Field experimental evidence on hiring discrimination in the German apprenticeship market[1]

Ekkehard A. Köhler
University of Siegen
School of Economic Disciplines
Economic Education
Kohlbettstrasse 17
D-57072 Siegen
ORCID 0000-0003-1466-9891

Dilara Wiemann
University of Siegen
School of Economic Disciplines
Economic Education
Kohlbettstrasse 17
D-57072 Siegen

This version: July 28, 2025

Abstract

Are apprenticeship applicants discriminated against, conditional on the quality of their school performance or conditional on their perceived ethnic background? If so, does prior work experience, business knowledge certificates or even volunteering increase the chances to receive an answer? To explore these questions, we fielded a randomized controlled correspondence trial (RCT) to measure how different foreign names (Turkish, Russian, Arab, Hebrew versus German) and individual performance indicators (school grades, the level of economic education and internship experiences) affect employers' responsiveness to apprenticeship application inquiries. Our findings, complemented by an employer survey, reveal three key insights: First, significant discrimination persists despite excellent academic credentials from applicants with foreign names. Second, evidence for statistical discrimination emerges as employers interpret identical productivity signals differently based on perceived backgrounds. Third, we find that discrimination varies significantly regionally as well as within sectors. The largest effects occur in rural and less densely populated areas and within the industry and skilled trades sector.

JEL Codes: C93, J71, J15, A21, I 24
Keywords: RCT, field experiment, correspondence studies, child discrimination, Ethnic discrimination, Gender discrimination, labor economics, educational economics

[1] AEA RCT Reg. Nr. AEARCTR-0012833- January 12 2024, IRB approval ER_35_2021 University of Siegen https://doi.org/10.1257/rct.12833-2.1
We thank Ayça Akçakoca, Johannes Effenberger and the staff of the Center for Economic Education at the University of Siegen for their assistance. We are grateful to the participants of the 2024 CESIfo Workshop on Institutional Economics. We received valuable comments from Elliot Ash, Christine Laudenbach, John G. Matsusaka and most notably Henning Hermes.

1. Introduction

We know that applicants with foreign-sounding names have a smaller chance to receive an answer to their applications on the job market. But does this also apply for teenagers with (and without) excellent grades? Does it matter if a migrant is from a cultural distant (or closer) descent? In the present article, we investigate how between 2023 and 2025 employers in Germany respond to apprenticeship inquiries from students with and without foreign-sounding names (by Turkish, Russian, Arab, Hebrew or German names) to signal a perceived migration background. All students are about to graduate with a vocational secondary school qualification[2] (i.e. GSCE equivalent). We address the following research questions:

i) To what extent does the perceived migration background— signaled out by the prospective applicant's name—negatively affect the likelihood of receiving responses to apprenticeship inquiries in the German vocational training market?
ii) Do additional qualifications (certificates exhibiting a level of economic education, internship experiences, volunteering activities) mitigate potential negative effects for applicants with foreign-sounding names?
iii) What are the underlying mechanisms driving observed differences in response rates between applicants with and without foreign-sounding names?

From a labor economics perspective, examining participation in vocational training for adolescents with migration background is crucial, as it influences individual long-term careers and labor market integration. Recent data from the Federal Institute for Vocational Education and Training (BIBB) underline this importance: In 2022, the share of foreign students entering vocational training (31.3 percent) was significantly lower than that of German students (53.5 percent) (BIBB, 2024). Furthermore, by the end of 2021, applicants with migration backgrounds were substantially less likely to secure vocational training positions compared to their German counterparts without a migration background—29 percent versus 43 percent, respectively (BIBB, 2022).

Our study builds upon a well-established literature on hiring discrimination. The literature consistently reveals biases against applicants from minority backgrounds and those with foreign-sounding names who typically receive lower response rates despite having equivalent qualifications to applicants whose families do not have a migration background (Bertrand & Duflo, 2017; Kaas & Manger, 2011). Becker's (1957) theory of taste-based discrimination and the statistical discrimination models of Arrow (1973) and Phelps (1972) provide the theoretical foundation for understanding such biases. To measure these forms of labor market discrimination, correspondence studies have become the standard empirical approach (Bertrand & Mullainathan, 2004; Baert et al., 2017; Carlsson & Rooth, 2007; Zschirnt & Ruedin, 2016). For example, foundational empirical studies by Bertrand and Mullainathan (2004) in the United States and Carlsson and Rooth (2007) in Sweden have demonstrated that applicants with African or Middle Eastern names were significantly less likely to receive responses compared to their counterparts without foreign-sounding names, despite equivalent qualifications.

[2] "Fachoberschulreife"

Similar patterns have been observed across European labor markets, with studies in Germany (Kaas & Manger, 2012; Koopmans et al., 2019, Thijssen et al., 2021, Schneider et al., 2019), Greece (Drydakis and Vlassis, 2010), and the Netherlands (Blommaert et al., 2014), revealing systematic disadvantages for applicants with Turkish, Moroccan, or African migration backgrounds.

While the meta-study of Lippens et al. (2023) have shown that hiring discrimination is prevalent in almost any field experiment, we interact racial backgrounds with a vector of different productivity signals to disentangle racial versus human capital effects of migrant applicants. In addition, so far, the existing literature has mostly focused on general labor markets for adults, with limited attention to minors and the apprenticeship context. We are the first to field an experiment with school students who enter the job market for the first time. We use a mixed-methods research approach, combining nationwide correspondence studies with an employer survey to examine both the prevalence of discrimination against underage applicants with migration backgrounds along with the underlying mechanisms that drive employers' decision-making processes.[3]

We conducted randomized controlled correspondence studies over three campaigns from November 2022 to March 2025, resulting in approximately 50,000 observations. The experiment was pre-registered and received Institutional Review Board (IRB) approval. Vacant apprenticeship positions were drawn from the official job portal of the German Federal Employment Agency. Fictitious email inquiries from underage students were automatically sent to firms offering VET positions in three major sectors: public administration, commerce and services, and industry and skilled trades.

Each campaign introduced specific treatments in a block-randomized design:

- **Campaign 1 (Nov 2022 – Jan 2023)** served as a pre-study. Treatments varied across four dimensions: (1) perceived migration background (German, Turkish, Russian names), (2) gender (male/female), (3) academic performance (high GPA versus satisfactory GPA), and (4) economic knowledge (signaled by the European Business Competence License (EBC*L) certificate). This allowed us to test whether merit-based signals (grades, economic literacy) can offset the perceived disadvantages of a migration background.

- **Campaign 2 (Feb – May 2024)** extended the design to include applicants with German, Hebrew, and Arab names, and replaced the EBC*L signal with (4) prior work experience (internship experience versus no internship experience). Since economic literacy did not appear to have an effect, we tested if practical work experience has any positive effect on the response rates.

- **Campaign 3 (Feb – Mar 2025)** focused exclusively on German and Turkish names and varied (4) volunteering activity: participation in a German-Turkish Society, which is a nationwide club with local chapters (signaling proximity to Turkish culture), a science club (signaling interests in science,

[3] A person is considered to have a migration background if they themselves or at least one parent was not born with German citizenship. Specifically, this definition encompasses foreign nationals who themselves or whose parent(s) migrated, naturalized citizens (who themselves or whose parent(s) migrated), (late) repatriates, and the German-born descendants of these groups (Destatis, 2025).
The vocational education statistics record the citizenship/nationality of trainees; however a potential migration background cannot be explicitly identified. All trainees without a German passport are counted as foreign trainees. Young people who possess both German and non-German citizenship are not counted as foreign trainees/apprentices.

technology, engineering and mathematics), or no volunteering experience (control). This setup aimed to test if volunteering activities influence applicants' chances to receive a reply from employers if they have a foreign-sounding name.

After analyzing initial results from our pre-study and presenting them at a CESifo workshop in March 2024, we fielded a survey between April and June 2024. In this survey, we confronted employers with our experimental findings and asked about their perceptions and explanations for the differences in response rates. Further, we asked them to estimate the actual dropout and completion rates of apprentices with German versus foreign-sounding names and asked about their perceptions towards hiring apprentices with migration backgrounds. The survey was distributed to a stratified random sample of firms that were treated in the previous campaigns. The survey was designed to address three primary research questions:

1. Are statistical discrimination mechanisms reflected in employers' perceptions (e. g. regarding productivity and success rates)?
2. Are taste-based concerns evident (e. g. workplace integration, language)?
3. Do perceptions vary systematically across ethnic groups in ways that match the experimental outcomes?

The 772 survey responses reveal significant differences in employers' perceptions of applicant characteristics across perceived migration backgrounds, particularly regarding perceived perseverance ("how likely an apprentice is to remain in the position"), and varying levels of cultural distance. Also, hiring applicants with a migration background is often associated with higher opportunity costs, citing factors such as increased efforts for bureaucracy efforts, concerns about language proficiency, and anticipated higher dropout rates.

The experimental results from the correspondence studies reveal statistically significant differences in response rates between applicants with German names and those with Turkish, Arab, Hebrew or Russian names. These differences persist even when applicants exhibit strong academic records, prior work experience, or volunteering credentials. In detail, the results can be summarized as followed:

1. **Traditional human capital signals are ineffective in reducing discrimination:** High GPA, completed levels of economic education, and internships do not significantly improve response rates for students with foreign-sounding names. This suggests that employer decisions are not solely driven by productivity concerns or the ability to carry out employment tasks — supporting the presence of taste-based discrimination.
2. **Evidence of statistical discrimination**: Survey results reveal that employers interpret identical human capital signals differently based on the perceived ethnic backgrounds, raising concerns about group-level differences in dropout rates, language skills and social behavior. These results still hold when controlling for other factors, such as local unemployment rates or varying levels of public expenditures.
3. **Contextual variation**: We find that treatment effects between applicants with and without foreign-sounding names vary geographically and are less pronounced in urban metropolitan areas compared to densely populated regions. By contrast, the largest effects occur in rural and less densely populated

areas. Moreover, we observe that in East Germany employers tend to base decisions more on perceived ethnic background than on academic performance.

The remainder of this paper is structured as follows: Section 2 elaborates on the theoretical economic framework and the institutional context of Germany's apprenticeship system; section 3 presents our research design in detail; section 4 presents the main empirical results and robustness checks; and a discussion ensues in section 5. Section 6 concludes. Additional analyses and supplementary material are provided in the appendix.

2. Theoretical Framework & Institutional Background

2.1 Institutional Context: The German Apprenticeship System

The German dual apprenticeship system is an institutional arrangement that is particularly interesting for studying labor market discrimination against underage students. The German dual apprenticeship system combines work-based training with formal education and thereby prepares students both vocationally and educationally.[4] Students commonly begin apprenticeships between the ages of 15 and 18, and the apprenticeships involving both practical training at a company and theoretical education at vocational schools typically last three years. Upon completion, apprentices receive nationally recognized qualifications. These signal standardized skills to employers and facilitate labor market mobility. Compared to higher academic education, companies can align training programs directly with specific industry needs, and many apprentices continue working with their training companies after completion. Traditionally, the German apprenticeship system has been highly regarded in line with its internationally-recognized reputation, as skilled workers constitute the foundation for technically sophisticated, high-quality production and Germany's success in exporting premium-priced goods (Streeck, 1991). Furthermore, apprenticeships play a large role in Germany's labor market. Approximately 40 percent of each cohort of middle school graduates enters the vocational training system, making it the primary pathway to skilled employment for a substantial portion of the workforce (Destatis, 2023). The high participation rate contributes to Germany's relatively low youth unemployment rate compared to other countries without vocational training systems (Zimmermann et al., 2013).

