819	5,819	5,819	5,819
R ² (within)	.08	.08	.08	.08	.08	.08
R ² (between)	.47	.50	.57	.56	.56	.56
R ² (gesamt)	.09	.09	.09	.09	.09	.09

In der dritten Spalte wird deutlich, dass der Berufsstatus (ISEI) besser zur Aufklärung der Diskriminierung beiträgt. Der Koeffizient von .02 impliziert, dass ein Anstieg um eine Standardabweichung in der Berufsstatusvariablen (2.44

Punkte auf der ISEI-Skala) zu einem 4,4 Prozent höheren Rücklauf führt. Somit bleiben nach Kontrolle für Berufsstatus keine statistisch signifikanten Nachteile für Muslime, Schwarze und Menschen aus nicht westlichen Herkunftsregionen bestehen. Ähnlich verhält es sich, wenn man die Erwerbslosenquote (Spalte 4) betrachtet: Der negative Koeffizient von $-.02$ impliziert, dass ein Anstieg um eine Standardabweichung in der Erwerbslosenrate (2.4 %) zu 4,5 Prozent niedrigerem Rücklauf führt. Auch in diesem Modell bleiben keine Nachteile für Muslime, Schwarze und Migranten bestehen. Allerdings können Berufsstatus und Arbeitslosigkeit selbst schon Folge von Diskriminierung sein.

In der Spalte 5 prüften wir Hypothese 2 und somit das theoretische Argument, ob Wissen über abweichende kulturelle Werte und Normen mögliche Ursachen von *präferenzbasierter Diskriminierung* sind. Der negative Koeffizient von $-.05$ impliziert, dass ein Anstieg um eine Standardabweichung in der Wertedistanz vom deutschen Durchschnittswert (1 Punkt auf der standardisierten Wertedistanzskala) zu 5,1 Prozent niedrigerem Rücklauf führt. Unter Berücksichtigung von Wertedistanz zu Deutschland bleiben also keine statistisch signifikanten Nachteile für Muslime, Schwarze und Migranten bestehen.

In den sechs verschiedenen Modellen finden wir somit sowohl für statistische als auch für präferenzbasierte Diskriminierung Belege. Allerdings sind sowohl der Berufsstatus als auch die Arbeitslosenquote problematische Indikatoren, denn beide können auch das Produkt von Arbeitsmarktdiskriminierung sein. Es ist nicht ausgeschlossen, dass statistische Diskriminierung ein zirkulärer Prozess ist, bei dem Arbeitgeber die relative Position auf dem Arbeitsmarkt als Indikator für Produktivität nutzen und Bewerber mit durchschnittlich niedrigerem Status weniger häufig oder unter ihren Qualifikationen beschäftigen. Dies wiederum würde höhere Arbeitslosigkeit oder niedrigeren Berufsstatus zur Folge haben, was wiederum zu mehr statistischer Diskriminierung führen kann. Beide Indikatoren können gleichzeitig aber auch Ergebnisse von vorangegangener präferenzbasierter Diskriminierung sein, denn Berufsstatus und Arbeitslosenquote korrelieren deutlich stärker mit Wertedistanz ($r = -.90$ und $r = .81$; beide $p < .001$) als mit dem durchschnittlichem Bildungsniveau ($r = .70$ und $r = -.37$; $p < .001$). Da wir nicht eindeutig feststellen können, in welchem Maß der Arbeitsmarktstatus durch vorangegangene präferenzbasierte beziehungsweise statistische Diskriminierung

beeinflusst wurde, eignen sich weder der Berufsstatus noch die Arbeitslosenquote, um die zwei theoretischen Konzepte der statistischen und der präferenzbasierten Diskriminierung voneinander zu unterscheiden, das durchschnittliche Bildungsniveau ist hierfür deutlich besser geeignet.

In Modell 6 wird deshalb das durchschnittliche Bildungsniveau (statt Berufsstatus oder Arbeitslosenquote) gemeinsam mit Wertedistanz in das Modell aufgenommen. Die Ergebnisse zeigen, dass sich der Effekt für die Wertedistanz nur minimal reduziert und weiterhin signifikant bleibt ($-.42$; $p < .001$), während sich der Effekt für das Bildungsniveau halbiert und an statistischer Bedeutung verliert ($.08$, *n.s.*). Dies führt uns zu der Schlussfolgerung, dass kulturelle Präferenzen einen stärkeren Einfluss auf Diskriminierung haben als die statistische Diskriminierung aufgrund von Unterschieden im durchschnittlichen Bildungsniveau. Diese Befunde werden auch durch eine zusätzlich durchgeführte Mediationsanalyse bestätigt (nicht dargestellt).

