

Ethnic Discrimination in Germany's Labour Market – a Field Experiment*

Leo Kaas[†] Christian Manger[‡]

January 25, 2011

forthcoming in *German Economic Review*

Abstract

This paper studies ethnic discrimination in Germany's labour market with a correspondence test. We send two similar applications to each of 528 advertisements for student internships, one with a Turkish-sounding and one with a German-sounding name. A German name raises the average probability of a callback by about 14 percent. Differential treatment is particularly strong and significant in smaller firms at which the applicant with the German name receives 24 percent more callbacks. Discrimination disappears when we restrict our sample to applications including reference letters which contain favourable information about the candidate's personality. We interpret this finding as evidence for statistical discrimination.

JEL classification: C93, J71

Keywords: Ethnic discrimination, Hiring discrimination, Correspondence test

*We are grateful to Anja Achtziger, Gerald Eisenkopf, Lisa Green, Till Grossmass, Thomas Hinz, Normann Lorenz, Anne Madden, three referees and an editor for numerous helpful comments and remarks. We also thank Fabian Fink, Jutta Obenland and Christiane Ralf for research assistance.

[†]Department of Economics, University of Konstanz, Box D145, 78457 Konstanz, Germany. E-mail: leo.kaas@uni-konstanz.de

[‡]Department of Economics, University of Konstanz, 78457 Konstanz, Germany. E-mail: christian.manger@uni-konstanz.de

1 Introduction

Discrimination against individuals with respect to ethnicity, gender, or religion has a wide impact on labour market outcomes, including job opportunities, promotions and earnings. The extent to which a society is plagued by discrimination is hard to measure, however. On the one hand, empirical studies based on field data can deliver measures for earnings inequality but they cannot unveil discriminatory practices in the hiring process, for example. Moreover, field data are not collected in a controlled environment, so that the researcher has typically much less information about worker characteristics than is available to the employing firm. Hence it is difficult to disentangle the effects of actual productivity differences from employer discrimination. On the other hand, laboratory experiments on discrimination can be conducted in fully controlled settings. What is measured there, however, is the behaviour of subjects in a sterile environment; how far the findings of such experiments extend to employer–worker interactions in real–world labour markets is unclear. Field experiments are a compromise between these approaches, combining the advantages of controlled experiments with a field context.¹ With regard to measuring hiring discrimination, the correspondence test method is a sensible way to measure the initial response of employers to varying characteristics of artificial applicants.

This paper describes a correspondence test conducted in the German labour market for student internships. We examine the hiring opportunities of individuals with a Turkish migration background. In 2009 Germany had about 2.5 million persons (3 percent of its population) with a Turkish ethnic background (Statistisches Bundesamt (2010)). Predominantly in the 1960s, migrants from Turkey came to Germany to enlarge its labour force. Forty years later, the children and grand–children of these workers, born and raised in Germany, represent a substantial share of Germany’s workforce. Uhlendorff and Zimmermann (2009) show that these workers need significantly more time to find new jobs than natives which gives rise to higher unemployment rates for workers with migration background. While the individual ethnic identity has some effects on labour supply, particularly on participation (Constant and Zimmermann (2009)), on search effort and on reservation wages (Constant

¹See Harrison and List (2004) for a survey on field experiments.

et al. (2009)), the role of worker ethnicity for labour demand has not been explored. It is evident, however, that a full integration of migrant workers into the German labour market necessitates equal employment opportunities.

Our experiment generates a snapshot of ethnic discrimination in one particular subsection of the labour market. In particular, we send more than one thousand applications to firms that offer internships for students of economics and management science. In practice such internships serve as important prerequisites for access into regular jobs. Although they are not well paid, a student who has successfully completed an internship gains job experience and significantly improves his employment opportunities after graduation.² Today, the completion of at least one internship is commonly expected and is often considered as a “foot in the door”. We also believe that the high-skill labour market is a particularly challenging one for students with migration backgrounds, since second generation immigrants are under-represented in high-skill jobs, whereas they account for a substantial share of current university students.³

To each of 528 job advertisements, we send two similar applications, one with a Turkish sounding name (“Fatih Yildiz” or “Serkan Sezer”) and one with a typical German name (“Dennis Langer” or “Tobias Hartmann”).⁴ Importantly, the name is the only distinguishing characteristic of the applicant with Turkish ethnical background. That is, all applicants have German citizenship and they were born and educated in Germany, and all of them specify “German” as their mother tongue.

²Gault, Redington, and Schlager (2000) find that business graduates in the U.S. gain faster access into employment and obtain higher salaries when they have relevant internship experience. We are not aware of related studies for Germany.

³On the one hand, Liebig (2007, Table 9) shows that second generation immigrants are under-represented in high-skill professions such as legislators, senior officials, managers, or professionals. For instance, compared to natives, an average second generation immigrant is 63% less likely to be employed as a professional and 58% more likely to work as a sales worker. On the other hand, Isserstedt et al. (2009, Chapter 15) report that 11% of the students at German universities had a migrational background in 2009 (compared to 23% of the German population in the age group 20–24), not including foreign students with school education outside Germany.

⁴The preparation of bogus applications and the deception of employers are unavoidable necessities to analyze discrimination in a controlled field environment. To minimize the cost on employers, we promptly and politely reply to a positive response by withdrawing the applicant’s further interest in the position. See Riach and Rich (2004) for a discussion of the ethical considerations involved in the correspondence test method.

With that design, we are able to isolate the effect of ethnicity from possible language effects. We create two slightly different types of applications with similar grades, soft skills, and photographs. For every job vacancy, applicant names are randomly assigned to the two different applications. Furthermore, the amount of information provided by the students varies between the two different applications. In particular, one application type contains reference letters stating favourable information about the candidate’s personality traits such as conscientiousness and agreeableness. We use this variation to explore the effect of statistical discrimination versus taste-based discrimination.⁵

The field experiment shows that an application with a German-sounding name is on average 14% more likely to receive a callback. Discrimination is more pronounced among smaller firms: firms with less than 50 employees give “Dennis” and “Tobias” about 24% more callbacks than “Fatih” and “Serkan”. We also find evidence that a reasonable fraction of the differential treatment can be attributed to statistical discrimination: while there is almost no difference in callback probabilities for the application that is equipped with personality information (37.4% with a German name vs. 36.9% with a Turkish name), the absence of such information in the other application gives rise to significant differences in callback probabilities (41.8% with a German and 32.5% with a Turkish name).

