

Discrimination against female migrants wearing headscarves

Doris Weichselbaumer
Johannes Kepler University Linz
doris.weichselbaumer@jku.at

JEL classification: C93; J15; J71

Keywords: Discrimination, Muslim religion, Headscarf, Hiring, Experiment

Abstract

In Western countries, as a result of a large influx of Muslim migrants, the Muslim headscarf has become the subject of major controversy. This paper presents the results of an experiment examining the effect of wearing this headscarf in the labor market. The experiment exploited specifics of the German job application process, in which job seekers typically attach their pictures, and sometimes also reference letters, to their résumés. Attaching photos allowed the experimenter to show an applicant wearing a headscarf, and reference letters allowed favorable personal characteristics to be indicated. Strongly negative reactions were found for applicants wearing headscarves; the presence or absence of reference letters also had effects on the level of discrimination.

I thank Sandra Leitner, Barbara Mühlbacher, Sylvia Sadzinski, Mario Schnalzenberger, Julia Schuster and Sandra Stoll for their excellent research assistance, and Sigrid Betzelt, Claudia Gather, Akin Hizarci, Monika Huesmann, Nuray Karakoc, Sengül Karakoc, Gertraude Krell, Friederike Maier, Thomas Staudacher, Rudolf-Winter-Ebmer and members of the “Türkische Gemeinde, Berlin” for their invaluable help in setting up the experiment and/or their helpful advice and comments. This research was funded by the Austrian Science Fund (FWF), National Research Network S103, Austrian Center for Labor Economics and the Analysis of the Welfare State.

Introduction

In Western countries, Anti-Muslim sentiments and Islamophobia have grown substantially since the beginning of the 21st century. In the aftermath of September 11, 2001, an increase in Anti-Muslim hate crimes has been noted particularly in the U.S. (ADC-RI 2008; FBI). Further, the U.S. Equal Employment Opportunity Commission has reported a rise in religion-related complaints by Muslims since 2001, which suggests that Muslims feel increasingly discriminated against in employment matters. Negative stereotypes depict Muslims as “religious fanatics, terrorists, hostile, evil, barbaric, wild, backward, disorganized people who mistreat and oppress women” (Ghumman and Jackson 2010, p. 6). Such negative stereotypes can, of course, harm Muslims in their work and social lives.

Islamophobia, however, is not confined to the U.S., but has also become more common in European countries, such as Germany, which this study examined more closely. Most Muslims in Germany are members of immigrant communities, and the majority has a Turkish background. During the shortage of labor in the 1960s and early 1970s “guest workers” from Turkey were actively recruited. However, contrary to Germany’s initial intentions, these “guest workers” came to stay. Today they and their descendants form an integral part of the German population.

Estimates of the proportion of Muslims in the German population range from 1.9% to 5% (Spielhaus 2013). According to a representative survey (Bertelsmann Stiftung 2015) that was published just before the recent influx of Syrian Muslim migrants into Germany, 57% of Germany’s non-Muslims perceived Islam as a threat. 40% of the respondents said that they did not feel at home in their own country due to the country’s purported “Islamization”, and 24% believed that Germany should no longer permit Muslims to immigrate. Prejudice against Muslims has also been documented by the World Values Survey, which routinely asks respondents whom, out of a list of people, they would not like to have as neighbors. In 2006, 24.8% of Germans were against the idea of Muslims as neighbors (World Values Survey 2006).¹ Given that other groups do not necessarily trigger such a negative response by Germans (in comparison, “people of a different race”, “foreign workers”, and “people of a different

¹ The World Values Survey USA included this question for the last time in 1999, when 10.7% of the US-American respondents said that they did not want Muslims as neighbors.

religion” were rejected by 7.2%, 13.3%, and 4.3% of the respondents, respectively), it appears that Muslims currently constitute a particularly disliked “outgroup” in Germany.

The Muslim headscarf or hijab is particularly controversial.² As Helbling (2014) showed for 6 European countries (one of them Germany), attitudes of non-Muslims are significantly more negative towards the headscarf than towards Muslims in general. While only about one quarter of survey respondents opposed Muslims, nearly 60% disagreed with the practice of women wearing headscarves. According to another survey, 60% of Germans support a ban of Muslim headscarves in public places (Van der Noll 2010). Indeed, several European countries have implemented such laws (Byng 2010). In some German federal states, teachers were banned from wearing headscarves in 2003, but in March 2015 the Constitutional Court in Karlsruhe ruled that an absolute ban is incompatible with religious freedom and unconstitutional. However, with the recent influx of Muslim migrants in Europe, the discussion on whether women should be allowed to veil their bodies in Muslim style reached a new peak. For example, in the summer of 2016, banning of the “burkini” (full-body swimwear worn by Muslim women) in coastal France stoked the debate, even though France’s highest administrative court soon overturned the ban.³

Given the high prominence of the topic in the public, it is surprising that no economic study has so far examined the effect of wearing the Muslim headscarf in the labor market. The reason for this gap in the literature is, of course, that empirical data on women wearing headscarves is difficult to come by. Hence, a field experiment was conducted that examined the employment chances of females with backgrounds of migration wearing Muslim headscarves. The method of sending out fictitious job applications (often called “correspondence testing”) that was applied in this study is well established and was most famously used by Bertrand and Mullainathan (2004) to examine race discrimination. Recent reviews by Bertrand and Duflo (2016) and Neumark (forthcoming) illustrate the significance of correspondence studies in the experimental literature on labor market discrimination.

The novelty of the current study is that it makes use of the specifics of the German application process, in which job seekers need to submit a large amount of material to be considered serious applicants. In particular, photographs are an integral

² The hijab is a veil that Muslim women wear to cover their head and chest.

³ Since late 2015 German politicians have called increasingly for the niqab (a Muslim cloth that covers the face as part of a hijab) to be banned in public. The niqab, however, is not worn in Turkish communities, which were examined in this study.

part of German job applications. This particularity of the application process allows the use of photographs to indicate the wearing of a headscarf.

Photographs have recently been discovered as valuable indicators of race in experimental settings. For example, Doleac and Stein (2013) as well as Ayres et al. (2015) examined how race affects market outcomes when selling an object online. In both experiments, online advertisements featured photographs that showed a dark- or light-skinned hand holding an object for sale. Both studies found significantly negative effects for African-American sellers. In a non-experimental study, Pope and Sydnor (2011) examined the effect of photographs in online peer-to-peer lending and found that pictures of whites yield better results than those of blacks.

In economic field experiments on labor market discrimination, photographs have rarely been used because most studies have been conducted in countries where they are typically not included in résumés. Exceptions are correspondence tests that measured the effect of physical attractiveness on employment chances (e.g. López Bóo et al. 2013, Rooth 2009, Ruffle and Shtudiner 2015). Weichselbaumer (2003) used photographs to indicate different personality types of female applicants and compared employment chances of those appearing “masculine” and those appearing “feminine”.

Aside from photographs, reference letters from previous employers are also often included in German applications. Reference letters describe the tasks an employee performed in a firm and give information about the employee’s personality. This study made use of this custom and included such letters (apparently from previous employers) in a fraction of all applications (see also Kaas and Manger 2012). These characterized the personality of the candidate (e.g., her friendliness and diligence) in a positive manner. This design therefore allowed measurement of the extent to which women wearing headscarves were able to reduce discrimination if they revealed positive information about their personalities. The study thus followed Bertrand and Duflo (2016), who recommended that correspondence testing experiments should not only demonstrate different treatment of groups, but also study the dynamics of discrimination.

This paper is structured as follows: The next section discusses possible reasons why the Muslim headscarf is the subject of so much debate. It is followed by an overview of empirical studies on the discrimination of Muslims and women who wear headscarves. Subsequently, the experimental design of the study is described, and the experimental results are presented. A concluding discussion follows.

