

WORKING PAPERS IN ECONOMICS AND STATISTICS

RAPPORT NR 01 | 2023

Language Proficiency and Hiring of Immigrants

Evidence from a New Field Experimental Approach

MAGNUS CARLSSON, STEFAN ERIKSSON OCH DAN-OLOF ROTH

Language Proficiency and Hiring of Immigrants

Evidence from a New Field Experimental Approach

MAGNUS CARLSSON, STEFAN ERIKSSON OCH DAN-OLOF ROTH

Language Proficiency and Hiring of Immigrants

EVIDENCE FROM A NEW FIELD EXPERIMENTAL APPROACH

RAPPORT NR 01, INSTITUTIONEN FÖR NATIONALEKONOMI OCH STATISTIK,
LINNÉUNIVERSITETET, VÄXJÖ, 2023

© MAGNUS CARLSSON, STEFAN ERIKSSON OCH DAN-OLOF ROTH

Skriftserieredaktör: Hans Grönqvist

ISBN: 978-91-8082-009-7 (pdf)

DOI: <https://doi.org/10.15626/ns.wp.2023.1>

Language Proficiency and Hiring of Immigrants: Evidence from a New Field Experimental Approach[†]

Magnus Carlsson[‡], Stefan Eriksson[§], and Dan-Olof Rooth^{**}

February 9, 2023

Abstract. Labor markets in advanced economies have undergone substantial change in recent decades due to globalization, technological improvements, and organizational changes. Due to these developments, oral and written language skills have become increasingly important even in less skilled jobs. Immigrants – who often have limited skills in the host country language upon arrival – are likely to be particularly affected by the increase in language requirements. Despite this increase in literacy requirements, little is known about how immigrants’ language proficiency is rewarded in the labor market. However, estimating the causal effect of immigrants’ language skills on hiring is challenging due to potential biases caused by omitted variables, reverse causality, and measurement error.

To address identification problems, we conduct a large-scale field experiment, where we send thousands of fictitious resumes to employers with a job opening. With the help of a professional linguist, we manipulate the cover letters by introducing common second-language features, which makes the resumes reflect variation in the language skills of real-world migrants. Our findings show that better language proficiency in the cover letter has a strong positive effect on the callback rate for a job interview: moving from the lowest level of language proficiency to a level similar to natives almost doubles the callback rate. Consistent with the recent development that language proficiency is also important for many low- and medium-skilled jobs, the effect of better language skills does not vary across the vastly different types of occupations we study. Finally, the results from employer surveys suggest that it is improved language skills *per se* that is the dominant explanation behind the language proficiency effect, rather than language skills acting as a proxy for other unobserved abilities or characteristics.

JEL classification: F22, J15, J24

Keywords: Labor migrants, Language proficiency, Language skills, Human capital, Field experiment

[†] We are grateful for valuable comments from Gordon Dahl, Pierre Deschamps, Mette Foged, Kevin Lang, Fabian Lange, Louis-Pierre Lepage, Conrad Miller, José Montalban, and Anders Stenberg as well as seminar participants at Bocconi University, Boston University, Stockholm University, Uppsala University, Frischsenteret, the Nordic Summer Institute in Labor Economics in Bergen, the Research Institute of Industrial Economics (IFN), Linnaeus University, the European Association of Labour Economists (EALE) 2022 Conference in Padova, and the Swedish Conference in Economics 2022. We thank Maria Rydell for her work with the fictitious resumes, and Olof Axman, Johanna Finnström, Axel Malmcrona, Towe Nilsson, and Ellen Westman Persson for excellent research assistance. Financial support from the Swedish Research Council is gratefully acknowledged. The field experiment has an ERB-approval (2019-06417); see the attached supplementary material to this submission showing the pre-registered ERB-application containing the experiment’s research design, hypothesis, model specification, and statistical power.

[‡] Department of Economics and Statistics, Linnaeus University; email: Magnus.Carlsson@lnu.se

[§] Department of Economics, Uppsala University; email: Stefan.Eriksson@nek.uu.se

^{**} Corresponding author. The Swedish Institute for Social Research, Stockholm University; IZA; Linnaeus University; email: Dan-Olof.Rooth@sofi.su.se

1. Introduction

Western labor markets have undergone substantial change in recent decades. Globalization as well as advances in technology and the organization of firms have eliminated many routine jobs, while high-skill jobs in general and low-skill jobs in the service sector have expanded (Autor and Dorn, 2013; Goos, Manning, and Salomons, 2009). Job tasks have also changed within many occupations (Atalay et al., 2020; Autor, Salomons, and Seegmiller, 2021; Deming and Noray, 2020). In particular, there is evidence that oral and written communication skills have become substantially more important not only for high-skill jobs but also for many common low- and medium-skill jobs (Corbel et al., 2022; Farrell et. al 2021; Karlsson, 2009). A typical workday in occupations such as construction work, health care, truck driving, mechanic, and retail sales often involves using advanced IT technology and communicating intensively with managers, colleagues, and customers.⁵

Immigrants are likely to be particularly affected by these changes since they often have limited skills in the host country language, which could present a major challenge for their labor market integration. Despite higher literacy requirements and unsatisfactory integration of immigrants in Europe and North America, little is known about how language proficiency is rewarded in the labor market. This is the important question that we shed light on in this paper. Using a field experimental approach, we measure *how much* the probability of being hired increases for an immigrant as host country language proficiency gradually improves. Specifically, we manipulate the language proficiency level of cover letters of fictitious resumes, which are sent to employers with a job opening.

Identifying the causal effect of language skills on hiring is challenging. Estimates obtained from studies using observational data on labor market outcomes for immigrants with different

⁵ For example, a truck driver nowadays must be able to not only drive but also read and understand safety signs, instructions on how to handle dangerous goods, vehicle manuals and shipping instructions and interact with other drivers and recipients of goods.

levels of literacy potentially suffer from biases caused by omitted variables and reverse causality. In addition, bias from measurement error is typically a major concern because language skills are often self-reported or based on low-stakes literacy tests. By conducting a large-scale field experiment in which we have full control over the information available to recruiting employers, we overcome these challenges and measure the causal effect of the language manipulation. We focus on the Swedish labor market, which is an ideal case to study since Sweden's immigrant population faces a problematic labor market situation, similar to the case in many other European countries, and few immigrants have knowledge of the Swedish language prior to immigration (Brell et al., 2020).

We start by presenting several indicators that support the notion that oral and written language skills are important in all jobs, including low-skill jobs. First, using a large database of job advertisements (with more than six million vacancies published on the website of the Swedish Public Employment Service), we find a threefold increase in the share of advertisements expressing requirements on literacy skills (such as “good knowledge of Swedish”) from 2007 to 2019, and in 2019, this share is very similar for high-, medium- and low-skilled occupations. Second, we conduct surveys of employers, which show that good proficiency in Swedish is seen as crucial by recruiting firms; approximately 80 percent of recruiters state that having good knowledge of Swedish is important at their workplace, again showing very little variation across different types of occupations. Finally, these findings are also consistent with US O*NET survey data, where responses from workers show extensive use of both oral and written communication, which again are found to be at a similar level across a range of occupations.

In our field experiment, a professional linguist specialized in Swedish as a second language was employed to manipulate cover letters of fictitious resumes to obtain different levels of language proficiency and reflect variation in the language skills of real-world migrants. The

cover letters increasingly introduced common second-language features, e.g., mistakes and simplifications, resulting in four distinct levels of language skills, where the highest level is similar to what could be expected from a native. We use the same standardized cover letter and language proficiency manipulations for all occupations and estimate the effect of language proficiency using the pooled sample of observations. Thus, any difference in the effect of language proficiency between occupations cannot be attributed to occupations receiving different treatments.

To verify the language signal, we asked employers recruiting at the same website as used in the field experiment to rate the level of language proficiency of the cover letters. The results show that the rating increases linearly across the four levels, and the proficiency rating of the cover letter almost doubles from the lowest to the highest level.

Between February 2020 and April 2021, we conducted the field experiment by applying our research design to the real labor market. We investigated labor immigrants from four European countries, Germany, Poland, Romania, and the UK, which are the four most common countries of origin for European migrants to Sweden during the last ten years. Our choice to study migrants from these countries should reduce the risk that ethnic discrimination interferes with the results. In the experiment, we sent 3,154 fictitious resumes to employers with advertised vacancies in 17 different occupations, randomly varying the level of language proficiency in the cover letter. For each resume, we measured whether the applicant received an invitation to a job interview by the employer (which we refer to as a callback), which is used as the outcome variable.

Our findings show that an increased level of language proficiency of the cover letter has a strong positive effect on the callback rate: moving from a low level of proficiency to a level without second-language features (i.e., similar to a native) almost doubles the callback rate from approximately 8 to 15 percent. This means that improving one's language proficiency from

‘poor’ to ‘great’ corresponds to the effect of an additional six years of work experience (Eriksson and Rooth, 2014). We also group the included occupations based on different characteristics, such as occupational skill level and the share of immigrants in an occupation, and conduct the same analysis, confirming that the effect does not vary across different types of occupational groups. Hence, the effect of language proficiency appears universal across the labor market, as suggested by the content analysis and the employer surveys.

An important question, not the least for policy, is how these causal estimates should be interpreted. Immigrants with better language skills could be preferred by employers because language skills constitute productive human capital or, alternatively, because they are a proxy for better unobserved skills. Our employer surveys point toward the human capital story being the dominant explanation. A clear majority of recruiters (87 percent) report that poor language skills are the major challenge when recruiting immigrants, while few report that it is cultural differences, a lack of knowledge about Swedish society, or difficulties in assessing foreign competence. Furthermore, a large majority of recruiters (>80%) state that having good knowledge of Swedish is important at their workplace.

Our results show that having poor skills in the host country language is a substantial hurdle for immigrants’ labor market integration. Especially troublesome is the fact that language skills are found to be important in many typical immigrant entry jobs, such as cleaners, waiters, and warehouse staff workers. Several indicators of language skills, including results from the PIAAC (Programme for the International Assessment of Adult Competencies conducted by the OECD), show that a large share of immigrants to the EU have very low language skills (Brell et al., 2020). In terms of policy, our results highlight how important it is that effective language training programs are available to immigrants and how crucial it is that the participants complete these programs.

The key contributions of our paper are (i) the modification of the correspondence testing methodology to arrive at clean identification of the causal effect of language skills on labor market outcomes and (ii) to demonstrate that language proficiency is important across all types of jobs, including low-skill occupations.⁶

Our contributions build on the merits of several important papers, which have documented a positive association between proficiency in the host country's language and key labor market outcomes, such as employment and wages (Dustmann, 1994, and Dustmann and Soest 2001; 2002). Chiswick and Miller (2007; 2015) survey this literature and show that estimates of the impact of good language skills on wages are located in a wide interval, which spans from 5 to 35 percent. The substantial variation in estimates across studies suggests that biases of the types discussed above could be a major problem in observational studies. In response to these identification problems, other studies have applied quasi-experimental and instrumental variable approaches (Dustmann and Fabbri, 2003; Bleakley and Chin, 2004; Yao and van Ours, 2015).