Our results can be compared with those from other field studies that explore ethnic discrimination in other countries.⁶ Across these studies, the measured degree of differential treatment varies remarkably with the respective context. In the U.S. labour market, Bertrand and Mullainathan (2004) show that applications with White-sounding names receive 50% more callbacks for interviews than those with African-American-sounding names. They find that the racial gap is uniform across

⁵Several studies explore the determinants of prejudice against foreigners or persons with migration background. For instance, Fertig and Schmidt (2010) find that education is a key determinant for such attitudes. Dustmann and Preston (2001) discuss the interplay between the ethnic composition of neighborhoods and the local attitudes towards ethnic minorities: A high concentration of minorities may lead to perceptions of threat and alienation, but can also reduce unrealistic negative preconceptions.

⁶There are several field studies that test discrimination against other characteristics, e.g. Neumark (1996), Goldin and Rouse (2000)) and Petit (2007) for gender, Banerjee et al. (2009) for caste and religious groups in India, Rooth (2007) for obesity and Weichselbaumer (2003) for sexual orientation.

occupation, industry, and employer size. A similarly huge difference in callback rates is documented by Drydakis and Vlassis (2007) who analyze the labour market opportunities of Albanians in Greece. They find that Albanians not only face a 43% smaller chance of access to occupations, but they are also significantly less likely to be registered with insurance coverage. For Arabs in Sweden, Carlsson and Rooth (2007) find that every fourth employer discriminates against the minority. Wood et al. (2009) conduct a correspondence test in Britain, finding that there are considerable gaps in callbacks between whites and several different ethnical groups. For the German labour market, Goldberg et al. (1996) conducted various field experiments to analyze ethnic discrimination of migrants, also finding substantial differences in callback rates. However, the legal framework has changed since 1994 when their experiments were conducted.⁷ Further, Goldberg et al. analyze the situation of migrants, that is, workers that were born in Turkey and with Turkish mother tongue, whereas we focus on German citizens with a Turkish migrational background.

Compared to these other studies, the quantitative extent of discrimination in our experiment seems to be small. But there are several reasons why international comparisons between such numbers are difficult. First, our applications contain much more information than those in the studies cited above. In particular, in Germany it is common practice to submit not only a resume, but also copies of all school and university certificates; these certificates provide detailed hard-evidence information about various skills. Second, we focus on a high-skill segment of the labour market; it is unclear whether ethnic discrimination in Germany is stronger in other segments of the labour market.⁸ Third, individual attitudes to ethnic minorities depend also on a country's immigration policy.⁹

⁷In August 2006, the German government passed an equal-treatment law ("Allgemeines Gleichbehandlungsgesetz") which prohibits discrimination in hiring and promotion (among others) and which encompasses several personal characteristics (including ethnicity). As this legislation does not entail any specific requirements on employment policies, it is unclear whether and how hiring processes have changed since 2006.

⁸Carlsson and Rooth (2007) find large differences across occupations, with differences in callback rates varying from 10 percent (computer professionals) to over 100 percent (shop sales and cleaning). But even in their study, there are several high-skill segments of the labour market with much higher discrimination rates than in our study. Large callback differences across occupations are also observed by Wood et al. (2009).

⁹As Bauer, Lofstrom, and Zimmermann (2000) show, natives in countries that receive many

There are also a few laboratory experiments on racial/ethnic discrimination. Fershtman and Gneezy (2001) examined behaviour in a trust game where one person has the option to pass part of his/her money to an unknown partner whose name signals a different ethnic background (of "Ashkenazic" or "Eastern" origin, among undergraduate students in Israel). They find evidence for considerable mistrust against men of Eastern origin. Castillo and Petrie (2010) conduct a public-good experiment with group formation to investigate whether people discriminate by race (black/white) or by gender. To explore statistical discrimination, they also vary information on average past contributions to the public good before individuals decide group membership. They find that race becomes irrelevant once private information becomes available, which suggests that discrimination is mainly statistical. For a survey of laboratory experiments on discrimination, see Anderson, Fryer, and Holt (2006).

The rest of this paper is structured as follows. Section 2 describes the experimental design and Section 3 presents and discusses the results. Section 4 concludes.

2 Experimental Design

The Vacancies

The experiment focuses on a specific segment of the labour market, in particular the market for internships for students in economics and business. This restriction allows us to completely automate the application process by sending serial letters and to eliminate potential bias caused by individually written and adjusted applications.

We also restrict our study to internships within Germany.

We only apply for internships with a duration ranging from 6 weeks to 6 months and consider all reasonable vacancies posted at large internet job sites (such as *monster.de* and *jobscout24.de*).¹⁰ The field experiment was conducted in two waves, the

non-economic migrants are mainly worried about increasing crime rates. In contrast, natives in countries that select migrants by skill worry more about labour market competition. Such country-specific effects may limit the transferability of results.

¹⁰Weitzel et al. (2010) report that 72% of the new hires of the 1000 largest firms in Germany in 2009 can be attributed to internet job search compared to only 13.7% attributed to job advertisements in print media. For the U.S., Stevenson (2009) reports that, in the year 2001, 69% of the unemployed with college degrees (all age groups) were searching for jobs online. Given the current

first one covering vacancies posted in December 2007 and January 2008, the second one considering positions posted in December 2008. Although all firms explicitly search for students in economics or business programmes, the internships are quite heterogeneous. This concerns employer characteristics (size, sector, location) and the division within the firm (human resources, marketing, finance or controlling). Most of the vacancies are at firms with 500 or more employees. Large firms and banks are the most relevant employers for graduates of economics and business, and they are more likely to post their vacancies on large internet job sites. Further, there are only few vacancies from East Germany since most large corporations have their headquarters in West Germany.