Despite its institutional strengths, the system faces growing challenges. In 2024, 69,400 apprenticeship positions (12.8 percent of the total supply of 556,100 positions) remained unfilled, while demand increased by 0.8 percent to 557,100 (BIBB, 2024). This mismatch between employer demand and apprentice supply is particularly pronounced in the skilled trades sector, which reported nearly 20,000 vacancies (12.7 percent). Similarly, in the industry and commerce sectors almost 42,000 unfilled positions (13.5 percent) remained unfilled (BIBB, 2024).

These vacancies exist alongside differences in qualifications among young adults. Among Germans aged 20 to 34, 12.7 percent have no completed qualification, while this figure is three times higher for foreign adults

[4] The dual system derives its name from the combination of two learning environments: vocational schools (*Berufsschulen*) and training companies. Apprentices typically spend 3-4 days per week at their training company and 1-2 days at vocational school, creating an integrated learning experience that bridges theoretical knowledge with practical application. This arrangement is governed by the Vocational Training Act (*Berufsbildungsgesetz*), which establishes uniform standards across regions and industries.
Training programs are highly standardized, with approximately 330 recognized occupational profiles (*Ausbildungsberufe*) across various sectors. Each profile specifies required competencies, duration of training, and examination standards. This standardization ensures that qualifications are nationally recognized and portable across employers, reducing friction in labor market mobility.

(BiBB, 2024, p.281). This gap is even more pronounced for certain groups, with approximately 40 percent of adolescents with Turkish citizenship in Germany lacking formal qualifications (BIBB, 2023).[5] Given that 28 percent of German adolescents had a migration background in 2022—with half belonging to the second generation born in Germany—understanding barriers to apprenticeship access for these groups has significant policy implications.

Regional variation represents another important dimension for understanding discrimination in the German apprenticeship market. Wößmann et al. (2024) document significant regional disparities in educational opportunities that correlate with socioeconomic background. This spatial heterogeneity likely extends to discriminatory practices in hiring. Urban centers with established immigrant populations, such as certain metropolitan areas of North Rhine-Westphalia, may exhibit lower discrimination due to greater employer exposure to diversity and intercultural competence. Conversely, in regions with less demographic diversity, limited intergroup contact may reinforce stereotypes and strengthen statistical or taste-based discrimination against minority applicants. These regional differences in employer behavior could create geographic inequality in apprenticeship access that could manifest existing socioeconomic disparities.

2.2 Theoretical Foundations of Labor Market Discrimination

The analysis of discrimination in labor markets has produced two major theoretical frameworks that guide our investigation and provide the foundation for interpreting our empirical findings. Gary Becker's (1957) theory of taste-based discrimination conceptualizes employer bias as a preference that effectively imposes additional psychological costs on hiring members of certain groups, irrespective of their actual productivity. In this model, employers hold prejudices against certain groups and act as if hiring members of these groups imposes additional costs beyond actual productivity considerations.

This framework can be especially relevant to apprenticeship markets for several reasons. Apprenticeships typically involve extended training relationships, up to three years in the German system. This could increase the perceived psychological costs for employers with discriminatory preferences. Second, in the apprenticeship context, customer and coworker interactions are often intensive, especially in smaller firms or when apprentices have significant customer contact, which might make employers more sensitive to perceived preferences of these stakeholders. Third, the firm investment in training of the student may exacerbate risk aversion toward candidates from minority groups, even when they hold objectively equal qualifications. Recent figures show that companies have significant training costs, averaging approximately €21,000 per apprentice annually (Wenzelmann & Schönfeld, 2022). This makes it more important that employers accurately assess the expected productivity of applicants. However, as applicants are typically minors between the age 15 and 18, they have no – or only very limited – previous work experience. This makes traditional productivity signals, such as previous work references and experiences, scarce.

[5] In Germany until 2024, it was usually not possible to have two passports. Often, the second generation, i. e. the children of immigrants, keep the passport of their parents' country of origin for practical or even emotional reasons.

Second, complementary to taste-based explanations, Arrow (1973) and Phelps (1972) developed models of statistical discrimination that explain how profit-maximizing employers might rationally discriminate based on group characteristics when individual productivity is difficult to assess (Aigner and Cain, 1977). Phelps (1972) argues that employers might discriminate against minorities, even if these groups have similar skill sets to the majority, due to poorer quality of information about minority groups. Arrow (1998) adds that if minorities expect to face discrimination, they may invest less in their own success in the labor market, while majorities, expecting preferential treatment, may invest more. This can lead to a self-fulfilling prophecy where employers rationally discriminate against minorities who are perceived to have invested less.

Recent theoretical work by Echenique and Li (2025) provide a theoretical framework for understanding discrimination in apprenticeship markets through their theory of rationally inattentive statistical discrimination. In contrast to traditional theories, the model suggests that employers face cognitive constraints when processing applications. Ultimately, they must strategically allocate their limited attention across candidates. Even if all groups are identical ex-ante, a discriminatory equilibrium can emerge in which employers rationally allocate more attention to the majority group application. Only little attention is given to minority applicants, unless they demonstrate exceptional qualification. This attention allocation creates a self-reinforcing cycle: Minority candidates anticipate that they receive less attention and therefore invest rationally less in acquiring relevant skills, for example. In contrast, majority candidates expect to receive more attention in the application process, and thus they have more incentives to invest in developing their skills and in applications. This differential investment then justifies the employers' asymmetric attention allocation.
In the context of the apprenticeship market, it can be assumed that the attention allocation mechanism plays a role too because employers often receive many applications from young candidates with limited work history. This causes costs to process the information, i. e. screening the applications, so that this described attention allocation mechanism can be triggered.

2.3 Research Hypotheses

The German apprenticeship market provides a compelling institutional setting for studying hiring discrimination as apprenticeship hiring differs from standard labor market and its recruitment. Hence, employers may apply different selection criteria than when hiring fully trained adult workers. Based on the empirical and theoretical foundations, we developed the following research questions and testable hypotheses:

i) To what extent does a perceived migration background (signaled by name that is commonly associated in Germany with either a Turkish, Russian, Arab or Hebrew cultural origin) affect the likelihood of receiving responses to apprenticeship inquiries in the German vocational training market?

ii) Do additional qualifications (certificates documenting economic competencies, internship experiences, volunteering activities) mitigate potential discrimination effects for applicants with foreign-sounding names?

iii) What are the underlying mechanisms driving any observed differences in response rates between applicants with and without foreign-sounding names?

H1a: Applicants with German-sounding names will receive significantly higher response rates than equally qualified applicants with foreign-sounding names.

H1b: The magnitude of discrimination will vary by ethnic group, with stronger effects for Turkish and Arab names compared to Russian and Hebrew sounding names, reflecting different levels of perceived cultural distance from German society.

H2a: Strong school performance (high GPA) will reduce discrimination against applicants with foreign-sounding names.

H2b: Additional qualifications (economic education certificate), prior work experience or participation in volunteering activities will increase response rates for all applicants, with potentially stronger effects for applicants with a foreign-sounding name.

By testing these hypotheses, we aim to disentangle the complex nature of different discrimination dynamics that interact with the characteristics of vocational training.

3. Research Design and Empirical Strategy

We use a mixed-method research design, combining correspondence studies with a complementary employer survey. This approach allows us to both measure the prevalence of discrimination and explore the underlying mechanisms prevailing employers' decision-making processes.

We conducted three correspondence studies between 2023 and 2025, systematically varying applicant characteristics while holding constant other relevant factors. These field experiments involved sending inquiries to employers advertising apprenticeship positions across three business sectors, namely public administration, commerce and services, and industry and skilled trades.

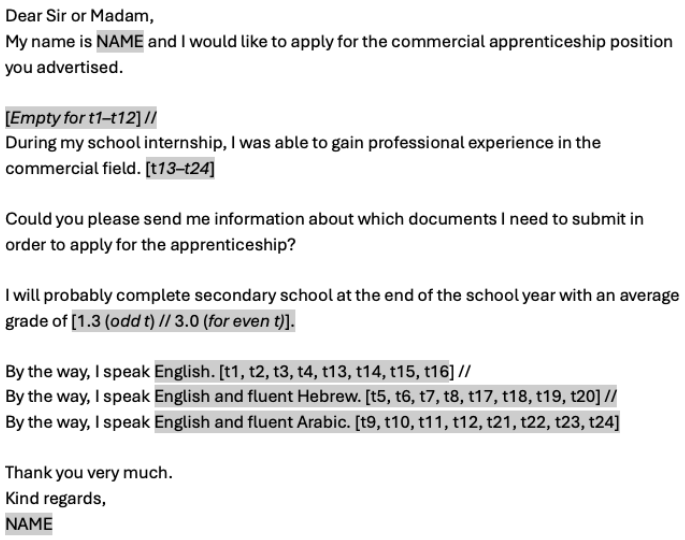
- **Campaign 1 (Nov 2022 – Jan 2023)** served as a pre-study with almost 19,000 email inquiries sent to companies that reported training positions to local job centers in German cities. Treatments varied across four dimensions: (1) migration background (German, Turkish, Russian), (2) gender (male/female), (3) academic performance (high GPA versus satisfactory GPA), and (4) economic competencies (signaled by the EBC*L certificate). This allowed us to test whether merit-based signals (grades, economic literacy) can offset the perceived disadvantages of having a migration background.

- **Campaign 2 (Feb – May 2024)** extended the design to include applicants with German, Hebrew, and Arab names, and replaced the EBC*L signal with (4) prior work experience (internship versus no internship). Since we did not find that economic literacy had a significant effect, we tested if practical work experience had any positive effect on the response rates to students with a migration background.

Across the first two campaigns, we sent a total of 40,214 inquiries of which 1,333 failed to reach their intended recipients due to invalid e-mail addresses. This resulted in a final sample of 38,882 successfully delivered

emails, with 21,919 (56.4 percent) receiving a reply from employers and 16,962 (43.6 percent) receiving no response.

As a robustness check of our preliminary results, we carried out a **Campaign 3** from February to March 2025. It focused exclusively on German and Turkish names and varied (4) volunteering activity: participation in a German-Turkish Society, which is a nationwide club with local chapters (signaling cultural proximity), a science club (signaling interests in science, technology, engineering and mathematics), or no volunteer experience (control). This setup aimed to test if volunteering activities influence the chances to receive a reply with a foreign-sounding name. All email inquiries were systematically dispatched at times aligned with the application deadlines for the yearly start of apprenticeships, and responses were collected over a period exceeding three months for each campaign. The email templates used in our correspondence studies were developed in collaboration with HR managers and employees from German job centers who have experience in apprenticeship admissions processes. We pre-tested the email format before the first wave to ensure it was both realistic and effective. The high overall response rate (56.1 percent for the first two campaigns) suggests that our inquiries were perceived as authentic by employers.