A recent strand of literature has instead turned to analyzing the effects of *language training programs* for immigrants on labor market outcomes. As an important step forward for identification, researchers have utilized that sometimes programs have randomized participation using lotteries when programs are oversubscribed, while others have utilized quasi-random variation in who participates originating from unexpected introductions or test score cutoffs. Foged et al. (2022), Heller and Slungaard Mumma (2022), and Sarvimäki and Hämäläinen (2016) find large effects on earnings, while Lochmann et al. (2019) document positive effects on labor force participation. A challenge in these studies is separating the effects

⁶ See, e.g., Bertrand and Mullainathan (2004), Carlson and Eriksson (2019), and Carlsson and Rooth (2007) for previous applications of the correspondence testing methodology.

of language training from other components of the programs, i.e., an estimate may capture the overall effect of a program rather than the pure language effect.

The rest of the paper is organized as follows. Section 2 documents evidence from several sources showing that language skills are important in the workplace, also in low- and medium-skill occupations. Section 3 describes the experimental design in detail. Section 4 explains how the experiment was conducted. Section 5 presents and discusses the results, while Section 6 concludes.

2. The importance of language proficiency for essentially all jobs

Language research shows how changes in the workplace have created new literacy demands on workers (Corbell et al., 2022; Farrell et al., 2021; Karlsson, 2009). While related research previously focused on the literacy demands of high-skill jobs, the new line of qualitative research explicitly describes literacy practices in everyday work for a large number of medium- and low-skill occupations. Job tasks involving routine forms of reading and writing are nowadays typically replaced by computers and IT systems. To be employable, workers must therefore increasingly possess uniquely human nonroutine literacy skills, which cannot be replaced by computers. The types of literacy skills needed in some of the low- and medium-skill jobs included in our field experiment are well illustrated by literacy job profiles. For example, a warehouse staff worker needs to be able to report problems to managers, contact transport or courier workers to request information or pick up, write damage reports, answer queries from customers, and read training material (forklift licensing, health and safety regulations, etc.).⁷

⁷ This literacy job profile is taken from the College Education Commission⁷ in New Zealand, see www.skillshighway.govt.nz. This website contains several other literacy job profiles as well. For example, a mechanic needs to be able to write notes to other technicians about work that has been done, explain to customers what work has been completed, listen to customers' description of problems with a vehicle, and ask customers questions to identify problems with a vehicle. A shop sales assistant needs to be able to write notes to managers or colleagues on work done, work still to do, and any problems that have arisen; use active listening skills to

In this section, we provide further (quantitative) evidence supporting the notion that good language skills are required in essentially all occupations in today's literacy-intense labor market. To this end, we first carry out content analysis of historical job advertisements, where we search for words that reflect that language proficiency is an important requirement when hiring. We then turn to the results from our employer surveys, where the respondents express how important language skills are to perform the tasks of the jobs in their firms. Finally, we present evidence from O*NET, now asking workers about everyday language use, which shows that the patterns we document for Sweden can also be seen in the US.

2.1 Content analysis of job advertisements

We retrieved historical job advertisements from a database at the Swedish Public Employment Service, which contains all vacancies published on their website. This is the largest vacancy website in Sweden, with approximately six million job advertisements published in 2007-2019.

We then investigated whether requirements for literacy skills are mentioned in the advertisements using a three-step procedure. First, we searched for the term "svenska" ("Swedish"), which appears in 28 percent of the job advertisements. If "svenska" is mentioned, we had to determine whether it refers to language skills, since it could refer to something else (e.g., "the Swedish market" or "a Swedish product"). In a second step, we therefore searched the text adjacent to the term "svenska" for words that would suggest that it refers to literacy skills requirements, e.g., "the job requires good knowledge of Swedish". In practice, this implied inspecting a subsample of advertisements containing "svenska" (N=1,000) and thereby obtaining a list of words frequently occurring when "svenska" refers to language skills, e.g., *speech, fluent, knowledge, understand, and communicate*. As a final step, we then searched all

understand and identify customer needs; make suggestions on improving work processes; and explain store policy to customers (on pricing, warranty, and returned and damaged goods).

advertisements again and identified which of them contained “svenska” + “literacy-related words”. This analysis shows that the extended search of “svenska” + “literacy related words” occurs in 23 percent of all advertisements, and, hence, when “Swedish” is mentioned in a job advertisement, it refers to requirements of literacy skills in over 80 percent (23/28) of these cases. We verified our procedure by manual inspection of two subsamples of advertisements and found that essentially no job advertisements were incorrectly classified.⁸ Appendix Section A1 describes our procedure in more detail.

Figure 1 illustrates the results of the content analysis using the pooled sample of advertisements in the 17 occupations included in the field experiment.⁹ In 2007, approximately 10 percent of job advertisements required literacy skills (“svenska” + “literacy related words”), while by 2019, the corresponding fraction had increased more than threefold to approximately 35 percent. This suggests that Swedish proficiency has become an increasingly important skill for workers to get hired.¹⁰

Figure 2 illustrates the results of the same content analysis when we group the 17 occupations included in the experiment by skill level (i.e., high-, medium-, and low-skill). The increase in the fraction of job advertisements with literacy skills requirements is strikingly similar across the three occupation categories.¹¹ Hence, the result for 2019 (close to when the field experiment started) shows a similar level of language proficiency requirements across all occupations.

⁸ We verified our procedure by taking two random subsamples of advertisements supposed to, in conjunction with “Swedish”, i) not mention and ii) mention requirements on literacy skills (N=1,000 in both cases). Hence, the first subsample includes advertisements containing “svenska” and none of the identified words, while the second subsample includes advertisements containing “svenska” and at least one of the identified words. A manual inspection of the texts in these advertisements confirm that there are essentially no advertisements that are incorrectly classified.

⁹ Conducting the content analysis including all occupations gives a similar picture (Appendix Figure A1).

¹⁰ Of course, this analysis cannot rule out alternative explanations. For example, employers could include language proficiency requirements in job advertisements to intentionally create a barrier for immigrant job seekers in times of increased immigration (i.e., effectively engaging in discrimination). However, the graphs have a similar positive slope throughout the period, while immigration increased dramatically in 2014-2015 during the war in Syria and the refugee crisis in Europe, which speaks against this as the main explanation.

¹¹ The same result is obtained if we consider the 17 occupations in the experiment separately; see Appendix Figure A2.

2.2 What do employers state about language proficiency requirements?

To investigate employers' stated preferences for language proficiency, we conducted two extensive employer surveys (Survey 1: N=382; Survey 2: N=437). Survey 1 targeted the same occupations and employers included in the field experiment, often smaller firms without HR departments, while the second survey targeted an independent sample of occupations and employers consisting mostly of large firms with HR departments. The employers were approached after the field experiment was completed and were asked to participate in a survey on the importance of language skills when hiring.

One of the questions in the surveys asked employers "How important is good knowledge of Swedish at your workplace?", which they responded to on a Likert scale. Appendix Figure A3 shows that the overall distributions across the response alternatives 1-7 are strikingly similar in the two surveys, despite the different target groups. Alternative 7, a response stating that good knowledge of Swedish is of utmost importance at their workplace, is by far the most frequent answer in both surveys, making up approximately 50 percent of the responses. Calculating the share of employers who responded 5 or greater (which we interpret as a clear indication that they view good knowledge of Swedish as important) amounts to over 80 percent.

Figure 3 presents this share separately for each of the occupations in our field experiment (i.e., using the answers in Survey 1). In essentially all occupations, more than 80 percent of the employers responded that good knowledge of Swedish is important at their workplace, again supporting the notion that being proficient in Swedish is almost a universal requirement. Two occupations deviate with a fraction far lower than 80 percent: computer professionals (49%) and waiters/waitresses (56%). Open comments to this question hint that English is sometimes the corporate language for computer professionals and that English is an alternative for waitering jobs.

2.3 Evidence from O*NET data

We also use O*NET data to illustrate the importance of language proficiency in essentially all types of jobs. O*NET describes the skills needed in different occupations in the US labor market and is collected by asking *workers* (instead of employers) questions about job requirements and characteristics.¹² We retrieved O*NET data from 2019 for questions concerning language skills and linked it to the 17 occupations in our field experiment using occupational codes in the Standard Occupational Classification (SOC).

Figure 4 shows the responses by occupation to questions on how often workers have face-to-face discussions at work and the demands for oral and written comprehension in everyday work. The responses from US workers are consistent with the findings for Swedish employers, i.e., that language use is frequent and that there is little variation in its use across occupations.

Taken together, the evidence presented from language research, the content analysis of job advertisements, our employer surveys, and the O*NET data on workers' perceptions paints a coherent picture of language proficiency being an important requirement in essentially all occupations when hiring. We next turn to the field experiment, where we investigate *how* important language skills are for firms' hiring decisions.

3. Design of the resumes in the field experiment

It has proven difficult to convincingly identify the return to language skills with observational data. With such data, it is challenging to distinguish the impact of language skills from other characteristics of job seekers that employers, but not researchers, may observe and to establish whether answers reflect real or perceived perceptions of the importance of language skills.

¹² See <https://www.onetonline.org/> for details.

Attenuation bias (especially if language skills are self-reported) and reverse causality (i.e., whether better labor market outcomes lead to better language skills) create additional problems.

We address these identification problems by using an experimental approach to measure the effect of language skills. The key feature of our approach is to manipulate the language proficiency level of cover letters of fictitious resumes, which are sent to employers with a job opening, and then measure the employers' responses. The responses from the employers (a dummy measuring callbacks) constitute the dependent variable, and because of random assignment of treatment, confounding factors can be ruled out, and any difference in the callback rate can credibly be ascribed to differences in language skills. Another important feature of our experiment is that it is carried out in a real environment and that employers are not aware of being part of an experiment, which should increase the external validity of the results. While this type of field experiment is an established method to study discrimination in hiring (see Bertrand and Duflo, 2017, for an overview), it has, to our knowledge, not been applied before to estimate the effect of language skills. Before conducting the experiment, we applied for and obtained ethical approval for the project; the preregistered application documents the research design, hypothesis, model specification, and statistical power.

3.1 Standardized resumes

The resumes were designed for young adults who search for a job four years after having immigrated to Sweden from one of four EU countries: Germany, Poland, Romania, and the UK. In the period from 2005 to 2015, these immigrant groups are the most common migrants to Sweden from the EU in the age group 25-35.¹³ The country of origin was randomly assigned to each resume. Most common second-language errors in written Swedish are general in character

¹³ Including four countries of origin is a compromise between the risk of including too few countries and, hence, making the study less generalizable and keeping the study feasible since each name/origin needs a separate "identity" (e.g., an email address).

and not specific to the country of origin (in contrast to spoken Swedish, e.g., ‘British-Swedish’), which means that we can use the same language manipulations irrespective of the applicant’s country of origin.

