



UvA-DARE (Digital Academic Repository)

Is there evidence for statistical discrimination against ethnic minorities in hiring? Evidence from a cross-national field experiment

Thijssen, L.; Coenders, M.; Lancee, B.

DOI

[10.1016/j.ssresearch.2020.102482](https://doi.org/10.1016/j.ssresearch.2020.102482)

Publication date

2021

Document Version

Final published version

Published in

Social Science Research

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Thijssen, L., Coenders, M., & Lancee, B. (2021). Is there evidence for statistical discrimination against ethnic minorities in hiring? Evidence from a cross-national field experiment. *Social Science Research*, 93, Article 102482. <https://doi.org/10.1016/j.ssresearch.2020.102482>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Is there evidence for statistical discrimination against ethnic minorities in hiring? Evidence from a cross-national field experiment

Lex Thijssen^{a,*}, Marcel Coenders^{b,c}, Bram Lancee^d

^a Department of Sociology, Utrecht University, Padualaan 14, 3584, CH Utrecht, the Netherlands

^b Netherlands Institute for Social Research, The Hague, the Netherlands

^c Department of Interdisciplinary Social Science, Utrecht University, Utrecht, the Netherlands

^d Department of Sociology, University of Amsterdam, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:

Diagnostic personal information
Ethnic discrimination
Field experiment
Group productivity
Hiring
Statistical discrimination

ABSTRACT

While statistical discrimination theory is often proposed as an important explanation for ethnic discrimination in hiring, research that empirically scrutinizes its underlying assumptions is scant. To test these assumptions, we combine data from a cross-national field experiment with secondary data indicative of the average labor productivity of ethnic communities. We find little evidence that adding diagnostic personal information reduces discrimination against ethnic minorities. Furthermore, we do not find an association between language similarity or the socioeconomic resources of the ethnic community and hiring discrimination. However, our findings show that discrimination is related to the socioeconomic development of the country of ancestry. Finally, the impact of these indicators of group productivity is generally not moderated by the amount of diagnostic personal information. Taken together, these findings question several core assumptions of statistical discrimination theory.

1. Introduction

Study upon study has shown that ethnic minorities are being discriminated against in hiring (Quillian et al. 2017, 2019; Zschirnt and Ruedin 2016).¹ Research further indicates that some ethnic minority groups face higher levels of discrimination than others (Ahmad 2019; Booth et al. 2012; Weichselbaumer 2017; Zschirnt 2019). While its existence has been frequently demonstrated, the mechanisms generating (differences in) discrimination rates against ethnic minorities remain strongly debated (Bertrand and Duflo 2017; Neumark 2018; Quillian 2006).

Scholars have suggested that – in addition to economically-irrational ethnic tastes and intergroup conflicts (e.g. Becker 1957; Blumer 1958) – economic rationality and information deficiencies may also explain ethnic discrimination in hiring (Baumle and Fossett 2005; Guryan and Charles 2013). According to economic models of statistical discrimination theory (Aigner and Cain 1977; Arrow 1973; Phelps 1972; Schwab 1986) – originally proposed by Kenneth Arrow and Edmund Phelps – economically maximizing employers systematically prefer ethnic majority over minority job applicants due to imperfect information in the recruitment process

* Corresponding author. Netherlands Institute for Social Research, The Hague, the Netherlands.

E-mail address: lexthijssen@gmail.com (L. Thijssen).

¹ In this study, we focus on ethnic minorities with a migrant background – that is, those whose parents or themselves were born abroad.

and “objective” skill disparities between ethnic groups.²

The large body of research on statistical discrimination theory (Bertrand and Duflo 2017; Guryan and Charles 2013) has primarily focused on two assumptions for discrimination of ethnic minorities. One strand of research focuses on whether hiring discrimination is affected by information deficiencies and, more specifically, by the (lack of) personal information on labor productivity. Yet, only a few studies find that adding diagnostic personal information eliminated discrimination against ethnic minorities (Baert and Vujčić 2016; Kaas and Manger 2012). By contrast, most studies find no effect of information (Agerström et al., 2012; Bertrand and Mullainathan 2004; Gaddis 2015; Koopmans et al. 2018; Nunley et al., 2015; Oreopoulos 2011; Vernby and Dancygier 2019). Although these findings appear to be largely inconsistent with statistical discrimination theory, these differences may be due to differences in field experimental designs, the selected minority groups, or national contexts.

The second strand of research has been concerned with assessing how discriminatory hiring practices are related to the group beliefs of employers. Surprisingly, however, statistical discrimination theory’s assumption that the average group labor productivity, as well as its variance, affect employers’ hiring practices has received much less scholarly attention (Midtbøen 2014; Pager and Karafin 2009; Quillian 2006; Schaeffer et al. 2016). A small but growing number of qualitative studies examined whether and how group beliefs of employers are linked with the negative recruitment outcomes of ethnic minorities (Bonoli and Hinrichs 2012; Friberg and Midtbøen 2018, 2019; Imdorf 2017; Midtbøen 2014; Moss and Tilly 2001; Neckerman and Kirschenman 1991; Pager and Karafin 2009; Waldinger and Lichter 2003). These studies find that employers base their hiring decisions on more “objective” assessments of performance and risk as well as on prejudiced attitudes towards ethnic minorities. While these qualitative studies provide valuable insights, research is lacking that quantitatively assesses whether actual hiring outcomes can be linked to group indicators of productivity and which of these indicators are most important for explaining ethnic discrimination in the labor market (cf. Quillian 2006).

In this study, our aim is to empirically scrutinize the assumptions of statistical discrimination theory. Using original data from a cross-national comparative field experiment in Germany, the Netherlands, Norway, Spain, and the United Kingdom (Lancee et al., 2019a, 2019b), we seek to advance previous research in three ways. First, to test whether ethnic discrimination is indeed lower once job applicants include more diagnostic personal information in application materials, we randomly varied the ethnic origin and three types of diagnostic information signaling productivity. Specifically, we experimentally manipulated whether or not the average final grade was mentioned, whether or not job applicants describe themselves as a person with strong social skills, and whether or not applicants list additional skills and extra responsibilities in their prior job. Studying multiple information manipulations in five national contexts enables us to test whether (and which type of) diagnostic personal information leads to a reduction of ethnic discrimination.

Second, we undertake – to the best of our knowledge – the first empirical investigation of the relationship between discrimination and group averages of labor productivity in multiple countries. In each of the five countries in our study, we examine thirty-one ethnic groups with varying socioeconomic backgrounds matched with three indicators of group-level labor productivity: (1) a measure combining aggregated information about educational and employment outcomes of ethnic minority groups in the country of destination (i.e. ethnic community) (OECD 2010), (2) a measure of similarity between the language of the destination country and the dominant language in the country of origin (Holman et al., 2011; Wichmann et al. 2018), and (3) a measure capturing the average level of socioeconomic development in the country of origin – that is, the Human Development Index (HDI) (United Nations Development Programme, 2018). These indicators allow for differentiating between the productivity of the ethnic community and productivity linked to the country of ancestry. By distinguishing between these proximate (indicator 1) and more distant (indicators 2 and 3) indicators of group productivity, we study the relative importance of different aspects of labor productivity in explaining ethnic discrimination in hiring (cf. Friberg and Midtbøen 2018; Midtbøen 2014).

Third, previous research suggests that the effect of group productivity on discrimination of individuals is contingent on the availability of diagnostic information about individual productivity (Arrow 1973; Phelps 1972; Zschirnt and Ruedin 2016). In particular, it has been suggested that once economically-rational employers have more and reliable personal information to infer the individual productivity of job applicants, they rely less on group statistics (Crawford et al., 2011; Guryan and Charles 2013; Rubinstein et al. 2018). Consequently, employers’ concerns regarding individual productivity depend on the amount of diagnostic personal information available and, likewise, the strength of the effect of diagnostic personal information can be affected by the content of the beliefs about the ethnic minority group. However, this idea has not been explicitly tested in prior research. Hence, we contribute to the literature by integrating insights of two separate strands of research and empirically testing whether the effect of group productivity is smaller if resumes contain more diagnostic personal information.

The question we set out to answer is: *to what extent is discrimination against (different) ethnic minority groups in Germany, the Netherlands, Norway, Spain, and the United Kingdom affected by the (independent and/or interactive) effects of (the amount of) diagnostic personal information and group averages of labor productivity?* In answering this question, we focus on a specific group of job seekers. First, we examine job seekers at the start of their working careers (aged 23–25 years, \pm 4 years work experience) because prior research indicates that prolonged unemployment spells at the beginning of people’s career are associated with an increased risk of unemployment later in life (Luijckx and Wolbers 2009). Current ethnic disadvantages might accordingly have important consequences for future labor market inequalities. Second, we study ethnic minority candidates who were raised and obtained all their education and

² As we discuss below, proponents of statistical discrimination theory assume that employers act on the basis of “objective” group information or “objective” stereotypes. Generally, psychologists (and sociologists) are skeptical about this claim. In this literature, however, there is debate as to whether group beliefs or stereotypes are largely disconnected from reality or could also stem from people’s direct and indirect observations of ethnic groups in society (Dixon 2017; Jussim et al., 2009; Koenig and Eagly 2014).

previous job experience in the country of study. We, therefore, investigate ethnic minority applicants who should be in the same position to successfully realize their preferred career path as their majority counterparts. Uncertainty about educational degrees and work experience obtained abroad is therefore not an issue in this study (cf. Oreopoulos 2011).