The Muslim headscarf

The particularly strong rejection of the headscarf – also in comparison to Muslims in general – may be due to a number of causes. First, the headscarf is often considered as a symbol not of Islam in general, but of Muslim fundamentalism (Shadid and van Koningsveld 2005), which is frequently associated with intolerance and terrorism. Second, the headscarf has been interpreted as a sign of apparent unwillingness on the part of Muslim migrants to integrate, which has been used by the media to claim that national integration policies have failed (Byng 2010). Finally, the headscarf has been considered as an expression of women’s oppression and the rejection of gender equality (Shadid and van Koningsveld 2005). Opponents of the headscarf maintain that Muslim women wear it because they are forced to by their parents, by (male) family members or by religious leaders. They argue that the liberal state should not accommodate the cultural traditions of minority groups if these infringe upon women’s rights, freedom and autonomy. However, it is debatable whether the headscarf or hijab is in conflict with these values or actually supports them. Many Muslim feminists argue that they wear the headscarf by choice and see it as an empowering practice. Banning the headscarf would then inhibit their autonomy as women. In any case, Westerners rejecting the headscarf as an indicator of the oppression of women may explain some of the observed difference in acceptance of Muslims in general and of the headscarf in particular.

One theoretical framework that is able to analyze dynamics between social groups and may thus help to explain why women wearing the headscarf are especially discriminated against in the labor market is Social Identity Theory. This theory (e.g., Tajfel 1981) argues that people gain self-esteem by attributing a higher value to their own social group (“ingroup”) than to others (“outgroup”). Along this line, Akerlof and Kranton (2000) pointed out that social identity is crucial to our understanding of who should do what job in the economy. For example, men may fear for their male identity if women do the same kind of work as they do. A “distaste” for women (Becker 1971) in male professions and a gender-segregated labor market may be the result. Similarly, non-migrants may feel threatened and fear for their presumed “superiority” if migrants, particularly Muslims, enter their ranks. To prevent this from happening, employers may exclude migrants from their immediate workforce. Particularly women wearing headscarves may experience this rejection because they signal a greater identification with the “Muslim outgroup”. Social Identity Theory suggests that individuals who

identify more strongly with an outgroup are perceived as more “distant” by members of the ingroup. As a result, they suffer from increased levels of discrimination. For example, Derous et al. (2009) found that Dutch and American students evaluated résumés with Arab names more negatively if these signaled a high Arab identification. Also, people whose “stigmas” are visible typically experience more discrimination than others.

Existing empirical evidence of unfavorable treatment of Muslims

So far, little research has focused on labor market discrimination against Muslims.⁴ A sizeable number of correspondence testing experiments have been conducted to examine discrimination against migrants from Muslim countries, but in most cases these did not explicitly examine the effects of religion. In correspondence testing experiments, fake application letters that signal identical qualifications but different demographic characteristics are sent to companies. If, despite identical qualifications, one applicant is invited for interview more often than the other, this is considered discrimination. Experiments like these have been lauded by the OECD and the National Academies’ National Research Council’s Committee on National Statistics as “an important and useful means of measuring discrimination” (Blank et al. 2004, p. 7). In the context of discrimination based on migration background in OECD countries, correspondence tests have examined the employment possibilities of migrants from the Middle East and Turkey, and of Arabs, North-Africans and Pakistanis (for an overview see: Rich 2014; Zschirnt and Ruedin 2016). These studies typically indicated the background of a migrant via the name given in the job application, and documented high levels of discrimination against migrants in most cases.

Only a few correspondence tests focused more explicitly on the Muslim religion of migrants and did more than simply assign a common name from a Muslim country to one of the applicants: Adida et al. (2010) and Pierné (2013) examined the employment chances of migrants with Muslim and Catholic religious affiliations in France. As signals of religion they used work and volunteering experience with Catholic and Muslim organizations in their applicants’ résumés. Both studies found that Muslim

⁴ Experiments testing discrimination outside the labor market have often used a “lost-letter” design and found that finders are more likely to post a lost letter if the addressee was a “local” and not a “Muslim” (e.g., Ahmed 2010).

migrants fared worse than Catholic migrants. Wright et al. (2013) listed involvement in various religious organizations on campus in applicants' résumés and found that in New England Muslim applicants received one-third fewer responses than the control group. Acquisti and Fong (2015) tested the effect of personal information posted online by U.S. American job candidates. In this study, religious affiliation was not signaled through the résumé but through manipulating the "religion" field in a social network profile (Christian, Muslim or no specific denomination) that employers could search for in the web. This manipulation led to significant bias against the Muslim candidate and favoritism for the Christian candidate in Republican but not in Democratic areas.⁵

The effect of wearing a headscarf or a hijab on employment chances has so far been barely researched. In particular, no correspondence study has tested its effect. The studies that do exist were conducted in the fields of psychology and management studies. Unkelbach et al. (2010) authored one of the few laboratory studies that examined the effect of wearing a hijab on personnel decisions.⁶ Students from a German university were presented grades and photos of more than a hundred applicants with different ethnicities, some of them wearing a hijab. The students were instructed to make quick decisions to reject half of the applicants right away while keeping the rest for in-depth evaluation. Results showed that applicants with a hijab were more likely to be rejected, but also the response latencies differed: Wearing a headscarf sped up negative and slowed down positive responses. This implies that participants find it easier to associate women without a hijab with employability than women who wear a hijab.⁷ For the U.S., Ghumman and Jackson (2008) conducted a web-based experiment with students from a Midwestern university. Participants were shown résumés and pictures of male or female job applicants with different religious identifiers (no religious identifier, Christian, Jewish, Muslim). For women, Muslim religion was indicated by the Muslim headscarf, while a turban was used for Muslim men. Interestingly, while male applicants who wore a Muslim turban received the lowest employability ratings of all groups,

⁵ Hedegaard and Tyran (forthcoming) recently examined coworker discrimination between Non-Muslims and Muslims, and found that members of both groups are willing to give up a sizeable amount of their earnings to avoid working with the other group.

⁶ In another study, Unkelbach et al. (2008) examined the effect of wearing a Muslim turban or hijab in a computer shooting game. The authors found that participants shot more often at targets with a Muslim headgear.

⁷ In this respect, the experiment was similar to the Implicit Association Test (IAT) (Bertrand et al. 2005).

women with a Muslim headscarf scored the highest of all groups.⁸ The disadvantage of laboratory experiments, of course, is that their results may not be transferrable to the real world. Participants in a typical laboratory study are often not representative of those making hiring decisions in real life. In particular, students who are often participants in laboratory studies may be more liberal than the average personnel manager. Further, participants in the laboratory do not face any consequences of their fictitious employment decisions: for example, they do not have to work with an individual that they indicated to be “hirable” but whom they dislike. As a result, choices in the laboratory may be systematically less discriminatory than in real life.

Experiments where testers with different demographic backgrounds apply for jobs in person are called “audit studies”. As they are conducted in the field, they more closely resemble real-life employment decisions. Since in audit studies employers are unaware of being part of an experiment, they give unbiased responses. Ghumman and Ryan (2013) sent female confederates to the stores of two shopping malls in the U.S. Midwest to ask for job openings. In half of the trials, the confederates wore a headscarf. While there was no difference in the likelihood of being told about job openings, the applicants with headscarves were significantly less likely to receive a callback and they were also less often given the possibility to complete a job application. In a similar study, King and Ahmad (2010) did not only examine the likelihood of being hired, but also the interpersonal behavior towards applicants.⁹ In their experiment, differences with respect to job recommendations and job callbacks did not reach statistical significance, which may be due to the small sample size.¹⁰ The authors did, however, identify significant effects concerning differences in interpersonal treatment: When the testers wore headscarves, interactions with managers were shorter, and the testers reported higher levels of negativity received from the potential employer. This result suggests that formal discrimination with respect to job opportunities may only be the tip of the iceberg: a lot of discrimination may occur through personal interaction.

Audit studies are of great interest because they examine job opportunities of disadvantaged groups in the field rather than in the laboratory. However, this method

⁸ In another experiment Ghumman and Jackson (2010) asked Muslim women, some of which wore a hijab, to estimate their chances of receiving a job offer – assuming they applied to a given job and held all the necessary qualifications. The authors found that women who wore a hijab had lower expectations of receiving a job offer than those who did not wear a headscarf.