The starting point for constructing the fictitious resumes was a number of real resumes obtained from the Swedish Public Employment Service. These were adjusted to suit our purposes and to appear realistic and have qualifications for a typical jobseeker. The result of this process was a standardized template resume (without any occupation-specific information at this stage), consisting of a cover letter on the first page and skeleton for a detailed CV on the second page. This standardized format of the resumes is an important feature because it means that we use the *same* treatment across occupations, which facilitates estimating the effect of language proficiency using a pooled sample of observations. Thus, any difference in the effect of language proficiency between occupations cannot be attributed to occupations receiving different treatments.

The cover letter introduces the job applicant and summarizes his or her professional background. The baseline cover letter is written in error-free Swedish and starts with a brief presentation, which explicitly mentions the country the applicant has immigrated from in the second sentence.¹⁴ The cover letter also states that the applicant has been in Sweden for four years and has had a job in Sweden similar to the one applied for during the last two years. In addition, the cover letter states that the applicant has a few years of work experience from the country of origin and contains some additional personal information, such as leisure time interests. The details and structure of the baseline cover letter are shown in Appendix Section A4.

¹⁴ For example: “My name is Otto and I am 25 years old. I am from Germany and have lived in Sweden for approximately 4 years”.

The CV part of the resume contains an email address and a telephone number (with a voice mail), which the employers could use to contact the applicants. These were registered at a large internet provider and phone company; each of the names we used in the experiment had a separate email address and telephone number.¹⁵ The CV part also includes other information that is expected to be found there.¹⁶

We used both male and female names for the applicants. A resume was assigned a female name if the job was in a female-dominated occupation and a male name if the job was in a male-dominated occupation. We used one name for each gender and country of origin (i.e., eight names in total). The names are among the most common names for each country.¹⁷

It should be emphasized that the applicant's other human capital and personal information described in this section is not altered when we randomize the level of language proficiency, which we turn to next.

3.2 Manipulating the language proficiency of the resumes

A researcher in linguistics at Stockholm University, specialized in Swedish as a second language, was employed to manipulate the level of language proficiency in the standardized cover letter. Cover letters in Sweden are typically longer than what is common in countries such as the UK and the US. This makes the Swedish setting advantageous when manipulating language skills and conducting an experiment such as the current one.

¹⁵ Since employers never, or at least very seldom, contact job seekers by postal mail, we added a fictitious address for the job applicant to the CV, which was in the same city as the job applied for (Stockholm, Gothenburg, or Malmö). The addresses were located in similar suburbs (in terms of the socioeconomic characteristics of the residents) not too far from the city centers.

¹⁶ This information was address and detailed information on work experience, education (from the country of origin), computer skills, driver's license, and in relevant cases occupation-specific certificates.

¹⁷ The names used were Otto and Heidi Müller, Jakub and Zuzanna Kowalski, Andrei and Elena Popescu, and Jack and Emily Wilson. We tested the names with students and nationals (Germans, etc.) living in Sweden to ensure that they signal the correct country of origin and gender.

Using the baseline cover letter with Swedish without second-language features as the starting point, the professional linguist introduced a rich set of common second-language features. The end result was four versions of cover letters, where the level of language skills varied from L1 (lowest) to L4 (highest – the baseline). The first level down from the baseline (L3) uses a less advanced Swedish and contains approximately 20 second-language features. The cover letter with the second lowest level of language skills (L2) then adds approximately another 20 second-language features compared to L3. The cover letter with the lowest level of language proficiency (L1) is characterized by many and clear second-language features: Approximately 60 second-language features are included in this letter compared to the baseline, and the text is written in a much simpler type of Swedish. Nevertheless, it should be noted that even the cover letter with the lowest level of language skills (L1) is fairly straightforward to read and understand. This level of Swedish is typically achieved by individuals who have completed the most advanced level of the introductory *Swedish for Immigrants* course (SFI D)¹⁸. Although the number of additional features introduced for each level of language skills is approximately the same, linguistically, they are not alike. The first changes relate to second language features such as incorrect word order in complex sentence construction. The latter changes are more related to simplifications and an overuse of frequent verbs and expressions. Appendix Section A3 provides a detailed language theoretical explanation of all of the changes made in the cover letters. Importantly, the cover letters do not contain any simple spelling errors, which are easily identified and corrected by common word processing software and could signal traits other than language skills.¹⁹

¹⁸ Swedish for Immigrants (SFI) courses are offered to all immigrants to Sweden and are cost-free for participants. The courses offered are called A, B, C and D. In 2020, approximately one-third of recent immigrants had completed SFI D within three years after starting SFI (see <https://www.skolverket.se/publikationer?id=8251>).

¹⁹ A pilot study showed that introducing spelling errors signals that the applicant is lazy or sloppy and is not interested in the job.

In Table 1, we provide examples of sentences in the cover letter at different levels of language skills to give an idea of how the different proficiency levels are constructed. The table shows one sentence from the first (Column 1) and one from the last (Column 2) paragraph of the cover letter.

To verify whether the theoretical manipulation of language proficiency is perceived by regular recruiters and to learn how differently the four levels of language skills are valued, we tested the manipulated cover letters on two samples of employers. To this end, we again use the two employer surveys described in Section 2.2. Starting with Survey 1, after some initial questions, the survey presented excerpts, in random order, from two resumes with language proficiency levels L1 and L3 to be evaluated. The question read: “You will now be asked to assess extracts from two cover letters. The cover letters are intended to be part of a resume for a typical job at your workplace. How do you assess the applicant’s knowledge of Swedish from the text above?” The rating was on a Likert scale from 1 to 7, with 1 being “very low rating” and 7 a “very high rating”. The results show that the text manipulated into the lowest level of proficiency (L1) received an average score of 2.9, while the more advanced cover letter (L3) received an average score of 4.7, see Figure 5 (the difference is significant at the one percent level; $p < .000$). The recruiters in Survey 2 were instead presented with excerpts from all four proficiency levels. For both proficiency levels L1 and L3, these employers scored the resumes .5 higher. Furthermore, the increase in the average score between each of the four levels L1-L4 was almost identical, i.e., it was perfectly linear (in all three cases, the difference between one proficiency level and the next level is significant at the one percent level, with $p < .000$).

It is striking that the ratings for the language levels that are common across the two surveys (L1 and L3) are graded very similarly in both samples of employers, although the respondents are quite different in terms of factors such as hiring experience and firm size. An even more striking result, which we document in Figure 6, is that the ratings of L1 and L3 in Survey 1

(where we have information on occupation) are similar across essentially all the occupations in the field experiment.

Overall, these results show that employers perceive the different language skill levels in the cover letters and, more importantly, rate the four levels of language skills as having clear proficiency differences in a very linear way.

3.3 Occupations

Our main objective when choosing which occupations to include in the field experiment was to obtain a representative picture of the labor market. We included the 17 most common occupations included on the vacancy website, covering a substantial part of the labor market. Of these, 13 are low-/medium-skill occupations (requiring primary or secondary education), and 4 are high-skill occupations (which typically require postsecondary education).²⁰ The low-/medium-skill occupations are accounting clerks, chefs, cleaners, customer service workers, HR assistants, mechanics, office assistants, personal assistants, preschool teachers, shop sales assistants, truck drivers, waiters/waitresses, and warehouse staff workers. Five of these can be classified as low-skill occupations: cleaners, customer service workers, shop sales assistants, waiters/waitresses, and warehouse staff workers. The high-skill occupations are accountants, business salespersons, computer professionals, and IT support technicians.

There are two main reasons why not all occupations could be included in the experiment. First, managerial jobs and some other jobs that are very high-skill and/or specific are difficult to include because job applicants in these occupations are typically expected to be easy to find on business and employment-oriented online platforms (e.g., LinkedIn).²¹ Second, smaller

²⁰ When classifying the occupations, we used the ISCO classification. All high-skill jobs are category 3, while all low- and medium-skill jobs are categories 4-9.

²¹ Employers in these occupations may become suspicious if they get applicants (like our fictitious ones) who cannot be found on these platforms.

occupations with few posted vacancies were not worth including in the experiment since there is a substantial fixed startup cost associated with each included occupation.²²

As stated above, we used the same standardized template resume with language proficiency levels L1-L4 across all the occupations included in the experiment. However, we did add a few occupation-specific sentences to the cover letter when describing previous experience of occupation-specific job tasks. In these sentences, the linguist introduced common second-language features depending on the assigned level of language proficiency in the cover letter using the same approach as described above. The details on how the occupation-specific text enters the cover letter are shown in Appendix Section A4.

4. The field experiment and descriptive statistics

The experiment was conducted in the three largest cities in Sweden (Stockholm, Gothenburg, and Malmö) between February 2020 and April 2021. During this period, all job advertisements in the selected occupations found on the website of the Swedish Public Employment Service were collected.²³ In total, 3,124 employers were sent fictitious resumes by e-mail.²⁴ Invitations to job interviews and other positive responses were received by e-mail and telephone (voice mailbox²⁵). To minimize any inconvenience to the employers, we promptly declined invitations to job interviews.

Table 2 presents descriptive statistics of the field experiment. Approximately 74 percent of resumes were sent to firms advertising a job in a low- or medium-skill occupation (Panels A

²² For example, CVs with occupation-specific education and employment history had to be designed, and we had to formulate and express experience of relevant previous job tasks in the cover letter (for the four language levels). We concluded that the benefit associated with including an occupation in terms of additional observations does not outweigh the cost if the number of advertisements in an occupation is small.

²³ The website of the Swedish Public Employment Service is the main job search portal in Sweden.

²⁴ As documented in the application for ethical approval, we also sent refugee migrant resumes to the employers, which are not considered in this paper.

²⁵ Using their phone numbers, we also tracked and recorded employers that did not leave a message when they called.

and B), while the remaining 26 percent concerned high-skill occupations (Panel C). The number of sent resumes for each occupation and the share of the total sample varies somewhat across occupations (Columns 1 and 2). The variation is likely to reflect factors such as the size of an occupation and the occupation-specific local labor market. As shown in Column 3, the callback rate varies more substantially across occupations, being the highest among mechanics (25%) and lowest for office assistants (2%), which could reflect occupation-specific labor market tightness. The majority of the jobs were located in Stockholm (60%), followed by Gothenburg (25%), and then Malmö (15%) (not shown in the table).

The experiment was conducted during the COVID-19 pandemic. However, as is well known, Sweden decided not to shut down the labor market as many other countries did, and in many occupations, presence at the workplace was required throughout the pandemic. A recent government report (SOU 2022:30) stated that production was back at prepandemic levels in Q3 2020, and there was even a shortage of workers in 2021. We do not find any evidence of the pandemic affecting the callback rate during the experiment, which was at a rather stable and constant level throughout the experiment (see Appendix Section A5).