2. Theoretical background

Statistical discrimination theory posits that discrimination “is based on rational decisions by maximizing actors who are guided by empirically informed assessments of productivity and risk” (Baumle and Fossett 2005:1251). The literature has focused on two mechanisms that explain why employers discriminate against ethnic minorities: (1) employers are faced with (more or less) information uncertainties in hiring processes (information uncertainty); and (2) employers base their decisions on their information about the productivity of ethnic groups (group productivity). Below we discuss whether and how these different mechanisms are backed up by theoretical and empirical insights.

2.1. The effect of diagnostic personal information

One branch of research has focused on the impact of information uncertainty on hiring outcomes (Bertrand and Duflo 2017; Neumark 2018). Information uncertainty arises, for example, because CVs and cover letters contain too little or unreliable information about the job applicant, and there is a lack of time and monetary resources to conduct an extensive assessment of the full pool of job applicants (e.g. screening, training, and dismissal costs) (Baumle and Fossett 2005; Midtbøen 2014). To overcome these information deficiencies and a way to minimize the risk of making wrong hiring decisions, employers rely on group membership as a readily available and inexpensive proxy for the (unobserved) productivity of job applicants (Arrow 1973; Phelps 1972). Employers could rely on both the mean productivity of groups as well as its variance. Discrimination is expected to be higher when employers have little information about the productivity of job applicants; conversely, if employers have perfect information about the productivity of the job seeker, they will not discriminate on the basis of ethnicity (Guryan and Charles 2013).

A growing number of studies have investigated whether discrimination is related to information uncertainty, and specifically tested whether discrimination is lower when more individual information was available (Bertrand and Duflo 2017; Neumark 2018). Support for this line of reasoning was found in studies on racial and ethnic discrimination in laboratory experiments (Lane 2016), the rental housing market (Auspurg et al. 2019; Flage 2018), and the sharing economy (Tjaden et al. 2018).

Studies on employment discrimination, however, find much less support for this argument. For example, several studies compared the level of discrimination among applicants with lower and higher quality educational credentials (Bertrand and Mullainathan 2004; Gaddis 2015; Nunley et al., 2015; Oreopoulos 2011), but find no evidence of lower discrimination rates when job applicants signal higher levels of competence and commitment. Besides manipulating the quality of educational qualifications, studies experimentally varied information about productivity-relevant characteristics of job applicants, such as descriptions about one’s personality (Agerström et al., 2012), out of work activities (Baert and Vujčić 2016), language skills (Oreopoulos 2011), the average final grade in education (Koopmans et al., 2018; Thijssen et al., 2019), additional tasks and responsibilities in previous job(s) (Andriessen et al., 2010; Thijssen et al., 2019; Vernby and Dancygier 2019), or reference letters from previous employers (Kaas and Manger 2012; Koopmans et al., 2018). The majority of studies find no evidence that discrimination decreased with the addition of personal information (Agerström et al., 2012; Andriessen et al., 2010; Koopmans et al., 2018; Oreopoulos 2011; Thijssen et al., 2019; Vernby and Dancygier 2019); only a few studies indicate that adding personal information eliminated ethnic discrimination (Baert and Vujčić 2016; Kaas and Manger 2012).

One interpretation of these results is that the impact of information is limited; however, the absence of more corroborative evidence might also be due to differences in, for instance, the experimental manipulations used, the selection of ethnic minority groups, or the national context. To provide a more comprehensive test, we independently manipulated different types of information in application materials. Furthermore, rather than only examining which type of information can minimize ethnic biases in hiring, we also assess whether discrimination is highest when no additional information was included in the resume. In line with statistical discrimination theory, we hypothesize that (1): *Discrimination against ethnic minority job applicants is lower when job applicants reveal more diagnostic personal information in their application.*

2.2. The effect of group productivity

In contrast to information uncertainties, there is much less scholarly attention to the influence of the mean and variance of the labor productivity of ethnic groups. This is nonetheless a crucial aspect in Arrow (1973) and Phelps (1972) original formulations of statistical discrimination theory. In particular, it has been argued that if economically rational employers have information that members of ethnic minority groups have lower productivity than those of ethnic majority groups, they will discriminate against ethnic minority job seekers. Furthermore, employers striving to minimize surprises in their hiring procedure will select individuals from groups with lower variance in productivity, typically the majority group. In this study, however, we focus on the average productivity of groups as we lack more detailed group information to test the latter.

While some theorists have proposed and tested more flexible interpretations (Altonji and Pierret 2001; Bartoš et al., 2016), most argue that employers should act on the basis of “true stereotypes” – that is, group beliefs which are based on actual performance differences between ethnic groups (Aigner and Cain 1977; Schwab 1986). In the words of Baumle and Fossett (2005, p. 1254): “If the employer ... chooses applicants on the basis of race or some other group membership and group membership is in fact uncorrelated

with unmeasured productivity or risk, the employer is choosing arbitrarily among otherwise similar candidates ... “Indeed, this would contradict that employers are driven by pure economic incentives.

A small body of qualitative research has studied whether and how hiring practices of employers could be affected by the content of group beliefs about ethnic groups (Bonoli and Hinrichs 2012; Friberg and Midtbøen 2018, 2019; Imdorf 2017; Midtbøen 2014; Moss and Tilly 2001; Neckerman and Kirschenman 1991; Nievers 2010; Oreopoulos 2011; Pager and Karafin 2009; Waldinger and Lichter 2003). These studies find that there are (some) employers who have bigoted views towards ethnic minorities, systematically overestimate negative incidents with ethnic minority workers, and/or apply double hiring standards (Midtbøen 2014; Moss and Tilly 2001; Pager and Karafin 2009; Waldinger and Lichter 2003). However, this research also finds evidence that employers say that they prefer majority over minority job applicants because of average skill differences between groups. That is, employers often express concerns about the attachment to work, language proficiency, work ethic, commitment, and professional appearance of ethnic minorities, resulting in trouble avoidance and exclusionary practices (Imdorf 2017; Midtbøen 2014; Moss and Tilly 2001; Neckerman and Kirschenman 1991; Nievers 2010; Oreopoulos 2011; Pager and Karafin 2009). As summarized by Midtbøen (2015, p. 208): “Indeed, many employers display both negative attitudes and crude stereotypes of ethnic minorities, and they clearly express strategies for risk minimization”.

While these studies provide interesting insights, we still do not know whether and which group information is most relevant for explaining hiring discrimination against ethnic minorities. Indeed, various studies have shown that there is not always a clear link between what employers say they do and what they actually do (Pager and Quillian 2005). Relatedly, and equally important, previous research has also not been able to assess whether and to what extent employers’ information about the productivity of ethnic minority groups is consistent with “objective” skill differences between ethnic groups (Pager and Karafin 2009).

The question as to whether employers’ group beliefs can reflect “objective” skill disparities touches upon an old but ongoing debate among psychologists concerning the (in)accuracy of stereotypes about ethnic or racial groups (Allport 1954; Brigham 1971; Dixon 2017; Fiske 1998; Jussim et al. 2015). On the one hand, scholars view stereotypes as “poorly founded beliefs about members of the target group” (Quillian, 2006, p. 300) which exist without any real basis or “kernel of truth” (Brigham 1971; Fiske 1998; LaPiere 1936). On the other hand, a recent series of studies contend that stereotypes can correspond largely with observed differences between ethnic groups (Arkes and Tetlock 2004; Jussim et al. 2009, 2015; Koenig and Eagly 2014; Stevens et al., 2018) and use more neutral definitions of stereotypes such as “a general belief about groups” (Ashmore and Del Boca 1981; Jussim et al., 2015). For example, a recent overview article by Jussim and colleagues (2009, p. 221) concludes that “the scientific evidence provides more evidence of accuracy than of inaccuracy in social stereotypes. The most appropriate generalization based on the evidence is that people’s beliefs about groups are usually moderately to highly accurate, and are occasionally highly inaccurate”. Hence, it is apparent that this debate will not be settled soon (Dixon 2017) and that it is still unclear whether group beliefs or stereotypes are largely disconnected from reality or could also stem from people’s direct and indirect observations of ethnic groups in society.