Prüfungen zur Robustheit der Ergebnisse

Alle Analysen wurden zusätzlich auch als logistische Regressionen durchgeführt, es konnten jedoch keine substantiellen Unterschiede festgestellt werden (siehe Tabelle 2). Zudem haben wir die 34 nicht deutschen Herkunftsländer gewichtet, um unsere Ergebnisse auf mögliche Verzerrungen aufgrund der großen Stichprobe für türkische Namen und der geringeren Fallzahlen für die restlichen Herkunftsländer zu prüfen. Auch hier zeigen sich keine substantiellen Unterschiede. Weitergehende Analysen, bei denen man die Stichprobe getrennt auswertet nach Geschlecht (männliche versus weibliche Bewerber), nach Berufen mit hohem Kundenkontakt (Verkäufer/in, Hotelfachfrau/-mann, Anlagenmechaniker, medizinische und zahnmedizinische Fachangestellte) und niedrigem Kundenkontakt (Koch/Köchin, Industriekauffrau/-mann, Anlagemechaniker, Mechatroniker) oder nach Berufen mit höherem Bildungsabschluss (Industriekauffrau/-mann, Mechatroniker, Hotelfachfrau/-mann, medizinische Fachangestellte) und Berufen mit niedrigerem Schulabschluss (Koch/Köchin, Anlagemechaniker, Verkäufer/in, zahnmedizinische Fachangestellte), deuten alle in die gleiche Richtung: Im Gegensatz zum durchschnittlichen Bildungsniveau (das als Indikator für statistische

Diskriminierung gilt), kann die Wertedistanz (die als Indikator für präferenzbasierte Diskriminierung gilt) die Unterschiede zwischen den Gruppen aufklären und stellt somit die bessere Erklärung für unterschiedliche Diskriminierungstendenzen dar.

Zusammenfassung und Diskussion

Die Ergebnisse unserer Studien zeigen eindeutig, dass Menschen mit Migrationshintergrund in Deutschland bei der Suche nach einem Arbeitsplatz diskriminiert werden. Diese Diskriminierung unterscheidet sich jedoch stark zwischen Herkunftsländern, der Religionszugehörigkeit und dem Phänotyp. So werden vor allem Muslime und Menschen mit schwarzem Phänotyp diskriminiert.

Der statistischen Diskriminierungstheorie zufolge sollte sich die Diskriminierung verringern, wenn Arbeitgebern mehr produktivitätsrelevante Information über einen Bewerber zur Verfügung steht, da sie dadurch die Eignung für die Stelle besser einschätzen können und sich nicht auf Gruppendurchschnitte berufen müssen. Auch wenn Arbeitgeber in unserer Studie auf einzelne produktivitätsrelevante Indikatoren wie gute Noten reagieren, lässt sich mit Blick auf die deskriptiven Befunde festhalten, dass Migranten im Vergleich zu deutschen Bewerbern keinen unmittelbaren Nutzen aus besseren Noten, guten Referenzen oder dem Signal eines unbefristeten Arbeitsvertrags ziehen. Ein zentraler Befund ist außerdem, dass die Diskriminierung besser durch durchschnittliche Unterschiede in Werten als durch Gruppenunterschiede im Bildungsniveau erklärt werden kann – was einen Hinweis auf präferenzbasierte Diskriminierung ist. Das erklärt auch, warum insbesondere Bewerber mit Ursprüngen in west- und südeuropäischen Ländern, aber auch ostasiatischen Ländern so hohe positive Rückmelderaten haben: Hier ist die Wertedistanz zu Deutschland besonders niedrig. Für Bewerber mit Ursprüngen in afrikanischen und überwiegend muslimisch geprägten Ländern, die wiederum eine sehr hohe Wertedistanz zu Deutschland aufweisen, spiegelt sich diese Distanz auch in den besonders niedrigen Rückmelderaten wieder. Im Hinblick auf die eingangs erwähnten ökonomischen Theorien finden wir in unserer Studie also mehr Belege für präferenzbasierte Diskriminierung als für statistische Diskriminierung.