⁹ In this experiment, Muslim attire included not only a black hijab, but also a black abbaya (robe).

¹⁰ 21% (38%) of applicants with (without) religious attire were recommended a position, 24% (35%) received job callbacks, N=81.

also has severe disadvantages (for an overview see Heckman and Siegelman 1993; Heckman 1998). For example, audit studies have necessarily small sample sizes, and often cover only a few localities.¹¹ Correspondence testing, in contrast, not only allows larger amounts of data to be collected, but can also cover a larger geographical area (e.g., multiple cities of a country). More importantly, because in audit studies job seekers talk to managers in person, it is difficult to avoid a “confederate bias”. This means that confederates may be inclined to behave in ways that foster discrimination. Experimenters try to avoid this problem by giving the confederates a script to adhere to. Ghumman and Ryan (2013) have even observed and compared confederates’ behaviors under different conditions and found no significant differences. However, it may be difficult to fully control and assess all aspects of an individual’s behavior. In correspondence testing experiments, matching all aspects of an application except demographics is not a problem. Furthermore, as letters of applications of individuals are strictly standardized, the automated procedure does not leave any room for a confederate or experimenter bias. The method of correspondence testing therefore circumvents the problems typically associated with audit studies (Bertrand and Mullainathan 2004). Of course, correspondence testing experiments measure only the level of discrimination at the first stage of hiring. However, as Rich (2014) showed on the basis of previous studies, the vast majority of discrimination in hiring occurs at this initial stage (between 86-94%).

Advantages of the German setting

The study presented here involved a correspondence testing experiment conducted to investigate whether wearing a headscarf significantly affects the employment chances of women with a Turkish background. To isolate the effect of the headscarf, also the effect of migration background was explicitly examined. The study made use of the large amount of application material that job seekers typically submit in German-speaking countries. In particular, it is common to include not only a letter of application, a résumé, and a diploma or certificate, but also a photograph in an application. Often also one or more reference letters are attached to an application. In other countries, such as the U.S. and the U.K., it is not usual to include photographs in résumés precisely for the reason

¹¹ For example, Ghumman and Ryan (2013) studied 112 trials in two shopping malls, and King and Ahmad (2010) based their analysis on 81 interactions collected in public shopping areas in the metropolitan D.C area.

that they can enable discrimination. As a result, attaching photographs would be considered awkward and raise suspicion. In Germany, however, attaching one's photograph is the norm. Photographs are increasingly used in economic studies (e.g., Doleac and Stein 2013, Ayres et al. 2015, Pope and Sydnor 2011) as a clear yet unsuspecting visual indicator of the characteristics of an individual. In a correspondence study, they provide an ideal opportunity for indicating the applicant wearing a headscarf.

Other information that is typically given in German applications includes date of birth, marital status and citizenship. Aside from photographs, also diplomas and certificates are included in German applications. For example, German high school diplomas detail students' school grades for ten subjects over four semesters in addition to final examination results for four subjects (every subject is evaluated on a 0-15 scale). Finally, they give five different measures of the overall performance of a student in the final years; for instance, the measure "overall points" takes a value between 280 and 840. As this example indicates, German diplomas and certificates give very rich and detailed information regarding students' educational achievements.¹²

Obviously, the necessity to prepare fictitious diplomas makes correspondence testing experiments very time-consuming in the German setting.¹³ An experiment in Germany can thus usually not cover as many jobs as one conducted in a country where no diplomas are required. Further, varying variables of minor interest, for example, years of job experience, is not efficient in the German setting, because this may require the fabrication of an entirely new document for a diploma. However, the advantage of the large amount of information needed in the German setting is that it leaves little ambiguity about an applicant's productivity. This makes statistical discrimination less likely. Phelps (1972) and Arrow (1973) stated that statistical discrimination occurs if information about an individual's productivity is lacking and employers therefore resort to group averages to infer an individual's productivity. Of course, covering all possible aspects of an individual's productivity is impossible, even in a German application. Given the detailed information in a German application (examination results, age, marital status, etc.), however, the level of ambiguity is relatively low compared to other countries.

¹² German certificates of successful completion of an apprenticeship as an office clerk, which are also used in this study, cover five subjects evaluated on a 0-100 scale.

¹³ Legally, the procedure is covered according to Klose and Kühn (2010).

Examination results for different fields of study, which in Germany are provided in diplomas and certificates, are indicative of the cognitive abilities of a candidate. Thus, in this study, educational achievements were held constant for different job applicants. However, statistical discrimination may also be driven by employers' beliefs about the personality traits of an applicant. In particular, it may be that stereotypes ascribe migrants – in particular those who wear headscarves – less favorable personal characteristics. Reference letters are frequently attached to German applications and give information about the personality of a worker. In this study, they were therefore used to test whether statistical discrimination with respect to personality traits exists (Kaas and Manger 2012). The details of the experimental setup are described in the following section.

Experimental design

Identities tested. In this experiment, job applications for three fictitious female characters with identical qualifications were sent out in response to job advertisements: one applicant had a German name, one a Turkish name, and one had a Turkish name and was wearing a headscarf in the photograph included in the application material. To indicate a Turkish background – or a lack thereof – first and last names that are common in Turkey (Meryem Öztürk)¹⁴ and Germany (Sandra Bauer), were used in the résumés.

All applications sent in this experiment included photographs of the same person, who was wearing a headscarf in some of the cases. The three different identities tested in this experiment are shown in *Figure 1*; these are Sandra Bauer (bareheaded), Meryem Öztürk (bareheaded) and Meryem Öztürk with a headscarf.¹⁵ For the fabrication of suitable photographs, a “model” was hired who could pass equally well as German and as Turkish. To ensure that the portrayals of the Turkish women included in the study were realistic, representatives of the Turkish community in Berlin (“Türkische Gemeinde e.V.”) were consulted, who helped with the final selection of photographs. A very modern binding of the headscarf was chosen to signal that the applicant was a young, modern woman who could easily fit into a secular environment. Given that her

¹⁴ Meryem is the Turkish name for „Mary“ and is therefore likely to resonate well with the German Christian population.

¹⁵ Of course, it could also be interesting to test the employment chances of a German woman without migration background who wears a headscarf, thereby signaling that she has converted to Islam. However, given the few women that fit this category in real life, this question is of little empirical relevance.

throat was not covered, it was unlikely that she was particularly strict with respect to her religion – let alone a religious radical. By choosing this representation of a Muslim woman, discrimination against her should be minimized.

Apart from the name and photograph, the applications of the different identities were strictly identical (e.g., with respect to schooling and job training in Germany and German citizenship). Nothing but her name connected Meryem Öztürk to Turkey. There was not even an indication in the résumé that Meryem Öztürk spoke Turkish (in the résumé only English was mentioned in addition to German). Thus, Meryem Öztürk would be perceived as a second (or later) generation migrant, even though she will occasionally be called “migrant” below for simplicity. Similarly, the job application contained no signal of religion aside from the headscarf, which Meryem Öztürk may actually wear only out of cultural tradition.

Randomized application procedure. Most correspondence testing experiments send multiple applications, and some even mail a whole battery of comparable résumés to one company. However, this may be problematic in certain circumstances. As Weichselbaumer (2015) illustrated, sending multiple applications to one firm can lead to seriously biased results, at least in Germany, because the results of correspondence testing experiments have been widely discussed in Germany’s popular media. The method is therefore increasingly well known among personnel managers. Consequently, employers may notice the similarities between fictitious applications, suspect that they are being tested, and present themselves as more minority-friendly than they usually are. The problem is aggravated particularly by the large set of application documents required in Germany. To maintain strict comparability of applicants within the experiment, all documents such as résumés and school reports must be matched. Paired fictitious applications therefore stand out even more from the heterogeneous pool of non-experimental applications.

To avoid detection, this experiment thus followed Ahmed et al. (2013) and Weichselbaumer (2015) in that only one profile was sent to each firm. This has the advantage that all applications sent can be virtually identical except for the variables of interest. It also allows photographs of the same person to be used for all identities. By showing the same “model” in the same pose for all identities, “beauty” was

automatically held constant across ethnic backgrounds.¹⁶ The identity of the applicant (German name, Turkish name/no headscarf, Turkish name/headscarf) was randomly assigned to each application by variation of name and photograph.

