The ADIS study: a large-scale correspondence test on labor market discrimination in Germany - Technical Report

Veit, Susanne; Yemane, Ruta

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Veit, Susanne; Yemane, Ruta

Working Paper

The ADIS study: A large-scale correspondence test on labor market discrimination in Germany - Technical Report

WZB Discussion Paper, No. SP VI 2018-103

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Veit, Susanne; Yemane, Ruta (2018) : The ADIS study: A large-scale correspondence test on labor market discrimination in Germany - Technical Report, WZB Discussion Paper, No. SP VI 2018-103, Wissenschaftszentrum Berlin für Sozialforschung (WZB), Berlin

This Version is available at:
http://hdl.handle.net/10419/178706

WZB
Berlin Social Science Center

Susanne Veit
Ruta Yemane

The ADIS study: A large-scale correspondence test on labor market discrimination in Germany – Technical Report

Discussion Paper

SP VI 2018-103

May 2018

Research Area
Migration and Diversity

Research Unit
Migration, Integration, Transnationalization

Susanne Veit, Ruta Yemane
The ADIS study:
A large-scale correspondence test on labor market discrimination in Germany – Technical Report
Discussion Paper SP VI 2018-103
Wissenschaftszentrum Berlin für Sozialforschung (2018)

Affiliation of the authors

Susanne Veit
WZB Berlin Social Science Center
E-mail: susanne.veit@wzb.eu

Ruta Yemane
WZB Berlin Social Science Center
E-mail: ruta.yemane@wzb.eu

The ADIS study: A large-scale correspondence test on labor market discrimination in Germany – Technical Report

Susanne Veit, Ruta Yemane

Zusammenfassung

Das Arbeitsmarktdiskriminierung "ADIS" Projekt ist eine großangelegte Korrespondenzteststudie, die deutschlandweit zwischen Oktober 2014 und April 2016 durchgeführt wurde. Während in der Vergangenheit mit Hilfe von Korrespondenzteststudien bereits eindrücklich nachgewiesen werden konnte, dass Bewerber mit Migrationshintergrund auf dem deutschen Arbeitsmarkt benachteiligt werden, geben diese Studien wenig Einblick in die zugrundeliegenden Ursachen und Mechanismen. Ziel der Studie war es daher, das Auftreten und das Ausmaß von Diskriminierung gegenüber qualifizierten Migranten zweiter Generation zu untersuchen und hierbei insbesondere zwei Fragestellungen tiefergehend zu erforschen: Erstens sollen zugrundeliegende Mechanismen ethnischer Diskriminierung auf dem Arbeitsmarkt betrachtet werden. Hierbei wird die Rolle von Faktoren wie ethnischer Gruppe, Religion oder Phänotyp beleuchtet, die eventuell ethnische Hierarchien erklären können. Zum anderen sollen Annahmen zweier unterschiedlicher ökonomischer Erklärungsmodelle für diskriminierende Entscheidungen (präferenzbasierte und statistische Diskriminierung) untersucht werden.

Für die ADIS-Studie nutzen wir ein multidimensionales Forschungsdesign und messen das Ausmaß von Diskriminierung für insgesamt 35 verschiedene ethnische Gruppen, für die sowohl der Phänotyp und die Religionszugehörigkeit, als auch produktivitätsrelevante Indikatoren wie Zeugnisnoten und Referenzschreiben variiert werden. Die Studie basiert auf rund 7,557 Beobachtungen von männlichen und weiblichen Kandidaten, die sich deutschlandweit auf insgesamt acht verschiedene Berufe beworben haben. Im folgenden Bericht diskutieren wir das Forschungsdesign und geben eine detaillierte Beschreibung der Studienumsetzung sowie der Herausforderungen bei der Datenerhebung und der Operationalisierung der einzelnen Treatments.

Abstract

The Arbeitsmarktdiskriminierung "ADIS" project is a large scale correspondence study that was conducted in Germany between October 2014 and April 2016 with the aim to study labor market discrimination against second generation immigrants. In particular, the experiment was designed for two purposes: First, to decompose drivers of ethnic discrimination in the labor market and to understand the role of phenotype, country of origin, and religion in order to explain ethnic hierarchies; and second, to test economic theories of taste-based and statistical discrimination models, in order to explain discriminatory hiring decisions. Previous studies on ethnic labor market discrimination have provided strong evidence of its occurrence, but provide limited insight into the mechanisms behind it and the causes of variation in rates of discrimination against different ethnic, racial, and religious groups. In this report we describe our multidimensional design that tests rates of discrimination across 35 ethnic groups, for which we vary productivity relevant information (such as grades and reference letters) as well as phenotype and religious background of the applicants. The study is based on applications to 7,557 job vacancies with male and female applicants in eight professions across Germany. In this technical report we will discuss our research design in depth and give detailed insight into the implementation of the study and the challenges during data collection, with a main focus on the choice of the individual treatments and how they were operationalized.

Table of contents

List of figures

List of tables

1. Research questions and theoretical background

Ethnic discrimination is normatively undesirable and a source of social divisiveness. Empirical studies on labor market integration have shown that immigrants run a higher risk of being unemployed than natives, even after controlling for human capital characteristics (Uhlendorff & Zimmermann 2006; Kogan 2004; Burkert & Seibert 2007). However, it is difficult to ascertain on the basis of cross-sectional survey or register data whether or not the observed 'ethnic penalty' results from relevant characteristics that were not taken into account (unobserved heterogeneity), for example different career aspirations and differences in application behaviors between natives and immigrants, or from the discrimination against immigrant job applicants. Therefore, experimental studies like audit and correspondence studies are more suitable to detect labor market discrimination and to provide causal evidence on ethnic discrimination in hiring decisions. One experimental approach are audit studies in which pairs of trained actors attend real-world selection situations such as job interviews or flat viewings (Pager 2007, Pager et al. 2009). To detect discrimination, researchers compare success rates between minority and majority candidates. These audit studies have provided strong evidence of discrimination in different contexts, but are very costly and difficult to implement.

Therefore, in the past decades correspondence studies have become the gold standard and a widely used method to investigate discrimination in the labor and housing market (Carlsson & Rooth 2007, Ahmed & Hammarstedt 2008, Oreopoulos 2012, Wood et al. 2009). In correspondence studies, the researcher sends out comparable applications to real job openings and only varies the characteristics of interest (e.g., gender, ethnicity). Differences in callback rates provide causal evidence on unequal treatment. In contrast to audit studies with human actors, the researcher has full control over the applicants' characteristics and behaviors (Heckman & Siegelmann 1993; Heckman 1998). Furthermore, with the rise of online applications, especially in the housing and labor market, a researcher can send out thousands of applications at comparably low cost, both in terms of financial resources and time.

In Germany, Kaas and Manger (2011) conducted a correspondence study on ethnic discrimination in the internship market for students of economics who were top of their class. They sent two similar applications to 528 ads for student internships, one with a Turkish and one with a German name. The callback rate for candidates with German names was 40 percent, whereas the callback rate for applicants with Turkish names was 35 percent. Differential treatment was particularly strong and significant in smaller firms (less than 50 employees), where the difference between the German and the Turkish callback rate was 18 percentage points.

Likewise, Schneider, Yemane, and Weinmann (2014) conducted a correspondence test and sent 3,588 applications to 1,794 companies with vocational training vacancies for

office administrators or motor vehicle mechatronics. Each company received two applications, one from a candidate with a German name and the other from one with a Turkish name. In 20 % of the cases the applicant with a German name was invited while this was true in only 15 % of the cases for applicants with a Turkish name. The gap in callback rates was statistically significant for applicants who applied for motor vehicle mechatronic jobs, but not for office administrators. In two focus groups, employers were asked about their recruitment strategies. Their statements suggest that besides the formal qualification of the candidate, criteria such as customer tastes and personal preferences as well as in-group favoritism and risk avoidance play a major role in the recruitment process.

These studies show evidence of discrimination in the labor market. However, they explain very little about the actual mechanisms that are underlying an employer's decision to discriminate against immigrant applicants and the interplay of different aspects of ethnicity (e.g., language and phenotype, but also cultural and socioeconomic characteristics of the country of origin). Therefore, in addition to examining whether immigrants are discriminated against when applying for jobs, we need to better understand the underlying drivers of ethnic discrimination in order to tackle it effectively and improve minorities' chances on the labor market.

In economics, there are two influential theories on ethnic labor market discrimination: theories on taste-based discrimination and theories on statistical discrimination. Hiring situations are always fraught with uncertainty and risk, as employers usually have to judge an individual's productivity based on the application material provided. Application documents, however, provide only imperfect information about the applicants' true productivity. According to statistical discrimination theories, employers therefore rely on group averages in productivity when judging individual candidates. That is, they use observable characteristics like ethnic background or gender to make predictions about unobservable characteristics, which in fact may be related to differences in productivity (Arrow 1998, Phelps 1972).

In contrast, taste-based discrimination is solely based on an animus against certain groups where individuals or groups are "willing to pay" (Becker 1957) so as not to employ, work with, or buy from these groups. Taste-based discrimination is practiced by employers who refuse to hire certain groups, by employees (co-workers) who refuse to work with certain groups, or by consumers who refuse to buy from certain groups. That means that even when an employer has no personal dislike of a group, his or her discriminatory behavior may still be driven by preferences, namely the assumed preferences of employees or customers (Becker 1957).

The empirical proof of whether or not discriminatory decisions are driven by dislike (taste-based discrimination) or by concerns about productivity (statistical discrimination), however, is very difficult to obtain. Negative expectations of a group's

average level of performance, for example, may themselves result from a dislike of this group and from associated negative group stereotypes. In addition, most previous correspondence tests varied only a limited number of experimental treatments and typically tested discrimination only for one specific immigrant group. In Germany, for example, previous studies compared invitation rates between candidates with German- and Turkish-sounding names, which made it impossible to carve out the roles of group stereotypes and average group characteristics.

To avoid these shortcomings and to go beyond previous studies, we employed an unpaired research design that 1) investigates discrimination across a large number of ethnic groups that differ from one another with regard to cultural and socioeconomic characteristics of the immigrant group as well as with regard to countries of origin and 2) tests the joint impact of several characteristics of candidates, some of which are linked to ethnic origin and some of which are independent of ethnic origin. In particular, the study seeks to advance our understanding of two key issues:

- Ethnic hierarchies: The extent and causes of variation in discrimination across ethnic groups.
- Taste-based vs. statistical discrimination: The relative importance of ethnic preferences and dislikes versus risk aversion and uncertainty about candidates' true ability in explaining ethnic discrimination in hiring decisions.

2. Research design

Based on the two criteria mentioned above we came up with a unique research design and conducted our large-scale correspondence study between October 2014 and April 2016. Prior to this, we ran a pilot study from September 2013 to April 2014 and adjusted our research design based on issues and open questions that arose during the pilot. After six months of solving these issues and fine tuning the research design, we were able to run our main study for which we send out 7,588 applications in total. To provide causal evidence on ethnic discrimination and to be able to disentangle statistical from taste-based mechanisms, we designed a correspondence test in which we sent applications from fictitious job candidates to vacancies in eight occupations advertised all over Germany. In addition to varying applicants' ethnic background, we varied several other ethno-cultural and performance-related characteristics which will be introduced in this chapter.