5. The effect of language skills

To analyze the average treatment effect of language skills on hiring, we estimate the following equation with the linear probability model (i.e., ordinary least squares):

$$Callback_i = \alpha + \beta * Language_i + \varepsilon_i \quad (1)$$

Here, the $Callback_i$ is an indicator that equals one if application i resulted in a positive response and the variable $Language_i$ measures the level of language proficiency – either as a continuous variable taking values 0-3 or as dummy variables for the level of language skills L1-L4. The

main parameter of interest is β , which measures the effect on the probability of receiving a callback, while α is the intercept.

5.1 Main results

Figure 7 shows that language proficiency has a positive effect on the callback rate, and this effect appears to be essentially linear – the callback rate increases almost with the same amount for each marginal step of language skills, as illustrated by the regression line. The magnitude of the effect is striking: Resumes with language skill level L1 have an average callback rate of approximately 8 percent, while resumes with language skill level L4 have a callback rate almost twice as high, at approximately 15 percent.

Table 3 presents a formal analysis of the effect of language proficiency. In Panel A, we include the language variable using a continuous variable ranging from 0 to 3 (L1-L4). A marginal increase in language skills raises the callback rate by approximately two percentage points (significant at the 1% level).²⁶ To obtain a sense of the magnitude of this effect, we can compare it to the effect of work experience estimated in Eriksson and Rooth (2014). This comparison shows that the full increase in the callback rate from better language proficiency is equivalent to the rise in the callback rate from having six more years of work experience.

Panel B presents estimates from a nonlinear specification, using a dummy variable for each level of language skills. The results are very similar to those obtained for the linear specification: increasing the level of language skills from ‘poor’ (L1) to ‘great’ (L4) increases the callback rate from 7.7 to 14.7 percent (i.e., $7.7 + 7.0$).

²⁶ There is no evidence that the language effect differs between applicants from Western and Eastern Europe. The language effect is positive and highly significant in both groups (Eastern Europe: .017, $p=.018$; Western Europe: .029, $p=.001$). An F-test of equal language coefficients across the two groups is not rejected ($p=.18$).

5.2 Language effects across occupation groups²⁷

The evidence presented in Section 2 indicates that language proficiency is important in essentially all occupations. Therefore, we expect that the positive effect of language skills should not differ much between different types of occupations. The following analyses will determine whether our main results are sensitive to which occupations we include in the analysis. Since each occupation on its own has too few observations, we group occupations to gain precision. The occupations are grouped based on the following characteristics: their skill level, historically less communication in Swedish to conduct work tasks, the share of immigrants (capturing typical immigrant jobs), and the share of women/men. The four different occupational groupings are distinct and have little overlap, as shown in Appendix Section A6.

Table 4 presents the effect of language proficiency for each of the occupational groupings presented above – a total of nine different language effects. This is achieved by extending the model in Table 3 by adding interaction effects between the continuous language skill variable and (dummies measuring) the different characteristics of the occupations. Since the language variable itself is omitted, we can directly compare these occupation-type language effects (the interactions) with the language coefficient of the main regression in Table 3.²⁸

We start by investigating whether the effect of language proficiency is the same irrespective of the occupations' skill level. In Panel A in Table 4, the language coefficients for the different occupational skill levels – low, medium and high – are all positive. Somewhat surprisingly, the coefficient for high-skill occupations is the smallest in magnitude, but an F-test does not reject the null hypothesis that the three coefficients are equal ($p=.61$).

²⁷ The analyses in this section are not included in the preregistered research plan and, hence, should be seen as explorative.

²⁸ Specifically, the language coefficient in Table 3 is the weighted average of the language coefficients in each panel in Table 4.

Some of the included low- and medium-skill occupations have historically been common for employing immigrants, often requiring less communication in Swedish to perform their work tasks. This was true for cleaners, warehouse staff workers, mechanics, truck drivers, and chefs (e.g., Karlsson, 2009). The regression in Panel B in Table 4 includes interaction variables between the continuous language variable and dummies for applying to a job where communication in Swedish at work historically was less required (compared to all other occupations grouped together). As in the previous analysis, both language coefficients are positive, with the estimate for occupations historically requiring less Swedish having a larger coefficient (.028 vs. .020). However, we cannot reject the null hypothesis that the two coefficients are equal ($p=.38$). In light of the evidence presented in Section 2, indicating a strong requirement for language proficiency in these occupations in today's labor market, this result is not surprising.

Relatedly, immigrants are overrepresented in some occupations, potentially reflecting lower requirements for proficiency in Swedish in those occupations. We therefore group the occupations into two groups depending on whether the share of immigrants employed in the occupation is above or below the median in the data (which is 15% in our data). Interestingly, the effect of language proficiency is actually larger in occupations with more immigrants (.026 vs .018), although the null hypothesis that the effect is equal in the two groups of occupations is not rejected ($p=.39$); see Panel C.

Finally, we analyze whether the effect of language proficiency is similar across female- and male-dominated occupations (remember that we sent only female applicants to female-dominated jobs and male applicants to male-dominated jobs). We conduct this analysis in Panel

D in the same way as in Panels A to C. Again, we estimate positive language coefficients, which are strikingly similar for male- and female-dominated occupations (.024 and .022).²⁹

In addition to the analysis presented in Table 4, we use the approach described in Section 2.1 to identify whether requirements for literacy skills are mentioned in the job advertisement and then include interaction variables between the continuous language variable and dummies for whether the job had a language requirement. We find a very similar (positive and significant) language effect among employers that mention and do not mention requirements on literacy skills in the job advertisement (see Appendix Section A7). Hence, both employers who include and employers who do not include a language proficiency requirement in their job advertisements seem to place a high value on language proficiency when hiring.

To learn whether the effect of language proficiency is universal across the labor market, we have grouped and analyzed the seventeen different occupations included in the experiment in nine different ways. Our experimental results nicely mirror the findings from the surveys in the sense that the requirement for language proficiency applies to all different types of occupations when applying for jobs in the Swedish labor market.

5.3 Increased human capital or a proxy for other (unobserved) skills?

An important question, not the least for policy, is *why* the callback rate increases with the level of language proficiency reflected in the resume. One potential explanation is that employers expect applicants with resumes reflecting better language proficiency to have higher productivity at work. This story is consistent with the human capital model, where language skills *per se* matter for productivity. However, another potential explanation is that employers view good language skills as a proxy for other individual characteristics that are unobserved by

²⁹ The F-test of the null hypothesis that the effect is equal in the two groups of occupations is not rejected ($p=.82$).

employers prior to hiring. They may, for example, believe that immigrants with better language skills have higher general abilities or are better culturally integrated into Swedish society – factors that could improve overall productivity in the workplace.

To help us interpret our results in relation to these two explanations, we use questions from our employer surveys. In Survey 1, we asked employers to choose among factors that they “see as a challenge in hiring people born outside the Nordic countries”. In this question, employers could select multiple factors that would support either human capital or the “proxy” story. Figure 8 shows the frequencies for the different response alternatives. The most commonly selected factor is by far “Lack of knowledge of the Swedish language”, while rather few employers selected the items “Lack of knowledge of the Swedish society”, “Difficulty in assessing foreign competence”, and “Risk of culture clashes in the workplace”.³⁰ Thus, “knowledge of the Swedish language” is perceived by employers to be by far the most important factor when hiring immigrants, which speaks in favor of the human capital explanation dominating the “proxy” explanation.

Additional support in favor of the human capital story is given by the question “How important is good knowledge of Swedish at your workplace?”. The answers to this question are likely to reflect to what extent language skills *per se* are important. Therefore, the fact that a large majority of employers (>80%) respond that language skills are important at their workplace is further evidence of the human capital story.

Overall, although we cannot rule out that language as a proxy for other unobserved abilities has some explanatory value, the answers to the survey questions are more supportive of the human capital story being the main explanation behind our results.

³⁰ The question also contained the response alternatives “Complicated government regulations” and “Other”, but these alternatives were chosen by very few respondents.

6. Conclusions

Western labor markets have undergone substantial changes in recent decades due to globalization, technological advances, and organizational changes. In particular, there is evidence that oral and written communication skills have become substantially more important in most jobs.

A group that is particularly affected by these changes is immigrants, who often have limited skills in the host country language. The increased requirements for oral and written language skills also in typical immigrant entry jobs present a major challenge for their labor market integration, which is an issue of key policy relevance in Europe and North America, where immigrants' labor market integration is currently a significant concern.

Despite the importance of this issue, we know little about how language skills are rewarded by employers, which is the main question we investigate. To this end, we conducted a large-scale field experiment in which a professional linguist manipulated the cover letters of fictitious resumes to obtain different levels of language proficiency and reflect variation in the language skills of real-world migrants. In total, we sent 3,154 fictitious resumes to advertised vacancies in 17 different occupations, randomly varying the level of language skills.

Our findings show that better language skills have a strong positive effect on the callback rate: improving language skills increases the callback rate linearly, and moving from a low level to a level similar to that of natives almost doubles the callback rate from 8 to 15 percent. This is a substantial effect and corresponds to the effect of an additional six years of work experience. Importantly, the effect of language skills is equally important across very different types of occupations. Thus, language proficiency is important not only in high-skill jobs but also in low- and medium-skill jobs and in occupations where immigrants are overrepresented. The responses to our employer surveys speak in favor of increased human capital rather than language as a proxy for other (unobserved) abilities being the main explanation of our findings.

Hence, learning the native language appears to be important for successful labor market integration, which is further supported by the large effects found in studies of language training programs.

Our results imply that poor skills in the host country language constitute a substantial hurdle for newly arrived immigrants, especially since our results apply to many typical immigrant entry jobs, such as cleaners, warehouse staff workers, and waiters/waitresses. Hence, creating more low-skill jobs may not be the solution to the current integration problems in Europe and North America.

A critique sometimes raised against correspondence testing experiments of discrimination (see Bertrand and Duflo, 2017) is that even if the degree of unequal treatment is internally valid, this does not necessarily have consequences for employment in labor market equilibrium. The argument, raised by Heckman (1998), is essentially that as long as there is a sufficient number of nondiscriminating employers who hire job seekers in the disadvantaged group, the employment rate of this group will not be affected. The key question in our context is whether there is a sufficient number of employers who will hire immigrant workers with poor language skills. Since we find that language proficiency is an important hiring requirement for most firms we study, immigrants cannot shy away from language requirements when looking for jobs using formal search methods (e.g., using the Employment Service). However, immigrants with poor language skills might find jobs informally within their ethnic network (see, e.g., Lazear, 1999, and Lang, 1986). Within an ethnic network, it may be sufficient if one or a few workers at each workplace are proficient in Swedish (e.g., those who communicate with customers), while a common foreign language is used for communication among the employees.