Occupations. This study focused on the employment chances of office workers, that is, secretaries, accountants and chief accountants. These three occupations were chosen for a number of reasons. First, although all of these are office jobs, they vary in terms of status: the job of chief accountant is most highly regarded, followed by accountant and secretary in this order. This allows an interesting comparison of results. Second, these jobs cover a large fraction of female employees. As is well known, in many Western countries the majority of female workers are active in only a few occupations. The professions tested in this experiment are among the most typical for women in Germany. In particular, according to the 2011 German census, company management and organization (which covers various office jobs, especially secretarial work) is the most common occupational group for females and covers a significant share of women with migration background (Zensus 2011). A larger fraction of females with migration background is in the low-paying occupations cleaning and sales, but this may be because classic white-collar jobs, such as those tested in this study, are often not available to them. If there is discrimination in white-collar jobs, crowding into low-paid jobs that no autochthonous Germans want to do will be the result. The other reasons for selecting office jobs are methodological. For female applicants, office jobs provide the researcher with a sufficiently large labor demand, which guarantees that a reasonably large data set can be collected. Furthermore, for these jobs written applications are required rather than call-ins, which are common in blue-collar jobs. Finally, for the chosen jobs it was possible to create convincing application material and to provide the necessary attachments.

Application material. Following German norms, each application used in this experiment consisted of the following: a letter of application, a résumé, a photograph, a high school diploma (certifying university entry qualification) and a certificate of successful completion of an apprenticeship as an office clerk. The applications for chief accountants also included a certificate of advanced training in accounting. As has been

¹⁶ As previous studies have shown, beauty considerably affects the likelihood of an applicant being hired (Ruffle and Shtudiner 2015; López Bóo et al. 2013; Rooth 2009).

argued before, the large amount of information given in German applications allows controlling for many different characteristics of an applicant. To test whether unequal treatment is due to statistical discrimination with regard to personality, the following procedure was implemented: only a fraction of the applications had a reference letter attached from the company where the applicant was trained.¹⁷ This letter described not only the tasks the applicant had performed and the areas in which she was trained, it also characterized her achievements and her personality in a positive manner (in particular her fast learning, friendliness, dedication to learning, diligence, reliability and promptness). This experimental design enabled testing of whether migrants fare relatively better if they can signal positive personal characteristics.¹⁸

The documents were created in strict adherence to rules defined by the German Federal Anti-Discrimination Agency that legally cover testing procedures (Klose and Kühn 2010). Since only one application was sent to each firm, the different identities' applications were identical in every respect except name (and thus also phone number and email address) and photograph. The identities were assigned randomly to the job openings. At the time of the experiment, Sandra Bauer and Meryem Öztürk were 27 years old. On average, Turkish migrants marry at a significantly younger age than German non-migrants. Without information to the contrary, employers would therefore expect the migrant to most likely be married and the "non-migrant" to be single. As marital status may affect employment chances, it came in handy that such information is typically provided in German applications. Within the experiment, all candidates were indicated to be single. Meryem Öztürk thus signaled that she did not resemble the "average female with a Turkish background" and was committed to the labor market. This way the measure for discrimination did not become confounded with effects of marital status.

After primary and secondary schooling, the applicants received their high school diplomas (Abitur) and undertook apprenticeships as office clerks. Afterwards, the job candidates had 4 years of job experience as office clerks or accountants. Applicants for the job of a chief accountant had additionally successfully completed an advanced training course that formed the precondition for such a job. All applications included

¹⁷ In Germany, personnel managers usually advise applicants to attach reference letters even if they are not particularly positive, arguing that incomplete applications would not be accepted. The results of this experiment show that, empirically, this is not the case.

¹⁸ In the experiment, reference letters were identical for all applications. Of course, in real life it is likely that discrimination also occurs with respect to job evaluation and that migrant workers who are equally productive as non-migrants receive less favorable reference letters.

information on IT skills and foreign language abilities (English). They also stated that the applicants held driver's licenses. The applicants' hobbies were painting and sports, and they volunteered for a cultural center for which they did the bookkeeping. With more recent migrants, employers are concerned about language proficiency, transferability of schooling and training acquired abroad,¹⁹ and labor market access. However, in this experiment the fictitious résumé did not leave any uncertainties in these dimensions. It was beyond question that all applicants would be fully proficient in the German language, have full labor market access and hold all the commonly accepted qualifications.

Procedure. For about one year, various popular online job portals were searched weekly for relevant job ads in the cities Berlin, Dresden, Frankfurt, Hamburg, Cologne, Munich and Stuttgart. The application documents were combined in one electronic file that was sent to the companies by email. The applications provided full contact information: an address, an email address and a cell phone number which would forward to voicemail. This allowed companies to get in touch with the fictitious applicants. Applications were restricted to openings for full-time jobs whose basic requirements matched the profile of the fictitious candidates. Because only one application was sent in response to each ad, it was necessary to make sure that the vacancies included in the study were comparable. Each ad was therefore carefully checked for its suitability and coded for how well it matched the standardized profile of the applicant. This information was used later as a control variable. If a personnel recruitment agency was involved in the selection process, the corresponding ad was omitted. To avoid detection, each company was contacted only once, even if it advertised multiple suitable jobs in the course of the experiment. All remaining companies that welcomed applications by email were contacted. Firms interested in an applicant could leave her a voicemail message or contact her by email or by regular mail. Invitations for interview and inquiries stating interest in the candidate were coded as positive responses (callback). If the applicant was invited for interview, the appointment was canceled within a day to avoid any inconvenience to the company.

¹⁹ Oreopoulos (2011) illustrated that foreign experience and foreign education is valued less by employers.

Results

In the course of the experiment, a total of 1,474 applications were sent in response to job advertisements of companies located in Germany (for summary statistics see *Table 1*). The overall callback rates for the three identities are presented in *Figure 2*. The applicant with the German name – Sandra Bauer – was the most successful. She received positive feedback from 18.8% of all companies she applied to, followed by the applicant with the Turkish name (Meryem Öztürk) without headscarf, who was contacted by 13.5% of the companies. This difference is statistically significant ($t=2.2633$, $p=0.012$, one-sided t -test). As hypothesized, the applicant with the Turkish name who wore the headscarf fared the worst – she received positive feedback from only 4.2% of all companies contacted. The difference to both other identities is strongly significant (Ms. Öztürk with headscarf versus Ms. Bauer: $t=7.3975$, $p=0.000$, one-sided; Ms. Öztürk with headscarf versus Ms. Öztürk without headscarf: $t=5.2395$, $p=0.000$, one-sided). This result implies that the candidate with the headscarf had to send 4.5 times as many applications as an identical applicant with a German name and no headscarf to receive the same number of callbacks for interview. This suggests that there is discrimination against female migrants – particularly if they wear a headscarf.

While callback rates differed between jobs (accountants are in higher demand than secretaries), discrimination was always highest against Meryem Öztürk when she was wearing a headscarf (see *Table 2*). However, as the relative callback rates (“ratio”) illustrate, some differences in the extent of her disadvantage emerged across occupations. In the profession of chief accountant, Meryem Öztürk with a headscarf had to send almost 7.6 times as many applications as Sandra Bauer to receive an invitation for interview. The corresponding number for secretaries is “only” 3.5. That the level of discrimination was highest in the profession of chief accountant is remarkable because within the experiment chief accountant was the occupation with the highest qualifications necessary. It is often assumed that discrimination decreases with increasing level of education. However, wearing a headscarf may be considered even more inappropriate for positions with higher occupational status.²⁰

²⁰ This result is also at odds with the notion that there is less discrimination when recruitment is difficult (Baert et al. 2015). According to Becker (1971), one would assume that discrimination can only occur if employers have a wide range of potential workers to choose from. The high callback rates for accountants and chief accountants indicate that positions in these occupations are more difficult to fill than those of

Table 3 gives the callback rates at the city level. Discrimination against the applicant wearing a headscarf was uniformly significant at the 5% level (except in Dresden, where hardly any relevant job openings were available); discrimination against the migrant without headscarf was significant only in Berlin.