In contrast to most previous correspondence tests that only look at a few ethnic groups, we compared responses between native German applicants and 34 different ethnic groups (2nd generation). In addition, to accommodate all these groups and the large number of treatments and to avoid detection of the experiment, we applied an unpaired

design, meaning that we only sent one application per vacancy. To implement our complex experimental design, we hired a programmer to create software that helped us to smoothly run the experiment. The software automatically searched for jobs on the website of the German employment agency and selected job ads that fit our criteria, assigned experimental treatments to applications, matched applications to job ads, automatically sent out E-mails, and allowed us to collect and store all responses from employers. In the following paragraphs, we provide more detailed information about the different building blocks of our research design.

2.1. Occupations

We tested for discrimination in eight occupations (see Table 1 and Appendix A1. for occupation profiles). All occupations are part of the German dual apprenticeship system. After completion of nine or ten years of general schooling, vocational education in a vocational school is combined with a practical apprenticeship in a company. This type of vocational training lasts for a period between 2 to 3.5 years, depending on the occupation. Students graduate after completing a final theoretical and practical exam and receive a diploma (*Gesellenbrief* or *IHK–Zeugnis*) after successful completion. A copy of this diploma is expected to be added to a formal application.

The eight occupations we chose (Table 1) belong to different occupational fields, ranging from the technical sector to the service, commercial, and health sector. Within each occupational field, we selected one occupation typically chosen by students who graduate from a lower secondary school (Hauptschule) after 9 years of education and one occupation mainly chosen by graduates from an intermediate secondary school (Realschule) after 10 years of education.

The occupations were chosen based on several criteria:

- number of vacancies per month (target goal was 1000 vacancies per occupation throughout the experiment),
- popularity among (2nd generation) immigrants and natives,
- level of customer contact

Table 1: Occupations and characteristics

Occupational field	Sex	Occupation	Customer contact	School leaving certificate
Technical	male	**Mechatronics fitter**	low	Medium secondary (10 years)
		Plant mechanic for sanitary, heating, ventilation & air conditioning systems	low	Lower secondary (9 years)
Services	male & female	**Hotel receptionist**	high	Medium secondary (10 years)
		Cook	low	Lower secondary (9 years)
Commerce		**Sales assistant**	high	Lower secondary (9 years)
		Industrial clerk	low	Medium secondary (10 years)
Health	female	**Medical assistant**	high	Medium secondary (10 years)
		Dental assistant	high	Lower secondary (9 years)

We made use of the fact that certain occupations are largely dominated by males or females and others are more equally distributed across gender. That is why we selected two male dominated occupations (mechatronics fitter & plant mechanic for sanitary, heating, ventilation & air conditioning systems), two female dominated (medical assistant & dental assistant), and four gender-balanced occupations (hotel receptionist, cook, sales assistant & industrial clerk).

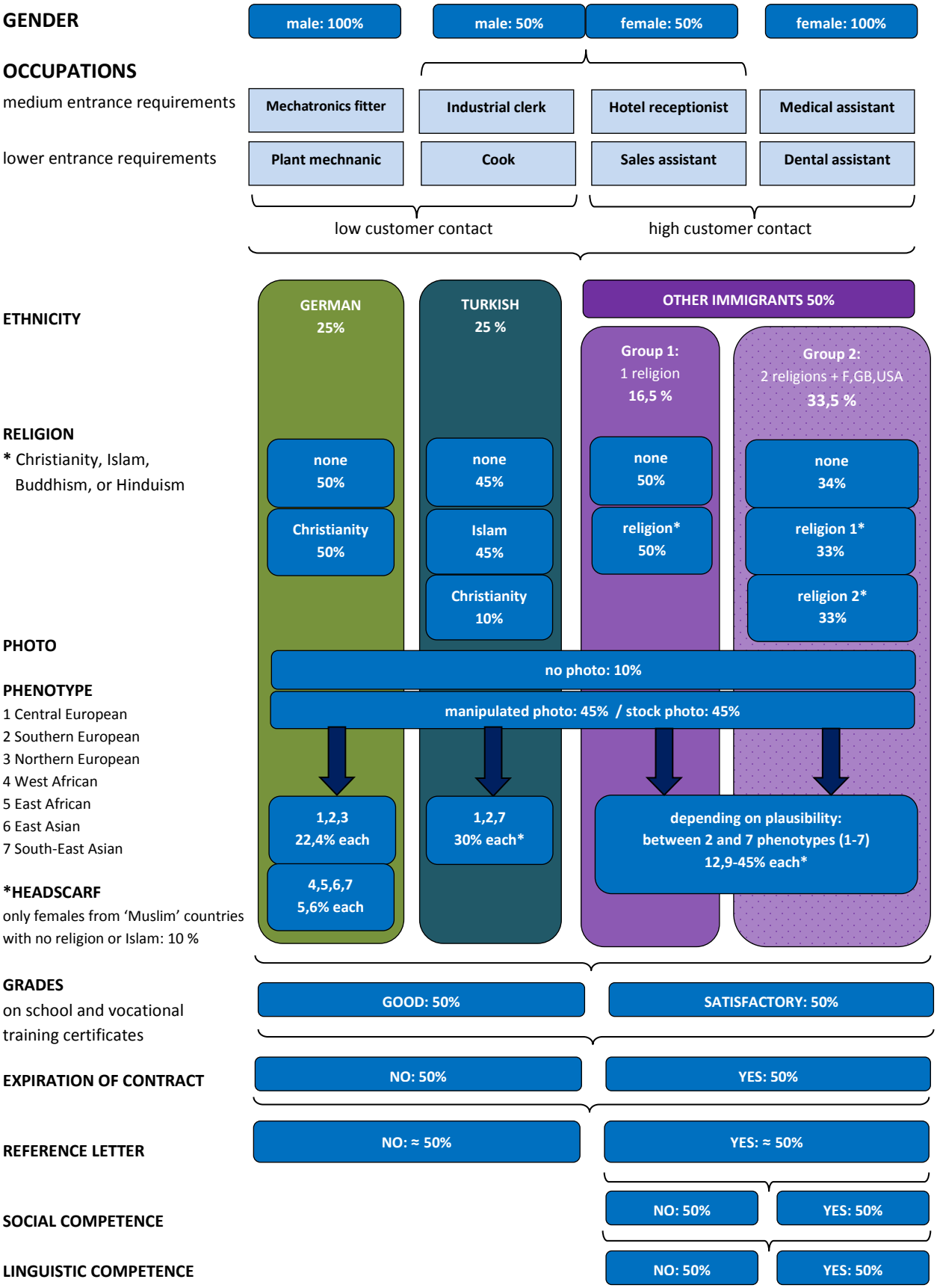
2.2. Treatments

We sent out applications for all eight occupations to vacancies advertised all over Germany. To investigate the drivers of discrimination we varied several characteristics of our fictitious candidates in the application documents. Most of the treatments were orthogonal, meaning that their assignment to profiles was random. For some treatments, assignment was fully random (e.g., expiration of contract or grades) and for others assignments, probabilities were determined by quota (e.g., ethnic group of the applicant). Other treatments, however, were non-orthogonal. Among the non-orthogonal treatments is gender, because in four occupations we applied with either only male or female profiles (see sections 2.1. and 2.2.1.). The other two non-orthogonal treatments are phenotype and religion (see sections 2.2.3. and 2.2.4.). For reasons of plausibility, we restricted the range of phenotypes that could be randomly assigned as

well as the religious affiliation an applicant could signal, depending on applicants' ethnicity.

In the following paragraphs, we provide short descriptions of all experimental treatments, and in Appendix A1. we added two examples of our experimental material (i.e., cover letters, CVs, and reference letters). Figure 1 provides a complete overview of all treatments and their probabilities.

Figure 1: Treatment overview with probabilities



2.2.1. Gender

We applied with male and female profiles. Gender was made salient in three ways: by applicants' names, by using the German gender-specific occupational designation in the subject line, and by the photo of the applicant, which was added to the CV in 90% of the cases.

Half of the total sample consisted of male and half of female applicants, but applicants' gender was varied only in applications to four of our eight occupations (see section 2.1. and Figure 1). Because of the gender distributions among employees in the remaining four occupations, we applied either only with female or only with male profiles.

2.2.2. Ethnicity and names

To test for ethnic discrimination, we experimentally varied the ethnic background of our fictitious job candidates. Ethnicity was signaled through applicants' name and language skills. In the skills section of the CV, we explicitly signaled immigrants' country of origin by indicating in addition to German mother tongue a second mother tongue, e.g., "Russian (mother tongue)", "Luganda (Uganda, mother tongue)", or "English (USA, mother tongue)".

In total, we applied with 35 different ethnic profiles. We chose countries of origin based on several criteria: First, we chose the most important immigrant groups in Germany by size of the respective immigrant population (Turks, Bosnians, Poles, Russians, and Italians). Second, we took into account the distribution of religious groups and favored countries with two major religions over mono-religious countries. Third, we included a set of countries with sufficient phenotypical heterogeneity. Fourth, we took into account the availability of register data, additional data sets and additional information about country characteristics (e.g., the German census, the World Values Survey Database (WVS), and the European Values Survey Database (EVS)). Finally, we wanted to include countries from as many regions of the world as possible.