Applications

All applicants are second-year students of age 21 or 22. Our applications are quite comprehensive compared to other field studies on hiring discrimination. In particular, each application contains a cover letter, a curriculum vitae, a high-school certificate and a certificate documenting university grades in the first year. In the German labour market, this amount of information is necessary to achieve a reasonable callback rate. In fact, most employers explicitly request copies of all these certificates. Omitting them would bias our results significantly since only quite unattractive employers would respond to an application that contains only a resume.

We create two slightly different types of resumes, labeled type A and type B, such that we can send two applications to each firm.¹¹ Students of both types were born, raised and educated in West Germany, but in different regions, one in the state of North Rhine–Westphalia, the other one in Baden–Wuerttemberg. After graduating from school, both skip military service, work at a summer job and then attend different universities. Both aim at a bachelor’s degree in business economics. At the time of the application, they are in their third semester and are applying for an internship during their fourth semester. The school and university certificates document grades between “good” and “very good”, so that the students range in

relevance of the internet, especially for young adults, it is unlikely that the focus on online job advertisements biases our results.

¹¹CVs of both types are contained in Appendix B.

the top 25 percent of their peer groups.

Both types are fluent in English and they have basic knowledge of one further foreign language. Since applications are identical for applicants with a Turkish and a German name, the applicant with the Turkish name is a native German speaker and he also does not report any command of Turkish in his CV. Both types have reasonably developed computer skills. Moreover, both applicants state in their CV that they had two minor part-time jobs, but while type B provides two letters of reference from previous employers, type A does not add any related documents. The two reference letters contain positive statements about the student's personality (affability, commitment, capacity for teamwork, conscientiousness). This variation in information is used to analyze the effects of statistical discrimination. In all other dimensions, applications of types A and B are rather similar; particularly, there are only minor deviations in individual school and university grades. Hence, the variation in information about personality is the decisive informational difference between these types.

Finally, all applications are completed by a type-specific photograph. While in many countries firms do not request or even oppose photographs in applications, they are still very common (and sometimes requested) in Germany. Omitting them would again bias the results. We select photographs that fit both a native German student as well as one with a Turkish migration background. Each resume type has its own unique photograph, while names are randomly assigned to the types.¹² This guarantees that the choice of the photograph has no overall effect on the callback rates of an individual name.

We systematically adjust certain details in the cover letter of every single application to match job-specific features relating to the sector of the firm or the division of the internship. This adjustment is performed automatically using serial letters. That is, we design for each type and each division a specific paragraph matching the basic requirements and interest for that division. For example, when applying for an internship in the human resource division, an applicant of type A would explain why he is interested in human resource management and he also states that he

¹²This random procedure indeed generated a balanced distribution of name-type combinations in nearly all of the subcategories listed in Table 1: with only two exceptions, the difference between the two combinations was less than 10 percent.

intends to pursue a master’s degree in human resource management after finishing his bachelor degree. Our approach to standardize serial letters reflects the trade-off between maximizing the callback rate by adjusting the application to the specific requirements of each post and generating unbiased letters (and also reducing our workload).

The Names

The first application that is supposed to be sent to each individual firm is randomly assigned a type (A or B) and a name (German or Turkish), while the second application then is assigned the complementary type and name.¹³ We choose “Dennis Langer” (first wave) and “Tobias Hartmann” (second wave) as names for the native German candidate. The first names as well as the surnames belong to the 30 most common ones for the birth years 1986 to 1988 in Germany. The name of the applicant from the ethnic minority is “Fatih Yildiz” in the first wave and “Serkan Sezer” in the second one. Both, first names and surnames, are common for male descendants of Turkish immigrants in Germany. It is also evident for every human resource manager to deduce the ethnic background from these names. We did not explicitly check for connotations of names regarding their social background, but we assured that the names do not contradict common sense, are very stereotype or exhibit other peculiarities (e.g. ruling out combinations between an Anatolian first name and a Kurdish surname). We also made sure that none of these names corresponds to a real person in Germany’s student web network (*studivz.de*).

The application process

We create an individual e-mail address for each name and prepare mobile phones with name-specific numbers. However, we do not answer incoming calls directly, but firms are redirected to the voice mail where they are politely asked to leave their names and contact information. Additionally, we made arrangements such that answers by regular mail to both candidates and addresses were redirected to us. Thus firms could contact the applicants via mail, e-mail and phone.

¹³In particular, we simulate an urn model to determine the type and name of the application that is sent first. All four possible combinations of types and names are used equally often.

When applying, all documents (cover letters, CVs, additional documents such as certificates) are automatically merged to one single pdf file and are e-mailed to the firm. After sending the first application we waited two days before sending the second one. A reasonable fraction of firms (especially larger ones that expect to receive many applications) do not accept applications by e-mail, but instead require the applicants to complete several pre-defined online forms. In these cases, the forms are filled out with the respective applicant's information and our documents were attached as pdf files whenever possible.

After applying for the vacancy, we registered callbacks in the subsequent four months. A callback is defined as any action of a firm that signals interest in the respective resume, including offers for interviews, direct job offers and leaving contact information on the voice mail. In contrast, automatic responses confirming the receipt of our applications are not considered as callbacks, as well as written requests for additional information which were answered whenever possible. For every reaction of a firm, be it a callback, rejecting the applicant, or a request for more information, we collect the date and the type of reaction. Within 24 hours of that reaction, we politely withdraw the candidate's further interest in the position.

3 Results

The application process at each of the 528 firms can have several outcomes, summarized in Table 1. Either the firm shows no positive reaction towards any of the candidates (column 1), or at least one applicant receives a callback (column 2). In the latter case, either both receive a callback (column 3), or the firm prefers one of the students, either the one with the German (column 4) or with the Turkish name (column 5). Column (6) calculates net discrimination as the difference in callbacks between applications with a German and with a Turkish name, expressed as a percentage of those observations where at least one candidate received a callback. This definition of net discrimination treats those cases where no candidate receives a callback as a non-observation. Riach and Rich (2002) discuss whether a negative answer (or no answer at all) for both candidates should be considered as equal treatment or as a non-observation. On the one hand, if a firm rejects both applicants (or does not even send an answer) this could be considered equal treatment; that

is, somebody reviewed both applications and found them not suitable for the job. On the other hand, it is also conceivable that the firm was not even considering the applications, for instance because the vacancy has already been filled.