In light of our results, if immigrants do not invest in learning the native language, either the unemployment rate among immigrants and/or employment through ethnic networks will increase, making the labor market more ethnically segregated. Both of these scenarios will have

political and social as well as economic consequences affecting both immigrants and society at large. To prevent these scenarios, it is important to design policies and implement measures to raise immigrants' language skills. This means that the results from the emerging literature on what constitutes effective language training courses seem crucial (e.g., Foged et al., 2022, and Heller and Slungaard Mumma, 2022).

References

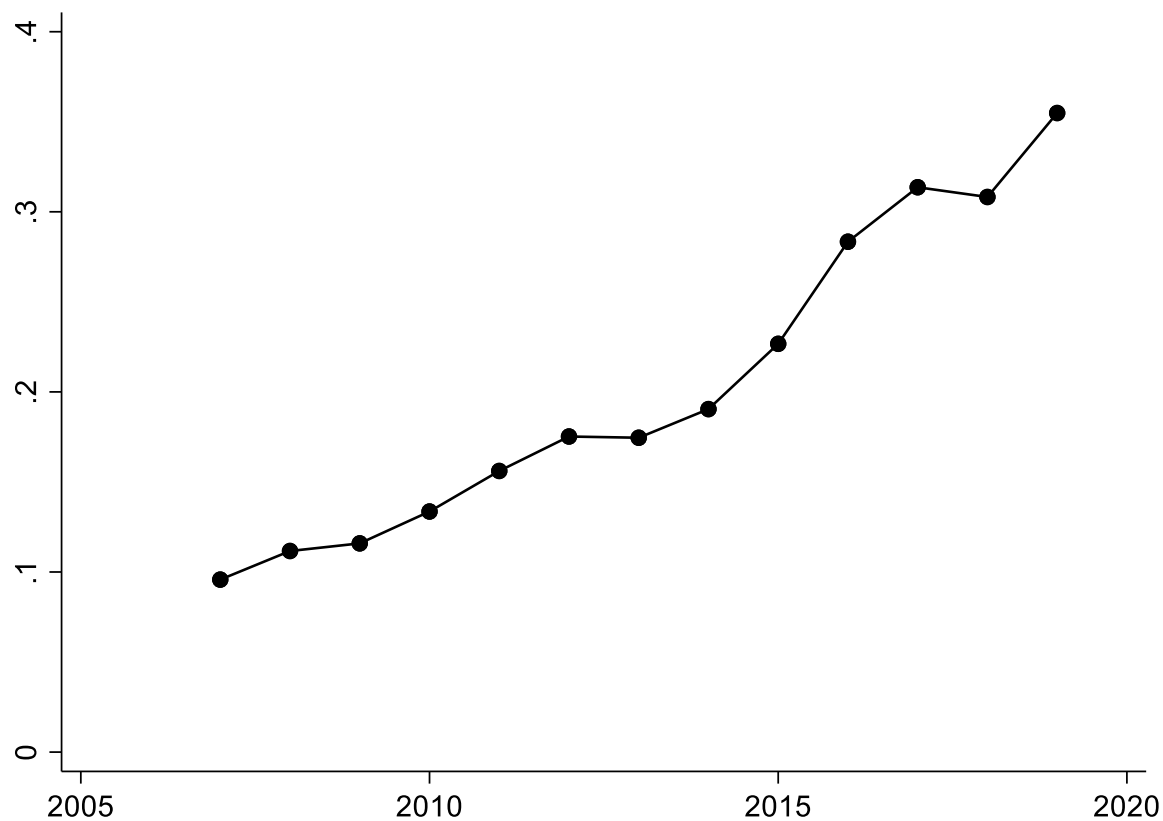
- Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The Evolution of Work in the United States. *American Economic Journal: Applied Economics*, 12(2):1–34.
- Autor, D. H., Salomons, A. and Seegmiller, B. (2021). New Frontiers: The Origins and Content of New Work, 1940-2018. *MIT Working Paper*, July 2021.
- Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5): 1553–97.
- Bertrand, M., and Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, 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.
- Bleakley, H., and Chin, A. (2004), Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86, 481-496.
- Brell, C., Dustmann, C., and Preston, I. (2020). The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives*, 34(1), 94-121.
- Carlsson, M., and Eriksson, S. (2019). Age discrimination in hiring decisions: Evidence from a field experiment in the labor market. *Labour Economics*, 59, 173-183.
- Carlsson, M., and Rooth, D. O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour economics*, 14(4), 716-729.
- Chiswick, B. R., and Miller, P. W. (2007). The economics of language: International analyses. Routledge, London.
- Chiswick, B. R., and Miller, P. W. (2015). International migration and the economics of language in *Handbook of the Economics of International Migration*, Vol. 1, 211-269, North-Holland.

- Corbel, C., T. Newman, and Farrell, L. (2022). Gig expectations: Literacy practices, events, and texts in the gig economy. *Written Communication*, 39(1): 66-96.
- Deming, D. J. and Noray, K. (2020). Earnings Dynamics, Changing Job Skills, and STEM Careers. *The Quarterly Journal of Economics*, 135(4):1965–2005.
- Dustmann, C., Fabbri, F. (2003), Language proficiency and labour market performance of immigrants in the UK, *Economic Journal*, 113, 695-717.
- Dustmann, C. and van Soest, A. (2002). Language and the earnings of immigrants. *Industrial Labor Relations Review*, 55 (3), 473–492.
- Dustmann, C. and van Soest, A. (2001). Language fluency and earnings: estimations with misspecified indicators. *Review of Economics and Statistics*, 83 (4), 663–674.
- Dustmann, C. (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics*, 7(2), 133–156.
- Eriksson, S., and Rooth, D. O. (2014). Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. *American economic review*, 104(3), 1014-39.
- Farrell, L., Newman, T., and Corbel, C. (2021). Literacy and the workplace revolution: a social view of literate work practices in Industry 4.0. *Discourse: Studies in the Cultural Politics of Education*, 42(6), 898-912.
- Foged, M., Hasager, L., Peri, G. Nielsen Arendt, J. Bolvig, I. (2022), Language Training and Refugees' Integration, *Review of Economics and Statistics* (forthcoming).
- Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *American Economic Review*, 99(2): 58–63.
- Heckman, J. J. (1998). Detecting discrimination. *Journal of economic perspectives*, 12(2), 101-116.
- Heller, B. and Slungaard Mumma, K. (2022). Immigrant integration in the United States: The

- role of adult English language training. *American Economic Journal: Economic Policy* (forthcoming).
- Karlsson, A-M. (2009). Positioned by reading and writing: Literacy practices, roles, and genres in common occupations. *Written Communication*, 26(1): 53-76.
- Lang, K. (1986). A language theory of discrimination. *The quarterly journal of economics*, 101(2), 363-382.
- Lazear, E. P. (1999). Culture and language. *Journal of political Economy*, 107(S6), S95-S126.
- Lochmann, A., Rapoport, H., and Speciale, B. (2019). The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review*, 113, 265-296.
- Sarvimäki, M., and Hämäläinen, K. (2016). Integrating immigrants: The impact of restructuring active labor market programs. *Journal of Labor Economics*, 34, 479-508.
- SOU (2022). Kommittén om stöd vid korttidsarbete. Korttidsarbete under pandemin – en utvärdering av stödets betydelse. Finansdepartementet. SOU 2022:30.
- Yao, Y., van Ours, J. C. (2015). Language skills and labor market performance of immigrants in the Netherlands. *Labour Economics*, 34, 76-85.

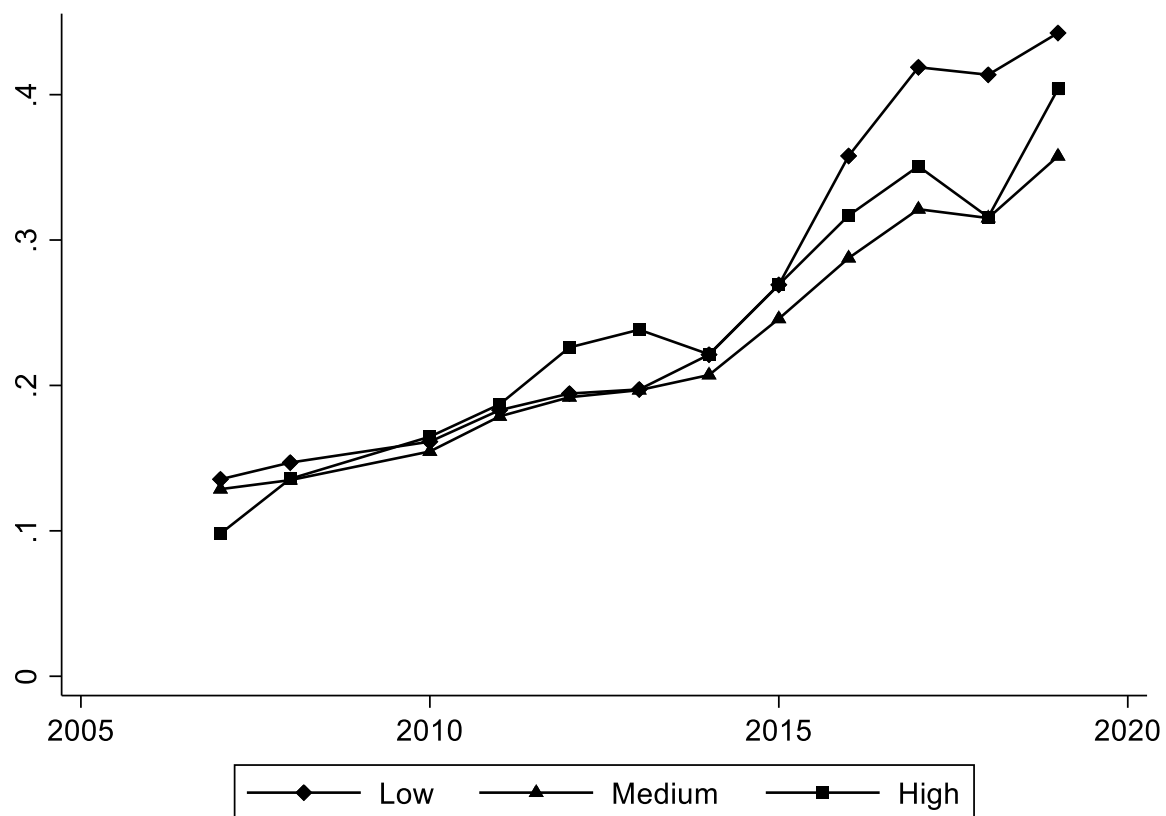
Tables and Figures:

Figure 1. Fraction of job advertisements with requirements for language proficiency.