Table 2: Names by ethnic group

Nr.	Ethnicity	Male names	Female names	Male/Female family names
1	**Albania**	Arben	Mergime	Dukagjini
2	**Bosnia and Herzegovina**	Ajdin	Belma	Kovačević
3	**Bulgaria**	Zlatan	Nevena	Dimitrov/a
4	**China**	Fu-Han	Ying	Lee
5	**Dominican Republic**	Pablo	Marisol	Martínez
6	**Egypt**	Karim	Dina	Hassan
7	**Ethiopia**	Mekbib	Abeba	Yerga
8	**France**	Guillaume	Claire	Durand
9	**Germany**	Florian	Lena	Müller
10	**Greece**	Giorgos	Konstantina	Papadopoulos/ou
11	**India**	Aarav	Priya	Singh
12	**Indonesia**	Dian	Putri	Bitang
13	**Iran**	Farhad	Soraya	Pahlavi
14	**Iraq**	Kathem	Rana	Hassan
15	**Italy**	Francesco	Valentina	Marino
16	**Japan**	Hiroto	Asuka	Sato
17	**Macedonia**	Dragomir	Vesna	Dimitrov/a
18	**Malaysia**	Aqil	Adilah	Bitang
19	**Mexiko**	Pedro	Lupita	Martínez
20	**Morocco**	Mehdi	Karima	Hassan
21	**Netherlands**	Jeroen	Maaike	De Vries
22	**Nigeria**	Amadi	Monifa	Achebe
23	**Pakistan**	Faisal	Shumaila	Bugti
24	**Poland**	Marek	Michalina	Wiśniewski/ka
25	**Rumania**	Andrei	Dana	Popescu
26	**Russia**	Sergej	Olga	Dimitrov/a
27	**South Korea**	Ji-Hun	Su-Min	Lee
28	**Spain**	Alvaro	Alma	Martínez
29	**Swiss**	Urs	Ursina	Mueller
30	**Trinidad and Tobago**	Toriano	Onika	Wilson
31	**Turkey**	Onur	Sevda	Yilmaz
32	**Uganda**	Wemusa	Kisakye	Ndikumana
33	**United Kingdom**	James	Fiona	Wilson
34	**USA**	Matthew	Ashley	Wilson
35	**Vietnam**	Danh	Linh	Nguyen

All applicants were born and raised in Germany and had German citizenship. To make ethnic origin salient, we chose distinct first names for our fictitious applicants (see Table 2). In order to find suitable first names for each country of origin, we conducted an online name search on websites of national name registers. For each country of origin, we chose the first names with the highest frequency in the birth year of our applicants (1992). In countries where such data were not available, we chose names that were frequently mentioned on internet websites that listed popular names. We carefully looked for names that could be considered neutral and had no religious connotations. Moreover, we avoided names of popular figures. Finally, we looked for names that are fairly easy to pronounce for German employers, so that we would not have biases in callback rates because employers would be hesitant to call an applicant whose name they wouldn't know how to pronounce. We chose family names based on the same procedure, except that we preferred family names that were common in as many countries of origin as possible. We did so because even though we had 70 different applicants in total, we only used two mail addresses in Bielefeld and Paderborn (see section 2.3). In order to reduce the number of family names that had to be registered under both addresses, we used the same family names (e.g. Lee in China and South Korea; Hassan in Iraq, Egypt and Morocco) where applicable.

According to our sampling plan, 25% of our applicants were German natives. Another 25% of our applicants were described as Germans with Turkish background (i.e., 2nd generation immigrants of Turkish origin). We chose to oversample Turkish names, as Turks are the biggest immigrant group in Germany and also because we wanted to compare our results with those of previous studies in Germany, which all used Turkish names.

We divided the remaining 33 countries of origin into two groups with different assignment probabilities (Group 1: 16.5% vs. Group 2: 33.5%). Group 1 represents countries with a single dominant religion (see Table 3). Each of the 14 ethnicities in Group 1 had an assignment probability of about 1.2%. Group 2 includes countries with two dominant religions as well as the United States, France, and the United Kingdom. Since the latter three countries all have a long standing immigration history and substantial colonial heritage, we were able to use a wider range of phenotypes for applicants with an ethnic background in these countries (see 2.2.4.). Group 2 had a higher assignment probability than Group 1 countries in order to make more detailed analyses of the effects of religion and phenotype possible, within immigrant groups as well as between them. Each of the 19 countries in Group 2 had an assignment probability of about 1.8%.

2.3. Religion

Following other correspondence studies of religious discrimination (e.g., Wallace et al. 2014) we indicated religious affiliation (Christian, Muslim, Hindu, or Buddhist) by applicants' civic engagement in a voluntary association. For half of the sample this engagement was secular (by referring to a *social* association: "Sozialverein Aktiv e.V.") and for the other half we indicated a *religious* affiliation: "Christlicher (or Islamischer/Buddhistischer/Hinduistischer) Sozialverein Aktiv e.V.".

Table 3: Distribution of religions by country of origin

NATIVES (25%)		TURKISH IMMIGRANTS (25%)		
Germany	Christianity	**Turkey**	Islam	Christianity (10%)

GROUP 1 (16, 5%)		GROUP 2 (33, 5%)		
Dominican Republic	Christianity	**Albania**	Christianity	Islam
Greece	Christianity	**Bosnia Herzegovina**	Christianity	Islam
Iran	Islam	**Bulgaria**	Christianity	Islam
Iraq	Islam	**China**	Buddhism	Christianity
Italy	Christianity	**Egypt**	Islam	Christianity
Japan	Buddhism	**Ethiopia**	Christianity	Islam
Mexico	Christianity	**India**	Hinduism	Islam
Morocco	Islam	**Indonesia**	Islam	Christianity
Netherlands	Christianity	**Macedonia**	Christianity	Islam
Pakistan	Islam	**Malaysia**	Islam	Buddhism
Poland	Christianity	**Nigeria**	Christianity	Islam
Romania	Christianity	**Russia**	Christianity	Islam
Spain	Christianity	**South Korea**	Christianity	Buddhism
Switzerland	Christianity	**Trinidad & Tobago**	Christianity	Hinduism
		Uganda	Christianity	Islam
		Vietnam	Christianity	Buddhism
		France	Christianity	
		UK	Christianity	
		USA	Christianity	

By including a range of countries of origin where both Christianity and Islam are plausible religions (e.g., Bosnia and Herzegovina, Egypt, and Nigeria) we replicated the innovation of using multiple religious backgrounds within one ethnic group, as has been done by Adida et al. (2010) and Pierné (2013).

As described before (see section 2.2.2.), 50% of our sample consist of 33 different ethnicities that represent either countries with a single dominant religion or countries with two dominant religions. The minimum threshold for a country to be listed with two dominant religions was that more than five percent of the population belongs to the second dominant religion (Pew Research Center 2015). Table 3 provides an overview of how ethnicities, religions, and their assignment probabilities are distributed. [1]

2.3.1. Phenotype

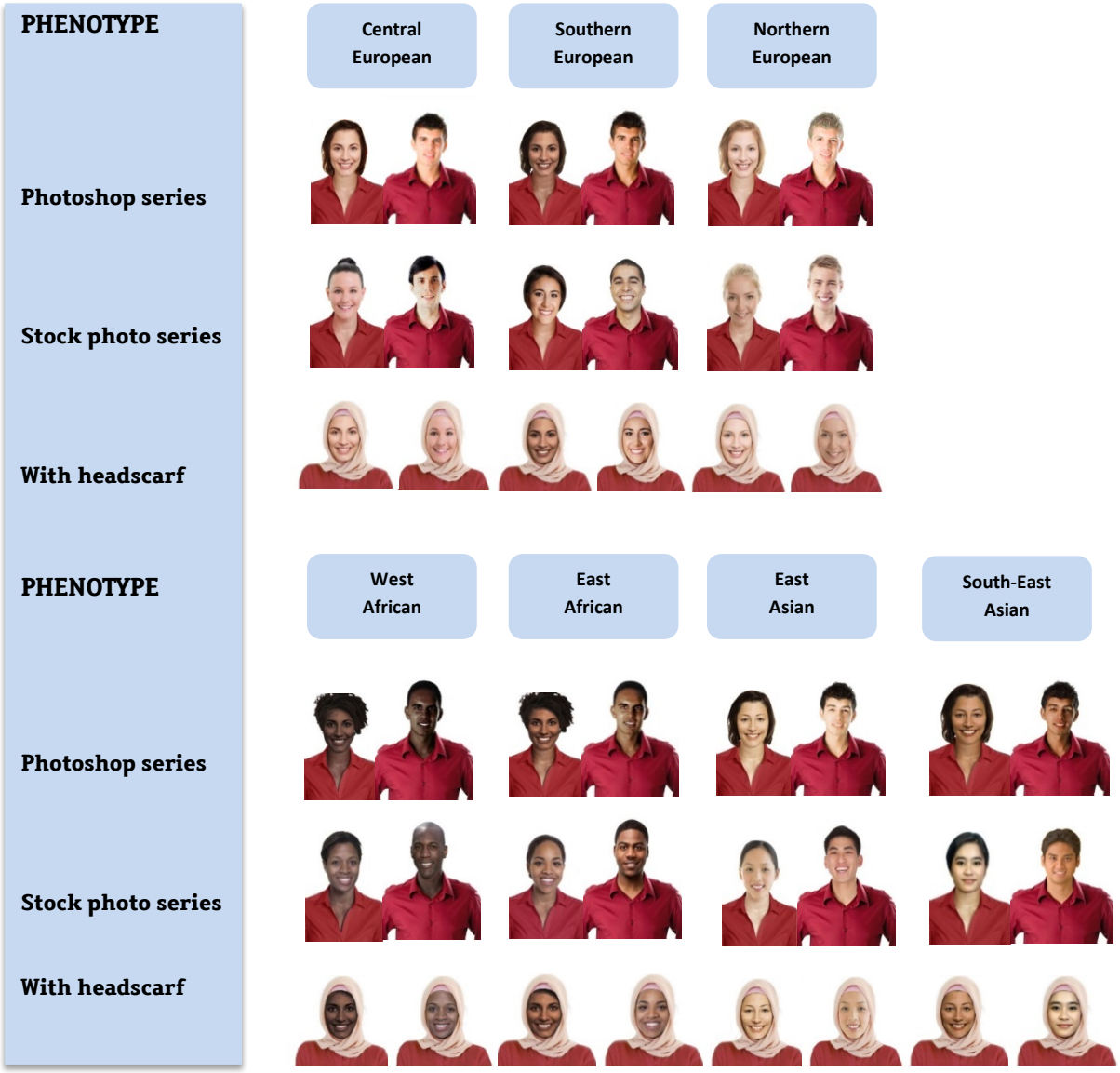
Sending résumé photos is the norm in Germany and widely considered as mandatory for an application in order to be viewed. We therefore added a photo to our CVs in 90% of all cases. Moreover, we took advantage of the necessity of adding photos by varying the phenotype of the person shown on the photo.

To adjust for different countries of origin, we prepared a set of résumé photos, and included seven phenotypes that are prototypical for seven regions of origin, ranging from Northern Europe to Central and Southern Europe to East and West Africa as well as East and South Asia. For each of the seven phenotypes we created two series of photos depicting a prototypical male or female person. Series 1 was created with Photoshop so as to control for attractiveness; series 2 consisted of real stock photos so as to increase external validity. To create series 1, we used stock photos of female and male Whites and then manipulated them with Photoshop to look darker or lighter by changing the skin tone and the color of the eyes and the hair. In addition, we changed the hair texture for African phenotypes. For the Asian phenotypes, we additionally changed the width of the eyes and slightly adapted the shape of the face. Overall, however, the facial characteristics of photos in series 1 were more Eurocentric than Asian or Afrocentric. To control for potential biases resulting from the applied procedure, we added a second series of photos that was based on stock photos of real people; where photos of Asians and Blacks also had other and more realistic ethnic markers than skin color.[2] Finally, for females in both photo series we created a version with headscarf. Altogether, we created 42 different photos in total (14 male and 28 female photos see Figure 2: Photo series).

[1] We made a small exception for applicants with Turkish names, allowing for a Christian affiliation in 10% of all cases.

[2] For an extensive literature review on the role of phenotypic variation and the impact on social perception see Maddox (2004).

Figure 2: Photo series



To ensure the comparability and plausibility of our photos, we conducted two pretests: one on attractiveness and competence judgments of individual photos and one on the plausibility of photos for members of different ethnic groups (see Textboxes 1 and 2).