Probabilities of a callback

A linear probability model was estimated to examine the probability of a positive callback more thoroughly. Since applications were sent to different firms, differences in invitation rates may have been due to differences in firm or job characteristics of the positions rather than discrimination, even though ads had been carefully checked for comparability. Consequently, to account for such potential differences, a number of characteristics of job vacancies and the corresponding firms were controlled for. For example, applications in particular cities, sectors, or months may be systematically more successful. Further, only a proportion of applications included a reference letter. Employers may consider application packages that include a reference letter more complete and treat them more favorably. The results of the linear probability model are presented in *Table 4*. The different specifications incrementally include the following control variables: occupational dummies (chief accountant, accountant and secretary), city dummies (Berlin, Dresden, Frankfurt, Hamburg, Cologne, Munich and Stuttgart), sector dummies (trade, public services, manufacturing and services), time dummies (in quarters), attachment of a reference letter (yes/no), variables concerning size and internationality of a company (“firm characteristics”) and a number of job characteristics that were derived from the requirements stated in the job advertisement – more specifically, whether the job required a special qualification that the standardized profile did not have (“special qualification requirement”, SQR) or whether the job involved team and/or customer contact.

The results in *Table 4* demonstrate that the effects found for the different identities are robust to adding these control variables. Hence, the unfavorable treatment of both job applicants with Turkish names cannot be explained by particular characteristics of either the specific firms or the jobs these women applied to in the randomized application design. In comparison to the German applicant, the bareheaded

secretaries, yet discrimination against women who wear headscarves was not reduced in these occupations.

applicant with the Turkish name (Öztürk) was consistently by 5-6 percentage points less successful, while the migrant applicant who also wore a headscarf was even 15 percentage points less likely to receive a callback for interview. A test for the equality of coefficients shows that the differences between coefficients are highly significant (at the 1% level). Interestingly, even though the inclusion of a reference letter is often considered standard practice in Germany, the results show that its attachment does not generally lead to a positive effect. In the current setting, this may be due to the fact that these letters were constructed to give a positive but not an excellent impression. Also, if a job ad required an unusual qualification (SQR) that the applicant did not have, this did not reduce her callback rate on a significant level. The reason is probably that the fake applicants held typical qualifications for a particular position. If they did not have a special qualification, only few others would.

Reasons for unequal treatment

This study also tested a set of hypotheses concerning the reasons for unequal treatment. As previously pointed out, unequal treatment may be due to statistical discrimination. In this experiment, application materials provided employers with a vast amount of information concerning the qualifications of an applicant (e.g., detailed school grades in various subjects during schooling and apprenticeship) rendering statistical discrimination less likely in that respect. However, Muslim migrants may also suffer from statistical discrimination if they are ascribed personality traits that are deemed less productive in the job. To test this, a written reference from a previous employer was attached to a proportion of the applications. It described the applicant and her personality in favorable terms. If statistical discrimination occurs due to incomplete information regarding personality, attaching an identical reference letter should help migrants more than non-migrants, as negative beliefs about migrants are counteracted by positive characterizations.

Another reason for discrimination was proposed by Becker (1971). He argued that not only employers, but also customers and coworkers may hold discriminatory preferences. If customers and coworkers do not accept migrant workers, these cannot become fully productive and are less profitable to an employer. To test this hypothesis, the variable “team or customer contact” (TCC) was created, which captures whether a job advertisement mentioned that such interaction was required at the job.

It is also possible that larger firms discriminate less because they follow a more standardized, and therefore more objective, evaluation process than small companies (Kaas and Manger 2012). For this reason, in the present study the dummy variable “small firm” measured whether a firm had 1-20 employees. Companies that operate at an international level may also behave differently towards migrant applicants because they may be more familiar with diverse groups of people. This was examined via the dummy “international firm”. Further, discrimination may be triggered by employers’ beliefs that migrants are less proficient in the German language. The variable “German” measured whether a job ad emphasized that proficiency in German was required for the job. In this experiment, fluency in German was signaled for all applicants by schooling and training received in Germany. Their callback rates should therefore have been unaffected by German language requirements. However, employers may use statements regarding German proficiency in job ads as subtle cues against migrant employees in general, irrespective of their language skills. Whether this is really the case was tested empirically. Another interesting question would be how the level of discrimination is affected by the attitudes of recruiters (Rooth 2010). However, it is difficult to obtain proxies for these. In Germany, clerical assistants usually handle the administration of the application process, and thus the identities (sex/migration background) of recruiters with decision-making power remain unknown.²¹ For this reason, a variable was constructed that examined whether a company actively seeks migrant workers. Some firms explicitly asked for cross-cultural competence in their job advertisements or signaled that they are an intercultural team – this was captured by the dummy variable “interculturalism”. It is likely that “intercultural firms” discriminate less. Finally, some job advertisements mentioned that a neat appearance was required for the job (“appearance”). If conventional looks are of great importance in a certain job, firms may be particularly unwilling to hire a migrant woman with a headscarf.

To test these hypotheses, the linear probability model was extended to cover not only the full set of control variables (now also including “German”, “Interculturalism” and “Appearance”), but also the interaction of each control variable with the variable “Öztürk” and with “Öztürk with headscarf”. The results are presented in *Table 5*. The interaction effects illustrate which variables affect migrants differently than

²¹ Since no information on decision makers was available in this study, it could not be tested whether individuals with migration background discriminate less or engage in reverse discrimination. Using a different methodology, Stirling Hedegaard and Tyran (forthcoming) found that both Muslims and Non-Muslims discriminated against coworkers of the other group.

autochthonous Germans, and therefore help to identify reasons for differential treatment. As can be seen from *Table 5*, few interaction effects are significant at conventional levels. As mentioned above, Meryem Öztürk with a headscarf was discriminated against to a greater degree in the occupations accountant and chief accountant. Otherwise, only the attachment of a positive reference letter increased the migrants' relative invitation rates. This suggests that some prejudice is indeed at work when no references are provided. Interestingly, the reason why this effect occurred in this experiment was not that the migrant improved her absolute chances by attaching a reference letter, but that the non-migrant applicant reduced hers. This is probably the case because the letter described the applicant as good but not excellent, which may be less than what is expected from a non-migrant (but fits the expectations for a migrant). Without reference letters, the non-migrant therefore benefits from positive stereotypes (positive statistical discrimination), while attaching them makes no difference to the migrant. Small firms may discriminate more, but these effects are – at best – significant at the 10% level. No effects were found concerning special qualification requirements (SQR), team or customer contact (TCC), internationality of a company, and German language requirements. The hypothesis that firms which expect a “neat appearance” from their applicants discriminate more against the headscarf than others was not confirmed. A marked effect, however, was identified for firms that asked for cross-cultural competence or signaled that they were an intercultural team, even if this effect was significant only at the 10% level and only for Ms. Öztürk with a headscarf. This indicates that, everything else being equal, Meryem Öztürk with a headscarf increases her relative callback chances by 37-39 percentage points if she applies to such a firm. Note, however, that only few firms fall into this category (about 2%).²²

Discussion

A heated debate is being led in the West about the apparently inferior position of women in Muslim (migrant) culture. However, little discussion takes place about how Muslim women are actually treated by the Western majority population. To test discrimination of Western employers against Muslim migrant women, this study investigated employment

²² Finally, also the hypothesis that the share of Turkish migrants in the population of a city affects the level of discrimination was tested (not shown). However, no significant effect was found with respect to discrimination against Meryem Öztürk with or without a headscarf. Results are available from the author.

opportunities of female office workers in Germany by sending identical applications to firms that differed only in the applicants' names and photographs. The results show that applicants with a Turkish name are treated unfavorably at a significant level, in particular when their photograph shows them wearing a Muslim headscarf. Everything else being equal, a female with a Turkish name who wears a headscarf has to send 4.5 times as many applications as an applicant with a German name and no headscarf to receive the same number of callbacks for interview. This level of unequal treatment is one of the highest that has been measured in the experimental discrimination literature (Hofer et al. 2013), and it is even higher for chief accountants and accountants. This massive rejection of the headscarf is all the more remarkable given the very modern and progressive binding used in the current experimental setting. That the headscarf shown in the application photograph did not cover the applicant's throat signaled that she is not particularly strict with respect to her religion. This should have minimized the level of discrimination measured. Discrimination is likely to be even higher against a more traditional binding of the headscarf.