Dear Sir or Madam,
My name is NAME and I would like to apply for the commercial apprenticeship position you advertised.

[*Empty for t1–t12*] //
During my school internship, I was able to gain professional experience in the commercial field. [*t13–t24*]

Could you please send me information about which documents I need to submit in order to apply for the apprenticeship?

I will probably complete secondary school at the end of the school year with an average grade of [1.3 (*odd t*) // 3.0 (*for even t*)].

By the way, I speak English. [t1, t2, t3, t4, t13, t14, t15, t16] //
By the way, I speak English and fluent Hebrew. [t5, t6, t7, t8, t17, t18, t19, t20] //
By the way, I speak English and fluent Arabic. [t9, t10, t11, t12, t21, t22, t23, t24]

Thank you very much.
Kind regards,
NAME

Figure 1: E-mail template of Campaign 2 with randomized information highlighted

Figure 1 shows the e-mail template from Campaign 2 which maintained a similar structure but varied the names (Arab, Hebrew, and German) and replaced the EBC*L certificate with relevant internship experience.

All e-mails were intentionally composed without grammatical errors, typos, or other mistakes that might be statistically associated with sender characteristics. This approach ensures a clear and unambiguous interpretation of any observed treatment differences. However, it is important to note a potential limitation: e-mails from students with migrant backgrounds displaying impeccable grammar might suggest an exceptionally high level of

education or effort. Consequently, any identified treatment effects can be interpreted as conservative estimates (Hermes et al., 2023).[6]

3.1 Data Collection and Outcome Measures

In our study, we use stratified randomization to generate the treatment groups. We randomized on the level of the cities and the counties to ensure a balanced distribution of certain characteristics across treatment groups. In addition, we executed balance tests within each wave, thereby checking for and ensuring comparability across treatment groups within each wave.

Our main variable of interest is the observation of receiving or not receiving a reply to an email inquiry. We constructed a binary indicator that takes the value 1 if an email inquiry received a reply, i. e. a manually written answer and no automated message, or the value of 0 if no such reply was received from the employer.

For each employer in our sample, we also gathered firm-level characteristics including the economic sector, company size (where available), as well as control variables that include geographic information (urban/rural and region within Germany, population density) and public finance indicators (public expenditures per civil servants per capita).

3.2 Employer Survey

To complement the first results of the correspondence studies, we fielded a survey among employers during the summer of 2024. We confronted the participants with the results of the field experiments and asked open ended questions on the mechanisms that may have caused the observed results. We used their answers to contextualize the reaction patterns of the field experiment. We also asked the participants to evaluate the challenges associated with apprentices from different ethnic backgrounds. Here we were interested in drop out probabilities, cultural proximity amongst others. In addition, we used these results to design the treatments of the third campaign during the winter of 2024. Again, we fielded a second survey in the summer of 2025 to contextualize the findings of this campaign.

We refrain from using the answers from these surveys as conclusive arguments of this study. The survey data is anecdotal evidence but by no means reflects a representative sample of the businesses involved in our study. The survey was distributed to a random sample of the full sample that previously posted apprenticeship positions. To explore differential perceptions across ethnic groups, we asked respondents to rate perceived challenges in cultural integration, perseverance, and language proficiency specifically for Arabic, Turkish, Russian, and Hebrew speakers. The survey also assessed familiarity with credentials like the EBC*L and the perceived value of practical work experience, which were treatment variables in our field experiments.

We find that employers express preferences for homogeneous work environments and report discomfort with cultural or ethnic diversity when explaining potential hiring biases, independent of productivity considerations.

[6] A main limitation of this approach is the inability to control for all environmental factors, which could lead to confounding factors influencing the result. This can be bounced ´ messages or the unobservable emotional state of the recipient.

In addition, based on statistical discrimination theory, employers cite assumptions about group-level differences in productivity, language skills, and completion rates to explain differential treatment.

4. Empirical Results

The following sections detail our data collection procedures and present our empirical results. Summary statistics and robustness checks for the results can be found in the appendix. All reported differences are statistically significant at conventional levels unless otherwise noted. We report 95% confidence intervals constructed using robust standard errors.

4.1. Evidence of Discrimination in the Apprenticeship Market

Effect of perceived origin of name on response rates

Our correspondence studies reveal discrimination against applicants with foreign-sounding names backgrounds in the German apprenticeship market. Figure 2 shows the mean response rates by perceived origin of names. Across all treatments, applicants with German-sounding names consistently received the highest response rates (67 percent) and significantly exceeded those with foreign-sounding names backgrounds. The magnitude of discrimination varies notably across different groups.

Applicants with Russian names received the second-highest response rates (57 percent), followed by Hebrew - named applicants (54 percent), and applicants with Turkish names (52 percent). Applicants with Arab-sounding names faced the most discrimination with average response rates of 40 percent. This pattern is in line with our hypothesis H1b, which predicted variation in discrimination across ethnic groups, with stronger effects for Turkish and Arab names compared to Russian and Hebrew names.

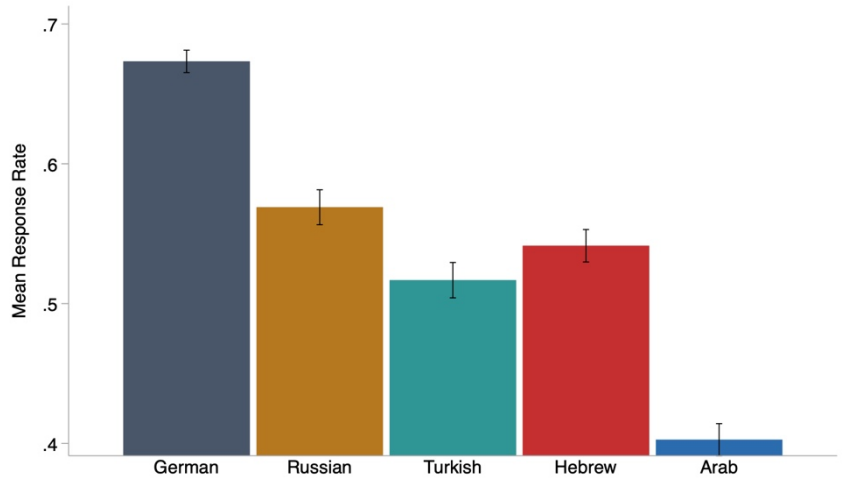


Figure 2: Mean response rates by perceived origin of name

Effect of Gender on Response Rates

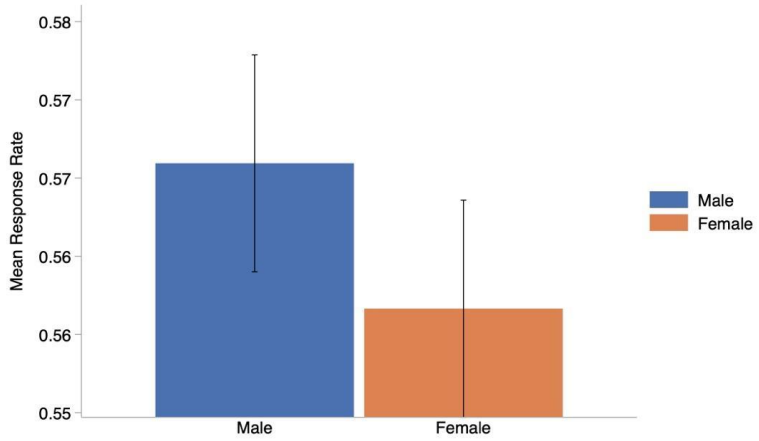


Figure 3: Mean response rates by gender

Figure 3 shows the overall effect of gender on response rates across the sample and shows a small advantage for male students (57.1 percent response rate) compared to female students (56.5 percent). However, this difference is not statistically significant.

Interaction effect of gender and perceived origin of names

Figure 4 examines the connection of gender and perceived origins. Most notably, Arab female applicants experience significantly lower response rates (approximately 37 percent) compared to Arab male applicants (approximately 44 percent), representing a 7 percentage point difference. This suggests that Arab female applicants face discrimination that relates to both ethnicity as well as gender. Interestingly, this gender effect is less pronounced or even reversed for other ethnic groups, with Russian and Hebrew female applicants receiving slightly higher response rates than their male counterparts.

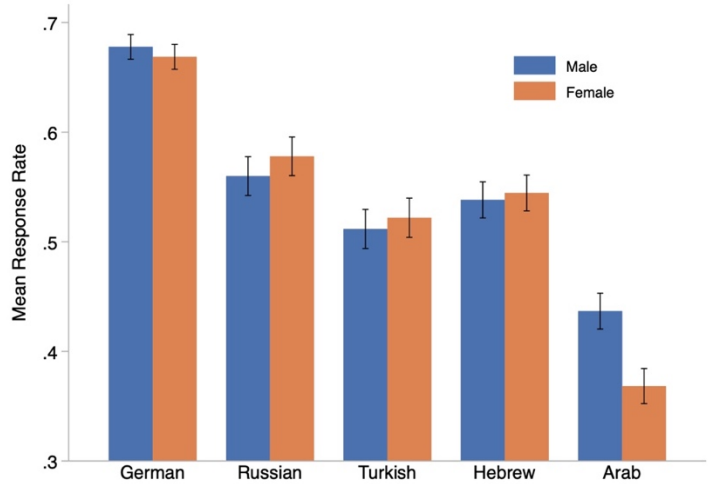


Figure 4: Mean response rates by gender and perceived origin of names

4.2 Mitigating factors

Effect of School performance on Response Rates

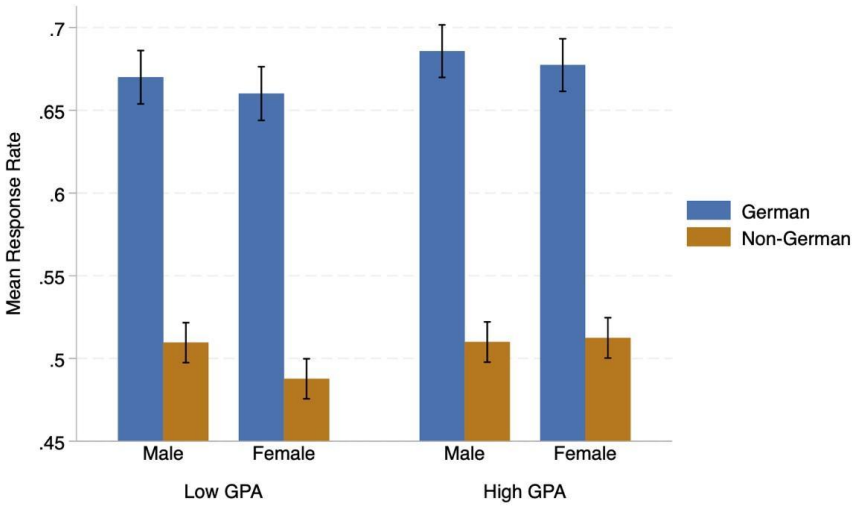


Figure 5: Mean response rates by gender, German vs. non-German names and school performance