Textbox 1: Phenotype assignment

We had to decide which photos to assign to which ethnicities. Our aim was to include as many phenotypes as possible for all ethnic groups in our experiment. However, even though all phenotypes could exist in all countries of origin, we had to stay within plausible boundaries to avoid unexpected combinations that could create doubts on the side of employers (e.g., a photo of a black person with a Chinese name). Drawing on photo series 1 (without headscarf), we explored which phenotypes were considered as implausible or unlikely for certain countries of origin. Therefore, we asked 38 people within our institution to indicate on a four point scale, from very unlikely to very likely, how plausible it was that a person with the picture shown on the questionnaire came from one of the 35 countries we used in our study. Based on the results, we determined the choice of plausible phenotypes (which could range between two and seven phenotypes) for each ethnicity. Whenever agreement fell below twenty percent (i.e., more than eighty percent of our respondents classified a photo as very unlikely for a certain ethnicity), phenotypes were excluded from the range of plausible phenotypes for the respective country of origin.

Textbox 2: Survey on attractiveness and competence

All photos (without headscarf) were tested for attractiveness and competence in a survey beforehand, and where necessary, adaptations were made. Initially we used a convenience sample of social science students from Humboldt University in Berlin. However, as minority photos received disproportionately high ratings, we decided to have the photos rated by a more representative sample as well. We therefore asked people in Berlin and Potsdam on busy shopping streets and train stations to take part in the paper/pencil survey. For each photo we asked roughly 25 people to rate how attractive they find the person on the photo on a ten point scale from very unattractive to very attractive. In a second question we asked how competent they believe the person is with regard to the labor market, also on a ten point scale from very incompetent to very competent.

The comparison between photos revealed that two pictures (North European for males and South European for females) had poor ratings and differed significantly from the other photos used in each series; we therefore replaced these two pictures.

Table 4: Distribution of phenotypes

Phenotype Ethnicity	Central European	North European	South European	East African	West African	East Asian	South Asian	No photo
Albania	22.5%	22.5%	22.5%				22.5%	10%
Bosnia & Herzegovina	30%	30%	30%					10%
Bulgaria	30%	30%	30%					10%
China						45%	45%	10%
Dominican Republic			30%	30%	30%			10%
Egypt			30%	30%	30%			10%
Ethiopia				45%	45%			10%
France	12.86%	12.86%	12.86%	12.86%	12.86%	12.86%	12.86%	10%
Germany	22.5%	22.5%	22.5%	5.625%	5.625%	5.625%	5.625%	10%
Greece	45%		45%					10%
India			45%				45%	10%
Indonesia[3]			30%			30%	30%	10%
Iran	30%		30%				30%	10%
Iraq			45%				45%	10%
Italy	30%	30%	30%					10%
Japan						45%	45%	10%
Macedonia	22.5%	22.5%	22.5%				22.5%	10%
Malaysia[4]			30%			30%	30%	10%
Mexico	30%		30%				30%	10%
Morocco[5]			22.5%	22.5%	22.5%		22.5%	10%
Netherlands	30%	30%	30%					10%
Nigeria				45%	45%			10%
Pakistan			45%				45%	10%
Poland	30%	30%	30%					10%
Romania	30%		30%				30%	10%
Russia	30%	30%	30%					10%
South Korea						45%	45%	10%
Spain	30%		30%				30%	10%
Switzerland	30%	30%	30%					10%
Trinidad and Tobago				45%	45%			10%
Turkey	30%		30%				30%	10%
Uganda				45%	45%			10%
UK	12.86%	12.86%	12.86%	12.86%	12.86%	12.86%		10%
USA	12.86%	12.86%	12.86%	12.86%	12.86%	12.86%	12.86%	10%
Vietnam						45%	45%	10%

[3] In all main analyses that we run in our papers, we deviated from the results of the pretest and dropped all applicants with Indonesian names and a South European phenotype from the analysis (33 cases). We did so because this phenotype is less common in Indonesia.

[4] For similar reasons, we dropped all applicants with Malaysian names and with a South European phenotype from the analysis (43 cases).

[5] For similar reasons, we dropped all applicants with Moroccan names and with a South Asian phenotype from the analysis (17 cases).

2.3.2. Grades

To test whether good grades have different effects for natives and immigrants, we varied the applicants' grades on school and vocational diplomas. Drawing on the actual distribution of grades in each occupation as listed by the IHK in their online examination statistics for 2014 (http://pes.ihk.de/berufsauswahl.cfm), we assigned two different grade levels to the applications. Half of all certificates provided information about good grades (equivalent to American B or German 2) and half of them provided information about satisfactory grades (equivalent to American C+ or German 3). For all applications as cooks, however, we had to lower the grades from B to C+ (German 3) and from C+ to C- (German 4), because in Germany, cook trainees on average have slightly lower grades than trainees in the other occupations that we tested.

2.3.3. Expiration of contract

At the end of the cover letter, all applicants mentioned that they could cancel their current (open-ended) contract within a four-week notice period. In 50% of the cases, applicants mentioned that their current work contract would expire in a few months, thus signaling a fixed-term contract. In 50% of the cases no further information was provided, suggesting that the applicant had an open-ended contract since tenured contracts are standard in Germany for the occupations we used in our study. This treatment was meant to signal performance, since most employers would offer a permanent contract to employees with high productivity.

2.3.4. Reference letter, social skills, and linguistic skills

As a replication of Kaas and Manger (2011), we included a reference letter with some of the applications (see Table 5), in which the current employer positively evaluates the applicant (see Appendix for an example). The probability of adding a reference letter was 50%, but throughout the course of the study we altered this probability[6].

Within the reference letters, we nested two additional independent treatments. First, half of the letters stressed the applicant's high social skills in terms of his or her willingness to help others and his or her kind, sincere and courteous personality, while the other half did not mention any social skills in particular. Second, all reference letters had a 50% probability to stress the applicant's linguistic skills in German, while

[6] We started and ended with a probability of 50%. However, for a period of two weeks in 2015 we always added a reference letter to the applications (04.06.–13.06.2015). In compensation, we reduced the probability of adding a reference letter to 25% for a period of four months (04.10.2015–30.01.2016).

50% of all reference letters did not mention linguistic skills. Combining these two treatments, four types of reference letters with equal assignment probability resulted (see Table 5).

Table 5: Distribution of reference letter

No reference letter (50%)	Reference letter (50%)	
	Basic reference letter (12.5%)	+ linguistic skills treatment (12.5%)
	+ social skills treatment (12. 5%)	+ social and linguistic skills treatment (12.5%)

2.4. Cover Story

For reasons of comparability and efficiency, all our fictitious applicants had the following cover story:

They had all been born on February 23rd, 1992 in the region where the respective job was advertised (see section 2.3.1.), but had moved to one of the two centrally located German cities of Bielefeld or Paderborn after primary school. After finishing their secondary school in Bielefeld, they completed vocational training in one of the eight chosen occupations in either Bielefeld or Paderborn. After completion of their vocational training, all applicants had been employed by the company where they had completed their apprenticeship and had gained between 32 and 70 months of work experience.[7] All candidates were based in Bielefeld or Paderborn during the application process. If an employer contacted an applicant via mail, a change of address order redirected all letters to the researchers institute in Berlin. We chose these two cities for practical reasons: First, because they are of comparable size and in the center of the country (though in the Western part of Germany). Second, because we could use two real addresses in both places, where people we knew would forward us all letters that were not redirected by our change of address order.

[7] Due to different periods of job training and a field phase of 20 months, the duration of work-experience differs slightly between occupations.

2.4.1. Place of birth

Companies in Germany that advertise low to medium skilled jobs often restrict job searches to local candidates only. To make applications to jobs all over Germany plausible, we therefore always used the city of the advertised job as the applicant's place of birth. We further mentioned in the cover letter that the candidate grew up in the region where the firm is based and that he or she would like to move back for personal reasons.[8] With regard to place of birth, the only restriction was that the respective city needed to have more than 15,000 inhabitants to make it realistic that the applicant could have been born in a local hospital and that a family of any ethnic background could have lived in that place without the employer's knowledge. In case a city had less than 15,000 inhabitants, we chose the closest city within a radius of 30 km that had a larger population (see Appendix A3. for a more detailed description of this procedure in R). In case there was no city nearby that fulfilled these criteria, we excluded the job offer.

2.4.2. Current employer

We decided to invent all the companies where our applicants were currently working to reduce the risk of detection. This was also necessary as we had to comply with German law that does not allow the "misuse" of names of real firms or people, even for scientific purposes. At the same time, we wanted to choose names that do not raise any suspicion. That's why we did an online search for typical names of firms in each sector and named them in a similar way (see Table 6).

All companies were based either in a busy main street or in industrial areas of Bielefeld or Paderborn. We ensured via a search on "Google Maps" that there was a building at the address that actually could be an office, workshop, or restaurant, respectively.

[8] We excluded vacancies within a radius of 50 km around Bielefeld and Paderborn as it wouldn't make sense for an applicant to "move back home" to an area that was within driving distance. This also made employers less likely to realize that the candidate's current employer was invented.

Table 6: Company names and logos by occupation

Occupation	Name	Logo
Plant mechanic	Neumann Sanitär, Heizung & Klimatechnik	Neumann Sanitär, Heizung & Klimatechnik
Mechatronics fitter	Mech Solutions	MECH Solutions
Hotel receptionist	Hotel Schlosshofstrasse	Hotel Schlosshofstraße
Cook	Restaurant Padersee	Restaurant Padersee
Industrial clerk	Neumann Produktion & Vertrieb	Neumann Produktion & Vertrieb
Sales assistant	Frischemarkt Kurt Neumann	Frischemarkt Kurt Neumann
Medical assistant	Hausärztlich- internistische Gemeinschaftspraxis Dres. Neumann	Hausärztlich-internistische Gemeinschaftspraxis Dres. Neumann
Dental assistant	Praxis Zahnwelten Paderborn	Praxis Zahnwelten PADERBORN

3. Field experiment

3.1. Implementation of the field experiment

Before starting the experiment, we tested our application documents and ran a pilot study (see Textbox 3).

Textbox 3: Pilot study

From 26.09.2013 to 08.04.2014, we conducted a pilot study. We manually sent out 288 applications all across Germany to test the credibility of our application material. After the pilot study, we decided to use an automatized procedure to increase our productivity and to reduce mistakes (for example: sending more than one application to the same employer). We also decided to leave out some redundant or ambiguous information on the CVs of the applicants, like information about the family (birthdate and birthplace of parents and siblings) and information about military service.

For the main study, we used a computer-assisted procedure to conduct the field experiment. The software we used was based on a cronjob, which schedules a command or script on a server to run automatically at a specified time and date. The software searched for vacancies in the target occupations on the job search website once a week (on Sundays). The software was programmed to automatically sort out job ads published more than seven days ago and those containing certain keywords, such as "part-time" or "recruitment agency". In addition, the software checked for similarities in contact details between recent and former job ads and automatically sorted out duplicates to minimize the risk of sending multiple applications to the same company.