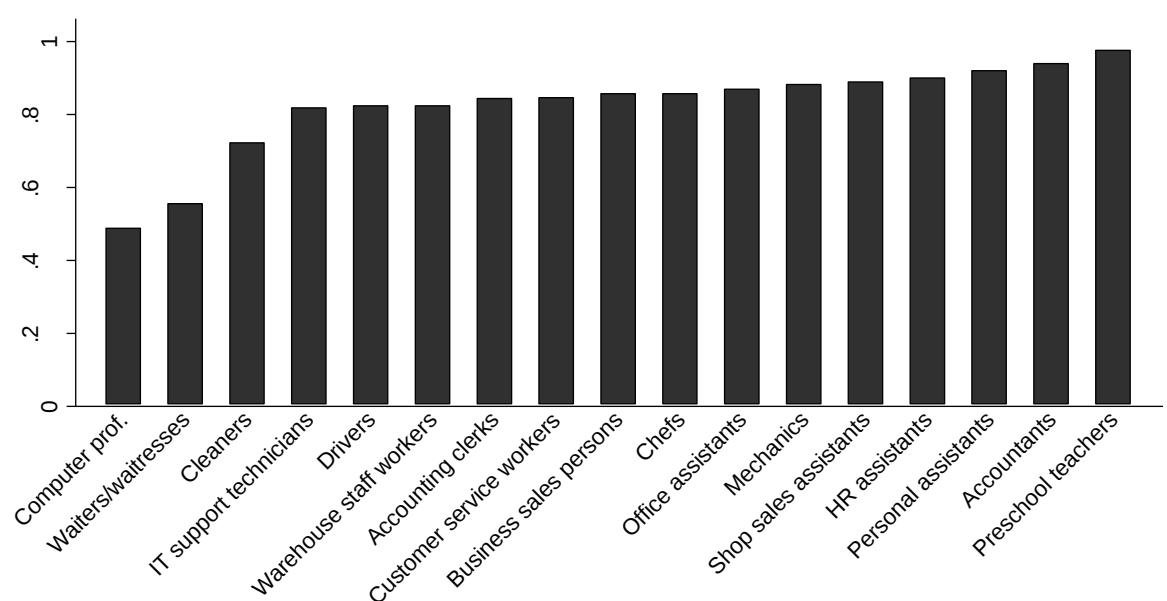
Notes: $N=1,311,091$. The figure is constructed from the job advertisements posted at the Swedish Employment Service from 2007 to 2019 in the 17 occupations included in the field experiment.

Figure 2. Fraction of job advertisements with requirements for language proficiency, by skill level.



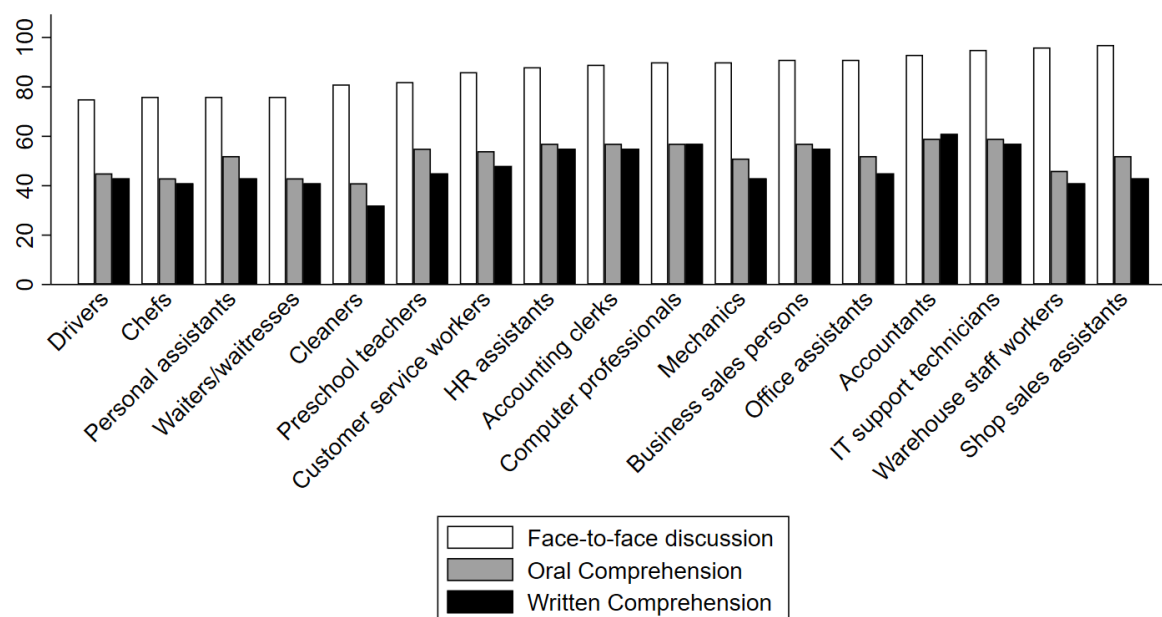
Notes: $N=1,311,091$. See also the notes below Figure 1.

Figure 3. Importance of language skills in Swedish.



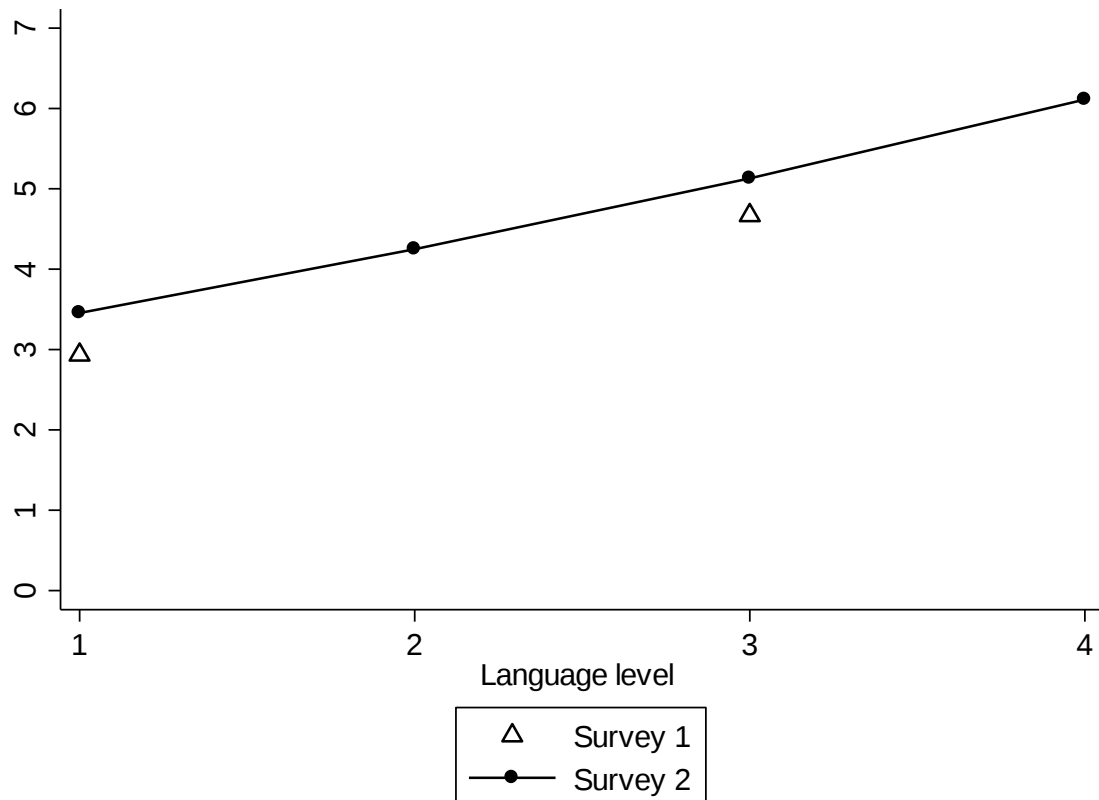
Notes: $N=382$. The employers participating in Survey 1 are a subsample of the employers included in the field experiment, and they were asked “How important is good knowledge of Swedish at your workplace?”, which they answered on a Likert-type scale from 1 to 7. Each bar shows the fraction of employers in the occupation that responded 5 or greater.

Figure 4. O*NET items on the importance of language skills.



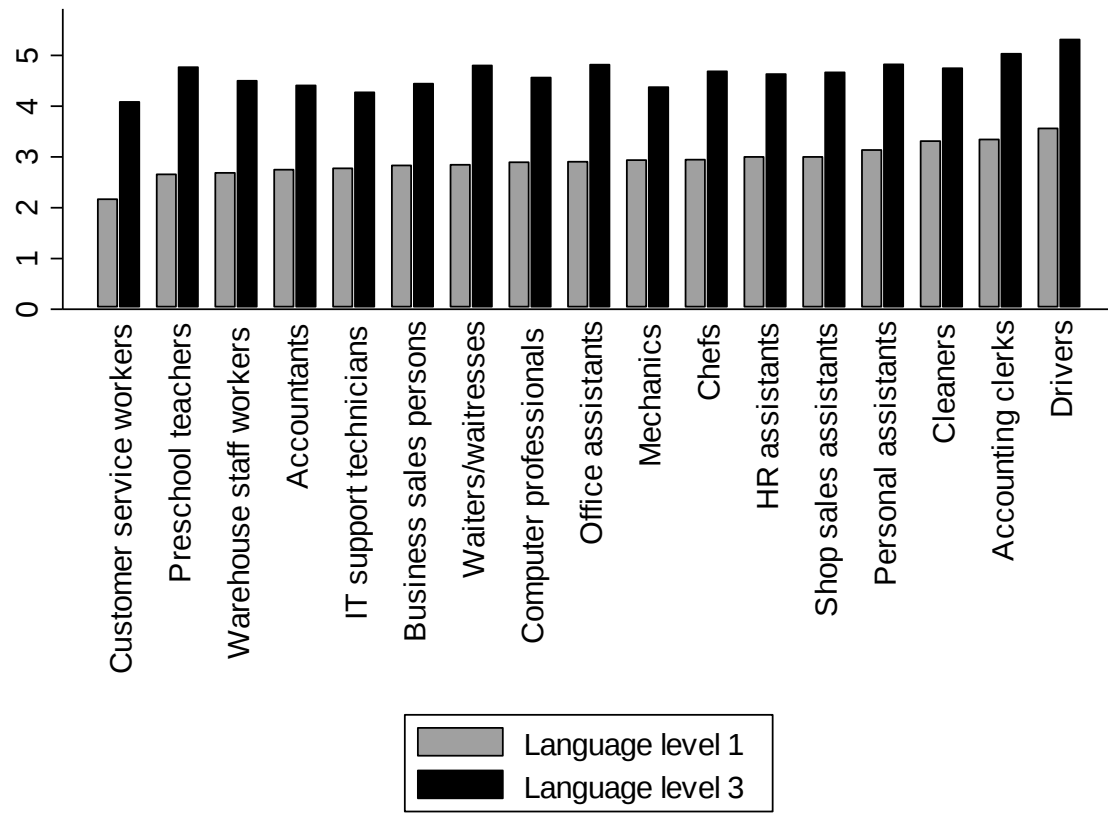
Notes: The survey question for face-to-face discussion was “How often do you have to have face-to-face discussions with individuals or teams in this job?”. The survey questions for oral and written comprehension concerned the needed level in this job concerning “The ability to listen to and understand information and ideas presented through spoken words and sentences” and “The ability to read and understand information and ideas presented in writing”. The variables in the O*NET data we use are indices that span 0-100. See <https://www.onetonline.org/> for details.