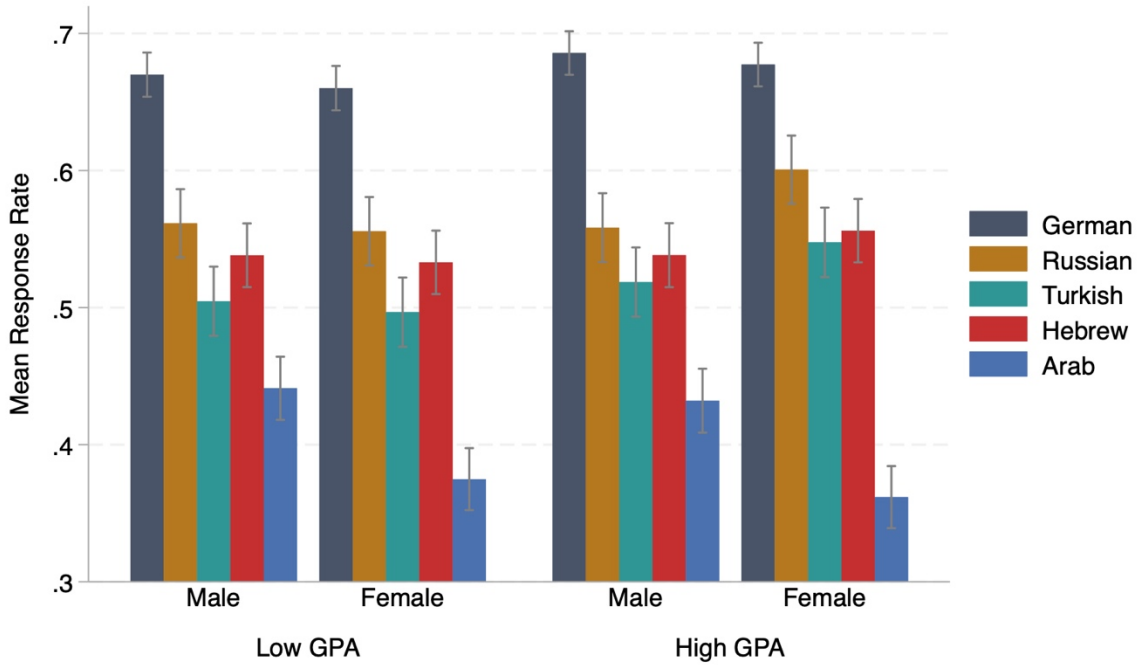


Figure 6: Mean response rates by gender, perceived origins of names and school performance

Our research examined whether good academic performance might mitigate ethnic discrimination. Figure 5 compares response rates across low and high GPA aggregated for the different names. The response rate

difference between high and low GPA is minimal for both applicants with German names (approximately 1 percentage point) and applicants with foreign-sounding names (approximately 1.5 percentage points), with neither difference reaching statistical significance.

Figure 6 further illustrates response rates across high and low GPA conditions for each group.[7] Surprisingly, academic performance has minimal impact on response rates: high-performing applicants receive only marginally higher response rates compared to lower-performing counterparts within the same ethnic group. This holds across all groups, indicating that even exceptional academic achievements do not substantially reduce the discrimination that applicants with foreign-sounding names face. This finding contradicts our hypothesis H2a, which predicted that strong academic performance would significantly reduce discrimination. Instead, the data suggest that employers may not consider GPA as an important productivity signal in apprenticeship selection decisions, or that stereotypes override positive signals of academic achievements.

Effects of Economic Education and Internship Experiences

We also tested whether economic education qualifications or relevant internship experience reduce ethnic discrimination. Like academic performance, we do not find that these additional qualifications moderate the negative effect of migration background on response rates (Figure 7 and Figure 8):

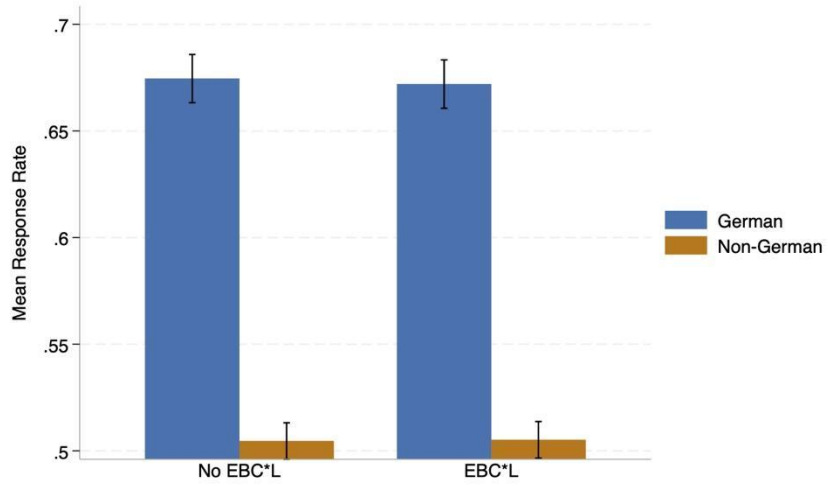


*Figure 7: Mean response rates by German and Non-German names and EBC*L certificate*

[7] In Germany, the grade scale includes values between 1.0 (letter grade "A" on the US grading scale) and 6.0 (i.e. letter grade "F" on the US grade scale). A GPA of 3.0 ("C") is considered low or satisfactory, while a GPA of 1.3 ("A") is considered high and very good.

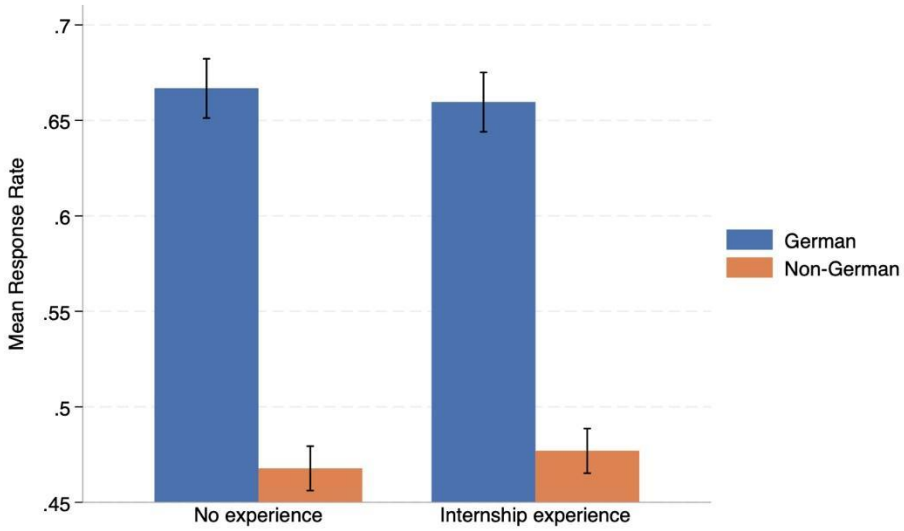


Figure 8: Mean response rates by German and Non-German names and internship experience

These findings contradict our hypothesis H2b, which predicted that additional qualifications would increase response rates, particularly for applicants with foreign sounding names.
The employer survey provides potential explanations for this pattern: only 11 percent of respondents were familiar with the economic education EBC*L certification. Among those, 42 percent questioned its relevance for apprenticeship qualification. Similarly, 27.4 percent of employers indicated that practical work experience like internships is relatively unimportant in terms of assessing candidates.

Collectively, these results suggest that conventional human capital investments and qualifications may have limited effectiveness in overcoming ethnic discrimination in the apprenticeship market. It appears that ethnicity-related signals influence initial screening decisions more than productivity-related signals.

4.3 Employer Survey Findings

The survey reveals that employers are concerned that hiring a person with a migration background would be more costly than a German native applicant. Such opportunity cost could occur due to higher perceived efforts of working with migrant employees compared to German employees, for example due to a lack of perseverance, higher integration efforts and language barriers for migrant employees.
We observe differences regarding the perceptions about different ethnic groups which can further inform the different treatment effects of the field experiment. For example, respondents perceive that among Russian, Hebrew, Turkish and Arab students, students with migration backgrounds from Arab countries would face the biggest challenge with cultural integration and language proficiency. This is followed by students with Turkish, Russians and Hebrew names

Due to the expected experimenter demand effect, we were not surprised to find that respondents generally rejected that apprentices with migration backgrounds lack social skills, or that hiring these apprentices could affect the company's image negatively.

In addition, the survey reveals limited recognition of additional qualifications: Only 11 percent of respondents were familiar with the EBC*L certification. Among these, 42 percent even questioned the relevance of the certificate for apprenticeship qualification. Similarly, almost a third of the respondents (27.4 percent) indicated that practical work experience like internships is relatively unimportant for assessing candidates. These findings are in line with the findings of the correspondence studies where additional qualifications had little to no significant effect on response rates.

The survey findings suggest that employer decision-making is a complex interplay that possibly includes elements of both, statistical and taste-based discrimination differing across ethnic groups. The predominance of concerns about language skills, integration challenges, and training costs points toward statistical discrimination based on perceived productivity differences. However, the systematic variation in perceptions across ethnic groups—particularly the consistently more negative evaluations of Arab and Turkish applicants—suggests that statistical discrimination may be amplified by stereotypes or implicit biases.

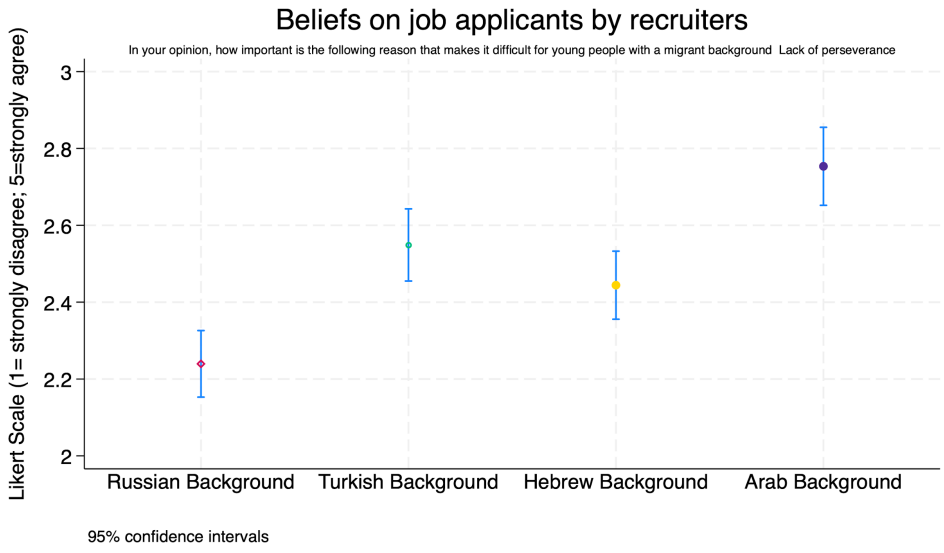


Figure 9: Perceived lack of perseverance among applicants with different migration backgrounds

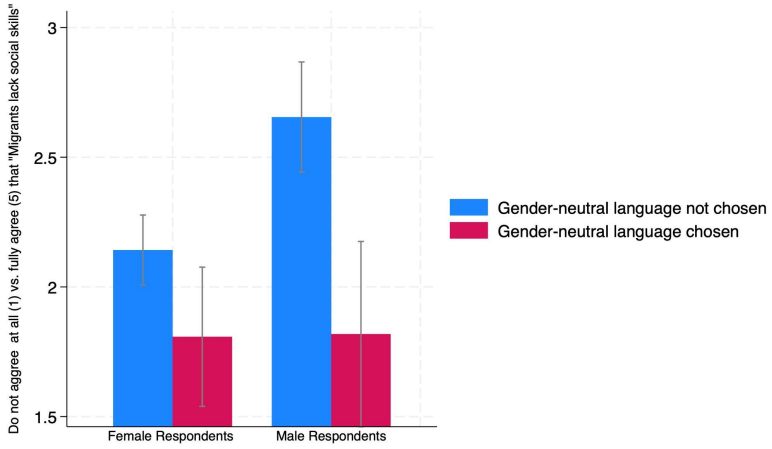


Figure 10: Mean response rates of female and male respondents on the question if the social behaviour of applicants with a migration background tends to have a negative influence on the social behaviour of other employees.