This debate about the (in)accuracy of group beliefs or stereotypes also indicates that it is insightful to investigate whether hiring outcomes are related to indicators of group productivity. In accordance with statistical discrimination theory, it can be expected that economically rational employers base their hiring decisions on information about the productivity of ethnic groups. In this study, we therefore complement earlier qualitative findings by testing whether and to what extent publicly available group information correlates with discrimination rates. In doing so, we follow statistical discrimination theory and assume that employers rely on (available) statistics indicative of the average labor productivity of ethnic groups. We specifically expect that (2): *Job applicants face higher levels of discrimination when they belong to ethnic groups with lower levels of labor productivity.*

As indicated by Friberg and Midtbøen (2018:1465): “The economic model may seem straight forward, but in reality the term on which it all hinges – skills – is rather vague and may refer to a wide variety of knowledge, characteristics and competencies that are not easily conceptualized or measured (Moss and Tilly 2001)”. Indeed, employers’ concerns about unobserved productivity might be related to, for example, professionalism, trustworthiness, the quality of social networks, communication problems due to language dissimilarity, trainability, or work attachment. Therefore, we test this hypothesis by looking at different aspects of the mean productivity of ethnic minority groups. Following previous studies on the socioeconomic integration of immigrants (Levels et al. 2008; Van Tubergen, Maas, and Flap 2004), we assume that employers hold more positive images regarding ethnic communities (1) with higher levels of education and labor participation in the country of study (i.e. higher levels of socioeconomic resources), (2) with an ancestry language that is more similar to that of the majority in the country of study (i.e. greater language similarity), or (3) that originated from countries of ancestry with a higher level of socioeconomic development. In the measurement section, we come back to this issue. It is, moreover, important to note that the size of skill disparities between ethnic groups is likely to be influenced by historical (taste-based) discrimination (Pager and Karafin 2009; Quillian 2006).

2.3. The interactive effects of diagnostic personal information and group productivity

Mostly, scholars have tested the effect of adding diagnostic information and group productivity in isolation and independently from one another. Strictly speaking, however, statistical discrimination theory states that the impact of group productivity is contingent on the availability of diagnostic information about individual job applicants. In particular, the more diagnostic personal information is available, the less uncertainty employers experience, and the less likely they are to resort to group information.

This line of reasoning is consistent with psychological models that assume that the effect of stereotypes and group images is dependent on the amount of diagnostic personal information (Crawford et al., 2011; Rubinstein 2018; Rubinstein et al., 2018). The diagnosticity and judgment task model contends that although group images can influence a person’s perception, individual information has an equally important and in most instances even a stronger impact than group images (Rubinstein 2018; Rubinstein et al.,

2018). This model specifically expects that only if individual information is absent or lowly diagnostic, people will rely on stereotypes; conversely, when individual information is highly diagnostic, stereotypes should play no part in decision making.

The diagnosticity and judgment task model thus shows strong similarities with statistical discrimination theory and provides an additional theoretical argument to expect that the presence of diagnostic information affects the relationship between (beliefs about) group productivity and hiring discrimination. Therefore, we derive the following hypothesis (3): *The effect of group productivity is weaker when job applicants reveal more diagnostic personal information in their application.*

3. Data and methods

3.1. Data

We make use of data from a cross-nationally comparative correspondence test on hiring discrimination against ethnic minorities in five countries (Germany, the Netherlands, Norway, Spain, and the United Kingdom) (Lancee 2019; Lancee et al., 2019a, 2019b). In this field experiment, fictitious applicants applied to real vacancies posted on online job boards in the period between November 2016 and April 2018.

To compare discrimination rates across countries, similar occupations were examined in all countries, namely: cook, payroll clerk, receptionist, sales representative, software developer, and store assistant. These occupations were chosen to have variation in interpersonal and educational skills. The CVs and cover letters were standardized cross-nationally. To construct realistic cover letters and CVs, resumes of real job seekers were used as examples, and experienced recruiters were asked to evaluate the newly constructed application materials. The cover letter provided reasons to apply for a job position, while the CV includes background information – such as the applicant's age (23–25), postal and e-mail address, telephone number – the educational degrees obtained, and information about previous job positions and employers (four years work experience). To take into account the country-specific labor market context, application documents were slightly adapted per country.

In contrast to many previous field experiments, we used an unpaired design (cf. Koopmans et al., 2018; Weichselbaumer 2017). This means that we applied with only one fictitious job application per job opening. The unpaired design provides increased possibilities of varying multiple experimental manipulations, lowers the risk of detection, and minimizes any inconveniences for employers and actual job applicants (Di Stasio and Lancee, 2020).

3.2. Measurements

3.2.1. Dependent variable

Responses from employers were tracked by matching mail, voice, or email messages to resumes. The dependent variable indicates whether the applicant received a positive response from an employer (callback) – that is, a message in which the employer clearly expressed his or her interest in the candidate (e.g. personal requests for additional information and (pre-)invitations to a job interview; all coded as one). Messages without concrete requests for additional personal information, rejections, or no messages are coded as zero.

3.2.2. Ethnicity

We distinguished between native-majority candidates and candidates with an ethnic minority origin. Furthermore, we varied the country of ancestry of candidates with a minority background. In total, 30 ethnic minority groups were simultaneously examined in all countries. This selection comprises the largest ethnic minority groups per country and groups of varying socioeconomic status (see Table S1 in the Supplementary Materials). Within each country, several groups were oversampled: 25 percent of all applicants had a native-majority origin and 25 percent of all applicants are members of one of the most sizeable or historically well-established minority groups. In each of the five countries, there is one ethnic majority group and 30 ethnic minority origin groups. Thus, at the ethnic community-level (level-2), we have 155 observations.

In correspondence tests on ethnic discrimination, it is important that employers can trace the ethnicity of the fictitious job candidate (Gaddis 2017). Accordingly, ethnic origin was conveyed by the candidate's first and last name (signaled in the cover letter and CV), the language skills mentioned in the CV (apart from mentioning the language of the country of study as their mother tongue, minority candidates also mentioned the language of the country of ancestry as a second mother tongue), and an additional passage in the cover letter of minority candidates stating that either their parents and/or him/herself were born abroad, but that the candidate completed all educational training in the country of study. Hence, on the basis of these different signals in resumes, employers should be able to identify the specific ethnic origin of job candidates.

3.2.3. Diagnostic personal information

To investigate the effect of adding diagnostic individual information, three features were experimentally varied across the resumes.

Grade. We randomly varied whether or not the average final grade in education was mentioned in the CV. Half of the applications mentioned no grade and half of the applicants mentioned a good grade, thereby indicating (good) cognitive skills and motivation.³

Performance. Half of the candidates had resumes without extra information about their labor skills and responsibilities in previous jobs. The other half had resumes that included an additional passage in the cover letter and extra information in the CV. In this passage, candidates describe themselves as someone who can perform under pressure, is motivated to acquire new skills, and has been assigned more responsibilities by the previous employers. Furthermore, bullet points were added to the CV to signal these additional responsibilities.

Social skills. Half of the candidates had resumes where no information was given about social skills, the other half had resumes in which these social skills were stressed in the cover letter. Applicants describe themselves as a pleasant and social person, who gets along well with others, a team player and someone who is attentive to other people's needs.

Information included. Based on the aforementioned three experimental manipulations, we constructed a variable that indicates whether extra information was added to resumes, varying between 0 (no information manipulations included) to 1 (one or more information manipulations included).

3.2.4. Control variables (at the individual-level)

We include the following controls which are included as dichotomous variables. We specifically differentiate between jobs that require less or more *interpersonal skills* (e.g. teamwork, having more customer or client contact) and jobs that require lower (ISCED < 4) or higher (ISCED ≥ 4) *educational skills*.

3.2.5. Ethnic community

In this study, we use three indicators to proxy the average productivity of ethnic communities: the socioeconomic resources of the ethnic community, language similarity, and the socioeconomic development of the country of ancestry. According to statistical discrimination theory, besides the mean, employers evaluate job applicants based on the variance of group-level productivity. Since our measures do not allow us to observe the group-level variance, we focus on the average productivity in this study.

Socioeconomic resources of the community in the country of destination. We used the most recent version of the Database on Immigrants in Organization for Economic Co-operation and Development countries (DIOC, 2010/2011) to measure the socioeconomic resources of ethnic minority groups in the country of destination. DIOC 2010/2011 contains information about demographic, educational, and labor market characteristics by country of birth. Based on the proportion of tertiary-educated (ISCED 5 A/5 B/6) and the proportion of employed in the working-age population (all persons aged between 15 and 64 years) of origin groups, we created an index for the socioeconomic resources of ethnic minority groups in the country of destination by means of a principal component analysis. It is important to realize that there may be feedback effects: educational levels and employment participation are likely to be influenced by discrimination that an ethnic minority group faces in the country of destination. For sixteen minority group observations in Germany information about education was missing. These observations were excluded from the analysis (see also [Table S2](#) in the Supplementary Materials).