Figure 5. Language proficiency ratings of cover letters.



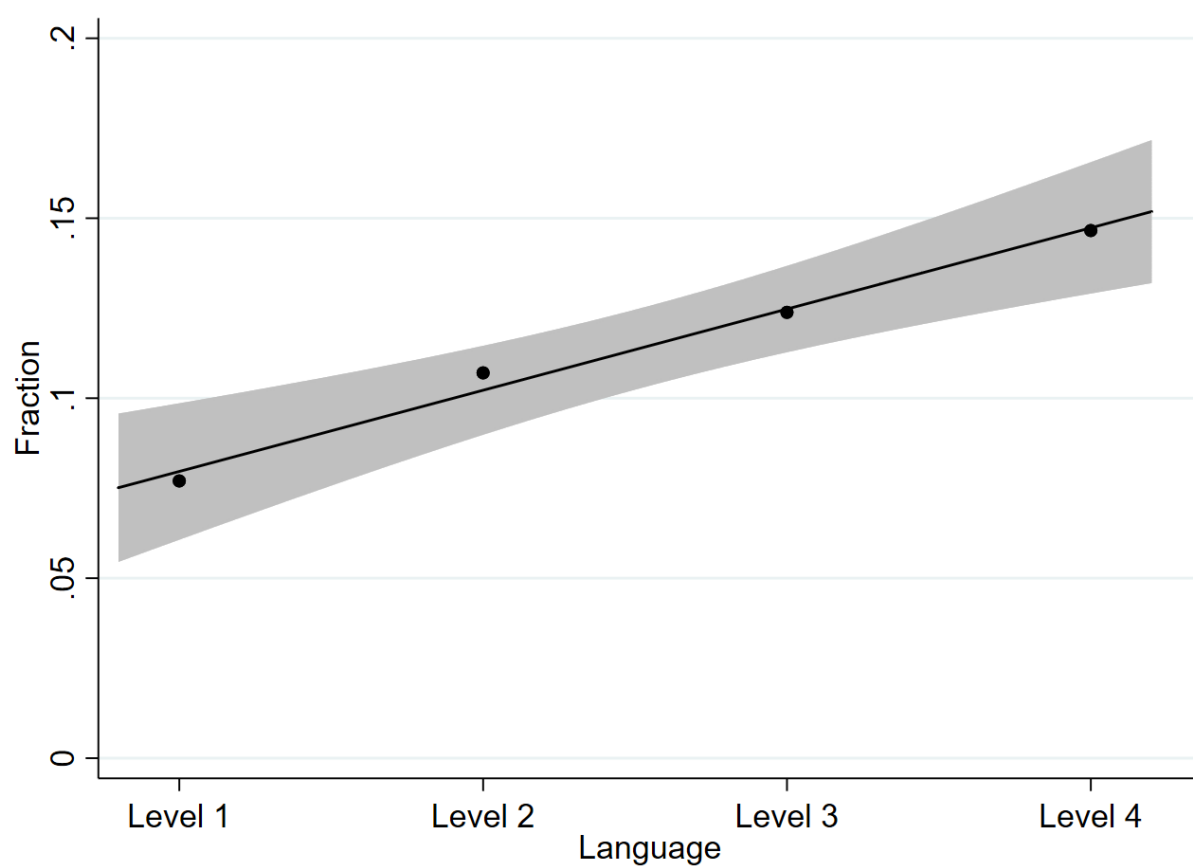
Notes: N=382 in Survey 1 and N=437 in Survey 2. The employers participating in Survey 1 are a subsample of the employers who were included in the field experiment, and they rated resumes with language levels 1 and 3. The employers participating in Survey 2 were an independent sample of employers and consisted mostly of large firms with HR departments, and they rated resumes with all four levels of language. The question in both survey experiments read: “You will now be asked to assess extracts from two cover letters. The cover letters are intended to be part of a resume for a typical job at your workplace. How do you assess the applicant’s knowledge of Swedish from the text above?” The rating was on a Likert scale from 1 to 7, with 1 being “very low rating” and 7 a “very high rating”.

Figure 6. Language proficiency ratings of cover letters, by occupation.



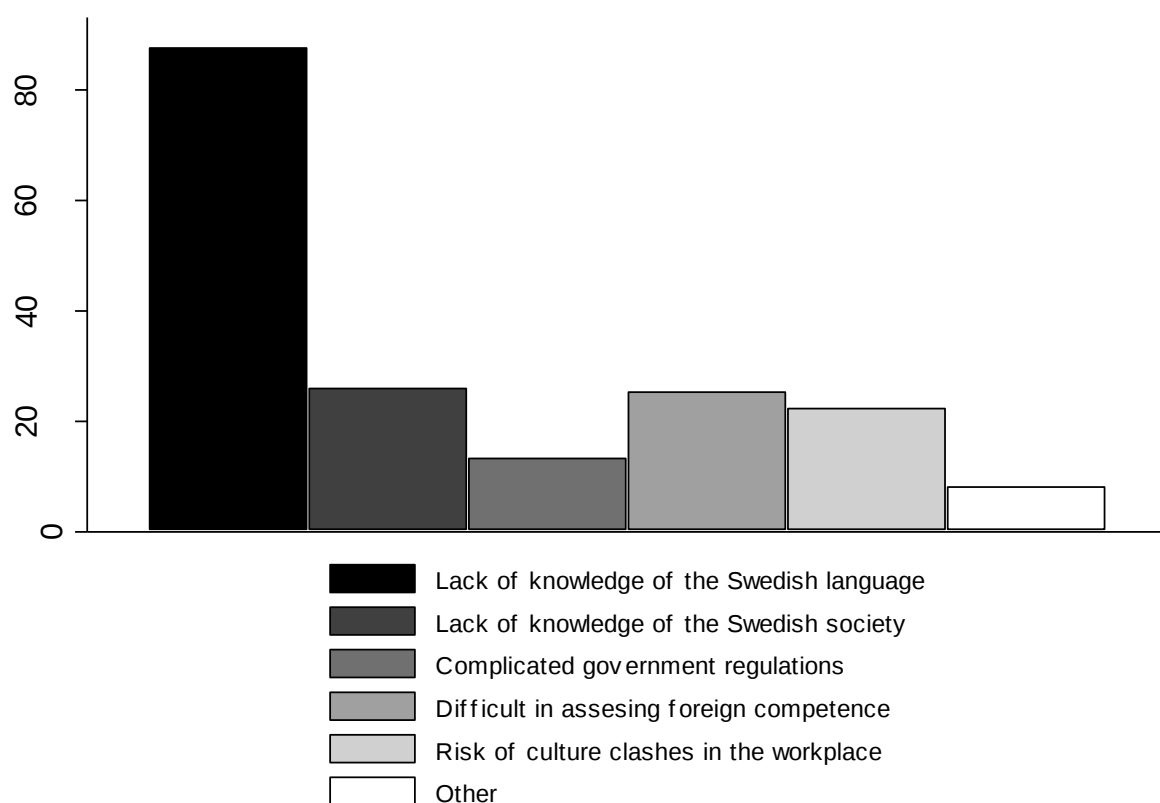
Notes: N=382 (Survey 1). See also the notes below Figure 5.

Figure 7. Callback rate by the level of language skills.



Notes: $N=3,153$ resumes. The four dots in the figure are the mean callback rates for the different levels of language skills, L1-L4. The shaded area shows the 95% confidence interval of the slope of the regression line in Table 3.

Figure 8. Factors seen as a challenge when hiring people born outside the Nordic countries.



Notes: $N=382$. This question was asked in Survey 1: "Which of the following factors do you see as a challenge in hiring people born outside the Nordic countries? Several options may be selected."

Table 1. Examples of text in the cover letter at different language proficiency levels.

Language proficiency level	(1)	(2)
L4 (highest)	Others usually describe me as open, easy to get along with and organized	I hope I have piqued your interest and will be happy to come and introduce myself more in an interview.
L3	Others usually describe me as open, easy to get along and organized	I hope I have piqued your interest and will be happy to come and introduce myself more in interview.
L2	Others describe me open, easy to get along and organized	I hope your interest and I happy to come and introduce myself on interview
L1 (lowest)	Others tell I open, agree well and organized	I like introduce me in interview

Notes: Column 1 shows text from the initial part and Column 2 from the final part of the cover letter. Own translations from Swedish.

Table 2. Descriptive results of the field experiment.

	Number of resumes	Share of sample (%)	Callback rate (%)
<i>Panel A: Low-skill occupations</i>	978	31.0	9.0
Cleaners	206	6.5	17.0
Customer service workers	155	4.9	12.3
Shop sales assistants	123	3.9	3.3
Waiters/waitresses	308	9.8	6.2
Warehouse staff workers	186	5.9	5.9
<i>Panel B: Medium-skill occupations</i>	1,354	43.0	12.0
Accounting clerks	192	6.1	5.7
Chefs	160	5.1	8.1
HR assistants	126	4.0	5.6
Mechanics	167	5.3	25.2
Office assistants	179	5.7	2.2
Personal assistants	193	6.1	20.2
Preschool teachers	163	5.2	8.0
Truck drivers	174	5.5	19.0
<i>Panel C: High-skill occupations</i>	821	26.0	13.5
Accountants	230	7.3	8.3
Business salespersons	211	6.7	11.4
Computer professionals	292	9.3	19.9
IT support technicians	88	2.8	11.4
<i>Panel D: All occupations</i>	3,153	100	11.5

Table 3. The effect of language proficiency on the callback rate.

(1)	
Panel A: Linear model	
Constant (L1)	.0796*** (.0086)
Language	.0226*** (.0050)
Panel B: Nonlinear model	
Constant (L1)	.0770*** (.0096)
L2	.0300** (.0151)
L3	.0468*** (.0147)
L4	.0696*** (.0158)

Notes: N= 3,153. In the regression in Panel A, Language is a continuous variable. In the regression in Panel B, there is one dummy variable for each language level (where L1 is the omitted category). The regressions include only the variables listed in the table. Standard errors in parentheses are robust.

*** $p < .01$, ** $p < .05$, * $p < .1$

Table 4. The effect of language proficiency on the callback rate, by occupational groupings.