We conduct a standard χ^2 test of the hypothesis that the two possible outcomes of unequal treatment (that is, columns (4) and (5)) are equally likely, where column (7) shows the test for the restricted sample which treats no-callback employers as non-observations.¹⁴ If the application with the German name is preferred significantly more often than the Turkish one, the H_0 of equal treatment is rejected. As this test considers only observations with differential treatment, observations of firms that either decline both applicants or callback both applicants are irrelevant. For completeness, column (8) shows the χ^2 test after inclusion of the no-callback employers. Here the difference between applications with German and Turkish names remains weakly statistically significant for the full sample, but for all subsamples the difference is statistically not significant.

The first row shows the aggregate results of the field study. Out of the 258 firms that accepted at least one application, 29.1% contacted only the German, and 19.0% only the Turkish candidate, while 51.9% contacted both. This corresponds to a callback rate of 34.7% for the Turkish student and of 39.6% for the German student. In other words, while the German candidate has to write 15 applications to obtain 6 callbacks, the Turkish candidate must send 17 applications for the same number of callbacks. This difference is significant at the 5%-level, but it is remarkably small compared to studies on employment discrimination of ethnic minorities in other countries, such as Albanians in Greece (Drydakis and Vlassis (2007)), Arabs in Sweden (Carlsson and Rooth (2007)) or Afro-Americans in the U.S. (Bertrand and Mullainathan (2004)); see Section 4 for further discussion.

Regarding the effect of firm characteristics on discrimination, we note that only four (8.9%) out of the 45 firms with less than 50 employees made an offer only to the Turkish candidate, while 12 (26.7%) preferred the German over the Turkish applicant. We would assume that smaller firms with fewer vacancies have a less standardized recruitment process. This leaves more scope for individual preferences

¹⁴The result is the same as the χ^2 of a McNemar test which considers the null hypothesis that a dichotomous and paired outcome variable (i.e. the reaction dummies to German and Turkish candidates) have the same distributions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	no callback	at least one	both	only G	only T	net discr.	χ^2 restr.	χ^2 total
all firms (528)	51.14 (270)	48.86 (258)	51.94 (134)	29.07 (75)	18.99 (49)	10.08 (26)	5.45**	2.74*
Firm Size								
large (> 500) (379)	51.19 (194)	48.81 (185)	49.19 (91)	29.73 (55)	21.08 (39)	8.65 (16)	2.72*	1.46
medium (59)	52.54 (31)	47.46 (28)	50.00 (14)	28.57 (8)	21.43 (6)	7.14 (2)	0.29	0.15
small (< 50) (90)	50.00 (45)	50.00 (45)	64.44 (29)	26.67 (12)	8.89 (4)	17.78 (8)	4.00**	1.46
Location								
South (292)	48.29 (141)	51.71 (151)	52.32 (79)	29.14 (44)	18.54 (28)	10.60 (16)	3.56*	1.84
East (40)	62.50 (25)	37.50 (15)	40.00 (6)	40.00 (6)	20.00 (3)	20.00 (3)	1.00	0.58
other (196)	53.06 (104)	46.94 (92)	53.26 (49)	27.17 (25)	19.57 (18)	7.60 (7)	1.14	0.54
Division								
Marketing (173)	43.35 (75)	56.65 (98)	54.08 (53)	29.59 (29)	16.33 (16)	13.26 (13)	3.76*	1.99
Controlling (91)	52.75 (48)	47.25 (43)	46.51 (20)	25.58 (11)	27.91 (12)	-2.33 (-1)	0.04	0.02
Finance (84)	48.81 (41)	51.19 (43)	46.51 (20)	32.56 (14)	20.93 (9)	11.63 (5)	1.09	0.63
Human Res. (122)	59.84 (73)	40.16 (49)	51.02 (25)	32.65 (16)	16.33 (8)	16.33 (8)	2.67	1.24
Consulting (32)	65.63 (21)	34.37 (11)	72.73 (8)	9.09 (1)	18.18 (2)	9.09 (-1)	0.33	0.07
other (26)	46.15 (12)	53.85 (14)	57.14 (8)	28.57 (4)	14.29 (2)	14.29 (2)	0.67	0.32
Industry								
Fin. Services (105)	41.90 (44)	58.10 (61)	40.98 (25)	34.43 (21)	24.59 (15)	14.75 (6)	1.00	0.71
Consulting (98)	56.12 (55)	43.88 (43)	69.77 (30)	16.28 (7)	13.95 (6)	2.33 (1)	0.08	0.02
Manufacturing (150)	55.33 (83)	44.67 (67)	56.72 (38)	26.87 (18)	16.42 (11)	10.45 (7)	1.69	0.72
IT & Telecom (31)	58.06 (18)	41.94 (13)	69.23 (9)	23.08 (3)	7.69 (1)	15.38 (2)	1.00	0.28
Public Services (11)	63.64 (7)	36.36 (4)	25.00 (1)	50.00 (2)	25.00 (1)	25.00 (1)	0.33	0.26
other (133)	47.37 (63)	52.63 (70)	44.29 (31)	34.29 (24)	21.43 (15)	12.86 (9)	2.08	1.29

Notes: This table shows the distribution of the firm responses, absolute numbers are in parentheses. Column (1) reports the fraction of firms that gave none of the candidates a callback, so the remainder in column (2) contacted at least one applicant. Firms that gave both candidates a positive reaction, column (3), are considered as equal treatment, while the rest preferred either the candidate with the German or the one with the Turkish name, columns (4) and (5). Net discrimination is calculated as (6)=(4)-(5). Columns (7) and (8) contain the χ^2 for equality between (4) and (5) for the restricted sample (2) and for the full sample (H_0 : Turkish and German candidates are equally likely to receive a callback at any of the paired observations). * denotes significance at the 10%-level, ** significance at the 5%-level.