Language similarity. Language similarity is an important predictor of destination language proficiency ([Van Tubergen and Kalmijn 2005](#)). For each ethnic community, we assign the similarity/distance between the destination language and the language in the country of ancestry, based on the Automated Similarity Judgment Program dataset (ASJP-dataset; [Wichmann et al., 2018](#)). The ASJP-dataset contains a measurement of lexical (dis)similarity of forty key words of almost all languages in the world. Because of the skewed distribution of the original variable (see [Table S3](#) in the Supplementary Materials), we created a dummy variable that differentiates between ethnic minority groups that score below the sample average (low language similarity, coded as zero) or above the average (high language similarity, coded as one).

Socioeconomic development of the country of ancestry. We use the Human Development Index as a proxy for labor productivity related to the country of ancestry. Specifically, this index summarizes the life expectancy, quality of education, and economic prosperity in the country of origin. Higher scores indicate higher (perceived) levels of labor productivity within ethnic minority groups originating from the same country of ancestry.

3.2.6. Control variables (at the ethnic community-level)

To control for systematic country differences, we include country fixed effects. Furthermore, to take into account unobserved differences between ethnic communities, we include region of ancestry fixed effects. We distinguish between the following regions (see also [Table S1](#) in the Supplementary Materials): Western Europe and the United States; Eastern Europe and Russia; Latin America; South Asia; South-East and East Asia; the Middle East and North Africa; and Sub-Saharan Africa.

³ In a few countries, the formulation of the grade manipulation was slightly adapted to reflect the country-specific standard practices. Because German application norms require to include copies of school leaving certificates from high school and vocational training in job applications, the addition of the average final grade in the CV in Germany is probably less strong than elsewhere. In addition, in the United Kingdom, it is always required to mention the average final grade in the CV. In the United Kingdom, the grade manipulation differentiates between mentioning a lower average final grade and mentioning a higher average final grade. Because of these and other small adaptations, we additionally conducted separate analyses by country as a robustness check.

3.3. Analytical strategy

The analytical sample consists of 12,450 individual-level observations (see also [Table S2](#) in the Supplementary Materials): 1630 job applications in Germany, 3396 in the Netherlands, 1403 in Norway, 3044 in Spain, and 2977 in the United Kingdom. At the ethnic community-level, we analyze 139 observations.

Descriptive statistics are presented in [Table S4](#) (in the Supplementary Materials). Correlations between the key independent variables at the ethnic community-level are presented in [Table S5](#) (in the Supplementary Materials). For means of interpretation, all continuous ethnic community-level variables are recoded so that zero represents the minimum score on these variables.

To test our hypotheses, we make use of hierarchical linear probability models in which fictitious job applicants ($N = 12,450$) are nested within ethnic communities ($N = 139$). The coefficients of linear probability models are straightforward to interpret and can be easily compared across models ([Mood 2010](#)). All models include a random intercept at the ethnic community-level and a random slope for the information variable(s). First, we test whether the effect of information (i.e. information included or grade, performance, and social skills) varies between minority and native-majority job candidates to test hypothesis one. Second, we examine whether group productivity is associated with callback to test hypothesis two by including socioeconomic resources of the ethnic community, language similarity, and socioeconomic development of the country of ancestry. Finally, we test hypothesis three by including the interaction terms of the indicators of group productivity and the information variable(s).

4. Results

[Table 1](#) presents the results of hierarchical linear probability models predicting callback, while controlling for required interpersonal skills, required educational skills, and country fixed effects. The findings in model 1 indicate that ethnic minority candidates have a 7.3 percentage point lower probability to receive a callback. Furthermore, we find no significant effect of adding more information about individual productivity. Thus, adding more diagnostic personal information to resumes does not lead to more callbacks. Model 2 includes the interaction term between minority background and information, allowing us to test the hypothesis that discrimination against ethnic minority job candidates is lower when resumes contain more diagnostic personal information about labor productivity (H1). We find no statistically significant interaction effect, however. Hence, rejecting hypothesis one, adding more diagnostic personal information about labor productivity does not reduce discrimination against ethnic minorities.

Results are similar when we estimate the effect of adding information for each treatment condition separately (i.e. grade, performance, and social skills) in models 3 and 4. In particular, adding grade, performance, or social skills does not lead to significantly more callbacks (model 3). Furthermore, in model 4, we find no significant interaction effects between minority and each of the three information treatment conditions. In sum, we find no empirical support for hypothesis one.

In [Table 2](#), we examine whether indicators of group productivity are related to callback (Hypothesis two). As shown in model 1, job applicants belonging to ethnic communities with lower socioeconomic resources receive significantly less callbacks. Thus, besides individual-level characteristics, employers' evaluations of job applicants correlate with the amount of socioeconomic resources of the ethnic community, which is in line with statistical discrimination theory. By contrast, model 2 shows no statistically significant effect of language similarity on callback. In model 3, we observe a statistically significant effect of the socioeconomic development in the country of ancestry, here measured with the Human Development Index. This means that employers' evaluations of job applicants are associated with the level of socioeconomic development in the country of ancestry.

Model 4 includes all three ethnic community-level variables simultaneously. It appears that the effect of the socioeconomic resources of the ethnic community is no longer significant with the inclusion of the level of socioeconomic development in the country of ancestry (the correlation is 0.436 between these two variables, see also [Table S5](#) in the Supplementary Materials). In model 4, we further observe a strong positive association between the socioeconomic development of the country of ancestry and callback. To put this in perspective, 21.7 percent of the job applicants belonging to an ethnic group with the lowest value on the level of socioeconomic development of the country of ancestry received a positive response from an employer, while this was the case in 32.1 percent of the cases for job applicants belonging to an ethnic group with the highest value.

The effect of the socioeconomic development of the country of ancestry could partly reflect (unmeasured) differences between regions of ancestry: perhaps that groups from certain regions of ancestry might face systematically higher levels of discrimination than others (e.g. due to perceived cultural dissimilarities)(cf. [Hagendoorn 1995](#)). As a robustness check, we therefore additionally control for seven regions of ancestry. The results remain unchanged (see also [Table S6](#) in the Supplemental Analysis). This finding provides evidence that a low level of socioeconomic development of the country of ancestry – notably, the crudest indicator of group productivity – is deeply scarring for job applicants, net of language similarity, the socioeconomic resources in the ethnic community, and the specific region of ancestry. Altogether, we find only weak support for hypothesis two.

To test hypothesis three, we interact the information-dummy variable with the indicators of group productivity (Model 5). However, none of the interaction terms is statistically significant, suggesting that the effect of group-level productivity is not different for job applicants who present more diagnostic information in their application.

Model 6 tests the interaction separately for each of the information conditions, yielding the same conclusions. Taken together, these results provide no support for hypothesis 3.

4.1. Sensitivity analysis

We performed several sensitivity analyses to assess the robustness of the results. First, we tested whether the findings might be

Table 1
Hierarchical linear probability models predicting callback by minority background and information included in the resume.

| | Model 1 | Model 2 | Model 3 | Model 4 |
|--|----------------------|----------------------|----------------------|----------------------|
| Constant | 0.297*** (0.031) | 0.310*** (0.039) | 0.298*** (0.028) | 0.304*** (0.030) |
| Minority (ref. majority) | -0.073*** (0.012) | -0.091** (0.029) | -0.073*** (0.012) | -0.081*** (0.019) |
| Information included (ref. no information included) | 0.007 (0.010) | -0.007 (0.019) | | |
| Minority*Information included | | 0.021 (0.023) | | |
| Grade included (ref. no grade included) | | | -0.000 (0.007) | -0.001 (0.013) |
| Performance included (ref. no performance included) | | | 0.011 (0.007) | 0.016 (0.013) |
| Social skills included (ref. no social skills included) | | | -0.000 (0.007) | -0.015 (0.009) |
| Minority*Grade included | | | | 0.002 (0.015) |
| Minority*Performance included | | | | -0.008 (0.016) |
| Minority*Social skills included | | | | 0.022 (0.013) |
| More interpersonal skills required (ref. less interpersonal skills required) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) |
| More educational skills required (ref. less educational skills required) | -0.035* (0.018) | -0.035* (0.018) | -0.035* (0.018) | -0.035* (0.018) |
| Country (ref. United Kingdom) | | | | |
| Spain | -0.060*** (0.012) | -0.060*** (0.012) | -0.060*** (0.012) | -0.060*** (0.012) |
| Germany | 0.289*** (0.016) | 0.289*** (0.017) | 0.289*** (0.016) | 0.289*** (0.016) |
| Norway | 0.077*** (0.015) | 0.077*** (0.015) | 0.077*** (0.015) | 0.077*** (0.015) |
| Netherlands | 0.275*** (0.016) | 0.276*** (0.016) | 0.275*** (0.016) | 0.275*** (0.016) |
| <i>Variance components</i> | | | | |
| Ethnic community-level | 0.001 | 0.001 | 0.001 | 0.001 |
| Individual-level | 0.178 | 0.178 | 0.178 | 0.178 |
| Slope (information) | 0.000 | 0.000 | | |
| Slope (grade) | | | 0.000 | 0.000 |
| Slope (performance) | | | 0.000 | 0.000 |
| Slope (social skills) | | | 0.000 | 0.000 |
| <i>Observations</i> | | | | |
| <i>N</i> Ethnic community-level | 139 | 139 | 139 | 139 |
| <i>N</i> Individual-level | 12,450 | 12,450 | 12,450 | 12,450 |