It is often assumed that the unfavorable labor market position of Muslim migrant women is the result of a lack of education, job experience or dedication to the labor market. Turkish migrants do indeed often have less favorable labor market characteristics than autochthonous Germans (see, e.g., Seebaß and Siegert 2011). However, in this experiment, the women with Turkish names had the exact same qualification as the women without migration history. Since in German applications an unusual amount of information is revealed, it was possible to control for a large set of variables that may affect hiring decisions (e.g., schooling outcomes and personal characteristics such marital status, age and citizenship). For example, the high school diploma alone included 41 measures of educational success, each on a scale of 0-15 or more. Of course, the fictitious applicants did not only attend the same schools, they also received exactly the same grades. As all applicants grew up in Germany, where they received all their schooling and training, it was obvious to employers that they were fluent in German. The many years of job experience signaled that they were eager to work. Since in Germany it is common to indicate one's family status, all applicants in this experiment identified as single, signaling to potential employers that they had no family obligations. Given these characteristics and credentials, the scale of discrimination is astounding. Attention discrimination (Bartos et al. 2016) may be one

explanation of the low regard for characteristics and achievements of the migrant applicants.

When applications also included reference letters that described the candidate in positive but not excellent terms, this reduced the level of discrimination. Interestingly, in this case, it was not that migrants improved their absolute chances by attaching a reference letter, but non-migrants did no longer benefit from positive stereotypes towards them (positive statistical discrimination). As a result, migrants did relatively better. In the experimental résumé, applicants were not given any spells of unemployment. Given the high level of discrimination against the woman with a headscarf, for her to hold such a résumé is in fact unlikely unless she is substantially more able than the native. This suggests that discrimination may still be underestimated in this experiment.

Critics may argue that in this study the woman wearing the headscarf was not discriminated against because of the headscarf per se but because headscarves are considered inappropriate for application photographs. However, as Unkelbach et al. (2010, p. 382) stated, it does not matter whether discrimination is due to wearing headscarves in application photos not being considered business savvy or due to a distaste for headscarves: “The hijab disadvantages the wearer, whether the underlying processes are explicit (i.e., it is directly used to discriminate between applicants) or implicit (i.e., it activates negative stereotypes that influence decisions about applicants.)”.²³ In any case, attaching a photograph with a headscarf is not an uncommon practice. Syed and Pio (2010) reported that, even in countries where photos are not commonly included in job applications, some headscarf-wearing women deliberately attach their photographs to screen for non-discriminatory companies. Thus, they avoid wasting time meeting employers who, on principle, do not hire women who wear headscarves.

To examine why and when discrimination occurs, the effects of various firm and job characteristics were examined in this study. For example, it may be that small firms discriminate more or that international firms discriminate less. However, firm and job characteristics helped little to explain unequal treatment. The dislike for migrant

²³ Tilcsik (2011) and Weichselbaumer (2015) argued similarly with respect to gays and lesbians who out themselves by mentioning volunteer work at a gay and lesbian organization in their résumés. If an applicant is treated unfavorably because an employer believes they violated a social norm by mentioning volunteer work at a gay and lesbian organization, this is still discrimination based on sexual orientation. Mentioning such an engagement can only be considered embarrassing if same-sex orientation itself is regarded as objectionable.

applicants identified by this study may be explained by Social Identity Theory. To German employers, migrants represent an outgroup, which may lead to their rejection. Social Identity Theory predicts that women wearing the headscarf should be particularly negatively affected, as they signal a strong identification with the Muslim outgroup and may be perceived as particularly distant. Exactly this was found in this study. Callback rates for women who wear the headscarf are meager, even in comparison to bareheaded migrant women. Obviously, women wearing headscarves are a social group that experiences particular rejection in the German labor market.

While in the late 20th century Muslim migrants to Germany came mainly from Turkey, Germany is currently facing a large influx of Muslim migrants that are predominantly from Syria. Many of the migrant women are wearing headscarves. If integration is to succeed, the existing discrimination in the labor market needs to be eliminated.

References

- Acquisti, Alessandro, and Christina M. Fong. 2015. "An Experiment in Hiring Discrimination via Online Social Networks." *Carnegie Mellon University Working Paper*.
- ADC-RI. American-Arab Anti-Discrimination Committee Research. 2008. *Report of Hate Crimes and Discrimination against Arab-Americans: 2003-2007*. Washington.
- Adida, Claire L, David D Laitin, and Marie-Anne Valfort. 2010. "Identifying Barriers to Muslim Integration in France." *Proceedings of the National Academy of Sciences* 107(52): 22384–90.
- Ahmed, Ali M. 2010. "Muslim Discrimination: Evidence from Two Lost-Letter Experiments." *Journal of Applied Social Psychology* 40(4): 888–98.
- Ahmed, Ali M., Lina Andersson, and Mats Hammarstedt. 2013. "Are Gay Men and Lesbians Discriminated against in the Hiring Process?" *Southern Economic Journal* 79(3): 565–85.
- Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–53.
- Arrow, Kenneth. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets*, eds. Orley Ashenfelter and Albert Rees. Princeton: Princeton University Press, 3–33.
- Ayres, Ian, Mahzarin Banaji, and Christine Jolls. 2015. "Race Effects on Ebay." *Rand Journal of Economics* 46(4), 891-917.
- Baert, Stijn, Bart Cockx, Niels Gheyle, and Cora Vandamme. 2015. "Is There Less Discrimination in Occupations Where Recruitment Is Difficult?" *Industrial and Labor Relations Review* 68(3): 467–500.
- Bartos, Vojtech, Michal Bauer, Julie Chytilova, and Filip Matejka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review* 106(6), 1437-1475.
- Becker, Gary S. 1971. *The Economics of Discrimination*. Second Edi. Chicago: University of Chicago Press.
- Bertelsmann Stiftung. 2015. *Religion Monitor. Understanding Common Ground. Special Study of Islam, 2015. An Overview of the Most Important Findings*. Gütersloh.
- Bertrand, Marianne, and Esther Duflo. 2016. "Field Experiments on Discrimination." *NBER Working Paper* (22014).
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94(4): 991–1013.
- Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95(2), 94-98.
- Blank, Rebecca M., Marilyn Dabady, and Constance F. Citro. 2004. *Social Sciences Measuring Racial Discrimination: Panel on Methods for Assessing Discrimination*. Washington: National Academies Press.
- Byng, Michelle D. 2010. "Symbolically Muslim: Media, Hijab, and the West." *Critical Sociology* 36(1): 109–29.
- Derous, Eva, Hannah-Hanh Nguyen, and Ann Marie Ryan. 2009. "Hiring Discrimination against Arab Minorities: Interactions between Prejudice and Job Characteristics." *Human Performance* 22(4): 297–320.
- Dickel, Doris. 2002. *Einwanderungs- und Asylpolitik der Vereinigten Staaten von Amerika, Frankreichs und der Bundesrepublik Deutschland*. Wiesbaden: Springer.