In addition, the software automatically assigned our experimental treatments to applications and vacancies (by taking some predefined quota and limitations related to plausible combinations of ethnicity, religion, and phenotype into account) and created all application documents, including cover letter, CV, certificates, and reference letters. In a second step, our research assistants manually checked the automatically created applications for mistakes. They also checked if the profile of our fictitious applicants fitted to the advertised job. In case of mistakes or a misfit, applications and thus job ads were sorted out and excluded from the experiment. Thereafter, a second manual check of applications was conducted to further minimize the risk of mistakes. As a result of the second check, all applications were either sorted out or identified as "ready to send". All applications identified as "ready to send" were automatically send to employers by e-mail on Saturdays. In case no e-mail address was provided, we printed the application documents and send them by mail on Fridays.

For the job search we used the vacancy database of the German employment agency (jobboerse.arbeitsagentur.de). Although not all employers publish their vacancies on this webpage, it is one of the most widely used databases for jobs that require vocational training in Germany. Most other online job vacancy databases copy job offers published by the employment agency. Therefore, we restricted our sample to this database only. We did not respond to newspaper or local ads as the search process is

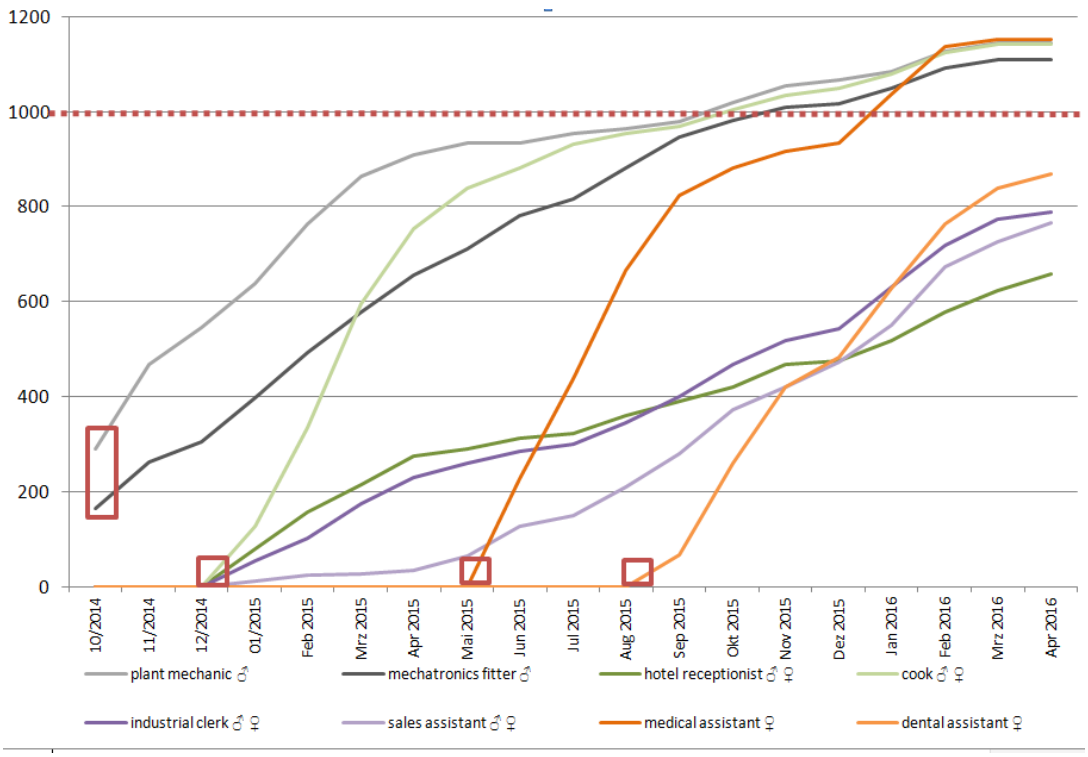
very time-consuming and cannot be performed systematically, and within the German context these ads are of minor importance in terms of numbers for the selected occupations.

We decided to exclude public employers, as they usually have a centralized application process for several vacancies at the same time. We also excluded applications that had to be filled in manually through an online platform, as well as vacancies that require religious affiliations (e.g., working in a religious hospice as a cook), a willingness to travel internationally, as well as any kind of leadership positions and seasonal or temporary work. We only selected vacancies for which employers were clearly identifiable (no anonymous employers or recruitment agencies) and applications via e-mail or postal mail were possible. We disregarded vacancies that required in-person or telephone contact. Since our fictitious applicants already had about three years of job experience, we did not apply to vacancies that were explicitly targeted at people who were looking for their first job. Since we never modified the automatically generated and assigned applications, we did not respond to specific questions mentioned in job ads, such as questions about desired salary. However, we coded such information in order to be able to control for this later on in the data analysis.

3.2. Sent applications

We planned to send about 1.000 applications per occupation, summing up to 8.000 applications in total. However, because of the high number of application documents that had to be created separately for each occupation (most importantly, certificates from school and occupational training for each ethnicity and gender at different grade levels), we didn't manage to start with all occupations at the same time. We instead started with only two occupations (plant mechanic and mechatronics fitter) in October 2014. In December 2014 we added four more occupations (cook, hotel receptionist, industrial clerk and sales), in May 2015 we started sending out applications for medical assistants, and finally in August 2015 we started sending out applications for dental assistants (see Figure 3).

Figure 3: Applications per month and occupation



In total we sent out 7,631 applications. 67 of them failed to arrive, in four cases the experiment was detected, and three cases were dropped from the analysis because of mistakes in the documents. For the data analysis a total of 7,557 applications remained.

As mentioned before, applications could be sent via e-mail or postal mail, depending on the sending request in the job advertisement. When both postal mail and e-mail applications were allowed, we sent the application via e-mail. Table 7 provides an overview of the share of applications sent via e-mail and postal mail. Most of the applications were sent via e-mail (81%); within the health sector, it was more common to send an application via postal mail.

Table 7: Type of delivery

Occupation	E-Mail	%	Mail	%
Plant mechanic	896	79	234	21
Mechatronics fitter	982	90	110	10
Hotel receptionist	563	86	93	14
Cook	919	82	203	18
Industrial clerk	738	95	40	5
Shop assistant	613	78	174	22
Medical assistant	724	64	416	37
Dental assistant	667	78	185	22
∑	**6,102**	**Ø 81**	**1,455**	**Ø 19**

Germany is a federal state and divided into 16 states. Nearly 1,400 applications were sent to vacancies in Bavaria. This is the economically most prosperous federal state in Germany and therefore has a high number of job offers. Only 15% of all applications were sent to vacancies in North Rhine-Westphalia; which was due to the fact that we excluded all vacancies in zip code regions within a 50 kilometers radius around Paderborn and Bielefeld. The distribution of applications by federal state also varies with occupations due to state-specific industries. Most applications for technical occupations were sent to North Rhine-Westphalia and Baden-Wuerttemberg, whereas for health occupations the highest number of applications was sent to Bavaria (see Table 8).

Table 8: Applications per state and occupation

State	Total	%	Plant mechanic	Mechatr. fitter	Hotel reception.	Cook	Industrial clerk	Sales assistant	Medical assistant	Dental assistant
Baden-Wuerttemberg	1,212	16	167	210	84	118	186	101	189	157
Bavaria	1,389	18	152	142	151	206	136	140	279	183
Berlin	190	3	30	21	11	13	8	20	55	32
Brandenburg	143	2	24	26	16	35	7	12	15	8
Bremen	66	1	9	9	8	12	2	7	9	10
Hamburg	141	2	15	7	10	13	4	26	49	17
Hesse	525	7	81	74	34	75	42	71	93	55
Lower Saxony	798	11	131	88	69	108	69	113	120	100
Mecklenburg-Western Pomerania	148	2	7	27	44	33	7	13	8	9
North Rhine-Westphalia	1,111	15	194	144	74	149	160	94	145	151
Rhineland-Palatinate	515	7	88	80	44	81	56	50	78	38
Saarland	111	1	29	15	7	15	6	9	23	7
Saxony	339	4	46	89	28	77	23	34	14	28
Saxony-Anhalt	168	2	27	34	16	49	11	21	3	7
Schleswig-Holstein	363	5	78	32	36	70	31	46	37	33
Thuringia	338	4	52	94	24	68	30	30	23	17
∑	**7,557**	**100**	**1,130**	**1,092**	**656**	**1,122**	**778**	**787**	**1,140**	**852**

3.3. Employer responses

Our fictitious job applicants provided contact details on their CVs, including an e-mail address, a mobile phone number, and a postal address. When employers called the phone number, a voicemail automatically answered after 30 seconds of ringing. Employer responses via postal mail were automatically redirected to the researchers' institute. Most employers first sent an e-mail, followed by contact attempts via phone.

The overall response rate was 70%, with some responses signaling interest while others didn't. In 30% of all cases we never received any response. We sorted all forms of employer responses (including non-responses) into nine response categories (see Table 9) and recorded them in the order of registration, since most employers tried to reach our fictitious job applicants several times. For all postal letters and e-mails, we could identify the employer who contacted us. The same holds true for most telephone calls. Nevertheless, we received 318 telephone calls for which we were not able to identify the caller.

Based on this information, we created a variable indicating employers' final response. In order to create a final response variable we created a loop that took the first response as a baseline and overwrote it with the second recent response if the second response was more positive or negative than the first one. Thereafter we did the same with the second and third response, the third and fourth one, etc. The process of overwriting former responses with more recent ones, however, was determined by the following principles (see Table 9 for a description of the categories "neutral", "positive" and "negative"):

- a previous neutral response is overwritten by any later response,
- no previous negative response can be overwritten by a later positive or neutral one, and
- no previous positive response can be overwritten by a later negative or neutral one.

We applied these principles to avoid mistakes in the process of determining the final response. We only 'responded' to positive contact attempts by withdrawing our application. After receiving a withdrawal, some employers sent a rejection letter and others repeated their invitation or asked for a callback. This also occurred after we had already received a rejection; some employers signaled interest in our application weeks or even months later. To accommodate all these different contact sequences, we opted for the procedure we just described. Once an applicant had received a response that was classified as positive (or negative) according to Table 9, the final response was also positive (or negative). A positive response could still become more positive but never less positive, neutral, or negative.

Table 10 provides an overview of the frequency of the final response types. Among all responses, 35% were invitations for a job interview and only 16% were rejections, suggesting that our application material was credible. Based on the nine response categories, we constructed a dummy variable differentiating between negative and positive responses. Thereby, positive responses ranged from "hang-up" to "invitation" and negative ones from "confirmation of receipt" to "rejection". In total, we received 50% positive responses and 50% negative responses, including all cases where an application never received any response from the employer.