Table 1: Callbacks conditional on names and firm characteristics.

of the human resource manager to influence hiring decisions. Indeed, discrimination is less prominent in larger firms, presumably since their recruitment processes follow pre-defined rules.

However, differences with respect to other firm characteristics are rather limited. We note that there are twice as many firms in East Germany¹⁵ that favour the candidate with the German name; however, there are both too few vacancies as well as too low callback rates to give rise to significant results. If we consider jobs in different divisions, differential treatment is weakly significant for jobs in the marketing department. But discrimination is not particularly strong in this division; instead there are simply many internships in marketing divisions, which implies that only the large sample is responsible for significant discrimination here. On the other hand, differences in callback probabilities are remarkably strong for internships in human resource departments, where the number of employers which favour the German candidate is twice as large as the number of employers preferring the Turkish one, but, due to the small sample size, the difference is not statistically significant.

Table 2 shows the callback rates for the different names conditional on the resume type. For applications of type B which include reference letters containing information regarding the candidate's personality, Turkish and German applicants achieve almost identical callback rates. However, for applications of type A (without personality information), the minority student receives only for 32.5% of his applications a callback, while the German student is successful in 41.8% of his applications. The difference in the number of callbacks between German and Turkish students of type A is significant at the 5%-level, while it is much smaller and non-significant for students of type B. We cautiously interpret this as evidence for statistical discrimination (Arrow (1973)): the difference in callbacks decreases with the provision of information about the applicant's character. Note however that "information" and any other type characteristics are perfectly correlated. Hence the observation that callback differences are smaller at applications of type B could, in principle, also be due to other characteristics of a type B application. However, it is not true

¹⁵ "East Germany" is defined as the states Berlin, Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt and Thuringia. "South Germany" includes the states Baden-Wuerttemberg, Bavaria, Hesse, Rhineland-Palatinate, and Saarland.

that type B is generally more attractive: in contrast, German applicants of type B receive *fewer* callbacks than German applicants of type A. More importantly, as we have argued above, the other differences between type A and type B applicants are rather minor and they should not be expected to generate the large observed difference in discrimination rates.¹⁶ It is worthy to note that there is no evidence that firms from different regions favour a certain type, although both types come from different regions within Germany.

callback rate	type A	type B
German name	41.8% (110 out of 263 applications)	37.4% (99 out of 265 applications)
Turkish name	32.5% (86 out of 265 applications)	36.9% (97 out of 263 applications)

Notes: Applications of type B contain two reference letters with information about the applicant’s personality, those of type A do not.

Table 2: Callback rates for different types.

Probit Estimation

In order to disentangle the effects of different employer and worker characteristics, we conduct several probit estimations with the callback dummy as the dependent variable. Table 3 summarizes the various configurations. Columns (1) to (5) show regression outputs using the full sample, while columns (6) and (7) omit firms that did not callback at least one of the applicants. In the basic model (1) we regress the callback only on a constant and a dummy for the Turkish name and find a weakly significant (at the 10 percent level) negative effect of the Turkish name on the callback probability. The interpretation of the result is that a Turkish name reduces the probability of a callback on average by about 14%. This effect is robust to adding a dummy for the location of the firm and the workers’s type in (2), a dummy for firms with less than 50 employees in (3) and dummies for the sector of

¹⁶The study program at the university of type A is slightly more business oriented. This could explain why type A is a bit more attractive for most employers. Nonetheless, it is rather implausible that this feature explains the strong difference in callback rates between ethnical types.

the firm as well as for the division of the internship in (4).¹⁷ The coefficient of the dummy for South Germany is (weakly) significantly positive, suggesting that it is easier to find an internship in regions with less unemployment and more job openings. There is no evidence that one of the two application types is more successful than the other in general. That is, the type dummy is always small and insignificant. However, once we add an interaction term between the name (ethnicity) and the type dummies in (5), this term catches all the disadvantage of the Turkish name, and the coefficient of the name dummy becomes small and insignificant. Again this confirms the conclusion from Table 2 that the difference in callbacks occurs predominantly at applicants of type A where personality information is not provided. The reduction in information seems to hurt applicants with a Turkish name. Again, we cautiously interpret this finding as evidence in support of statistical discrimination.

In columns (6) and (7) we restrict the sample to those job openings at which at least one applicant received a callback. Now the adverse impact of the Turkish name becomes much more prominent and is significant at the 1%-level. Note that the name coefficient even stays significant at the 5%-level when we add in column (7) the same controls as for the unrestricted sample, including the interaction term between type and name. In contrast to column (5), the interaction term plays no important role in the restricted sample. The explanation is that there is a large number of Turkish applicants of type A among those firms that make no offer to any candidate.

Other means of discrimination

Discrimination can manifest itself in several ways, not only in different callback rates. If a firm has no interest in any of the applicants, but the applicant with the German name receives a polite message declining him, while an applicant with a Turkish name is just ignored, we would consider this as discriminatory treatment, even though it takes no direct impact on the job search outcome.

¹⁷We tested several other variables but found all of them (and their interactions with the ethnicity dummy) insignificant. This includes a “first application” dummy, a dummy on “top employers” (obtained from a student survey), a dummy for the employer signing the “Charta of Diversity” at the time of the experiment (www.diversity-charter.org), and a dummy for employers located in a large city region.

Callback	(1)	(2)	(3)	(4)	(5)	(6)	(7)
constant	-0.26*** (0.06)	-0.34*** (0.08)	-0.36*** (0.08)	-0.04 (0.22)	-0.04 (0.22)	0.84*** (0.14)	0.94*** (0.34)
Turkish name	-0.13* (0.08)	-0.13* (0.08)	-0.13* (0.08)	-0.13* (0.08)	0.003 (0.11)	-0.33*** (0.12)	-0.39** (0.18)
Southern Germany		0.13* (0.08)	0.14* (0.08)	0.14* (0.08)	0.13 (0.08)	0.03 (0.12)	-0.03 (0.13)
type A		-0.0001 (0.08)	-0.0001 (0.08)	0.002 (0.08)	0.14 (0.11)	-0.03 (0.12)	-0.08 (0.18)
small firm			0.13 (0.10)	0.14 (0.12)	0.14 (0.12)	0.27 (0.17)	0.19 (0.18)
sector dummies				yes	yes		yes
division dummies				yes	yes		yes
Turk*Type A					-0.27* (0.16)		0.1 (0.25)
Observations	1056	1056	1056	1056	1056	516	516

Notes: Each column represents a probit regression with the callback dummy as dependent variable. Robust standard errors are in parentheses. The full sample contains all 1056 applications, while the restricted sample omits firms that did not callback at least one of the applicants. * denotes significance at the 10%-level, ** significance at the 5%-level, and *** at the 1%-level. Several other interactions have been tested and not found significant.