Notes: The dependent variable distinguishes between “no callback” and “callback”. Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Source: GEMM, 2019

driven by the results of a single country. We estimated all our models while excluding one country at a time (see also Tables S7–S16 in the Supplementary Materials). Overall, these additional analyses lead to the same substantial conclusions as the models discussed above. Compared to model 6 in Table 2, however, the negative interaction effect between grade and the level of socioeconomic development of the country of ancestry is statistically significant after excluding the observations from the Netherlands ($b = -0.130$, see Table S10). Also, compared to model 4 in Table 1, we observe a positive interaction effect between social skills and having a minority background when excluding the observations from Spain (see Table S13). However, this effect is small ($b = 0.037$) and due to a combination of a positive effect of social skills among ethnic minorities and a negative effect of social skills among ethnic majorities. This evidence is more in line with hypothesis 3.

Second, we estimated all our models by country and found cross-national similarities but also some differences (see also Tables S17–S26 in the Supplementary Materials).⁴ As for our tests of hypotheses one and two, we find roughly similar results as those presented earlier. Compared to model 4 in Table 1, however, we find a positive but relatively small interaction effect between social skills and having a minority background in the Netherlands ($b = 0.062$) and the United Kingdom ($b = 0.030$) (see Tables S19 and S25). In all countries, we find a positive and statistically significant effect of the level of socioeconomic development of the country of

⁴ To recap, please note that we can analyze only a limited set of ethnic minority groups in Germany due to missing values on educational attainment at the ethnic-community-level. This strongly limits the comparability of the results in Germany with those found in other countries.

Table 2
Hierarchical linear probability models predicting callback by indicators of group productivity and information included in the resume.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Constant | 0.288*** (0.031) | 0.279*** (0.035) | 0.272*** (0.029) | 0.263*** (0.033) | 0.254*** (0.031) | 0.264*** (0.031) |
| Minority (ref. majority) | -0.071*** (0.013) | -0.059** (0.020) | -0.037** (0.013) | -0.032 (0.020) | -0.032 (0.020) | -0.031 (0.020) |
| Information included (ref. no information included) | 0.008 (0.010) | 0.007 (0.010) | 0.007 (0.010) | 0.007 (0.010) | 0.017 (0.014) | |
| Socioeconomic resources ethnic community | 0.016*** (0.005) | | | 0.002 (0.006) | -0.009 (0.010) | -0.001 (0.009) |
| More language similarity (ref. less language similarity) | | 0.017 (0.018) | | 0.008 (0.015) | 0.018 (0.027) | 0.003 (0.020) |
| Socioeconomic development country of ancestry | | | 0.221*** (0.039) | 0.210*** (0.043) | 0.284*** (0.077) | 0.283*** (0.061) |
| Information included*Socioeconomic resources ethnic community | | | | | 0.013 (0.010) | |
| Information included*Language similarity | | | | | -0.011 (0.025) | |
| Information included*Socioeconomic development country of ancestry | | | | | -0.085 (0.077) | |
| Grade included (ref. no grade included) | | | | | | 0.001 (0.009) |
| Grade included*Socioeconomic resources ethnic community | | | | | | 0.006 (0.007) |
| Grade included*Language similarity | | | | | | 0.006 (0.015) |
| Grade included*Socioeconomic development country of ancestry | | | | | | -0.090 (0.049) |
| Performance included (ref. no performance included) | | | | | | -0.002 (0.011) |
| Performance included*Socioeconomic resources ethnic community | | | | | | -0.004 (0.008) |
| Performance included*Language similarity | | | | | | 0.028 (0.015) |
| Performance included*Socioeconomic development country of ancestry | | | | | | -0.026 (0.048) |
| Social skills included (ref. no social skills included) | | | | | | 0.011 (0.011) |
| Social skills included*Socioeconomic resources ethnic community | | | | | | 0.004 (0.007) |
| Social skills included*Language similarity | | | | | | -0.022 (0.015) |
| Social skills included*Socioeconomic development country of ancestry | | | | | | -0.028 (0.037) |
| More interpersonal skills required (ref. less interpersonal skills required) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) | -0.118*** (0.018) |
| More educational skills required (ref. less educational skills required) | -0.035* (0.018) | -0.035* (0.018) | -0.035* (0.018) | -0.035* (0.018) | -0.035* (0.018) | -0.035 (0.018) |
| Country (ref. United Kingdom) | | | | | | |
| Spain | -0.040** (0.013) | -0.059*** (0.012) | -0.069*** (0.011) | -0.065*** (0.013) | -0.065*** (0.014) | -0.065*** (0.014) |
| Germany | 0.299*** (0.016) | 0.291*** (0.017) | 0.277*** (0.014) | 0.280*** (0.015) | 0.280*** (0.015) | 0.280*** (0.015) |
| Norway | 0.088*** (0.014) | 0.078*** (0.015) | 0.071*** (0.011) | 0.073*** (0.011) | 0.074*** (0.011) | 0.074*** (0.011) |
| Netherlands | 0.283*** (0.014) | 0.277*** (0.016) | 0.264*** (0.013) | 0.267*** (0.013) | 0.267*** (0.013) | 0.267*** (0.013) |
| <i>Variance components</i> | | | | | | |
| Ethnic community-level | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| Individual-level | 0.178 | 0.178 | 0.178 | 0.178 | 0.178 | 0.178 |
| Slope (information) | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |
| Slope (grade) | | | | | | 0.000 |
| Slope (performance) | | | | | | 0.000 |
| Slope (social skills) | | | | | | 0.000 |
| <i>Observations</i> | | | | | | |
| N _{Ethnic community-level} | 139 | 139 | 139 | 139 | 139 | 139 |
| N _{Individual-level} | 12,450 | 12,450 | 12,450 | 12,450 | 12,450 | 12,450 |

Notes: The dependent variable distinguishes between “no callback” and “callback”. Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Source: GEMM, 2019

ancestry on callback except in Spain (although the sign is in the right direction). As for our test of hypothesis three, we observe a negative and significant interaction effect between information and the socioeconomic level of the country of ancestry in the United Kingdom ($b = -0.193$, see Table S25). In Spain and the United Kingdom, we observe negative and significant interaction effects between grade and the level of socioeconomic development of the country of ancestry (Spain: $b = -0.189$ in Table S24; United Kingdom: $b = -0.259$ in Table S26). Furthermore, we find a negative and significant interaction effect between social skills and language similarity in the Netherlands ($b = -0.085$, see Table S20). Lastly, we observe a negative and significant interaction effect between social skills and the level of socioeconomic development of the country of ancestry in the United Kingdom ($b = -0.104$, see Table S26). Overall, there seems to be more evidence in line with hypothesis three in the United Kingdom (and the Netherlands and Spain) than elsewhere.

Third, in our main analyses, we only included occupations and ethnic minority groups which were investigated in all countries simultaneously. Besides the thirty-one ethnic groups in this study, at least five more groups were investigated per country that were not always examined in the other countries of study. In addition, in some countries, several additional occupations were investigated in order to increase the total number of observations. Therefore, we verified whether we could replicate the main findings with this larger, unharmonized dataset (see also Tables S27 and S28 in the Supplementary Materials). In this unharmonized dataset, we neither observe strong differences between the results of this robustness analysis and those presented in the main text. Compared to model 4 in Table 1, however, we find a positive but relatively small interaction effect between social skills and having a minority background ($b = 0.021$, see Table S27).

Finally, we analyzed whether the religion treatment condition might have affected the results found (see also Di Stasio et al., 2019). The results of analyses excluding either all job applicants with religious affiliations or only those with Muslim affiliations are highly similar to those presented in the main text (see also Tables S29, S30, S33, and S34 in the Supplementary Materials). However, in the analysis with job applicants with religious affiliations, we find positive interaction effects between grade (respectively, social skills) and having a minority background (see also Table S31 in the Supplementary Materials). Also, we find negative and significant interactions effects between the socioeconomic development of the country of ancestry and the inclusion of information and grade (see also S32 in the Supplementary Materials). These results, therefore, warrant further attention in future research on the economic and non-economic drivers of ethnic discrimination in hiring.

5. Conclusions

In this paper, we set out to test two central assumptions of statistical discrimination theory (Arrow 1973; Phelps 1972). Using unique field experimental data on thirty-one ethnic groups in five European countries (Lancee 2019), we contribute to the literature by investigating whether ethnic discrimination is lower if job applicants include more diagnostic information about their productivity or if they are a member of an origin group that signals higher levels of labor productivity. Broadly, however, our analyses provide limited support for the underlying assumptions of statistical discrimination theory.