Table 9: Employer responses

Initial meaning	Response categories	Response dummy	Frequency	%
Positive	Invitation	**Positive** (signal of interest)	2,664	35
	Invitation on trial basis		50	1
	Request for callback		626	8
	Further questions		236	3
Neutral	Hang-up		214	3
	Confirmation of receipt	**Negative** (no signal of interest)	131	2
	No response		2,270	30
Negative	Position already taken		189	3
	Rejection		1,177	16
		∑	**7,557**	**100**

However, the frequency of different response types varied across occupations. As table 10 shows, dental assistants had the highest positive response rate across all occupations, whereas industrial clerks had the lowest.

Table 10: Responses by occupation

	Applications	% of all	Positive Response	%	Negative Response	%
Plant mechanic	1,130	15	649	57	481	43
Mechatronics fitter	1,092	14	598	55	494	45
Hotel receptionist	656	9	375	57	281	43
Cook	1,122	15	631	56	491	44
Industrial clerk	778	10	146	19	632	81
Sales assistant	787	10	299	38	488	62
Medical assistant	1,140	15	566	50	574	50
Dental assistant	852	11	526	62	326	38
∑	**7,557**	**100**	**3,790**	**Ø 50**	**3,767**	**Ø 50**

4. Ethical considerations

By their very nature, field experiments of discrimination involve deception, as employers must not be aware that their decisions are being recorded. Therefore, the condition of informed consent cannot be met. In our experiment, deception consisted in sending application documents from fictitious job candidates to real vacancies.

One justification for deception is that field experiments have crucial methodological advantages for detecting discrimination compared to all other methods (Riach and Rich 2002, Wood et al 2009). Carefully controlled field experiments constitute the most robust procedure for measuring the incidence of discriminatory behavior and for charting, over time, the effectiveness of equal opportunity legislation. Additionally, there is no expectation of privacy in the act of hiring workers, as national governments and international bodies have accepted the onus of ensuring equality of opportunity for all citizens by declaring discrimination in employment to be unlawful (Riach and Rich, 2002). Since privacy is not necessarily a legitimate expectation when public and commercial acts are involved (McClendon, 2012), one could regard field experiments of this sort as analogous to the 'mystery shopping' exercises frequently carried out by public bodies and NGOs.

Given the societal importance of monitoring equal opportunity legislation and improving labor market effectiveness, the methodological advantages of field experiments are considered to outweigh the potential negative consequences of violating research ethics with respect to informed consent. It is widely acknowledged that field experiments can provide "clear and convincing evidence" of discriminatory behavior that cannot be obtained in an unbiased way through any alternative method (Pager 2007). There has been a long tradition of such experiments in the United States and many other countries, which have secured ethical approval for the deception involved.

In addition, field experiments of hiring discrimination typically refrain from debriefing subjects after the completion of the study. Pager (2007) refers to the possible risk posed to human resource personnel or managers who are thought to be discriminating and as a consequence may fall under greater scrutiny by superiors. Also, harm may occur to participants who are debriefed in field-based experiments as awareness of the study results may generate negative emotions (such as embarrassment). For these reasons, in this study we did not debrief individual participants.

In the rare event (four times) that a participant raised suspicion about the application during the data collection or discovered the field experiment, we debriefed them immediately and withdrew the respective application from the study. In addition, we tried to keep the potential burden to employers and real job applicants as minimal as possible by withdrawing our applications within three days after receiving any kind of positive response, such as a request for callback, further questions, or invitations.

When we started our experiment, our research institute did not yet have a research ethics committee. We therefore applied for an ethics approval at the German sociological association (DGS). We waited for a decision several months and contacted the DGS several times. However, our application never received any response. The DGS, at this point in time, did not consider the review of research proposals as one of their tasks. Instead, they focused on the formulation and development of ethical guidelines for researchers in the field of sociology. Therefore, drawing on the expert report by Kühn, Liebscher and Klose (2013), we decided to start our experiment with no further delay. In the meantime, a research ethics committee was established in our research institute; partly in consequence of these difficulties. Since then, a similar study with an almost identical design was approved by the WZB ethics committee (GEMM project).

5. Bibliography

Adida, C. L., Laitin, D. D., & Valfort, M.–A. (2010). Identifying barriers to Muslim integration in France. Proceedings of the National Academy of Sciences of the United States of America, 107(52), 22384–22390. https://doi.org/10.1073/pnas.1015550107

Ahmed, A. M., & Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the Internet. Journal of Urban Economics, 64(2), 362–372. https://doi.org/10.1016/j.jue.2008.02.004

Arrow, K. 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

Becker, G. S. (1957). The economics of discrimination. Chicago: The Univ. of Chicago Press.

Burkert, C., & Seibert, H. (2007). Labour market outcomes after vocational training in Germany - equal opportunities for migrants and natives? IAB discussion paper, No. 2007,31. Retrieved from http://hdl.handle.net/10419/32679

Carlsson, M., & Rooth, D.–O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. Labour Economics, 14(4), 716–729. https://doi.org/10.1016/j.labeco.2007.05.001

Fix, M. E., & Struyk, R. J. (Eds.). (1993). Clear and Convincing Evidence. Measurement of Discrimination in America. Washington, D.C.: The Urban Institute Press.

Heckman, J. J. (1998). Detecting Discrimination. Journal of Economic Perspectives, 12(2), 101–116. https://doi.org/10.1257/jep.12.2.101

Heckman, J. J., & Siegelman, P. (1993). The Urban Institute Audit Studies: Their Methods and Findings. In M. E. Fix & R. J. Struyk (Eds.), Clear and Convincing Evidence. Measurement of Discrimination in America (pp. 187–258). Washington, D.C.: The Urban Institute Press.

Humphrey, C., Moizer, P., & Turley, S. (1993). The Audit Expectations Gap in Britain: An Empirical Investigation. Accounting and Business Research. (23), 395–411.

Kaas, L., & Manger, C. (2011). Ethnic Discrimination in Germany's Labour Market: A Field Experiment. German Economic Review, 13(1), 1–20. https://doi.org/10.1111/j.1468-0475.2011.00538.x

Kogan, I. (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

Kühn, K., Liebscher, D., & Klose, A. (2013). Expertise zu den rechtlichen Aspekten sozialwissenschaftlicher Testing-Verfahren im Rahmen der vom SVR geplanten Studie „Diskriminierung am Ausbildungsmarkt?" Expertise im Auftrag des SVR-Forschungsbereichs, Berlin.

Maddox, K. B. (2004). Perspectives on Racial Phenotypicality Bias. Personality and Social Psychology Review, 8(4), 383–401.

McClendon, G. H. (2016). Race and Responsiveness: An Experiment with South African Politicians. Journal of Experimental Political Science, 3(01), 60–74. https://doi.org/10.1017/XPS.2015.10

Oreopoulos, P., & Dechief, D. (2012). Why do some employers prefer to interview Matthew but not Samir? New evidence from Toronto, Montreal and Vancouver (Working Paper No. 95).

Pager, D. (2007). The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future. The ANNALS of the American Academy of Political and Social Science, 609(1), 104–133. https://doi.org/10.1177/0002716206294796

Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a Low-Wage Labor Market: A Field Experiment. American sociological review, 74(5), 777–799. https://doi.org/10.1177/000312240907400505

Pew Research Center. (2017). "The Future of World Religions: Population Growth Projections, 2010-2050".

Phelps, E. (1972). The Statistical Theory of Racism and Sexism. American Economic Review, 62(4), 659–661.

Pierné, G. (2013). Hiring discrimination based on national origin and religious closeness: Results from a field experiment in the Paris area. Journal of Labor Economics, 2(1), 4. https://doi.org/10.1186/2193-8997-2-4

Riach, P. A., & Rich, J. (2002). Field Experiments of Discrimination in the Market Place The Economic Journal, 112(483), F480-F518. https://doi.org/10.1111/1468-0297.00080

Schneider, J., Yemane, R., & Weinmann, M. (2014). Diskriminierung am Ausbildungsmarkt: Ausmaß, Ursachen und Handlungsperspektiven. Berlin. Retrieved from https://www.svr-migration.de/wp-content/uploads/2014/03/SVR-FB_Diskriminierung-am-Ausbildungsmarkt.pdf

Uhlendorff, A., & Zimmermann, K. F. (2006). Unemployment Dynamics among Migrants and Natives. IZA Discussion Papers, No. 2299. Retrieved from http://hdl.handle.net/10419/33769

Wallace, M., Wright, B. R. E., & Hyde, A. (2014). Religious Affiliation and Hiring Discrimination in the American South: A Field Experiment. Social Currents, 1(2), 189–207. https://doi.org/10.1177/2329496514524541

Wood, M., Hales, J., Purdon, S., Sejersen, T., & Hayllar, O. (2009). A test for racial discrimination in recruitment practice in British cities (Research Report No. 607). Retrieved from Department for Work and Pensions website: http://www.natcen.ac.uk/media/20541/test-for-racial-discrimination.pdf

APPENDIX

A1. Occupation profiles

Mechatronics fitters

Mechatronics fitters work in the installation and maintenance of mechatronics components and systems for manufacturers. They assemble and install sub-assemblies and components into mechatronic systems and program them. Moreover, they hand over plants and instruct users in their operation.

Plant mechanics

Plant mechanics work in the installation and maintenance of complex supply engineering systems. They carry out insulation, sealing and protective measures and test, adjust, optimize and maintain the functionality of supply engineering plants.

Hotel receptionists

Hotel receptionists mainly work in hotels, restaurants or bed and breakfast establishments and serve at the front desk where they welcome and advise guests. In addition to checking in and out guests, they handle correspondence and customer queries, operate the hotel cash desk and also serve food and beverages in the hotel restaurant or bar.

Cooks

Cooks work in the hotel and restaurant trade, e.g., in hotels, bed and breakfast hotels and restaurants where they cook and prepare meals from fresh. They also find employment in social institutions such as hospitals and old peoples' homes and canteens.

Sales assistants

Sales assistants assist and advise customers in stores and respond to queries. They operate the till, label and store goods and arrange and present goods in the sales area.

Industrial clerks

Industrial clerks work in a range of industrial branches and are employed in a number of sectors and areas, such as materials management, production, and sales. Tasks range from analyzing market potential and determining the requirements for products and services to customer service and organizing work processes. They are also employed in human resources or accountancy roles.

Medical assistants

Medical assistants work in general and specialist practices, hospitals and other institutions providing medical care. They assist in examinations, treatments and surgical operations, organize workflows, monitor schedule times and carry out administrative tasks.

Dental assistants

Dental assistants work in dentists' surgeries, in orthodontic surgeries, oral and dental surgeries and dental clinics where they look after patients before, during and after treatment. They also assist during treatments of patients and carry out hygiene measures. Further tasks include documentation of treatment courses, organizing surgery procedures and monitoring of payments.