Table 3: Probit regression with callback dummy as dependent variable.

Table 4 provides a first snapshot. The most noticeable difference in treatment of the two applicants takes place at firms that callback one candidate and do not respond to the second one; indeed, 28 firms showed interest in the applicant with the German name and ignored the Turkish applicant, while only 12 firms contacted the Turkish and ignored the German one. This difference is significant at the 5%-level. We would consider this as the strongest form of discrimination: the firm has a vacant post, it shows interest in the German candidate and does not even answer the Turkish one.

German / Turkish	callback	rejection	no reaction	Σ
callback	134	47	28	209
rejection	37	179	21	237
no reaction	12	20	50	82
Σ	183	246	99	528

Table 4: Differential treatment in answers.

As another type of discrimination, we check how long applicants with different names have to wait for the firm's decision, and what determines that waiting period. Most

firms react only a few workdays after we sent the application (see Table 5). A callback is, on average, received after 11 workdays, while a rejection takes on average 17 workdays. An applicant with a Turkish name waits on average slightly longer for a callback (11.3 workdays with a Turkish and 10.7 workdays with a German name). For rejections, the difference is even smaller (17.0 workdays with a Turkish and 17.4 workdays with a German name). Both differences are not significant. However, small firms react significantly faster in general. For example, a small firm needs on average 6.2 workdays less to callback an applicant than a large firm. We would assume that small firms are faster in general because there is often just one decision maker, whereas larger firms are more likely to have standardized recruitment processes where applications have to go through many hands. To complement the descriptive evidence of Tables 4 and 5, Appendix A contains the result of a survival analysis with a multinomial logit model.

	callback			rejection		
	German name	Turkish name	average	German name	Turkish name	average
all	10.7	11.3	11.0	17.4	17.0	17.2
small	6.1	7.0	6.5	9.9	12.2	11.0
medium	7.8	7.5	7.7	19.6	19.0	19.3
large	12.4	13.0	12.7	18.1	17.3	17.7

Table 5: Average reaction time in workdays.

The role of the business cycle

Our field study also permits us to explore the impact of the business cycle. We conducted the study in two waves, the first one in Winter 2007/08 and the second one in Winter 2008/09. The macroeconomic situation changed substantially between these two dates. In particular, the accelerating financial crisis strongly affected the real economy in late 2008. By then, most research institutes published forecasts for Germany’s GDP growth in 2009 at -3 percent or less. Although the German labour market remained remarkably robust during the course of 2009, we would expect that firms adjusted their hiring behaviour in January 2009 relative to the year before. However, the aggregate callback rates (as well as the name-specific

ones) were relatively stable over time (38.8% in 2007/08 versus 35.4% in 2009). Similarly, when we include a dummy for the second wave in our probit models, we did not find any significant coefficients.

4 Conclusions

We conducted a correspondence test, sending more than a thousand applications with randomly assigned German and Turkish names to firms advertising student internships. The difference in callbacks is significant but, compared to similar studies for other countries, relatively small. Several explanations can account for this result. First, we focus on a specific high-skill segment of the labour market. If competition for qualified students is intense, discriminating firms cannot survive the “war for talents” and are driven out of the market. On the other hand, in labour market segments with an excess supply of qualified workers, discrimination should be stronger as firms can choose their favourite candidate from among a large number of applicants. Moreover, in our experiment the student with the Turkish name is a German citizen with a migration background. He was born and raised in Germany, went to school in Germany and is now studying in Germany.¹⁸ We focus on these second and third generations of immigrants as they represent the largest ethnic minority in the German labour market. The fact that all applicants are observationally equivalent (except their name) permits us to isolate the name effect from any language effects. Conducting a similar study with non-German citizens with a mother tongue other than German should be expected to produce a larger difference in callback rates.¹⁹ Both applications were endowed with very good grades and interesting enough CVs so as to guarantee a reasonably high callback rate. It is conceivable that net discrimination is substantially larger for candidates with mediocre grades; for such applications negative stereotypes of human resource managers (such as “students with

¹⁸Aldashev et al. (2007) show that degrees obtained in Germany give rise to better earning prospects for workers with migration background than degrees obtained abroad.

¹⁹However, Aldashev et al. (2007) find that the wage profiles of high-skilled German citizens with migration background are only slightly higher than the ones of high-skilled foreigners. For Western Europe, Kahanec and Zaiceva (2009) show that predominantly foreign origin has a significant impact on the labour market success, while citizenship only affects the earnings of women. Thus, the effect of citizenship in our study is presumably limited.

Turkish migration background are under-performing”) become potentially more important. On the other hand, the rather good grades can be inconsistent with employers’ expectations about a Turkish applicant. Hence such an observation is more salient and takes a stronger impact on impression formation than it would do for a German candidate (see e.g. Sherman, Stroessner, Loftus, and Deguzmani (1997)). We would further expect that most firms with a standardized recruitment procedure use a threshold strategy. Any candidate that fulfills certain minimal criteria receives a callback, and these criteria apply equally for all candidates. Since larger firms have more often a standardized recruitment process, they also discriminate less than smaller firms. In this regard, an optimization of recruitment practices can help to provide equal opportunities for minority workers.²⁰

Our study gives only a first insight into the extent of ethnical discrimination in Germany’s labour market. There are several questions that should be explored in further research. First, measuring hiring discrimination in different segments of the labour market would provide more information on the effects of sector and firm characteristics. Second, varying the quality of the applications would also permit to measure group-specific returns to skills, as in Bertrand and Mullainathan (2004). Third, our experiment shows that provision of information about personality reduces the extent of discrimination. In many countries it is uncommon to use references at early stages of the recruitment process.²¹ Our result suggests that such conventions can potentially backfire on minority employees. Future experiments, also those conducted in the laboratory, should further illuminate the role of information about personality on recruitment decisions.