First, we find no strong evidence for the hypothesis that job applicants with a minority background are less discriminated against once applicants add (more) diagnostic information about their productivity (Bertrand and Duflo 2017; Guryan and Charles 2013; Kaas and Manger 2012; Neumark 2018). We find that the inclusion of (different) information treatments (grade, performance, and social skills) is not clearly related to a lower degree of ethnic discrimination. Although other types of information may have a stronger effect, our findings are in line with the majority of findings from previous field experiments on hiring discrimination (e.g. Agerström et al., 2012; Bertrand and Mullainathan 2004; Gaddis 2015; Koopmans et al., 2018; Nunley et al., 2015; Oreopoulos 2011; Thijssen et al., 2019; Vernby and Dancygier 2019). Hence, adding diagnostic information about individual productivity does not improve ethnic minorities' ability to shield themselves from discriminatory actions by employers, thereby contradicting the assumption of statistical discrimination theory that ethnic discrimination is largely due to lack of information about individual productivity.

Second, our results shed new light on the role of group characteristics in explaining group variations in discrimination rates. Whereas previous research assumed that ethnic minority groups are discriminated against because these groups have, on average, lower levels of socioeconomic resources (Aigner and Cain 1977; Arrow 1973; Baumle and Fossett 2005; Phelps 1972; Schwab 1986), our research design allowed us to test this empirically. In particular, we examined whether employers resort to indicators of group productivity even when they have information about the (place of) education and work experience of job applicants and they could know that all candidates were raised in the country of study. Our analysis shows that the probability to receive a callback is not associated with the level of socioeconomic resources of the ethnic community or the degree of similarity between the language of the destination country and the dominant language in the country of ancestry. Strikingly, however, the findings do show that a lower level of socioeconomic development in the country of ancestry results in less callbacks, even when accounting for unobserved heterogeneity between regions of ancestry. Thus, rather than on proximate indicators of group productivity, employers discriminate based on the socioeconomic development of the country of ancestry. While contradicting with a core assumption of statistical discrimination theory, these findings are in line with qualitative work by Midtbøen and Friberg showing that employers mainly use foreign names as proxies for abstract immigrant stereotypes or prejudices and are unable to distinguish between migrant generations (Friberg and Midtbøen 2018; Midtbøen 2014). If at all, employers seem to select on the basis of perceived skill differences between origin groups, not on actual differences (England and Lewin 1989; Quillian and Pager 2010; Tomaskovic-Devey and Skaggs 1999).

Finally, we tested whether the impact of group productivity reduced when employers have more diagnostic personal information about individual job seekers (Guryan and Charles 2013; Rubinstein 2018; Rubinstein et al., 2018). We find very limited support for this expectation, with only significant interaction effects between different information treatment conditions (i.e. mostly social skills and sometimes grade) and group productivity (i.e. mostly the socioeconomic development in the county of ancestry), particularly the

United Kingdom (and sometimes the Netherlands). Whether this is indicative of a systematic and meaningful cross-national pattern or a statistical artifact must be assessed in future cross-national research. Altogether, however, these findings do not strongly support statistical discrimination theory's assumption that employers rationally update their group beliefs with more reliable signals of individual productivity (see also Oreopoulos 2011; Pager and Karafin 2009). Rather, these findings could also be consistent with models of stereotype amplification, stressing that people are mostly inattentive to information that is disconfirming of their systematically biased expectations (Brewer 1988; Fiske 1998; Fiske and Neuberg 1990; Pager and Karafin 2009; Quillian and Pager 2010).

Naturally, this study has some limitations, some of which could be addressed in future research. First, while we find an association between the level of discrimination and the socioeconomic development of the country of origin, future research could test more directly whether this effect is mediated by employers' perceptions about skill differences between origin groups. A promising avenue for research could be combining the results of a field experiment with a survey among employers that includes questions about their economic motives and perceptions on skill differences between ethnic groups (cf. Pedulla 2016). Second, to test for the effect of group-level productivity, our study included the average level of productivity. However, according to statistical discrimination theory, employers could also evaluate job candidates based on the variance of group productivity (see e.g. Schaeffer et al., 2016). As such, it could be that findings are different when focusing on group level variance. Future research could be aimed at doing so. Third, whereas this study was able to demonstrate that aggregated patterns of discrimination are not clearly driven by "rational optimizing behavior and limited information" (Guryan and Charles 2013:418), one could argue that statistical discrimination is perhaps more prevalent in certain sectors or for certain jobs. By leveraging a targeted sampling strategy (cf. Malhotra et al. 2013), future field experiments could look for an upper-bound estimate of statistical discrimination by focusing on jobs or sectors where quality criteria are highly ambiguous, workers' true productivity can only be observed after a relatively long period of work, and/or hiring contexts where managers are more prone to risk avoidance (Arrow 1973; Friberg and Midtbøen 2018; Midtbøen 2015; Weichselbaumer 2017). Finally, our study calls for renewed thinking about alternative mechanisms generating ethnic discrimination in hiring – that is, mechanisms unrelated to economically-rational motives. Perhaps that certain ethnic minority groups face higher levels of discrimination because employers perceive more cultural and religious distance and conflict between the majority group and certain minority groups (Adida et al. 2016; Hagendoorn 1995; Di Stasio et al., 2019; Veit and Thijssen et al., 2019). In this sense, future research would do well to consider the degree to which employers' perceptions of skill differences between groups might overlap with (or are used to rationalize) perceived levels of cultural distance or conflict. This also resonates with the words of Phelps. Already in 1972, he wrote: "A sensitive person, I have been warned, might read this paper as expressing an impression on the part of the author that most or all discrimination is the result of beliefs that blacks and women deliver on the average an inferior performance. Actually, I do not know (nor claim to know) whether in fact most discrimination is of the statistical kind studied here." (Phelps 1972: 661).

In conclusion, the present study challenges the view that ethnic discrimination in hiring is largely driven by economic rationality and incomplete information about individual productivity. A low level of socioeconomic development in the country of origin is deeply scarring for job applicants of ethnic minority origins, despite being raised and having completed all their education in the country of destination. Mirroring findings found in previous qualitative research (Friberg and Midtbøen 2018; Midtbøen 2014; Pager and Karafin 2009), employers thus seem to select on the basis of very crude stereotypes about the overall skills of ethnic communities and tend to ignore signals which are more predictive of individual labor productivity. One might hence be tempted to conclude that evidence presented here is more in line with error discrimination theory – that is, "actions of employers who underestimate the average productivities of a group, and, based upon this mistaken belief, are unwilling to hire group members or will hire them only for a lower wage" (England and Lewin 1989:242). We welcome researchers to further derive and test empirically falsifiable hypotheses from error discrimination theory. Especially fruitful in this regard would be the development and assessment of hypotheses concerning the beliefs and behaviors of the central actors in hiring – that is, employers (Bills et al. 2017). Researchers should specifically explicate their assumptions about employers' knowledge, preferences, and decision-making and incorporate the unique features of the context in which hiring decisions are being made (e.g. different hiring phases, organizational characteristics, and labor market circumstances). By directly investigating the individual and contextual factors that affect employers' hiring decisions, research will be able to significantly enhance our understanding of the mechanisms generating ethnic discrimination in employment.

Funding

This research has been supported by funding from the European Commission in the Horizon 2020 project 'Growth, Equal opportunities, Migration and Markets (GEMM), grant number H2020 649255.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssrsearch.2020.102482>.

References

- Adida, Claire L., Laitin, David D., Valfort, Marie-Anne, 2016. *Why Muslim Integration Fails in Christian-Heritage Societies*. Harvard University Press, Cambridge.
- Agerström, Jens, Björklund, Fredrik, Carlsson, Rickard, Rooth, Dan-Olof, 2012. Warm and competent hassan = cold and incompetent eric: a harsh equation of real-life hiring discrimination. *Basic Appl. Soc. Psychol.* 34 (4), 359–366.