Figure 4: Example cover letter for mechatronics fitters

Bielefeld, 08.12.2017

Florian Müller, [redacted] 33605 Bielefeld

[redacted]
Stolberg

Bewerbung um eine Stelle als Mechatroniker

Sehr geehrter Herr [redacted],

mit großem Interesse habe ich Ihre Stellenausschreibung auf der Internetseite der Bundesagentur für Arbeit gelesen und möchte mich bei Ihnen bewerben. Ich bin ausgebildeter Mechatroniker mit Berufserfahrung. Seit dem Abschluss meiner Ausbildung im Januar 2012 arbeite ich als festangestellter Mechatroniker in meinem Ausbildungsbetrieb. Nun möchte ich meine Kenntnisse und Fertigkeiten erweitern und suche deshalb die berufliche Veränderung. Zudem möchte ich wieder in die Gegend von Stolberg ziehen, wo ich aufgewachsen bin.

In der MECH Solutions GmbH bin ich mit dem Bau, der Programmierung und der Wartung von Industrieanlagen betraut. Zu meinen Aufgaben gehört es durch die Anwendung von Prüfverfahren und Problemanalysetools den nahtlosen Ablauf der Produktion zu garantieren und die Qualität sicher zu stellen. Zudem habe ich Maschinen durch neuere Komponenten und für Produktwechsel umgerüstet. Ebenfalls habe ich Sonderanlagen beim Kunden zusammengebaut und diese nach Inbetriebnahme übergeben. Die Zufriedenheit des Kunden ist mir bei meiner Arbeit besonders wichtig.

Ich bin bereit zu reisen, am Wochenende und in Schichten zu arbeiten. Ich arbeite gerne im Team, kann aber durch meine gute Auffassungsgabe komplexe Probleme auch unter Zeitdruck selbstständig analysieren und lösen und arbeite mich schnell in ein neues Arbeitsumfeld ein. Auch soziale Verantwortung und gesellschaftliche Werte sind mir wichtig, daher bin ich aktives Mitglied im Christlichen Sozialverein AKTIV e.V. und organisiere jedes Jahr ein Sommerfest für Kinder. Ich freue mich auf die Chance, Sie bei einem persönlichen Gespräch von meinen Fähigkeiten überzeugen zu können. Derzeit befinde ich mich in ungekündigter Anstellung (Kündigungsfrist 4 Wochen zum Monatsende), mein Vertrag läuft jedoch in wenigen Monaten aus.

Mit freundlichen Grüßen

F. Müller

Florian Müller

Anlagen: Lebenslauf, Zeugnisse

Note: This cover letter signals ethnic background (German) and gender (male) by applicant's name, religion by engagement in a Christian association, and a fixed-term contract (all highlighted in yellow). In addition, the place where the applicant was raised is adapted based on the city of the advertised job (also highlighted in yellow: Stolberg).

Figure 5: Example CV for mechatronics fitters

Lebenslauf

Persönliche Daten

Name	Florian Müller
Anschrift	[redacted] Bielefeld
Telefon	[redacted]
E-Mail	florian.mueller.2302@gmx.de
Geburtsdatum	23.02.1992 in Stolberg
Nationalität	Deutsch
Familienstand	Ledig

Schule und Ausbildung

09/98- 07/02	Grundschule in Stolberg
08/02- 06/08	Realschule Bielefeld
	Abschlussnote: gut
08/08-01/12	Ausbildung zum Mechatroniker bei der MECH Solutions GmbH und am
	[redacted] Berufskolleg [redacted], Bielefeld
	Abschlussnote: gut

Berufstätigkeit

Seit 02/12	Mechatroniker bei der MECH Solutions GmbH

Besondere Kenntnisse und Interessen

Sprachkenntnisse	Deutsch (Muttersprache), Englisch (gute Schulkenntnisse)

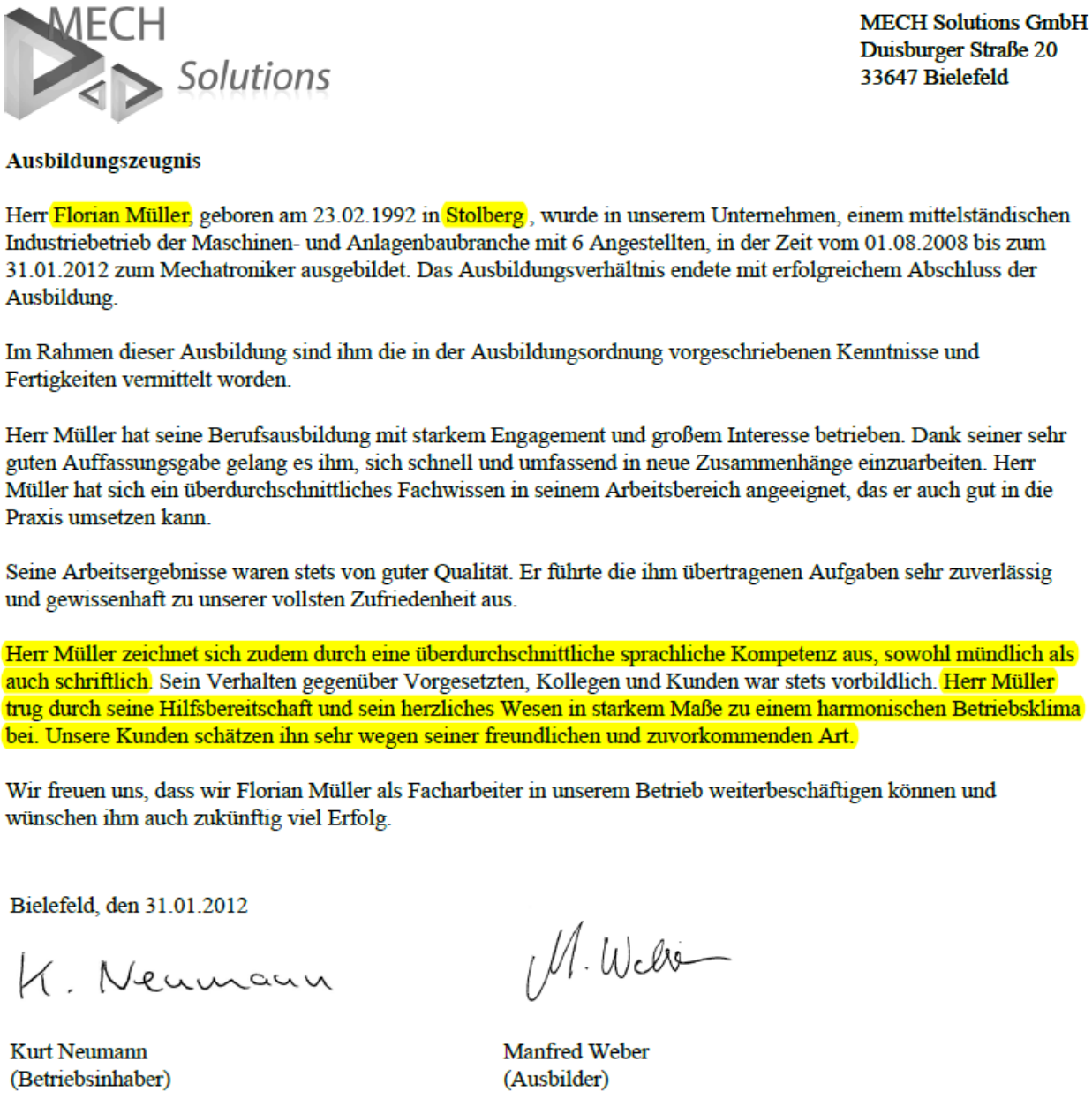
Example CV for mechatronics fitters (continued)

Programmierung	CNC, SPS (Siemens Step 5/7)
EDV	sicherer Umgang mit Microsoft Office, Internet-Recherche
Ehrenamt	Mitglied im Christlichen Sozialverein AKTIV e.V.
Sonstiges	Führerschein Klasse B

Bielefeld, 08.12.2017

Note: This CV signals ethnic background (German) by applicant's name, religion by engagement in a Christian association, and good grades (all highlighted in yellow). In addition, the place where the applicant was born and raised is adapted based on the city of the advertised job (also highlighted in yellow: Stolberg). Moreover, the photo on the CV shows a male person with a Northern European phenotype.

Figure 6: Example reference letter for mechatronics fitters



MECH Solutions GmbH
Duisburger Straße 20
33647 Bielefeld

Ausbildungszeugnis

Herr Florian Müller, geboren am 23.02.1992 in Stolberg, wurde in unserem Unternehmen, einem mittelständischen Industriebetrieb der Maschinen- und Anlagenbaubranche mit 6 Angestellten, in der Zeit vom 01.08.2008 bis zum 31.01.2012 zum Mechatroniker ausgebildet. Das Ausbildungsverhältnis endete mit erfolgreichem Abschluss der Ausbildung.

Im Rahmen dieser Ausbildung sind ihm die in der Ausbildungsordnung vorgeschriebenen Kenntnisse und Fertigkeiten vermittelt worden.

Herr Müller hat seine Berufsausbildung mit starkem Engagement und großem Interesse betrieben. Dank seiner sehr guten Auffassungsgabe gelang es ihm, sich schnell und umfassend in neue Zusammenhänge einzuarbeiten. Herr Müller hat sich ein überdurchschnittliches Fachwissen in seinem Arbeitsbereich angeeignet, das er auch gut in die Praxis umsetzen kann.

Seine Arbeitsergebnisse waren stets von guter Qualität. Er führte die ihm übertragenen Aufgaben sehr zuverlässig und gewissenhaft zu unserer vollsten Zufriedenheit aus.

Herr Müller zeichnet sich zudem durch eine überdurchschnittliche sprachliche Kompetenz aus, sowohl mündlich als auch schriftlich. Sein Verhalten gegenüber Vorgesetzten, Kollegen und Kunden war stets vorbildlich. Herr Müller trug durch seine Hilfsbereitschaft und sein herzliches Wesen in starkem Maße zu einem harmonischen Betriebsklima bei. Unsere Kunden schätzen ihn sehr wegen seiner freundlichen und zuvorkommenden Art.

Wir freuen uns, dass wir Florian Müller als Facharbeiter in unserem Betrieb weiterbeschäftigen können und wünschen ihm auch zukünftig viel Erfolg.

Bielefeld, den 31.01.2012

K. Neumann

Kurt Neumann
(Betriebsinhaber)

M. Weber

Manfred Weber
(Ausbilder)

Note: This reference letter signals linguistic and social skills (highlighted in yellow). In addition, the letter is adapted to applicant's name and place of birth (also highlighted in yellow).

Figure 7: Example cover letter for hotel receptionists

Bielefeld, 13.02.2015

Konstantina Papadopoulou
[redacted]

[redacted]

Bewerbung um eine Stelle als Hotelfachfrau

Sehr geehrte Frau [redacted],

mit großem Interesse habe ich Ihre Stellenausschreibung auf der Internetseite der Bundesagentur für Arbeit gelesen und möchte mich bei Ihnen bewerben. Ich bin ausgebildete Hotelfachfrau mit Berufserfahrung.
Seit dem Abschluss meiner Ausbildung im Juli 2011 arbeite ich festangestellt als Hotelfachfrau in meinem Ausbildungsbetrieb. Nun möchte ich meine Kenntnisse und Fertigkeiten erweitern und suche deshalb die berufliche Veränderung. Zudem möchte ich wieder in die Gegend von Wilnsdorf ziehen, wo ich aufgewachsen bin.