²⁰In November 2009, the government of France announced that 50 firms voluntarily participate in a trial scheme with anonymized applications that contain neither the name, the age, nor the address of the applicant. Motivated by the public reaction to this study, a similar pilot project is starting in Germany in autumn 2010 with five participating companies.

²¹In the UK, the “Employment Practices Data Protection Code” states that employers should only carry out pre-employment vetting (e.g. reference letters) on an applicant at a late stage in the recruitment process.

References

- ALDASHEV, A., J. GERNANDT, AND S. THOMSEN (2007): “Earnings Prospects for People with Migration Background in Germany,” ZEW Discussion Paper No. 07–031.
- ANDERSON, R., R. FRYER, AND C. HOLT (2006): “Discrimination: Experimental Evidence from Psychology and Economics,” in *Handbook on the Economics of Discrimination*, ed. by W. Rodgers III. Edward Elgar, Northampton, MA.
- ARROW, K. J. (1973): “The Theory of Discrimination,” in *Discrimination in Labor Markets*, ed. by O. Ashenfelter, and A. Rees. Princeton University Press.
- BANERJEE, A., M. BERTRAND, S. DATTA, AND S. MULLAINATHAN (2009): “Labor Market Discrimination in Delhi: Evidence from a Field Experiment,” *Journal of Comparative Economics*, 37, 14–27.
- BAUER, T. K., M. LOFSTROM, AND K. F. ZIMMERMANN (2000): “Immigration Policy, Assimilation of Immigrants and Natives’ Sentiments towards Immigrants: Evidence from 12 OECD-Countries,” *Swedish Economic Policy Review*, 7, 11–53.
- BERTRAND, M., AND S. MULLAINATHAN (2004): “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination,” *American Economic Review*, 94, 991–1013.
- CARLSSON, M., AND D.-O. ROTH (2007): “Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data,” *Labour Economics*, 14, 716–729.
- CASTILLO, M., AND R. PETRIE (2010): “Discrimination in the Lab: Does Information Trump Appearance?,” *Games and Economic Behavior*, 68, 50–59.
- CONSTANT, A., M. KAHANEC, U. RINNE, AND K. ZIMMERMANN (2009): “Ethnicity, Job Search and Labor Market Reintegration of the Unemployed,” IZA Working Paper No. 4660.
- CONSTANT, A., AND K. ZIMMERMANN (2009): “Work and Money: Payoffs by Ethnic Identity and Gender,” DIW Discussion Paper No. 908.

- DRYDAKIS, N., AND M. VLASSIS (2007): “Ethnic Discrimination in the Greek Labour Market: Occupational Access, Insurance Coverage, and Wage Offers,” Working Paper No. 0715, University of Crete, Department of Economics.
- DUSTMANN, C., AND I. PRESTON (2001): “Attitudes to Ethnic Minorities, Ethnic Context and Location Decisions,” *Economic Journal*, 111, 353–373.
- FERSHTMAN, C., AND U. GNEEZY (2001): “Discrimination in a Segmented Society: An Experimental Approach,” *Quarterly Journal of Economics*, 116, 351–377.
- FERTIG, M., AND C. M. SCHMIDT (2010): “Attitudes towards Foreigners and Jews in Germany: Identifying the Determinants of Xenophobia in a Large Opinion Survey,” *Review of Economics of the Household*, forthcoming.
- GAULT, J., J. REDINGTON, AND T. SCHLAGER (2000): “Undergraduate Business Internships and Career Success: Are They Related?,” *Journal of Marketing Education*, 22, 45–53.
- GOLDBERG, A., D. MOURINHO, AND U. KULKE (1996): “Labour Market Discrimination against Foreign Workers in Germany,” International Migration Papers 7, International Labour Office, Geneva.
- GOLDIN, C., AND C. ROUSE (2000): “Orchestrating Impartiality: The Impact of Blind Auditions on Female Musicians,” *American Economic Review*, 90, 715–741.
- HARRISON, G., AND J. LIST (2004): “Field Experiments,” *Journal of Economic Literature*, 42, 1009–1055.
- ISSERSTEDT, W., E. MIDDENDORFF, M. KANDULLA, L. BORCHERT, AND M. LESZCZENSKY (2010): *Die Wirtschaftliche und Soziale Lage der Studierenden in der Bundesrepublik Deutschland 2009*. Bundesministerium fuer Bildung und Forschung, Berlin.
- KAHANEC, M., AND A. ZAICEVA (2009): “Labor Market Outcomes of Immigrants and Non-Citizens in the EU: An East-West Comparison,” *International Journal of Manpower*, 30, 97–115.
- LIEBIG, T. (2007): “The Labour Market Integration of Immigrants in Germany,” OECD Social, Employment and Migration Working Papers No. 47.