We believe that the heterogenous beliefs about the attributes of applicants with foreign-sounding names can be explained to varying biases. Men who do not choose the option to receive a gender-neutral questionnaire reported significantly different results compared to individuals who did choose gender neutral questionnaires or female respondents (Figure 10).

4.4 Robustness Analyses

4.4.1 Population density and regional effects

We find a modest moderating influence of population density on discrimination, as illustrated in Figure 11. While students with a foreign-sounding name always receive fewer responses compared to students with German names, the relative disadvantage decreases in more densely populated areas. However, the difference between ethnicities remains significant throughout. This indicates that population density moderates, but does not eliminate, this discrimination.

In addition, we can conclude that discrimination is most prevalent in less densely populated areas. In the least densely populated areas (50 people per square km), applicants with German names receive response rates of approximately 69 percent, compared to 53 percent for applicants with foreign-sounding names. In the most densely populated areas (3300 people per square km), response rates decline to about 63 percent for applicants with German names and 49 percent for applicants with foreign-sounding names.

This suggests that while urbanization slightly reduces the magnitude of discrimination, disparities persist even in Germany's most densely populated regions. This finding contradicts expectations that greater exposure to diversity in urban areas significantly reduces discriminatory behaviors.

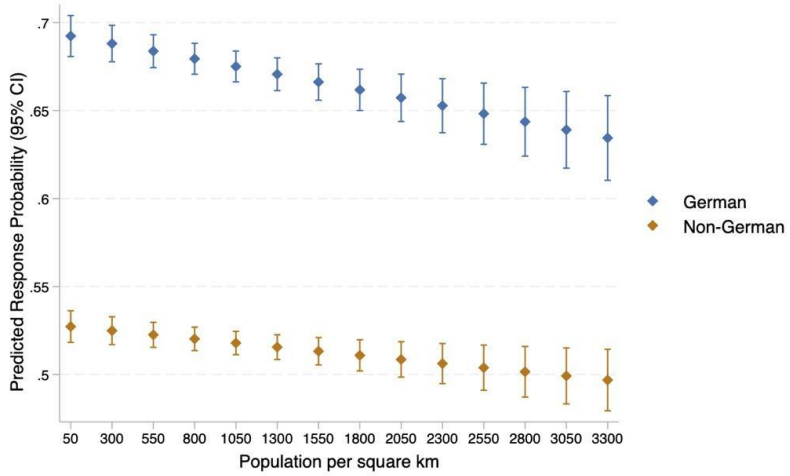


Figure 11: Predicted response probability German vs. Non-German applicant (Logit; N=33.331)

In the next step we analyze if discrimination varies between municipalities of different sizes. This is meaningful to further explore the behavior within large cities and smaller cities over 500,000 inhabitants. Note that we lose around more than 13,000 observations by this restriction. However, this step is meaningful since we expect different behavior patterns between rural and urban areas as shown in Figure 11.

Based on a panel of around 19,800 observations we find no significant different predictive marginal effects in cities with more than 150,000 inhabitants as shown in Figure 12. The relationship between population size and response probability shows that discrimination decreases in more populous areas. While German applicants maintain stable response rates of approximately 67 to 70 percent regardless of regional population, applicants with migration backgrounds see higher response rates in larger areas— rising from about 53 percent in smaller regions to 55 percent in the most populated areas. This suggests that greater diversity exposure in urban areas may moderately reduce discrimination.

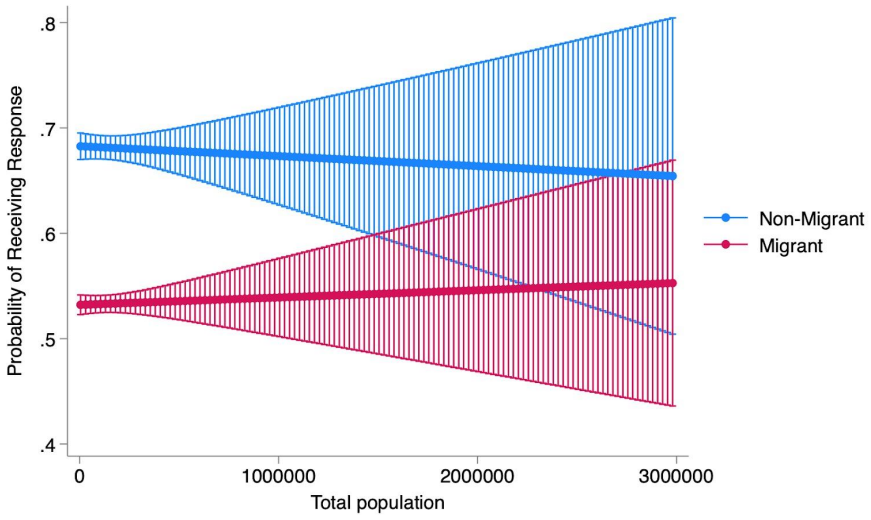


Figure 12: Predicted response probability German vs. Non-German applicant by population size (Logit; N=19.800)

To further explore the heterogeneity of this effect, we aggregated the observations across all sectors on the level of counties and cities. Jurisdictions with a higher treatment effect compared to the estimated treatment effect of -.155 appear on the blue (positive) side of the spectrum (Figure 13). This allows us to further understand local heterogeneity. Note that we find a lower response rate for the foreign-names treatments (red) in almost every state of the Federal Republic of Germany. This also implies that local labor market conditions, aging population

and other effects might be correlated with our results.

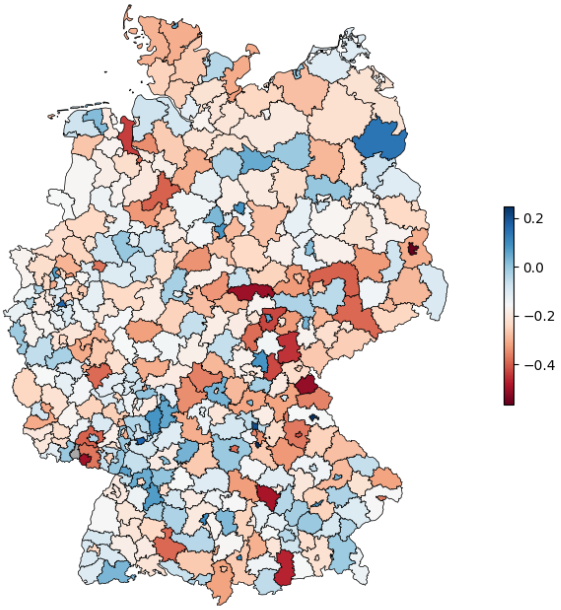


Figure 13: Regional variation of Treatment Effects by counties

Firm size effect

Figure 14 shows response rates by firm size, revealing a strong positive relationship between organizational scale and responsiveness to apprenticeship inquiries. The smallest firms (1-5 employees) demonstrate the lowest response rates. In these firms, migrant applicants have a very low reply rate of 32.9 percent versus German applicants with 52.2 percent. The difference amounts to 19.3 percentage points, and it is highly significant.

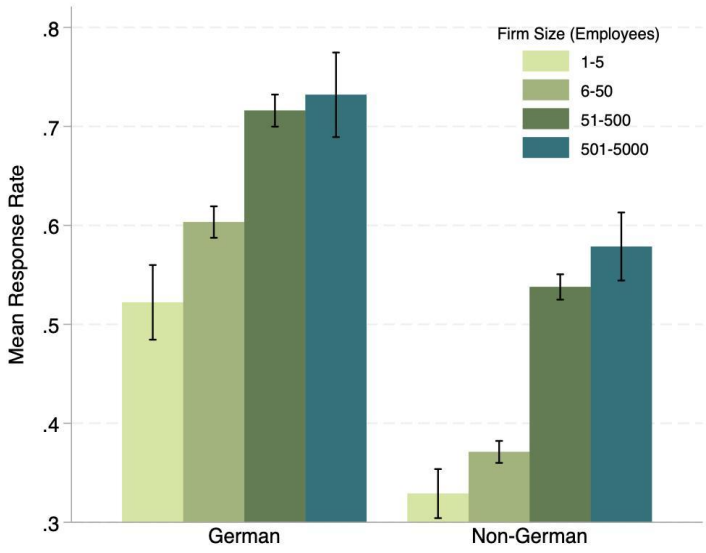


Figure 14: Response rates by firm size

On the other hand, the largest firms (501-5,000 employees) have the highest response rates but still have significant differences between applicants: While applicants with German names have an average response rate of 73.1 percent, the response rate for applicants with foreign-sounding names amounts to 57.8 percent.

This supports the hypothesis that discrimination is stronger in smaller firms, perhaps due to less professional hiring structures, greater autonomy, or less accountability. In other words, informal recruitment structures—such as responses written by secretaries or owners—may lead to higher marginal costs or perceived risks in hiring migrants. Smaller firms, often with limited HR infrastructure and higher relative costs for training investments, appear more selective in their initial responses to potential apprentices.

In contrast, larger firms (with more than 500 employees) show a smaller, though still substantial, response gap of approximately 15 percentage points. These findings align with theories suggesting that larger organizations typically have more formalized hiring processes, greater resources for managing inquiries, and potentially stronger commitments to diversity and equal opportunity. This suggests that while professional HR procedures may mitigate bias to some extent, they do not fully eliminate it. Thus, institutional structures alone may not fully explain discriminatory behavior in hiring decisions.