Unsere Ergebnisse zeigen, dass Arbeitgeber starke Annahmen über die Unterscheide in Werte- und Normvorstellungen einzelner Migrantengruppen haben und sich dies auch in ihrem Verhalten widerspiegelt. So bevorzugen Arbeitgeber diejenigen Bewerber, die ihnen am ähnlichsten sind, also ähnliche Werte teilen wie die Mehrheitsbevölkerung in Deutschland. Dieser Befund deckt sich mit den Aussagen von Adida, Laitin und Valfort (2016), die in ihrer Studie gezeigt haben, dass Arbeitgeber Unterschiede in Wertevorstellungen als Gründe für ihr diskriminierendes Verhalten gegenüber Muslimen angeben. Die Frage, ob und inwieweit Arbeitgeber sich auf Unterschiede in den Wertvorstellungen berufen, um ihre diskriminierenden Präferenzen zu verschleiern, oder ob sie tatsächliche Produktivitätseinbußen durch kulturelle Differenzen innerhalb ihrer Belegschaft befürchten und ob es solche Produktivitätsnachteile gibt, kann aufgrund unseres Studiendesigns nicht beantwortet werden. Zukünftige Studien sollten diese Fragen mit alternativen Forschungsansätzen genauer beleuchten.

Literaturverzeichnis

- Adida, Claire L., David D. Laitin, und Marie-Anne Valfort. 2016. *Why Muslim Integration Fails in Christian-Heritage Societies*. Cambridge: Harvard University Press.
- Agerström, Jens, Fredrik Björklund, Rickard Carlsson, und Dan-Olof Rooth. 2012. "Warm and Competent Hassan = Cold and Incompetent Eric: A Harsh Equation of Real-Life Hiring Discrimination." *Basic and Applied Social Psychology* 34(4): 359-66.
- Ahmed, Ali M., Lina Andersson, und Mats Hammarstedt. 2013. "Are Gay Men and Lesbians Discriminated against in the Hiring Process?" *Southern Economic Journal* 79 (3): 565-85.
- Aigner, Dennis J., und Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30 (2): 175-187.
- Arrow, Kenneth J. 1973. "The theory of discrimination." *Discrimination in labor markets* 3(10): 3-33.
- Arrow, Kenneth J. 1998. "What Has Economics to Say About Racial Discrimination?" *Journal of Economic Perspectives*. 12(2): 91-100. <https://doi.org/10.1257/jep.12.2.91>
- Auspurg, Katrin, Thomas Hinz, und Laura Schmid. 2017. "Contexts and Conditions of Ethnic Discrimination: Evidence from a Field Experiment in a German Housing Market." *Journal of Housing Economics* 35: 26-36.
- Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, und Filip Matějka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *The American Economic Review* 106(6): 1437-1475.
- Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: Univ. of Chicago Press.
- Bertrand, Marianne, und Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*: 94: 991-1013.

- Blommaert, Lieselotte, Marcel Coenders, and Frank Van Tubergen. 2013. "Discrimination of Arabic-Named Applicants in the Netherlands: An Internet-Based Field Experiment Examining Different Phases in Online Recruitment Procedures." *Social Forces* 92 (3): 957–982.
- Burkert, Carola, and Holger Seibert. 2007. "Labour market outcomes after vocational training in Germany - equal opportunities for migrants and natives?" *IAB discussion paper*, 2007 (31). Retrieved from <http://hdl.handle.net/10419/32679>
- Drydakis, Nick. 2014. "Sexual Orientation Discrimination in the Cypriot Labour Market. Distastes or Uncertainty?" *International Journal of Manpower* 35(5): 720–44.
- Edo, Anthony, Nicolas Jacquemet, and Constantine Yannelis. 2013. "Language Skills and Homophilous Hiring Discrimination: Evidence from Gender- and Racially-Differentiated Applications." *CES Working Papers* 2013-58. Paris: CES.
- EVS. 2015. "European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)." GESIS Data Archive, Cologne. ZA4804 Data file Version 3.0.0. doi:10.4232/1.12253.
- Gaddis, S. Michael. 2015. "Discrimination in the Credential Society: An Audit Study of Race and College Selectivity in the Labor Market." *Social Forces* 93(4): 1451–79.
- Inglehart, Ronald, and Pippa Norris. 2003. The True Clash of Civilizations. *Foreign Policy* 135: 62–70.
- Kaas, Leo, and Christian Manger. 2012. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment." *German Economic Review* 13 (1): 1–20.
- Kogan, Irena. 2004. "Last Hired, First Fired? The Unemployment Dynamics of Male Immigrants in Germany." *European Sociological Review* 20(5): 445–461. <https://doi.org/10.1093/esr/jch037>
- McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27 (January): 415–44.
- Oreopoulos, Philip. 2011. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3 (4): 148–171.

- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62 (4): 659–61.
- Schneider, Jan, Ruta Yemane, und Martin Weinmann. 2014. *Diskriminierung am Ausbildungsmarkt. Ausmaß, Ursachen und Handlungsperspektiven*. Berlin: Sachverständigenrat Deutscher Stiftungen für Integration und Migration.
- Uhlendorff, Arne, und Klaus F. Zimmermann. 2006. "Unemployment Dynamics among Migrants and Natives." *IZA Discussion Papers*, No. 2299. Retrieved from <http://hdl.handle.net/10419/33769>
- Wallace, Michael, Bradley RE Wright, und Allen Hyde. 2014. "Religious Affiliation and Hiring Discrimination in the American South: A Field Experiment." *Social Currents* 1 (2): 189–207.
- Weichselbaumer, Doris. 2015. "Testing for Discrimination against Lesbians of Different Marital Status: A Field Experiment." *Industrial Relations: A Journal of Economy and Society* 54 (1): 131–61.
- Weichselbaumer, Doris. 2016. "Discrimination against Female Migrants Wearing Headscarves." *IZA Discussion Papers* No. 10217.
- WVS. 2015. "World Value Survey 1981–2014 Official Aggregate v.20150418, 2015. World Values Survey Association (www.worldvaluessurvey.org)." Aggregate File Producer: JDSystems, Madrid.
- Veit, Susanne, und Ruta Yemane. 2018. "The ADIS study: A large-scale correspondence test on labor market discrimination in Germany: Technical Report". WZB Discussion Paper SP VI 2018-103.

Anhang

Tabelle A1. Herkunftsländer und Namen

Herkunftsländer	Männliche Vornamen	Weibliche Vornamen	Familiennamen
Albanien	Arben	Mergime	Dukagjini
Ägypten	Karim	Dina	Hassan
Äthiopien	Mekbib	Abeba	Yerga
Bosnien Herzegowina	Ajdin	Belma	Kovačević
Bulgarien	Zlatan	Nevena	Dimitrov/a
China	Fu-Han	Ying	Lee
Deutschland	Florian	Lena	Müller
Dominikanische Republik	Pablo	Marisol	Martínez
Frankreich	Guillaume	Claire	Durand
Griechenland	Giorgos	Konstantina	Papadopoulos/ou
Indien	Aarav	Priya	Singh
Indonesien	Dian	Putri	Bitang
Irak	Kathem	Rana	Hassan
Iran	Farhad	Soraya	Pahlavi
Italien	Francesco	Valentina	Marino
Japan	Hiroto	Asuka	Sato
Malaysia	Aqil	Adilah	Bitang
Marokko	Mehdi	Karima	Hassan
Mazedonien	Dragomir	Vesna	Dimitrov/a
Mexiko	Pedro	Lupita	Martínez
Niederlande	Jeroen	Maike	De Vries
Nigeria	Amadi	Monifa	Achebe
Pakistan	Faisal	Shumaila	Bugti
Polen	Marek	Michalina	Wiśniewski/ka
Rumänien	Andrei	Dana	Popescu
Russland	Sergej	Olga	Dimitrov/a
Schweiz	Urs	Ursina	Mueller
Spanien	Alvaro	Alma	Martínez
Südkorea	Ji-Hun	Su-Min	Lee
Trinidad und Tobago	Toriano	Onika	Wilson
Türkei	Onur	Sevda	Yilmaz
Uganda	Wemusa	Kisakye	Ndikumana
UK	James	Fiona	Wilson
USA	Matthew	Ashley	Wilson
Vietnam	Danh	Linh	Nguyen