- Ahmad, Akhlaq, 2019. When the name matters: an experimental investigation of ethnic discrimination in the Finnish labor market. *Socio. Inq.* 1–29.
- Aigner, Dennis J., Cain, Glen G., 1977. Statistical theories of discrimination in labor markets. *Ind. Labor Relat. Rev.* 30 (2), 175–187.
- Allport, Gordon Willard, 1954. *The Nature of Prejudice*. Addison-Wesley, Reading.
- Altonji, Joseph G., Pierret, Charles R., 2001. Employer learning and statistical discrimination. *Q. J. Econ.* 116 (1), 313–350.
- Andriessen, Iris, Nievers, Eline, Faulk, Laila, Dagevos, Jaco, 2010. Liever mark dan mohammed? Onderzoek naar arbeidsmarktdiscriminatie van Nietwesterse migranten via praktijktests. Den Haag: Sociaal en Cultureel Planbureau.
- Arkes, Hal R., Tetlock, Philip E., 2004. "Attributions of implicit prejudice, or 'would jesse jackson "fail" the implicit association test? *Psychol. Inq.* 15 (4), 257–278.
- Arrow, Kenneth J., 1973. The theory of discrimination. In: Ashenfelter, O., Rees, A. (Eds.), *Discrimination In Labor Markets*. Princeton University Press, Princeton, pp. 3–33.
- Ashmore, Richard D., Del Boca, F.K., 1981. Conceptual approaches to stereotypes and stereotyping. In: Hamilton, D.L. (Ed.), *In Cognitive Processes in Stereotyping and Intergroup Behavior*. Erlbaum, Hillsdale, NJ.
- Auspurg, Katrin, Schneck, Andreas, Hinz, Thomas, 2019. Closed doors everywhere? A meta-analysis of field experiments on ethnic discrimination in rental housing markets. *J. Ethnic Migrat. Stud.* 45 (1), 95–114.
- Baert, Stijn, Vujić, Sunčica, 2016. Immigrant volunteering: a way out of labour market discrimination? *Econ. Lett.* 146, 95–98.
- Bartoš, Vojtěch, Bauer, Michal, Chytilová, Julie, Matějka, Filip, 2016. Attention discrimination: theory and field experiments. *Am. Econ. Rev.* 106 (6), 1437–1475.
- Baumle, Amanda K., Fossett, Mark, 2005. Statistical discrimination in employment: its practice, conceptualization, and implications for public policy. *Am. Behav. Sci.* 48 (9), 1250–1274.
- Becker, Gary, 1957. *The Economics of Discrimination*. University of Chicago Press, Chicago.
- Bertrand, Marianne, Dufló, Esther, 2017. "Field experiments on discrimination." pp. 309–93. In: Banerjee, A., Dufló, E. (Eds.), *Handbook of Economic Field Experiments*. Elsevier B.V. Amsterdam.
- Bertrand, Marianne, Mullainathan, Sendhil, 2004. Are emily and greg more employable than lakisha and jamal? A field experiment on labor market discrimination. *Am. Econ. Rev.* 94 (4), 991–1013.
- Bills, David B., Di Stasio, Valentina, Gërçhani, Klarita, 2017. The demand side of hiring: employers in the labor market. *Annu. Rev. Sociol.* 43, 291–310.
- Blumer, Herbert, 1958. Race prejudice as a sense of group position. *Pac. Socio Rev.* 1 (1), 3–7.
- Bonoli, Giuliano, Hinrichs, Karl, 2012. "Statistical discrimination and employers' recruitment: practices for low-skilled workers. *Eur. Soc.* 14 (3), 338–361.
- Booth, Alison L., Leigh, Andrew, Elena, Varganova, 2012. Does ethnic discrimination vary across minority groups? Evidence from a field experiment. *Oxf. Bull. Econ. Stat.* 74 (4), 547–573.
- Brewer, Marilyn B., 1988. "A dual process model of impression formation." pp. 1–36. In: Srull, T.K., Wyer Jr., R.S. (Eds.), *Advances in Social Cognition*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ.
- Brigham, John C., 1971. Ethnic stereotypes. *Psychol. Bull.* 76 (1), 15–38.
- Crawford, Jarret T., Lee, Jussim, Madon, Stephanie, Cain, Thomas R., Stevens, Sean T., 2011. The use of stereotypes and individuating information in political person perception. *Pers. Soc. Psychol. Bull.* 37 (4), 529–542.
- Dixon, John, 2017. "Thinking ill of others without sufficient warrant? transcending the accuracy-inaccuracy dualism in prejudice and stereotyping research. *Br. J. Soc. Psychol.* 56 (1), 4–27.
- England, Paula, Lewin, Peter, 1989. Economic and sociological views of discrimination in labor markets: persistence or demise? *Socio. Spectr.* 9, 239–257.
- Fiske, Susan T., 1998. Stereotyping, prejudice, and discrimination. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (Eds.), *The Handbook of Social Psychology*, vol. 2. McGraw-Hill, New York, pp. 357–411.
- Fiske, Susan T., Neuberg, Steven L., 1990. A Continuum of Impression Formation, from Category-Based to Individuating Processes: Influences of Information and Motivation on Attention and Interpretation. Academic Press, New York.
- Flage, Alexandre, 2018. Ethnic and gender discrimination in the rental housing market: evidence from a meta-analysis of correspondence tests. *J. Hous. Econ.* 41, 251–273.
- Friberg, Jon H., Midtbøen, Arnfinn H., 2018. Ethnicity as skill: immigrant employment hierarchies in Norwegian low-wage labour markets. *J. Ethnic Migrat. Stud.* 44 (9), 1463–1478.
- Friberg, Jon Horgen, Arnfinn, H., Midtbøen, 2019. The making of immigrant niches in an affluent welfare state. *Int. Migrat. Rev.* 53 (2), 322–343.
- Gaddis, S. Michael, 2015. Discrimination in the credential society: an audit study of race and college selectivity in the labor market. *Soc. Forces* 93 (4), 1451–1479.
- Gaddis, S. Michael, 2017. Racial/ethnic perceptions from hispanic names: selecting names to test for discrimination. *Socius: Sociological Research for a Dynamic World* 3, 1–11.
- Guryan, Jonathan, Charles, Kerwin Kofi, 2013. Taste-based or statistical discrimination: the economics of discrimination returns to its roots. *Econ. J.* 123 (572), 417–432.
- Hagendoorn, Louk, 1995. Intergroup biases in multiple group systems: the perception of ethnic hierarchies. *Eur. Rev. Soc. Psychol.* 6 (1), 199–228.
- Holman, Eric W., Brown, Cecil H., Wichmann, Soren, Müller, André, Velupillai, Viveka, Hammarström, Harald, Sauppe, Sebastian, Jung, Hagen, Bakker, Dik, Brown, Pamela, Belyaev, Oleg, Urban, Matthias, Mailhammer, Robert, List, Johann-Mattis, Egorov, Dmitry, 2011. "Automated dating of the world's language families based on lexical similarity. *Curr. Anthropol.* 52 (6), 841–875.
- Imdorf, Christian, 2017. Understanding discrimination in hiring apprentices: how training companies use ethnicity to avoid organisational trouble. *J. Vocat. Educ. Train.* 69 (3), 405–423.
- Jussim, Lee, Cain, Thomas R., Crawford, Jarret T., Kent, Harber, Cohen, Florette, 2009. "The unbearable accuracy of stereotypes." pp. 199–227. In: Nelson, T.D. (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination*. Psychology Press, New York.
- Jussim, Lee, Crawford, Jarret T., Rubinstein, Rachel S., 2015. Stereotype (In)Accuracy in perceptions of groups and individuals. *Curr. Dir. Psychol. Sci.* 24 (6), 490–497.
- Kaas, Leo, Manger, Christian, 2012. "Ethnic discrimination in Germany's labour market: a field experiment. *Ger. Econ. Rev.* 13 (1), 1–20.
- Koenig, Anne M., Eagly, Alice H., 2014. "Evidence for the social role theory of stereotype content: observations of groups' roles shape stereotypes. *J. Pers. Soc. Psychol.* 107 (3), 371–392.
- Koopmans, Ruud, Veit, Susanne, Yemane, Ruta, 2018. *Ethnische Hierarchien in Der Bewerberauswahl: Ein Feldexperiment Zu Den Ursachen von Arbeitsmarktdiskriminierung*. WZB Berlin Social Science Center, Berlin.
- Lancee, Bram, 2019. Ethnic discrimination in hiring: comparing groups across contexts. Results from a cross-national field experiment. *J. Ethnic Migrat. Stud.* 1–20.
- Lancee, Bram, Birkelund, Gunn Elisabeth, Coenders, Marcel, Di Stasio, Valentina, Reino, Mariña Fernández, Heath, Anthony F., Koopmans, Ruud, Larsen, Edvard N., Polavieja, Javier G., Ramos, María, Hannah, Soiné, Lex, Thijssen, Veit, Susanne, Yemane, Ruta, 2019a. *The GEMM Study: A Cross-National Harmonized Field Experiment on Labour Market Discrimination: Codebook*. University of Amsterdam, University of Oslo, University of Utrecht, Charles III University of Madrid, WZB Berlin Social Science Center.
- Lancee, Bram, Birkelund, Gunn Elisabeth, Coenders, Marcel, Di Stasio, Valentina, Reino, Mariña Fernández, Heath, Anthony F., Koopmans, Ruud, Larsen, Edvard N., Polavieja, Javier G., Ramos, María, Lex, Thijssen, Veit, Susanne, Yemane, Ruta, Zwier, Dieuwke, 2019b. *The GEMM Study: A Cross-National Harmonized Field Experiment on Labour Market Discrimination: Technical Report*. University of Amsterdam, University of Oslo, University of Utrecht, Charles III University of Madrid, WZB Berlin Social Science Center.
- Lane, Tom, 2016. Discrimination in the laboratory: a meta-analysis of economics experiments. *Eur. Econ. Rev.* 90, 375–402.
- LaPiere, Richard, 1936. Type-rationalizations of group Antipathy. *Soc. Forces* 15, 232–237.
- Levels, Mark, Dronkers, Jaap, Kraaykamp, Gerbert, 2008. "Immigrant children's educational achievement in western countries: origin, destination, and community effects on mathematical performance. *Am. Socio. Rev.* 73 (5), 835–853.
- Luijckx, Ruud, Wolbers, Maarten H.J., 2009. The effects of non-employment in early work-life on subsequent employment chances of individuals in The Netherlands. *Eur. Socio Rev.* 25 (6), 647–660.