- NEUMARK, D. (1996): “Sex Discrimination in Restaurant Hiring: An Audit Study,” *Quarterly Journal of Economics*, 111, 915–942.
- PETIT, P. (2007): “The Effects of Age and Family Constraints on Gender Hiring Discrimination: a Field Experiment in the French Financial Sector,” *Labour Economics*, 14, 371–391.
- RIACH, P., AND J. RICH (2002): “Field Experiments of Discrimination in the Market Place,” *The Economic Journal*, 112, F480–F518.
- (2004): “Deceptive Field Experiments of Discrimination: Are they Ethical?,” *Kyklos*, 57, 457–470.
- ROOTH, D.-O. (2007): “Evidence of Unequal Treatment in Hiring against Obese Applicants: A Field Experiment,” IZA Working Paper No. 2775.
- SHERMAN, J., S. STROESSNER, S. LOFTUS, AND G. DEGUZMANI (1997): “Stereotype suppression and recognition memory for stereotypical and nonstereotypical information,” *Social Cognition*, 15, 205–215.
- STATISTISCHES BUNDESAMT (2010): *Bevoelkerung und Erwerbstaetigkeit. Bevoelkerung mit Migrationshintergrund - Ergebnisse des Mikrozensus 2009*. Fachserie 1, Reihe 2.2, Statistisches Bundesamt, Wiesbaden.
- STEVENSON, B. (2009): “The Internet and Job Search,” in *Studies of Labor Market Intermediation*, ed. by D. H. Autor. University Of Chicago Press.
- UHLENDORFF, A., AND K. ZIMMERMANN (2009): “Unemployment Dynamics among Migrants and Natives,” IZA Working Paper No. 2299.
- WEICHELBAUMER, D. (2003): “Sexual Orientation Discrimination in Hiring,” *Labour Economics*, 10, 629–642.
- WEITZEL, T., W. KNIG, A. VON STETTEN, A. ECKHARDT, AND S. LAUMER (2010): *Recruiting Trends 2010*. Centre of Human Resources Information Systems, Frankfurt.

WOOD, M., J. HALES, S. PURDON, T. SEJERSEN, AND O. HAYLLAR (2009):
“A Test for Racial Discrimination in Recruitment Practice in British Cities,”
Department for Work and Pensions Research Report No. 607.

Appendix A: Multinomial logit

We conduct a multinomial logit with the outcomes 1 (callback), 2 (applicant receives a rejection) versus 0 (no reaction) for every application and every workday, beginning at the application day until the workday of the firm’s reaction. We complemented the controls of the previous probit estimation by the number of workdays that have passed without a reaction (t and t^2) and by an interaction term between t and the firm-size dummy (see Table 6). The main results of the probit analysis can also be observed in the multinomial logit: applicants with a Turkish name are less likely to receive a callback, and firms in South Germany are more likely to give a positive response. In contrast, there are no such effects for rejections. That is, applications from the Turkish applicant do not receive significantly less (or more) rejections than candidates with a German name. Table 6 also confirms that large firms take more time to answer an application. These results remain robust for various model specifications (including employer, occupation, and region characteristics). We also checked other interaction terms (in particular those involving the name dummy), but found all of them to be insignificant.

reaction	1	2
constant	-2.59*** (0.16)	-3.48*** (0.17)
t	-0.11*** (0.14)	-0.03*** (0.008)
t^2	-0.00008 (0.0001)	-0.00025*** (0.00008)
Turkish name	-0.187* (0.103)	-0.03 (0.09)
largefirm	-0.76*** (0.16)	0.036 (0.156)
$t \cdot$ largefirm	0.068*** (0.015)	0.019*** (0.007)
typeA	-0.018 (0.102)	0.022 (0.092)
South	0.187* (0.104)	0.038 (0.093)
Sample	full	full
Observations	1056	1056

Notes: Multinomial logit with the outcomes 1: callback, 2: rejection (vs 0: no reaction at workday t). Standard errors are in parentheses. * denotes significance at the 10%-level, ** significance at the 5%-level, and *** at the 1%-level. Several other interactions have been tested and not found significant.

Table 6: Multinomial Logit.

Appendix B: Curriculum Vitae of two candidates

Figures 1 and 2 show the curriculum vitae of types A and B in the second wave of the study (2008/09). As explained in the text, both types were assigned with German and Turkish names, randomly for each application. Information about addresses and schools has been removed.

Serkan Sezer

xxxxxxxx 17 • xxxxx xxxxxxxxxxxx • Tel.: 0151 20522033

Lebenslauf



■ Persönliche Daten

Name: Serkan Sezer
Anschrift: xxxxxxxxxxx 17
xxxxx xxxxxxxxxxxxxxxx
Mobil: 0151 20522033
Email: Serkan_Sezer@gmx.de
Geburtsdaten: 07.01.1988
Geburtsort: xxxxxxxx
Familienstand: ledig
Nationalität: deutsch

■ Ausbildung

Seit Oktober 2007 Universität xxx
Wirtschaftswissenschaften (Bachelor)
1997 - 2007 Gymnasium xxx
Abitur (Note: 1,6)
2004 - 2005 Auslandsaufenthalt:
xxx High School, xxx, Texas

■ Praktische Erfahrungen

Juni 2007-August 2007 Aushilfe bei E!Motion GmbH, xxx
Seit Oktober 2008 Tutor Mathematik I

■ Sprachen

- Englisch (fließend in Wort und Schrift)
- Französisch (gute Kenntnisse)

■ EDV-Kenntnisse

- Systemadministration
- MS Office (Excel, Word, Powerpoint)
- Photoshop

Figure 1: Type A applicant with a Turkish name.

LEBENS LAUF



Persönliche Daten

Name	Tobias Hartmann
Wohnsitz	xxx Str. 2 Xxxxx xxxxxxxxx
Telefon	0152 / 24183570
e-Mail	Tobias_Hartmann@gmx.com
Geburtsdatum	13. Juni 1987
Geburtsort	xxx
Staatsangehörigkeit	deutsch
Familienstand	ledig

Schul Ausbildung

1997 - 2003	Gymnasium xxx
2003 - 2004	xxx High School, xxx, England
2004 - 2007	Wirtschaftsgymnasium xxx, Abitur: 1,7

Studium

Oktober 2007	Beginn des Bachelor-Studiums der Volkswirtschaftslehre an der Universität xxx
--------------	---

Sonstiges

Bisherige Tätigkeiten	Aushilfe bei w.w.w. Web Solutions (xxx), Juli - September 2007 Studentische Hilfskraft (seit Mai 2008)
Sprachkenntnisse	Englisch (verhandlungssicher) Französisch (Grundkenntnisse) Spanisch (Grundkenntnisse)
IT-Kenntnisse	Microsoft Office-Paket, SPSS, LaTeX, Mathematica
Hobbies	Sport (Handball, Skifahren), Musik (Saxophon)

Figure 2: Type B applicant with a German name.