Tabelle A2. Zuweisungshäufigkeit von Phänotyp und Religion nach Ethnizität

ETHNIZITÄT	PHÄNOTYP			RELIGION				Σ
	Weiß	Schwarz	Asiatisch	Christl.	Muslim.	Buddh./Hindu.	keine	
Albaner	69		18	40	15		32	87
Ägypter	40	55		38	32		25	95
Äthiopier		77		24	24		29	77
Bosnier	91			37	25		29	91
Briten	35	25	28	35			53	88
Bulgaren	116			49	30		37	116
Chinesen			115	41		40 B	34	115
Deutsche	1,313	214	179	871			835	1,706
Dominikaner	25	58		46			37	83
Franzosen	35	42	36	74			39	113
Griechen	77			39			38	77
Inder	58		41		34	33 H	32	99
Indonesier			72	23	22		27	72
Iraker	36		31		33		34	67
Iraner	46		27		40		33	73
Italiener	73			33			40	73
Japaner			85			37 B	48	85
Makedonier	76		20	41	30		25	96
Malaien			71		23	32 B	16	71
Marokkaner	16	29			21		24	45
Mexikaner	46		11	33			24	57
Niederländer	70			32			38	70
Nigerianer		125		46	38		41	125
Pakistaner	34		30		32		32	64
Polen	83			39			44	83
Rumänen	43		22	38			27	65
Russen	105			43	33		29	105
Schweizer	66			30			36	66
Spanier	53		25	41			37	78
Südkoreaner			129	45		48 B	36	129
Trinidadier		130		39		42 H	49	130
Türken	785		428		645		568	1,213
Ugander		93		33	25		35	93
US	36	28	31	56			39	95
Vietnamesen			117	29		39 B	49	117
Σ	3,427	876	1,516	1,895	1,102	271	2,551	5,819

Anmerkung: Die Zuordnung von Phänotypen zu Herkunftsländern basiert auf dem Ergebnis von Vorstudien; die Zuordnung von Religionen zu Herkunftsländern auf den jeweils größten Religionsgruppen in den Herkunftsländern (siehe Veit & Yemane 2018).

Tabelle A3. Mittleres Bildungsniveau, Erwerbslosigkeit, ISEI und Wertedistanz zu Deutschland nach Ethnizität

ETHNIZITÄT	Mittleres Bildungsniveau	Mittlere Erwerbslosigkeit	Mittlerer ISEI	Mittlere Wertedistanz
Albaner	2.06	.16	13.82	1.17
Ägypter	1.87	.19	15.56	1.84
Äthiopier	1.86	.15	15.44	1.29
Bosnier	1.51	.07	16.06	1.30
Briten	2.37	.05	27.31	.29
Bulgaren	2.13	.13	18.96	.57
Chinesen	2.61	.07	24.74	.35
Deutsche	1.84	.05	22.68	0
Dominikaner	2.26	.06	19.17	1.03
Franzosen	2.38	.05	24.61	.07
Griechen	1.47	.10	17.20	1.39
Inder	1.92	.07	19.80	.86
Indonesier	2.07	.08	17.55	1.98
Iraker	1.33	.22	9.088	1.88
Iraner	2.50	.10	20.24	1.83
Italiener	1.39	.07	18.03	1.01
Japaner	2.80	.05	23.34	.28
Makedonier	1.31	.11	12.86	1.39
Malaien	2.07	.08	17.55	1.74
Marokkaner	1.53	.13	12.23	2.52
Mexikaner	2.26	.10	19.17	1.07
Niederländer	2.06	.03	22.33	.04
Nigerianer	1.63	.18	9.84	2.03
Pakistaner	1.72	.11	12.65	1.94
Polen	1.90	.06	19.95	1.13
Rumänen	1.94	.05	20.16	1.49
Russen	1.82	.09	17.55	.76
Schweizer	2.32	.03	27.27	.03
Spanier	1.80	.07	19.95	.17
Südkoreaner	2.80	.05	23.34	.46
Trinidadier	2.26	.10	19.17	1.79
Türken	1.18	.12	12.14	2.20
Ugander	1.86	.153	15.44	2.22
US Amerikaner	2.48	.059	22.57	.84
Vietnamesen	1.84	.125	14.54	1.88

Tabelle A4. Bewerbungen nach Berufsgruppe und Geschlecht

BERUFSGRUPPE	WEIBLICH	MÄNNLICH	Σ
Koch/Köchin	408	441	849
Industriekaufmann/frau	287	319	606
Hotelfachmann/frau	235	274	509
Verkäufer/in	280	317	597
Anlagenmechaniker-SHK	0	961	961
Mechatroniker	0	929	929
Zahnmedizinische	566	0	566
Medizinische Fachangestellte	802	0	802
Σ	2,578	3,241	5,819