Im Hotel Schloßhofstraße bin ich am Hotelempfang und in den Hotelzimmern tätig. Im Mittelpunkt meines Arbeitsalltags steht die Betreuung der Gäste. Ebenso zählen das Ein- und Auschecken der Hotelgäste, sowie die Zimmerkontrolle und die Vor- und Nachbereitung der Hotelzimmer zu meinen Aufgaben. Darüber hinaus übernehme ich verwaltende Tätigkeiten wie Rechnungsstellung und Datenpflege und bin deshalb gut mit dem Hotelverwaltungs- und Reservierungssystem Fidelio vertraut.

Das Wohlergehen und die Zufriedenheit der Hotelgäste stehen für mich an erster Stelle. Daher arbeite ich stets serviceorientiert und sorgfältig. Ich arbeite gerne im Team, kann mich aber durch meine gute Auffassungsgabe auch selbstständig in neue Aufgaben und ein neues Arbeitsumfeld einarbeiten. Auch soziale Verantwortung und gesellschaftliche Werte sind mir wichtig, daher bin ich aktives Mitglied im Sozialverein AKTIV e.V. und organisiere jedes Jahr ein Sommerfest für Kinder.

Ich freue mich auf die Chance, Sie bei einem persönlichen Gespräch von meinen Fähigkeiten überzeugen zu können. Derzeit befinde ich mich in ungekündigter Anstellung (Kündigungsfrist 4 Wochen zum Monatsende).

Mit freundlichen Grüßen

K. Papadopoulou

Konstantina Papadopoulou

Anlagen: Lebenslauf, Zeugnisse

Note: This cover letter signals ethnic background (Greek) and gender (female) by applicant's name (highlighted in yellow). In addition, the place where the applicant was raised is adapted based on the city of the advertised job (also highlighted in yellow: Wilnsdorf).

Figure 8: Example CV for hotel receptionists

Lebenslauf

Persönliche Daten

Name	Konstantina Papadopoulou
Anschrift	[redacted] 33605 Bielefeld
Telefon	[redacted]
E-Mail	konstantina.papadopoulou.2302@gmx.de
Geburtsdatum	23.02.1992 in Wilnsdorf
Nationalität	Deutsch
Familienstand	Ledig

Schule und Ausbildung

09/98- 07/02	Grundschule in Wilnsdorf
08/02- 06/08	Realschule Bielefeld
	Abschlussnote: befriedigend
08/08-07/11	Ausbildung als Hotelfachfrau beim Hotel Schloßhofstraße und am
	[redacted] Berufskolleg [redacted], Bielefeld
	Abschlussnote: befriedigend

Berufstätigkeit

Seit 08/11	Hotelfachfrau im Hotel Schloßhofstraße

Besondere Kenntnisse und Interessen

Sprachkenntnisse	Deutsch (Muttersprache), Griechisch (Muttersprache), Englisch (gute Schulkenntnisse)

Example CV for hotel receptionists (continued)

EDV	Micros Fidelio Suite 8, sicherer Umgang mit Microsoft Office, Internet-Recherche
Ehrenamt	Mitglied im Sozialverein AKTIV e.V.
Sonstiges	Führerschein Klasse B

Bielefeld, 13.02.2015

Note: This CV signals ethnic background (Greek) by applicant's name and Greek as second mother tongue as well as satisfactory grades (highlighted in yellow). In addition, the place where the applicant was born and raised is adapted based on the city of the advertised job (also highlighted in yellow: Wilnsdorf). Moreover, the photo on the CV shows a female person with a Central European phenotype.

Figure 9: Example reference letter for hotel receptionists

Hotel
Schloßhofstraße

Hotel Schloßhofstraße
Schloßhofstr. 63
33615 Bielefeld

Ausbildungszeugnis

Konstantina Papadopoulou, geboren am 23.02.1992 in Wilnsdorf, wurde in unserem Hotel mit 20 Angestellten, in der Zeit vom 01.08.2008 bis zum 29.07.2011 zur Hotelfachfrau ausgebildet. Das Ausbildungsverhältnis endete mit erfolgreichem Abschluss der Ausbildung. Im Rahmen dieser Ausbildung sind ihr die in der Ausbildungsordnung vorgeschriebenen Kenntnisse und Fertigkeiten vermittelt worden.

Frau Papadopoulou hat ihre Berufsausbildung mit starkem Engagement und großem Interesse betrieben. Dank ihrer sehr guten Auffassungsgabe gelang es ihr, sich schnell und umfassend in neue Zusammenhänge einzuarbeiten. Frau Papadopoulou hat sich ein überdurchschnittliches Fachwissen in ihrem Arbeitsbereich angeeignet, das sie auch gut in die Praxis umsetzen kann. Ihre Arbeitsergebnisse waren stets von guter Qualität. Sie führte die ihr übertragenen Aufgaben sehr zuverlässig und gewissenhaft zu unserer vollsten Zufriedenheit aus.

Ihr Verhalten gegenüber Vorgesetzten, Kollegen und Kunden war stets vorbildlich. Frau Papadopoulou trug durch ihre Hilfsbereitschaft und ihr herzliches Wesen in starkem Maße zu einem harmonischen Arbeitsklima bei. Unsere Kunden schätzen sie sehr wegen ihrer freundlichen und zuvorkommenden Art.

Wir freuen uns, dass wir Konstantina Papadopoulou als Hotelfachfrau in unserem Betrieb weiterbeschäftigen können und wünschen ihr auch zukünftig viel Erfolg.

Bielefeld, den 29.07.2011

K. Neumann

M. Weber

Kurt Neumann
(Hotelleitung)

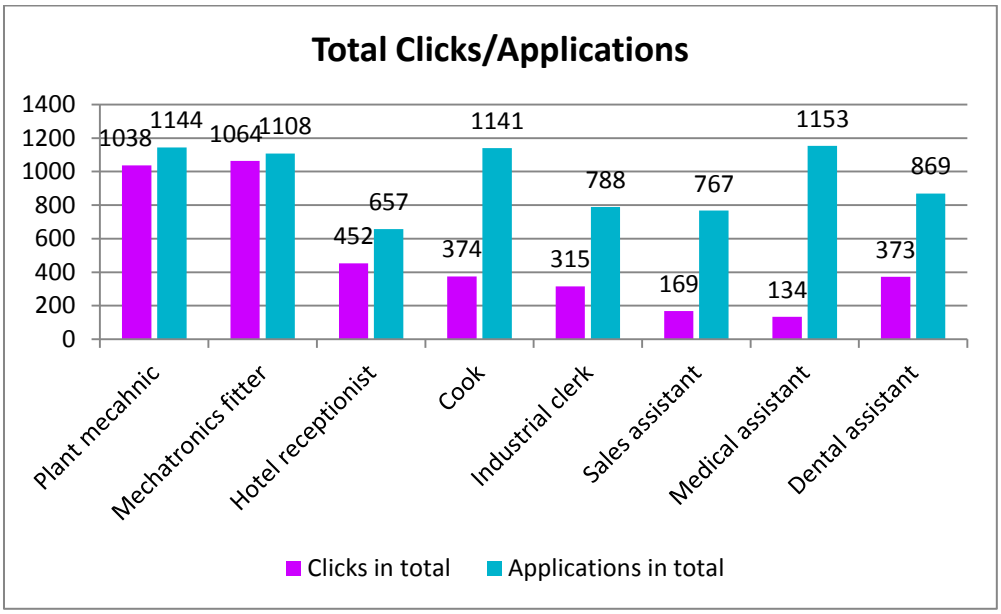
Manfred Weber
(Empfangsleiter)

Note: This reference letter signals social but no linguistic skills (highlighted in yellow). In addition, the letter is adapted to applicant's name, gender, and place of birth (also highlighted in yellow).

A2. Fictitious employers

We anticipated the possibility that prospective employers may decide to search for information about applicants' current employer or that they might even try to get in touch with the current employer. To preclude employers' suspicions about the application and the previous job of the applicants, we created webpages for our fictitious firms that were currently "under construction". The websites had a sophisticated design with different thematic tabs and looked similar to other websites of firms for the occupations in question. In the imprint, we gave contact information (e-mail address and telephone number) under which the current employer could be reached. If an employer decided to call the company, an automatic voicemail would answer and ask to leave a message. Overall, we received eight calls on the mailboxes of the fictitious employers, though none of them left a voice message. The e-mail address was always "info@companyname.de". No former employer was ever contacted via e-mail. However, employers (or other internet users) frequently visited the websites for plant mechanics and mechatronics fitters, but quite rarely for medical or sales assistants (see Figure 10). Visits were counted per click on webpage and not per clicks per IP address, so it is possible that we overestimate the number of visits on the webpage.[9] Furthermore, it is not possible to match clicks by employers with the CVs they received as we are not able to trace back individual IP addresses.

Figure 10: Total clicks per website



[9] The charts also include clicks that came from our own computers to check whether the webpage was online and running. Especially for the plant mechanics and mechatronics fitters, we checked the webpages more frequently in the beginning.

A.3 Place of birth

To determine the place of birth in a systematic way, we programmed an R script (Place of Birth: PoB). The PoB script takes a *German zip code (PLZ)* as input. The output is a list of towns that are geographically close to the input zip code and have more than x inhabitants, where x was originally set to 15,000, but changed to 20,000 in February 2015. We had to adjust this figure over time because in some regions, not all towns with 15,000 inhabitants had a hospital where the applicant could have been born.

To start the procedure, the location of the input town i (town of job vacancy) was geocoded via Google Maps, yielding the longitude and latitude of town i. The PoB script identified the *county* (*Kreis*) in which the zip code of town i is located. It then collected all towns with more than x inhabitants in the county. Then, all towns with more than x inhabitants in all adjacent counties were collected. These towns were combined to form the set of towns A_i, and all their locations were geocoded as well. Subsequently, the distance in kilometers $d(i,j)$ between zip code i and all towns $j \in A_i$ was calculated. The output was a list of the towns j,ordered by increasing distance $d(i,j)$. The list was limited to five towns. Should town i itself have more than x inhabitants, the output was simply town i.

We used two primary data sources:

The *Gemeindeverzeichnis–Informationssystem (*https://www.destatis.de/DE/ZahlenFakten/*LaenderRegionen/Regionales/Gemeindeverzeichnis/Gemeindeverzeichnis.html),* maintained by the *Statistisches Bundesamt*, contains information on every municipality in Germany, including the official municipality key (Amtlicher Gemeindeschlüssel, AGS), the zip code, population and municipality name. From the AGS, it was possible to infer the county in which a municipality is located.

To determine whether counties are adjacent to each other, the the Database of Global Administrative Areas (GADM. http://www.gadm.org) was used. GADM is a database of administrative boundaries for many countries in the world that is maintained by Robert Hijmans at UC Davis and colleagues.