Sector-specific effects

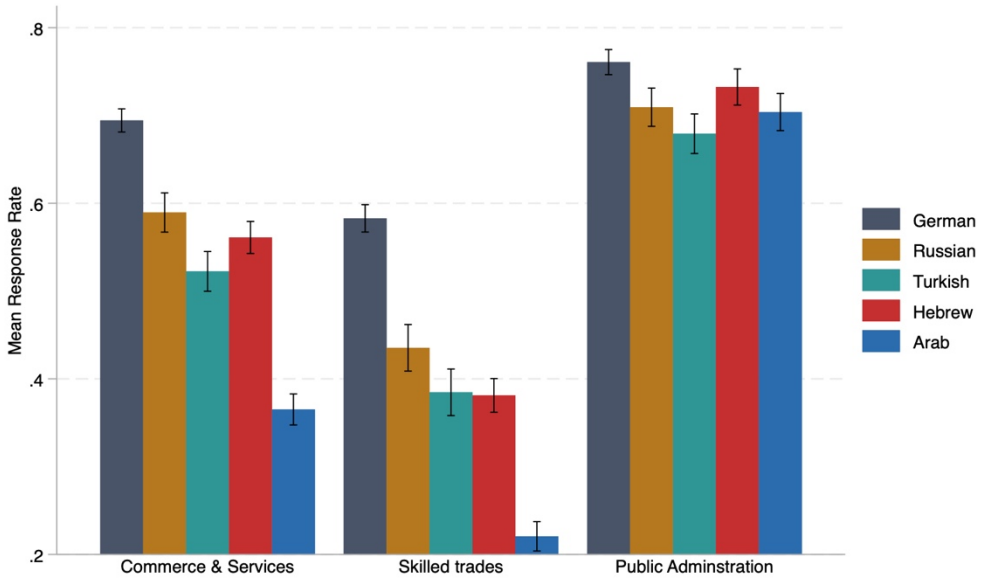


Figure 15: Response rates by sector

Figure 15 shows the variation across the sectors investigated, in both overall response rates and the magnitude of discrimination. Public administration demonstrates both the highest average response rates (approximately 70 percent for German applicants) and the smallest differences across groups. Notably, even applicants with Arab names receive response rates of around 65 percent in this sector, representing a much smaller ethnic gap compared to other industries.

In contrast, the industry and skilled trades sector exhibits the most severe discrimination, with applicants with Arab names receiving only a roughly 22 percent response rate compared to 58 percent for applicants with German names– a 36-percentage point gap. The commerce and services sector falls between these extremes, showing moderate levels of discrimination.

This sectoral variation likely reflects differences in institutional environments, with public administration facing stronger formal and informal pressures for equal opportunity. Additionally, differences in skill requirements, customer contact levels, and workplace cultures across sectors may contribute to these observations.

4.4.2 Robustness Check: Evidence from Campaign 3

We fielded a third campaign to explore the effects of volunteering in a club and to test the robustness of our previous findings. We conducted Campaign 3 from February to March 2025. This campaign focused on applicants with German versus Turkish names and introduced a new treatment dimension: volunteering activities that signal different degrees of cultural proximity.

Experimental Design and Theoretical Motivation

Campaign 3 employed a 2×2×2×3 factorial design, manipulating: (1) ethnicity (German vs. Turkish), (2) gender (male vs. female), (3) academic performance (high vs. satisfactory GPA), and (4) volunteering experience with three conditions:

- No volunteering (control condition)
- Science club participation ("Science club treatment" signals commitment)
- German-Turkish cultural association (“Deutsch Türkische Gesellschaft”, signals cultural proximity)

The theoretical motivation for Campaign 3 builds directly on our employer survey findings from Campaign 2 as we wanted to test whether signaling high effort and extracurricular commitment can mitigate employers’ concerns. Science club participation provides a suitable signal to testing whether perceived perseverance concerns can be addressed through individual signaling. Science competition involvement requires sustained commitment over months, independent research capabilities, and the persistence to complete complex projects—qualities directly relevant to apprenticeship completion. If statistical discrimination based on perceived dropout risk drives differences in response rates, then credible signals of perseverance should reduce discrimination against applicants with Turkish names.

The German-Turkish cultural association (DTG) treatment tests a different dynamic by signaling cultural proximity and familiarity with the Turkish community. If taste-based discrimination results from concerns about cultural distance, unfamiliarity, or discomfort with Turkish culture, then signals of cultural proximity and positive engagement with the Turkish community might have negative effects on applicants with Turkish names.

On the other hand, for applicants with German names, DTG participation signals cultural competence, openness to diversity, and positive attitudes toward Turkish culture—potentially making them more attractive to employers who value intercultural skills. For Turkish applicants, DTG involvement might signal cultural pride and community engagement while demonstrating integration into German society through voluntary participation in formal associations.

This design allows us to differentiate between statistical discrimination (addressed by perseverance signals through science club treatment) and taste-based discrimination (potentially addressed by cultural proximity signals through DTG participation).

Main Results: Robustness of Main Discrimination Findings

Campaign 3 confirms our main findings from Campaigns 1 and 2. Across all volunteering conditions, applicants with German names significantly outperform applicants with Turkish names (Figure 16). This consistency across different time periods (2022-2025), treatment variables (economic credentials vs. internships vs. volunteering), and sample compositions provides strong evidence for the robustness of ethnic discrimination in Germany's apprenticeship market.

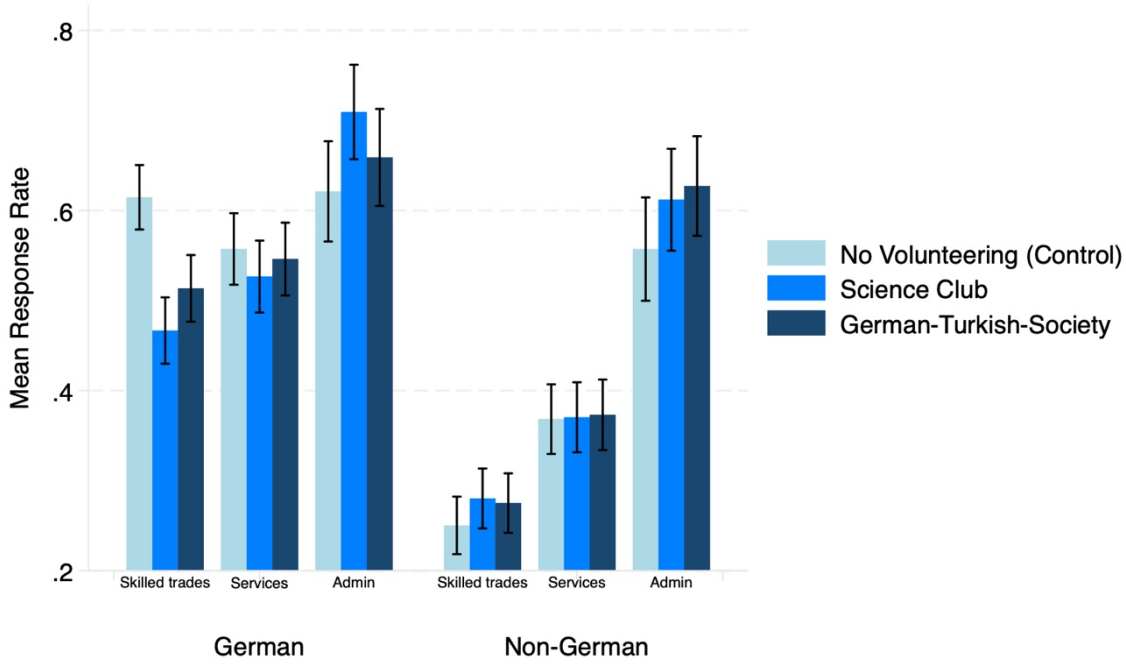


Figure 16: Mean response rates by sectors and volunteering treatment

Volunteering Effects: Testing perseverance vs. Cultural Signaling

The impact of volunteering activities is heterogenous across sectors:
In public administration, German applicants benefit substantially from cultural proximity signaling (DTG

treatment) as well as science club participation. Applicants with Turkish names receive small improvements across volunteering conditions, though differences are not statistically significant.
In the commerce and services sector, responsiveness to these signals is very small: volunteering activities have little impact for both German and Turkish names and response rates and remain relatively stable across all volunteering conditions. The difference in the response rate between applicants with Turkish and German names persists at approximately 18 percentage points regardless of volunteering. Thus, neither the perseverance (science club treatment) nor cultural proximity signals (DTG treatment) reduce discrimination in this sector significantly.

The industry and skilled trades sector reveal counterintuitive negative effects of volunteering for German applicants, as response rates decline across all treatments compared to the previous campaigns. Turkish applicants show small variation across treatments (25 to 28 percent). Surprisingly, perseverance signals fail to help applicants with Turkish names and harm applicants with German names.
It is surprising that participation in a science club cannot reduce differences in response rates for applicants with Turkish names even though this activity provides a credible signal of exactly the trait (perseverance) that employers claimed to be concerned about in our survey. If dropout risk were the primary concern, perseverance signals should improve outcomes. This contradicts the theory of statistical discrimination, which predicts that better information about individual characteristics should reduce reliance on group stereotypes.
It could also be assumed that the negative effects of both science club treatment and DTG participation for applicants with German names in the industry and skilled trades sector go back to an overqualification hypothesis. In that sense, science competition involvement may signal academic interests and capabilities that employers perceive as incompatible with craft apprenticeships. Instead, it may raise concerns about retention likelihood, job satisfaction among academically oriented students, and whether high-achieving students have better alternative opportunities elsewhere.

As we find asymmetric effects of DTG participation across applicant groups and sectors, it could be assumed that cultural proximity to Turkish communities functions only as a positive signal in specific institutional contexts where multiculturalism is explicitly valued. For example, for German applicants, DTG participation provides benefits in public administration but negative effects in craft sectors. This suggests that cultural proximity to Turkish communities may be valued in professional contexts with diverse clientele but may be viewed negatively in traditional sectors. However, for Turkish applicants, DTG participation shows no significant effects across all sectors, failing to reduce discrimination despite signaling positive cultural engagement and community integration.

Overall, the Campaign 3 results demonstrate that signaling effects are highly context-dependent, varying not only by sector but also by the intersection of signal type and applicant characteristics. Moreover, the failure of participation in volunteering activities to reduce discrimination underscores that discrimination in apprenticeship markets operates through complex dynamics, suggesting that there might not be a one-size-fits-all-solution to address these.

5. Discussion

Our correspondence study reveals statistically significant discrimination against underage apprenticeship applicants with foreign-sounding names in Germany. Three key findings emerge from our analysis of approximately 50,000 observations across three experimental campaigns. First, applicants with German names receive response rates of 67 percent, significantly exceeding those with migration backgrounds: for Russian (57 percent), Hebrew (54 percent), Turkish (52 percent), and Arab (40 percent) names. Second, traditional productivity signals—including high academic performance, economic education certificates, internship experience or volunteering—fail to reduce this difference in response rates. Third, discrimination varies significantly geographically and across contexts, with the largest response rate differences occurring in the industry and skilled crafts sectors, the smallest firms and least populated areas. We find the smallest differences in response rates in public administration, large firms and densely populated areas.

Our findings indicate that there is not a single explanation for why discrimination occurs but rather a complex interplay of different mechanisms. Nevertheless, it is striking that certain groups had extremely low chances of receiving a response to their information request.