- Malhotra, Neil, Margalit, Yotam, Mo, Cecilia Hyunjung, 2013. Economic explanations for opposition to immigration: distinguishing between prevalence and conditional impact. *Am. J. Polit. Sci.* 57 (2), 391–410.
- Midtbøen, Arnfinn H., 2014. The invisible second generation? Statistical discrimination and immigrant stereotypes in employment processes in Norway. *J. Ethnic Migrat. Stud.* 40 (10), 1657–1675.
- Midtbøen, Arnfinn H., 2015. The context of employment discrimination: interpreting the findings of a field experiment. *Br. J. Sociol.* 66 (1), 193–214.
- Mood, Carina, 2010. Logistic regression: why we cannot do what we think we can do, and what we can do about it. *Eur. Socio Rev.* 26 (1), 67–82.
- Moss, Philip, Tilly, Chris, 2001. *Stories Employers Tell. Race, Skill, and Hiring in America.* Russel Sage Foundation, New York.
- Neckerman, Kathryn M., Kirschenman, Joleen, 1991. Hiring strategies, racial bias, and inner-city workers. *Soc. Probl.* 38 (4), 433–447.
- Neumark, David, 2018. Experimental research on labor market discrimination. *J. Econ. Lit.* 56 (3), 799–866.
- Nievers, Eline, 2010. "Personeelselecteurs over Niet-Westerse Migranten En Discriminatie Op de Arbeidsmarkt." Pp. 52–85. In: *Discriminatiemonitor niet-westerse migranten op de arbeidsmarkt 2010.* Den Haag: Sociaal en Cultureel Planbureau.
- Nunley, John M., Adam, Pugh, Romero, Nicholas, Alan Seals, Richard, 2015. An examination of racial discrimination in the labor market for recent college graduates: estimates from the field. *Berkeley Electronic Journal of Economic Analysis and Policy* 15 (3), 1093–1125.
- OECD, 2010. *Database on Immigrants in OECD and Non-OECD Countries (DIOC).* OECD Publishing, Paris.
- Oreopoulos, Philip, 2011. Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. *Am. Econ. J. Econ. Pol.* 3 (4), 148–171.
- Pager, Devah, Karafin, Diana, 2009. Bayesian bigot? Statistical discrimination, stereotypes, and employer decision making. *Ann. Am. Acad. Polit. Soc. Sci.* 621, 70–93.
- Pager, Devah, Quillian, Lincoln, 2005. Walking the talk? What employers say versus what they do. *Am. Socio. Rev.* 70 (3), 355–380.
- Pedulla, David S., 2016. Penalized or protected? Gender and the consequences of nonstandard and mismatched employment histories. *Am. Socio. Rev.* 81 (2), 262–289.
- Phelps, Edmund S., 1972. The statistical theory of racism and sexism. *Am. Econ. Rev.* 62 (4), 659–661.
- Quillian, Lincoln, 2006. New approaches to understanding racial prejudice and discrimination. *Annu. Rev. Sociol.* 32, 299–328.
- Quillian, Lincoln, Heath, Anthony F., Pager, Devah, Midtbøen, Arnfinn H., Fleischmann, Fenella, Hexel, Ole, 2019. Do some countries discriminate more than others? Evidence from 97 field experiments of racial discrimination in hiring. *Sociological Science* 6, 467–496.
- Quillian, Lincoln, Pager, Devah, 2010. Estimating risk: stereotype amplification and the perceived risk of criminal victimization. *Soc. Psychol. Q.* 73 (1), 79–104.
- Quillian, Lincoln, Pager, Devah, Hexel, Ole, Arnfinn, H., Midtbøen, 2017. Meta-analysis of field experiments shows No change in racial discrimination in hiring over time. *Proc. Natl. Acad. Sci. Unit. States Am.* 114 (41), 10870–10875.
- Rubinstein, Rachel S., 2018. *Reliance on Individuating Information and Stereotypes in Impression Formation: the Diagnosticity and Judgment Task Model of Person Perception.* School of Graduate Studies Rutgers, The State University of New Jersey, New Brunswick, New Jersey (dissertation).
- Rubinstein, Rachel S., Lee, Jussim, Stevens, Sean T., 2018. Reliance on individuating information and stereotypes in implicit and explicit person perception. *J. Exp. Soc. Psychol.* 75, 54–70.
- Schaeffer, Merlin, Höhne, Jutta, Teney, Céline, 2016. Income advantages of poorly qualified immigrant minorities: why school dropouts of Turkish origin earn more in Germany. *Eur. Socio Rev.* 32 (1), 93–107.
- Schwab, Stewart, 1986. Is statistical discrimination efficient? *Am. Econ. Rev.* 76 (1), 228–234.
- Di Stasio, Valentina, Lancee, Bram, 2020. Understanding why employers discriminate, where and against whom: the potential of cross-national, factorial and multi-group field experiments. *Res. Soc. Stratif. Mobil.* 65, 1–11.
- Di Stasio, Valentina, Lancee, Bram, Veit, Susanne, Yemane, Ruta, 2019. Muslim by default or religious discrimination? Results from a cross-national field experiment on hiring discrimination. *J. Ethnic Migrat. Stud.* 1–22.
- Stevens, Sean T., Lee, Jussim, Stevens, Lillian A., Anglin, Stephanie M., 2018. "Cultural competence: a form of stereotype rationality." pp. 651–64. In: Frisby, C., O'Donohue, W.T. (Eds.), *Cultural Competence in Applied Psychology.* Springer, Cham, Switzerland.
- Thijssen, Lex, Lancee, Bram, Veit, Susanne, Yemane, Ruta, 2019. Discrimination against Turkish minorities in Germany and The Netherlands: field experimental evidence on the effect of diagnostic information on labour market outcomes. *J. Ethnic Migrat. Stud.* 1–18.
- Tjaden, Jasper Dag, Schwemmer, Carsten, Khadjavi, Menusch, 2018. Ride with me: ethnic discrimination, social markets and the sharing economy. *Eur. Socio Rev.* 34 (4), 418–432.
- Tomaskovic-Devey, Donald, Skaggs, Sheryl, 1999. An establishment-level test of the statistical discrimination hypothesis. *Work Occup.* 26 (4), 422–445.
- Van Tubergen, Frank, Kalmijn, Matthijs, 2005. destination language proficiency in cross national perspective: a study of immigrant groups in nine western countries. *Am. J. Sociol.* 110 (5), 1412–1457.
- Van Tubergen, Frank, Ineke Maas, Henk Flap, 2004. The economic incorporation of immigrants in 18 western societies: origin, destination, and community effects. *Am. Socio. Rev.* 69 (5), 704–727.
- United Nations Development Programme, 2018. *Human Development Indices and Indicators 2018 Statistical Update.* United Nations, New York.
- Veit, Susanne, Lex, Thijssen, 2019. Almost identical but still treated differently: hiring discrimination against foreign-born and domestic-born minorities. *J. Ethnic Migrat. Stud.* 1–20.
- Vernby, Kåre, Dancygier, Rafaela M., 2019. Can immigrants counteract employer discrimination? A factorial field experiment reveals the immutability of ethnic hierarchies. *PLoS One* 14 (7), 1–19.
- Waldinger, Roger, Lichter, Michael I., 2003. *How the Other Half Works: Immigration and the Social Organization of Labor.* University of California Press, Berkeley.
- Weichselbaumer, Doris, 2017. Discrimination against migrant job applicants in Austria: an experimental study. *Ger. Econ. Rev.* 18 (2), 237–265.
- Wichmann, Søren, Holman, Eric W., Brown, Cecil H., 2018. *The ASJP Database. Version 18.*
- Zschirnt, Eva, 2019. Evidence of hiring discrimination against the second generation: results from a correspondence test in the Swiss labour market. *J. Int. Migrat. Integrat.* 1–23.
- Zschirnt, Eva, Ruedin, Didier, 2016. "Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. *J. Ethnic Migrat. Stud.* 42 (7), 1115–1134.