- Doleac, Jennifer, and Luke Stein. 2013. "The Visible Hand: Race and Online Market Outcomes." *Economic Journal* 123, F469-F492.
- FBI. Federal Bureau of Investigation. "Hate Crime Statistics (Various Years)." <https://www.fbi.gov/about-us/cjis/ucr/hate-crime/> (May 8, 2015).
- Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics* 116(1): 351–77.
- Ghumman, Sonia, and Linda Jackson. 2008. "Between a Cross and a Hard Place: Religious Identifiers and Employability." *Journal of Workplace Rights* 13(3): 259–79.
- Ghumman, Sonia, and Linda Jackson. 2010. "The Downside of Religious Attire: The Muslim Headscarf and Expectations of Obtaining Employment." *Journal of Organizational Behavior* 31(1): 4–23.
- Ghumman, Sonia, and Ann Marie Ryan. 2013. "Not Welcome Here: Discrimination towards Women Who Wear the Muslim Headscarf." *Human Relations* 66(5): 671–98.
- Heckman, James J. 1998. "Detecting Discrimination." *Journal of Economic Perspectives* 12(2): 101–16.
- Heckman, James J., and Peter Siegelman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." In *Clear and Convincing Evidence: Measurement of Discrimination in America*, eds. Michael Fix and Raymond Struyk. Washington: Urban Institute Press, 187–258.
- Hedegaard, Morten Storling, and Jean-Robert Tyran. forthcoming. "The Price of Prejudice" *American Economic Journal: Applied Economics*.
- Helbling, Marc. 2014. "Opposing Muslims and the Muslim Headscarf in Western Europe." *European Sociological Review* 30(2): 242–57.
- Hofer, Helmut, Gerlinde Titelbach, Doris Weichselbaumer, and Rudolf Winter-Ebmer. 2013. *Diskriminierung von MigrantInnen Am Österreichischen Arbeitsmarkt. IHS Projektbericht*. Wien: Bundesministerium für Arbeit, Soziales und Konsumentenschutz.
- Kaas, Leo, and Christian Manger. 2012. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment." *German Economic Review* 13(1): 1–20.
- King, Eden B., and Afra S. Ahmad. 2010. "An Experimental Field Study of Interpersonal Discrimination toward Muslim Job Applicants." *Personnel Psychology* 63(4): 881–906.
- Klose, Alexander, and Kerstin Kühn. 2010. *Die Anwendbarkeit von Testing-Verfahren Im Rahmen Der Beweislast, § 22 Allgemeines Gleichbehandlungsgesetz*. Berlin: Antidiskriminierungsstelle des Bundes.
- López Bóo, Florencia, Martín A. Rossi, and Sergio S. Urzúa. 2013. "The Labor Market Return to an Attractive Face: Evidence from a Field Experiment." *Economics Letters* 118(1): 170–72.
- Neumark, David. forthcoming. "Experimental Research on Labor Market Discrimination." *Journal of Economic Literature*.
- 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): 148–71.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62(4): 659–61.
- Pierné, Guillaume. 2013. "Hiring Discrimination Based on National Origin and Religious Closeness: Results from a Field Experiment in the Paris Area." *IZA Journal of Labor Economics* 2(4): 1-15.

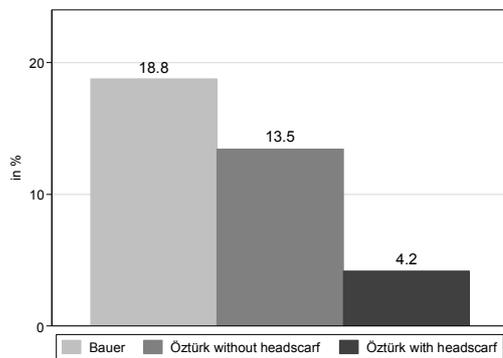
- Pope, Devin, and Justin Sydnor. 2011. "What's in a Picture? Evidence of Discrimination From Prosper.com." *Journal of Human Resources* 46(1): 53-92.
- Rich, Judith. 2014. "What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000." *IZA Discussion Paper* (8584).
- Rooth, Dan-Olof. 2009. "Obesity, Attractiveness, and Differential Treatment in Hiring. A Field Experiment." *The Journal of Human Resources* 44(3): 710–35.
- Rooth, Dan-Olof. 2010. "Automatic Associations and Discrimination in Hiring: Real World Evidence." *Labour Economics* 17: 523-534.
- Ruffle, Bradley J, and Ze'ev Shtudiner. 2015. "Are Good-Looking People More Employable?" *Management Science* 61(8): 1760–76.
- Seebaß, Katharina, and Manuel Siegert. 2011. *Migranten am Arbeitsmarkt in Deutschland. Integrationsreport*. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Shadid, Wasif, and Pieter Sjoerd van Koningsveld. 2005. "Muslim Dress in Europe: Debates on the Headscarf." *Journal of Islamic Studies* 16(1): 35–61.
- Spielhaus, Riem. 2013. *Muslimen in Der Statistik. Wer Ist Muslim Und Wenn Ja Wie Viele? Gutachten Im Auftrag Des Mediendienst Integration*. Berlin: Mediendienst Integration.
- Syed, Jawad, and Edwina Pio. 2010. "Veiled Diversity? Workplace Experiences of Muslim Women in Australia." *Asia Pacific Journal of Management* 27(1): 115–37.
- Tajfel, Henri. 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge: Cambridge University Press.
- Tilcsik, András. 2011. "Pride and Prejudice: Employment Discrimination against Openly Gay Men in the United States 1." *American Journal of Sociology* 117(2): 586–626.
- Unkelbach, Christian, Joseph P. Forgas, and Thomas F. Denson. 2008. "The Turban Effect: The Influence of Muslim Headgear and Induced Affect on Aggressive Responses in the Shooter Bias Paradigm." *Journal of Experimental Social Psychology* 44(5): 1409–13.
- Unkelbach, Christian, Hella Schneider, Kai Gode, and Miriam Senft. 2010. "A Turban Effect, Too: Selection Biases against Women Wearing Muslim Headscarves." *Social Psychological and Personality Science* 1(4): 378–83.
- Van der Noll, Jolanda. 2010. "Public Support for a Ban on Headscarves: A Cross-National Perspective." *International Journal of Conflict and Violence* 4(2): 191–203.
- Weichselbaumer, Doris. 2003. "Sexual Orientation Discrimination in Hiring." *Labour Economics* 10, 629-642.
- Weichselbaumer, Doris. 2015. "Testing for Discrimination against Lesbians of Different Marital Status: A Field Experiment." *Industrial Relations* 54(1): 131–61.
- World Values Survey. 2006. "Data." <http://www.worldvaluessurvey.org/> (May 6, 2015).
- Wright, Bradley R.E., Michael Wallace, John Bailey and Allen Hyde. 2013. "Religious Affiliation and Hiring Discrimination in New England: A Field Experiment." *Research in Social Stratification and Mobility* 34: 111-126.
- Zensus. 2011. "Data." <https://ergebnisse.zensus2011.de> (January 3, 2017).
- Zschirnt, Eva, and Didier Ruedin. 2016. "Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015." *Journal of Ethnic and Migration Studies* 42(7): 1115–34.

Figures

Fig. 1. Names and photographs, indicators for identity



Fig. 2. Callbacks by identity



Tables

Table 1: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Callback	1474	0,12	0,32	0	1
Bauer	1474	0,32	0,47	0	1
Öztürk	1474	0,34	0,47	0	1
Öztürk with headscarf	1474	0,34	0,47	0	1
Secretary	1474	0,60	0,49	0	1
Accountant	1474	0,24	0,43	0	1
Chief accountant	1474	0,17	0,37	0	1
Small firm	1467	0,10	0,30	0	1
International firm	1468	0,54	0,50	0	1
Reference letter	1474	0,32	0,47	0	1
Special qualification requirements	1474	0,17	0,38	0	1
German proficiency required	1474	0,24	0,43	0	1
Team and/or customer contact	1474	0,65	0,48	0	1
Interculturalism	1474	0,02	0,15	0	1
Appearance	1474	0,03	0,18	0	1
Berlin	1474	0,11	0,31	0	1
Dresden	1474	0,01	0,12	0	1
Frankfurt	1474	0,16	0,37	0	1
Hamburg	1474	0,23	0,42	0	1
Cologne	1474	0,18	0,39	0	1
Munich	1474	0,20	0,40	0	1
Stuttgart	1474	0,10	0,30	0	1
Branche dummy: Services	1474	0,65	0,48	0	1
Branche dummy: Trade	1474	0,19	0,40	0	1
Branche dummy: Public services	1474	0,04	0,20	0	1
Branche dummy: Production	1474	0,12	0,32	0	1

Table 2: Callback rates and relative callback rates by occupation

	All		Chief accountant		Accountant		Secretary	
	Callback rate	Ratio†	Callback rate	Ratio†	Callback rate	Ratio†	Callback rate	Ratio†
German name (bareheaded)	18.8% (N=474)		27.3% (N=77)		29% (N=107)		12.8% (N=290)	
Turkish name (bareheaded)	13.5% (N=498)	1.39**	19% (N=84)	1.44	19% (N=121)	1.53*	9.6% (N=293)	1.33
Turkish name with headscarf	4.2% (N=502)	4.48***	3.6% (N=83)	7.58***	5.8% (N=121)	5.00***	3.7% (N=298)	3.46***

Note: † the ratio (or “relative callback rate”) is defined as follows: callback rate German name/callback rate Turkish name (bareheaded or with headscarf).