To begin with, it seems that traditional productivity-related signals and pure merit-based selection do not mitigate discrimination. From an economic perspective, high-quality qualifications should reduce information asymmetries on the employer side and facilitate efficient matching between employers and applicants. While good grades do not increase the chances to receive an answer (for candidates with non-German sounding names) it seems that ethnic stereotypes override productivity-relevant signals.
Moreover, the German apprenticeship system itself shows economic characteristics that could amplify information asymmetries between employers and applicants. Unlike standard labor markets where applicants possess work experience and professional references, apprenticeship candidates are predominantly 15-18 years old with limited employment history. In addition, employers reported in the survey that grades are not meaningful information. Therefore, employers may face a lack of information at the initial screening process. This can favor decision-making in which stereotypes are used instead to evaluate potential applicants, which is referred to in behavioral economics as cognitive bias (Tversky & Kahneman, 1974).

Another aspect is that apprenticeships differ from standard employment in terms of the significant upfront investment requirements and the typical three-year commitment. According to recent data from the BIBB (Wenzelmann & Schönfeld, 2022), gross training costs average €21,000 per apprentice. After accounting for productivity gains of approximately €14,500, the net employer invest per apprentice over the training period amounts to €6,500. This implies that an apprentice can only generate a positive payoff for the employer's firm if the apprentice remains within the firm after the apprenticeship has ended.
Due to asymmetric information about the long-term intentions of the applicants, employers do not know ex ante whether a candidate will remain with the firm after completing their apprenticeship. For example, high performing apprentices may be more likely to leave the firm for university later on. We found that volunteering in a club that signals out interest in STEM ("Jugend forscht") reduced the response rate in the skilled trades sector. Employers may interpret this signal as overqualification and therefore as a negative signal that the

apprentice remains with the firm. This mechanism might also explain, why find that good grades are not significantly increasing the chances to receive an answer in our study. It may disincentive employers to hire candidates who signal high ambition, even if they are highly capable.
Even if we provided additional competences to increase quality of the candidate, we could not find any positive and significant effect on the reply rate. Here, we observed that basic business competences, as signaled out with the EBC*L certificate or previous work experience did not significantly increase the chances to receive an answer. All of the findings underscore that employers are rationally inattentive and are responsive only to those candidates that might be -in their own eyes - a good fit for the firm.

Survey respondents told us that applicants with migration background might find it more difficult to blend into the work environment. This is particularly plausible in smaller firms in regions with a low prevalence of migrants. Note that we found the largest treatment effects in firms up to 5 employees in the hinterland of the agglomerations.
To sum up and assuming that every employer faces a negative payoff of an apprentice, it might be plausible to conclude that a risk-averse employer tries to minimize the marginal costs of an additional apprentice by relying on stereotypes due to the lack of information as proxies for individual success probability. In that case, it could be argued that employers are actors exhibiting bounded rationality who try to minimize expected costs and perceived risk depending on the limited information available, which is ultimately a form of unintended statistical discrimination.
These considerations align with our sector-specific observations: Public administration demonstrates the smallest difference in response rates between Germans and non-Germans, typically involves lower training investments than industry and skilled crafts training which require specialized technical training with uncertain transferability across firms. Thus, sectors with different cost structures and skill requirements may favor different levels of risk aversion. This could reflect employers' rational concerns about returns on investment and opportunity costs.
At the same time, even if this behavior is economically rational for an employer, it creates larger social welfare losses by creating information mismatches and excluding potentially suitable candidates. As a result, if highly qualified candidates are systematically excluded from training opportunities because of foreign-sounding names, while employers simultaneously report difficulties in filling available positions (69,400 vacant positions in 2024), employers may forgo productivity gains and miss opportunities to address the pressing skills shortages. From an economic perspective, this is especially relevant in the German crafts and industrial sectors which is crucial to Germany's industrial competitiveness. Further welfare losses are created if excluded candidates experience reduced lifetime earnings prospects.

However, this complex interplay of specific market, institutional and organizational factors and discrimination creates complex dynamics that are difficult to address. Our findings suggest several potential directions for policy consideration, though each involves trade-offs that need careful evaluation. The ineffectiveness of productivity signals in reducing discrimination points towards possible benefits of more standardized assessment mechanisms or structured evaluation procedures. The regional variation we observe—with discrimination decreasing in areas with higher foreign population shares—suggests that contact effects may play a role, though the direction of causality remains unclear. The sectoral differences, from minimal gaps in public

administration to large disparities in crafts, indicate that institutional context matters. From an economic perspective, the coexistence of unfilled apprenticeship positions alongside systematic exclusion of qualified minority candidates suggests welfare losses exist. However, the optimal design and implementation of policy intervention to correct such issues require further research.

Limitations and future research

Several limitations should be acknowledged in interpreting these findings. Our correspondence study design captures only initial employer responses, and discrimination may also occur at subsequent stages of the hiring process that our methodology cannot detect. The use of error-free German language in our experimental e-mails provides conservative estimates of discrimination as it may not reflect real-world applications from candidates belonging to minority groups. Additionally, our employer survey represents only a subset of contacted firms. Moreover, in our future work, we will also analyze the content of the e-mails received, ideally in combination with information about who sent the email. However, due to the nature of this experimental design, it is often not possible to clearly identify the sender, her origin or position in the organization.

Future research should examine discrimination across the entire hiring process to identify where and how biases operate beyond initial screening. Further extended surveys and experiments with employers can help to identify more factors and concerns that drive selection processes on the employer side.
Moreover, comparative analysis across countries with different apprenticeship systems could identify institutional features that most effectively address discrimination. That may allow the evaluation of interventions designed to reduce discrimination in apprenticeship contexts and could provide valuable guidance for policies. More research on hiring that includes anonymous application procedures, standardized evaluation criteria, and transparency requirements could contribute to the understanding of discrimination dynamics.

6. Conclusion

This study examines hiring discrimination in Germany's apprenticeship market using three correspondence field experiments between 2022 and 2025, complemented by a survey with over 700 employers. Across all studies with more than 50,000 observations, we find significant discrimination against applicants with foreign sounding names. Applicants with German sounding names receive significantly more responses than equally qualified applicants with Turkish, Arab, Russian, or Hebrew names.

Three key findings emerge. First, discrimination persists despite productivity signals. Academic achievements, economic education certificates, internship experiences, and volunteer activities fail to improve response rates from minority applicants. This suggests that employer decisions are not solely driven by considerations of whether the prospective applicant is able to do the job well, indicating the presence of taste-based discrimination alongside statistical mechanisms. Second, we observe variation across organizational and regional contexts. Discrimination is most pronounced in industry and skilled crafts sectors and for small firms, while public administration and large firms show minimal differences. Discrimination is significantly more prevalent in rural areas than in urban areas.

The apprenticeship context may favor discrimination through several channels. Long-term training commitments of three years, high employer investments (€21,000 per apprentice), and information asymmetries may favor conditions conducive to both statistical and taste-based discrimination. The institutional structure of Germany's dual system, while providing valuable skills training, may facilitate discriminatory selection practices, particularly in smaller firms with informal hiring procedures. From an economic perspective, the coexistence of 69,400 unfilled apprenticeship positions annually with exclusion of qualified candidates suggests that welfare losses are created through these inefficiencies in the hiring processes.

Our results indicate that addressing discrimination in apprenticeship markets requires interventions that target both information problems, biases and underlying employer preferences, while accounting for the substantial institutional and regional variation that we find.

Literature

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. Industrial and Labor Relations Review, 30(2), 175-187.

Arrow, K. J. (1973). The Theory of Discrimination. *Discrimination in Labor Markets,* ed. O. Ashenfelter and A. Rees. New Jersey: Princeton University Press.

Arrow, K. J. (1998). What Has Economics to Say about Racial Discrimination? J*ournal of Economic Perspectives* 12 (2): 91–100. DOI: 10.1257/jep.12.2.91

Becker, G. S. (1957). *The Economics of Discrimination*. 2nd ed. Chicago: The University of Chicago Press.

Bertrand, M. and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments,*Volume 1, 309–393.

Bertrand, M. and Mullainathan, S. (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.

Blommaert L., Coenders M., van Tubergen F. (2014). Discrimination of Arabic Named Applicants in the Netherlands: An Internet-Based Field Experiment Examining Different Phases in Online Recruitment Procedures. Social Forces 92, 957–982.

Bundesinstitut für Berufsbildung (BIBB) (2022): Datenreport zum Berufsbildungsbericht 2022 - Informationen und Analysen zur Entwicklung der beruflichen Bildung. Bonn, available under: https://www.bibb.de/dokumente/pdf/Datenreport%202022_20102022_online.pdf

Bundesinstitut für Berufsbildung (BIBB) (2023): Datenreport zum Berufsbildungsbericht 2023 - Informationen und Analysen zur Entwicklung der beruflichen Bildung, Bonn, available under: https://www.bibb.de/dokumente/pdf/bibb_datenreport_2023_korr_11102023.pdf

Bundesinstitut für Berufsbildung (BIBB) (2024). Datenreport zum Berufsbildungsbericht 2024 - Informationen und Analysen zur Entwicklung der beruflichen Bildung. Bonn, available under: https://www.bibb.de/datenreport/de/189191.php

Carlsson, M., Rooth, D.-O., (2006). Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data, No 2281, *IZA Discussion Papers*, Institute of Labor Economics (IZA), https://EconPapers.repec.org/RePEc:iza:izadps:dp2281 .

Destatis (2023). Zahl der Woche Nr. 30 vom 25. Juli 2023 https://www.destatis.de/DE/Presse/Pressemitteilungen/Zahl-der-Woche/2023/PD23_30_p002.html

Destatis (2025). Migrationshintergrund Definition. available under: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Glossar/migrationshintergrund.html

Drydakis, N.; Vlassis, M. (2010). Ethnic Discrimination in the Greek Labour Market: Occupational Access, Insurance Coverage, and Wage Offers. *Manchester Schoo*l, 78:3, 201-218. .

Echenique, F., Li, A.,(2025). Rationally Inattentive Statistical Discrimination: Arrow Meets Phelps, *Journal of the European Economic Association*, 2025, https://doi.org/10.1093/jeea/jvaf002.

Hermes, H.; Mierisch, F.; Lergetporer, P.; Peter, F.; Wiederhold.. S. (2023). *Discrimination on the Child Care Market: A Nationwide Field Experiment*. Discussion Paper No. 16082, IZA Institute of Labor Economic

Kaas, L. and Manger, C. (2012). Ethnic Discrimination in Germany's Labour Market: A Field Experiment. *German Economic Review*, 13:1, 1-20.

Koopmans, R.; Veit, S.; Yemana, R. (2019). Taste or statistics? A correspondence study of ethnic, racial and religious labour market discrimination in Germany. *Ethnic and Racial Studies,* 42:16, 233-252.

Lippens, L., Vermeiren, S., Baert, S. (2023). The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments. *European Economic Review.* 151, https://doi.org/10.1016/j.euroecorev.2022.104315.

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

Schneider, J., Yemane, R.; Weinmann, M. (2014). *Diskriminierung am Ausbildungsmarkt. Ausmaß, Ursachen und Handlungsperspektiven*. Sachverständigenrat deutscher Stiftungen für Integration und Migration, Berlin.

Streeck, W.: (1991). On the Social and Political Conditions of Diversified Quality Production. In: Matzner E, Streeck W (eds) Beyond Keynesianism. Aldershot: Edward Elgar, 21-61.