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Callback rates by city

	1. Bauer	2. Öztürk	3. Ö. headscarf	N	(1)	(2)	(3)
	invit.rate	invit.rate	invit.rate		sig. 1-2	sig. 1-3	sig. 2-3
Berlin	24.1%	11.1%	5.4%	163	**	***	
Dresden	20%	0%	0%	21		*	-
Frankfurt	22.4%	15.4%	4.5%	242		***	***
Hamburg	13.9%	15.7%	3.6%	339		***	***
Cologne	14%	11.6%	2.5%	268		***	**
Munich	21.3%	14.1%	8.2%	293	*	***	*
Stuttgart	21.4%	10.6%	0%	148	*	***	***

Note: Results from a t-test: *** p<0.01, ** p<0.05, * p<0.1

Column (1) compares the callback rates of Ms. Bauer and Ms. Öztürk (without headscarf), (2) of Ms. Bauer and Ms. Öztürk with a headscarf, and (3) of Ms. Öztürk without and with a headscarf.

Table 4: Probability of a callback (Linear Probability Model)

Variables	(1) Callback	(2) Callback	(3) Callback	(4) Callback
Öztürk	-0.053*** (-2.59)	-0.054*** (-2.63)	-0.054*** (-2.64)	-0.057*** (-2.79)
Öztürk with headscarf	-0.146*** (-7.13)	-0.147*** (-7.24)	-0.148*** (-7.24)	-0.149*** (-7.29)
Chief accountant		0.074*** (3.19)	0.075*** (3.20)	0.077*** (3.28)
Accountant		0.090*** (4.43)	0.090*** (4.44)	0.088*** (4.30)
Reference letter			-0.012 (-0.51)	-0.012 (-0.51)
<i>Job characteristics</i>				
Special qualification requirements (SQR) (Yes = 1)				-0.031 (-1.37)
Team and/or customer contact (TCC) (Yes = 1)				0.004 (0.21)
<i>Firm characteristics</i>				
Small firm (Yes = 1)				0.004 (0.15)
International firm (Yes = 1)				0.001 (0.08)
Time dummies	no	yes	yes	yes
City dummies	no	yes	yes	yes
Sector dummies	no	yes	yes	yes
Constant	0.188*** (12.78)	0.129*** (3.84)	0.136*** (3.77)	0.139*** (3.58)
No. of observations	1,474	1,474	1,474	1,467
R ²	0.034	0.061	0.062	0.063
Adjusted R ²	0.033	0.051	0.051	0.049

Note: t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is a dummy for receiving a positive callback. The reference category for identity: Bauer; for occupation: secretary.

Table 5: Reasons for unequal treatment (Linear Probability Model)

Variables	(1) Callback	(2) Callback	(3) Callback	(4) Callback	(5) Callback
Öztürk	-0.060*** (-2.91)	-0.093*** (-3.74)	-0.072** (-2.38)	-0.088* (-1.84)	-0.053 (-1.01)
Öztürk with headscarf	-0.153*** (-7.45)	-0.179*** (-7.15)	-0.123*** (-4.07)	-0.183*** (-3.70)	-0.114** (-2.13)
Chief accountant	0.078*** (3.28)	0.078*** (3.26)	0.143*** (3.43)	0.081*** (3.38)	0.161*** (3.74)
Chief accountant*Öztürk			-0.047 (-0.82)		-0.075 (-1.28)
Chief accountant*Öztürk with headscarf			-0.144** (-2.53)		-0.157*** (-2.66)
Accountant	0.087*** (4.14)	0.086*** (4.09)	0.151*** (4.14)	0.088*** (4.18)	0.164*** (4.41)
Accountant*Öztürk			-0.057 (-1.14)		-0.077 (-1.49)
Accountant*Öztürk with headscarf			-0.133*** (-2.69)		-0.145*** (-2.82)
Reference letter	-0.010 (-0.42)	-0.072** (-2.06)	-0.067* (-1.94)	-0.076** (-2.17)	-0.074** (-2.11)
Reference letter*Öztürk		0.103** (2.35)	0.099** (2.25)	0.110** (2.47)	0.109** (2.44)
Reference letter*Öztürk with headscarf		0.079* (1.81)	0.076* (1.73)	0.085* (1.89)	0.082* (1.83)
Special qualification requirements (SQR)	-0.033 (-1.47)	-0.033 (-1.47)	-0.037* (-1.65)	-0.065* (-1.72)	-0.072* (-1.87)
SQR*Öztürk				0.047 (0.86)	0.054 (0.97)
SQR*Öztürk with headscarf				0.057 (1.06)	0.054 (1.00)
Team and/or customer contact (TCC)	0.004 (0.21)	0.003 (0.18)	0.004 (0.20)	-0.008 (-0.27)	-0.010 (-0.34)
TCC*Öztürk				0.031 (0.72)	0.033 (0.77)
TCC*Öztürk with headscarf				-0.005 (-0.11)	0.000 (0.01)
Small firm	0.007 (0.25)	0.009 (0.31)	0.013 (0.45)	0.086* (1.71)	0.093* (1.86)
Small firm*Öztürk				-0.114 (-1.57)	-0.120* (-1.66)
Small firm*Öztürk with headscarf				-0.112 (-1.63)	-0.118* (-1.72)

International firm	-0.000 (-0.01)	0.001 (0.03)	0.001 (0.06)	0.010 (0.31)	0.010 (0.32)
International firm*Öztürk				-0.020 (-0.47)	-0.021 (-0.49)
International firm*Öztürk with headscarf				-0.006 (-0.14)	-0.006 (-0.14)
German	-0.008 (-0.39)	-0.006 (-0.29)	-0.004 (-0.20)	0.013 (0.34)	0.037 (0.98)
German*Öztürk				-0.066 (-1.34)	-0.090* (-1.77)
German*Öztürk with headscarf				0.020 (0.41)	-0.020 (-0.39)
Interculturalism	0.134** (2.40)	0.131** (2.37)	0.135** (2.44)	-0.181 (-0.97)	-0.194 (-1.04)
Interculturalism*Öztürk				0.306 (1.49)	0.321 (1.57)
Interculturalism*Öztürk with headscarf				0.367* (1.81)	0.390* (1.92)
Appearance	0.017 (0.38)	0.022 (0.47)	0.018 (0.39)	0.003 (0.04)	0.029 (0.37)
Appearance*Öztürk				0.007 (0.05)	-0.019 (-0.15)
Appearance*Öztürk with headscarf				0.036 (0.35)	-0.019 (-0.18)
Time dummies	yes	yes	yes	yes	yes
City dummies	yes	yes	yes	yes	yes
Sector dummies	yes	yes	yes	yes	yes
Constant	0.136*** (3.45)	0.153*** (3.80)	0.123*** (2.94)	0.153*** (3.32)	0.113** (2.36)
No. of observations	1,467	1,467	1,467	1,467	1,467
R ²	0.067	0.071	0.078	0.079	0.087
Adjusted R ²	0.051	0.054	0.059	0.054	0.059

Note: t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is a dummy for receiving a positive callback. The reference category for identity: Bauer; for occupation: secretary.