Thijssen, L.; Lancee, B.; Veit, S., Yemane, R. (2021). Discrimination against Turkish minorities in Germany and the Netherlands: field experimental evidence on the effect of diagnostic information on labour market outcomes. *Journal of Ethnic and Migration Studies*, 47:6, 1222-1239.

Tversky, A.;, Kahneman, D. (1974).Judgment under Uncertainty: Heuristics and Biases. *Science*, 185,1124-1131.DOI https://doi.org/10.1126/science.185.4157.1124" \h

Wenzelmann, F.; Schönfeld, G. (2022).Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe. Ergebnisse der sechsten BIBB-Kosten-Nutzen-Erhebung. Bonn 2022

Wößmann et al. (2024). Ungleiche Bildungschancen: Ein Blick in die Bundesländer. ifo Schnelldienst 5 / 2024 (77).

Zimmermann K., Biavaschi C., Eichhorst W. et al. (2013). Youth Unemployment and Vocational Training, Foundations and Trends in Microeconomics: Vol. 9: No. 1–2, pp 1-157. http://dx.doi.org/10.1561/0700000058

Zschirnt, E., Ruedin, D. (2016). Ethnic discrimination in hiring decisions: A meta-analysis of correspondence tests 1990–2015, *Journal of Ethnic and Migration Studies*, 42(7), pp. 1–19, https://doi.org/10.10800/133693183X.20105.11303279.

APPENDIX

In the following, additional data and tests for section 4 are presented.

The following table shows the summary statistics of the data set.

Table 1: Summary Statistics

Variable	**Mean**	**Std. dev**	N
Reply	0.5613	0.4962	39,306
Female	0.5003	0.5001	40,213
High GPA	0.4988	0.5000	40,213
Foreign Name	0.6652	0.4719	40,213
Arab Name	0.3306	0.4704	21,257
Hebrew Name	0.3340	0.4717	21,257
Turkish Name	0.3329	0.4712	18,956
Russian Name	0.3328	0.4712	18,956
German Name	0.3348	0.4719	40,213
Campaign 1	0.4714	0.4991	40,213
Campaign 2	1	0	21,257
N	40,213		

Table 2: Marginal effects from logit regressions

Variables	(1) Basic Marginal effects	(2) Firm size Marginal effects	(3) Labor market Marginal effects
Migration background	-0.157***	-0.205***	-0.155***
Female applicant	-0.00695	-0.0136**	-0.00197
High GPA school performance	0.0156***	0.0191***	0.00844
	(0.00539)	(0.00683)	(0.00649)
Crafts & Manufacturing sector	-	-0.150***	-0.290***
Services sector	-	0	-0.164***
		(omitted)	(0.00815)
Public administration sector	-	0	0
		(omitted)-	(omitted)
Firm size category 1	-	-0.235***	-
		(0.0190)	
Firm size category 2	-	-0.185***	-
		(0.0159)	
Firm size category 3	-	-0.0403**	-
		(0.0162)	
Firm size category 4	-	-	-
Share of foreigners of total population	-	-	0.00390*** (0.000569)
Social Transfers per capita	-	-	3.65e-05*** (1.22e-05)
Unemployment rate in %	-0.0193*** (0.00286)	-	-0.0136*** (0.00375)
Youth unemployment rate in %	0.0149***	-	0.00873
Share of non-working population		-	0.00156 (0.00116)
Observations	32,858	19,538	21,287

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Marginal effects from logit regressions with robust standard errors in parentheses. Results show percentage point changes in response probability. We estimate marginal effects based on logit estimations to further explain the variance of responsiveness. The firm size category 1 implies 1 to 5 employees, category 2 6 to 50 employees, category 3 51 to 500 employees and category 4 501 to 5000.

Table 3: Treatment Overview

Campaign	Treatment	Name	Ethnicity	Gender	GPA	Other information
1	1	Leonie Schröder	German	Female	1.3	no EBC*L, knowledge of English language
1	2	Leonie Schröder	German	Female	3.0	no EBC*L, knowledge of English language
1	3	Lukas Becker	German	Male	1.3	no EBC*L, knowledge of English language
1	4	Lukas Becker	German	Male	3.0	no EBC*L, knowledge of English language
1	5	Zeynep Yildrim	Turkish	Female	1.3	no EBC*L, knowledge of English and Turkish language
1	6	Zeynep Yildrim	Turkish	Female	3.0	no EBC*L, knowledge of English and Turkish language
1	7	Yusuf Kaya	Turkish	Male	1.3	no EBC*L, knowledge of English and Turkish language
1	8	Yusuf Kaya	Turkish	Male	3.0	no EBC*L, knowledge of English and Turkish language
1	9	Olga Ivanova	Russian	Female	1.3	no EBC*L, knowledge of English and Russian language
1	10	Olga Ivanova	Russian	Female	3.0	no EBC*L, knowledge of English and Russian language
1	11	Ivan Smirnov	Russian	Male	1.3	no EBC*L, knowledge of English and Russian language
1	12	Ivan Smirnov	Russian	Male	3.0	no EBC*L, knowledge of English and Russian language
1	13	Leonie Schröder	German	Female	1.3	EBC*L, knowledge of English language
1	14	Leonie Schröder	German	Female	3.0	EBC*L, knowledge of English language
1	15	Lukas Becker	German	Male	1.3	EBC*L, knowledge of English language
1	16	Lukas Becker	German	Male	3.0	EBC*L, knowledge of English language
1	17	Zeynep Yildrim	Turkish	Female	1.3	EBC*L, knowledge of English and Turkish language
1	18	Zeynep Yildrim	Turkish	Female	3.0	EBC*L, knowledge of English and Turkish language
1	19	Yusuf Kaya	Turkish	Male	1.3	EBC*L, knowledge of English and Turkish language
1	20	Yusuf Kaya	Turkish	Male	3.0	EBC*L, knowledge of English and Turkish language

Campaign	Treatment	Name	Ethnicity	Gender	GPA	Other information
1	21	Olga Ivanova	Russian	Female	1.3	EBC*L, knowledge of English and Russian language
1	22	Olga Ivanova	Russian	Female	3.0	EBC*L, knowledge of English and Russian language
1	23	Ivan Smirnov	Russian	Male	1.3	EBC*L, knowledge of English and Russian language
1	24	Ivan Smirnov	Russian	Male	3.0	EBC*L, knowledge of English and Russian language
2	1	Marie Schmidt	German	Female	1,3	No internship experience
2	2	Marie Schmidt	German	Female	3,0	No internship experience
2	3	Leon Müller	German	Male	1,3	No internship experience
2	4	Leon Müller	German	Male	3,0	No internship experience
2	5	Yael Rosenberg	Hebrew	Female	1,3	No internship experience
2	6	Yael Rosenberg	Hebrew	Female	3,0	No internship experience
2	7	Ariel Rubinstein	Hebrew	Male	1,3	No internship experience
2	8	Ariel Rubinstein	Hebrew	Male	3,0	No internship experience
2	9	Habiba Mahmoud Al Numan	Arab	Female	1,3	No internship experience
3	10	Habiba Mahmoud Al Numan	Arab	Female	3,0	No internship experience
3	11	Mohamad Khalil Khaled	Arab	Male	1,3	No internship experience
3	12	Mohamad Khalil Khaled	Arab	Male	3,0	No internship experience
2	13	Marie Schmidt	German	Female	1,3	Internship experience
2	14	Marie Schmidt	German	Female	3,0	Internship experience
2	15	Leon Müller	German	Male	1,3	Internship experience
2	16	Leon Müller	German	Male	3,0	Internship experience
2	17	Yael Rosenberg	Hebrew	Female	1,3	Internship experience
2	18	Yael Rosenberg	Hebrew	Female	3,0	Internship experience
2	19	Ariel Rubinstein	Hebrew	Male	1,3	Internship experience
2	20	Ariel Rubinstein	Hebrew	Male	3,0	Internship experience
2	21	Habiba Mahmoud Al Numan	Arab	Female	1,3	Internship experience

Campaign	Treatment	Name	Ethnicity	Gender	GPA	Other information
2	22	Habiba Mahmoud Al Numan	Arab	Female	3,0	Internship experience
2	23	Mohamad Khalil Khaled	Arab	Male	1,3	Internship experience
2	24	Mohamad Khalil Khaled	Arab	Male	3,0	Internship experience
3	1	Leonie Schröder	German	Female	1.3	No volunteering
3	2	Leonie Schröder	German	Female	3.0	No volunteering
3	3	Lukas Becker	German	Male	1.3	No volunteering
3	4	Lukas Becker	German	Male	3.0	No volunteering
3	5	Zeynep Yildrim	Turkish	Female	1.3	No volunteering
3	6	Zeynep Yildrim	Turkish	Female	3.0	No volunteering
3	7	Yusuf Kaya	Turkish	Male	1.3	No volunteering
3	8	Yusuf Kaya	Turkish	Male	1.3	No volunteering
3	9	Leonie Schröder	German	Female	1.3	Volunteer at German-Turkish association
3	10	Leonie Schröder	German	Female	3.0	Volunteer at German-Turkish association
3	11	Lukas Becker	German	Male	1.3	Volunteer at German-Turkish association
3	12	Lukas Becker	German	Male	3.0	Volunteer at German-Turkish association
3	13	Zeynep Yildrim	Turkish	Female	1.3	Volunteer at German-Turkish association
3	14	Zeynep Yildrim	Turkish	Female	3.0	Volunteer at German-Turkish association
3	15	Yusuf Kaya	Turkish	Male	1.3	Volunteer at German-Turkish association
3	16	Yusuf Kaya	Turkish	Male	1.3	Volunteer at German-Turkish association
3	17	Leonie Schröder	German	Female	1.3	Volunteer at science club
3	18	Leonie Schröder	German	Female	3.0	Volunteer at science club
3	19	Lukas Becker	German	Male	1.3	Volunteer at science club
3	20	Lukas Becker	German	Male	3.0	Volunteer at science club
3	21	Zeynep Yildrim	Turkish	Female	1.3	Volunteer at science club
3	22	Zeynep Yildrim	Turkish	Female	3.0	Volunteer at science club
3	23	Yusuf Kaya	Turkish	Male	1.3	Volunteer at science club
3	24	Yusuf Kaya	Turkish	Male	1.3	Volunteer at science club

Table 3 displays the treatment variations across the campaigns. The first dimensions of the ethnicity, i. e. potential immigrant background as signaled by the inquirer's name signals. The second dimension is the gender of the inquirer. For example, the names 'Leonie Schröder' and 'Lukas Becker' are selected to signal German ethnicity of a female and a male student, respectively. 'Zeynep Yildirim' and 'Yusuf Kaya' suggest a Turkish migrant background of a female and male student, respectively. Given the context of the applicants' interest in upper secondary school after grade 10, an assumed age of 16 is assigned. The names chosen ranked among the most prevalent first names for girls and boys born in 2005 to 2007 in Germany and the respective country of origin. The third dimension varies the quality of the educational background of the inquirer. The students indicate that they are likely to graduate with a high GPA of 1.3 or a low GPA of 3.0, here used to signal differences in performance at school. The fourth dimension, depending on the campaigns, shows potentially relevant other qualifications, namely the EBC*L economic education certificate in campaign 1, work experience through internships in campaign 2 and volunteering experience in campaign 3.