Panel A: Occupational skill level		Panel C: Share of immigrants in occupation	
Language x low-skill	.0206*** (.0074)	Language x share immigrants above median	.0265*** (.0084)
Language x medium-skill	.0286*** (.0081)	Language x share immigrants below median	.0177*** (.0052)
Language x high-skill	.0163 (.0107)	Share immigrants above median	.0186 (.0262)
Medium-skill	.0174 (.0186)	p-value (test of equal Language coefficients)	.39
High-skill	.0524** (.0230)		
p-value (test of equal Language coefficients)	.61		
Panel B: Occupations historically requiring less Swedish		Panel D: Female- or male-dominated occupation	
Language x occupations historically requiring less Swedish	.0285*** (.0074)	Language x female dominated	.0216*** (.0064)
Language x all other occupations	.0200*** (.0058)	Language x male dominated	.0236*** (.0060)
Occupations historically requiring less Swedish	.0360 (.0317)	Female dominated	-.0554** (.0249)
p-value (test of equal Language coefficients)	.38	p-value (test of equal Language coefficients)	.82

Notes: $N = 3,153$. The regression in the table includes interaction variables between the continuous language variable and dummies for the skill level of the occupation (Panel A), whether the occupation historically have required less communication in Swedish to perform the job tasks (Panel B), whether the share of immigrants in the occupation is above or below the occupational median of 15% (Panel C), and female- and male-dominated occupation (Panel D). Note that the language variable itself is omitted in the regressions. The regressions include only the variables listed in the table. The dummy for historically requiring less Swedish is equal to 1 for chefs, cleaners, mechanics, truck drivers, and warehouse staff workers and 0 for the other occupations. The dummy for being above the median share of immigrants is equal to 1 for chefs, cleaners, mechanics, personal assistants, preschool teachers, truck drivers, warehouse staff workers, and waiters/waitresses and 0 for the other occupations. The dummy for female-dominated occupation is equal to 1 for accountants, accounting clerks, customer service workers, HR assistants, office assistants, personal assistants, preschool teachers, shop sales assistants, and waiters/waitresses and 0 for the other occupations. Appendix Table A2 shows how the different occupational groupings used in Panels A-D overlap. Standard errors in parentheses are robust.

*** $p < .01$, ** $p < .05$, * $p < .1$

Online Appendix for:

**“Language Proficiency and Hiring of Immigrants:
Evidence from a New Field Experimental Approach”**

Magnus Carlsson, Stefan Eriksson, and Dan-Olof Rooth

A1. Content analysis of job advertisements

A1.1 Details of the procedure for conducting the content analysis

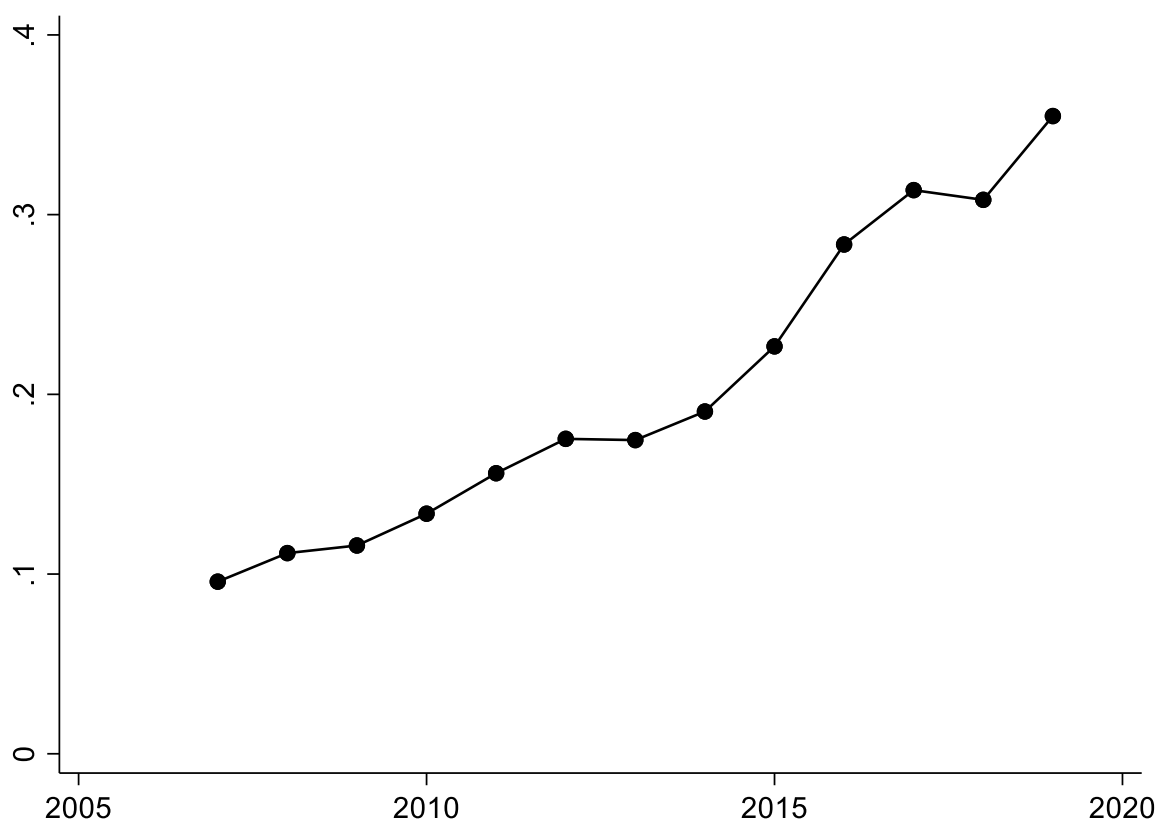
We retrieved historical job advertisements from a database at the Swedish Public Employment Service, which contains all vacancies published on their website. This is the largest vacancy website in Sweden with approximately six million job advertisements published in 2007-2019.

We searched all job advertisements in 2007–2021 (N=5,990,989) for the term “svenska” (“Swedish”), and it appears in 28% of the job advertisements. To determine the context of the term “svenska”, we then searched the text adjacent to “svenska” – 50 characters to the left and right – for words which we would expect to find if “svenska” refer to language skills. To figure out which words to search for we considered a random subsample of 1,000 advertisements containing “svenska” and identified the words which frequently occurred when it clearly referred to language skills (it turned out that in practice all the relevant words were identified already after having considered the first 50-100 advertisements). The following words were identified (translated from Swedish): *speech, oral, writing, written, fluent, good, language, mother tongue, command, requirements, knowledge, understand, speak, telephone, read, understood, know, correspondence, communicate, communication, communication skills, and contact, customer questions*. We then searched all advertisements again and identified which of them contain both “svenska” and at least one of the identified words in the text adjacent to “svenska”. This analysis shows that the extended search of “svenska” + ”literacy related words” occurs in 23 percent of all advertisements, and hence, when “Swedish” is mentioned in a job advertisement it refers to requirements of literacy skills in over 80 percent (23/28) of these cases. Finally, we verified our procedure by manual inspection of two subsamples of advertisements. We did this by taking two random subsamples of advertisements supposed to i) not mentioning and ii) mentioning requirements on literacy skills, respectively (N=1,000 in both cases). The first subsample are advertisements containing “svenska” and none of the

identified words. The second subsample are advertisements containing “svenska” and at least one of the identified words. A manual inspection of the texts in these advertisements confirm that there are essentially no advertisements which are incorrectly classified.

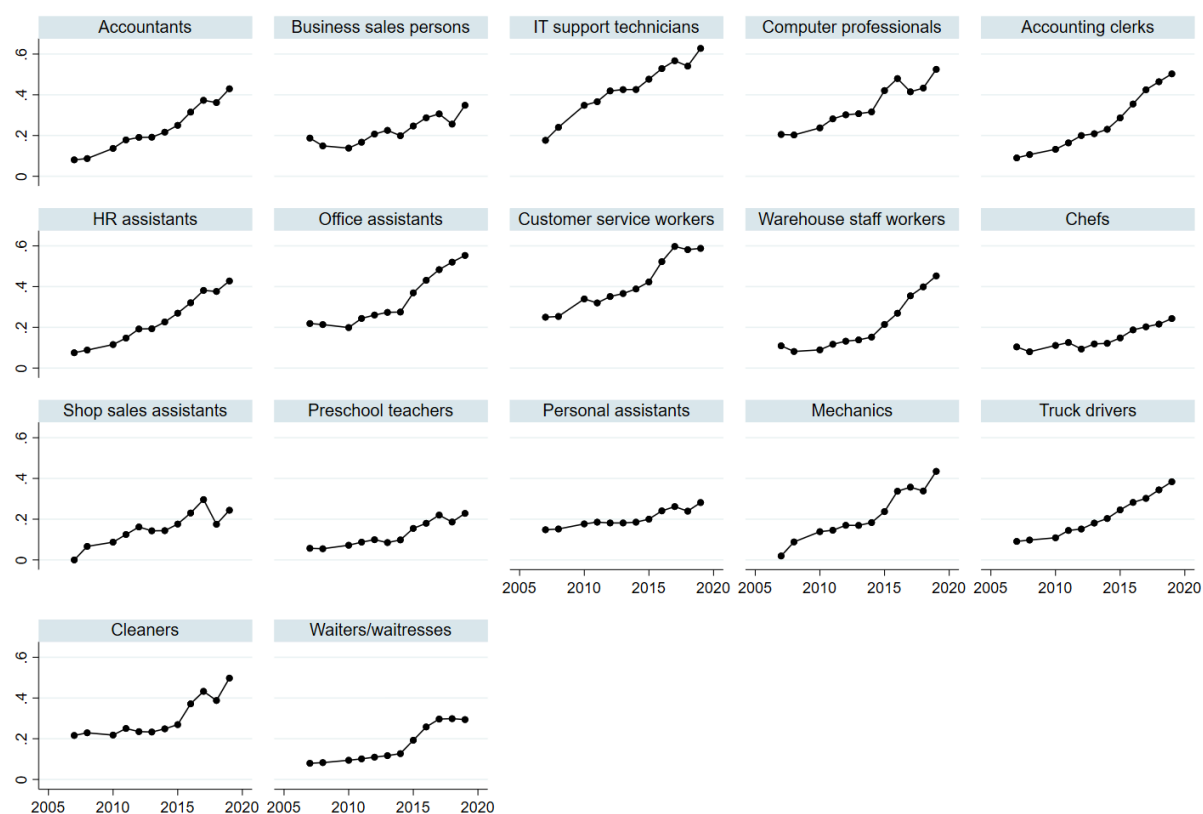
A1.2 Additional results for content analysis

Appendix Figure A1. Fraction of job advertisements with requirements on language proficiency. All occupations.



Notes: $N=5,990,989$. The figure is constructed from all job advertisements posted at the Swedish Employment Service from 2007 to 2019 in all occupation, i.e., it is constructed from the full sample of job advertisements.

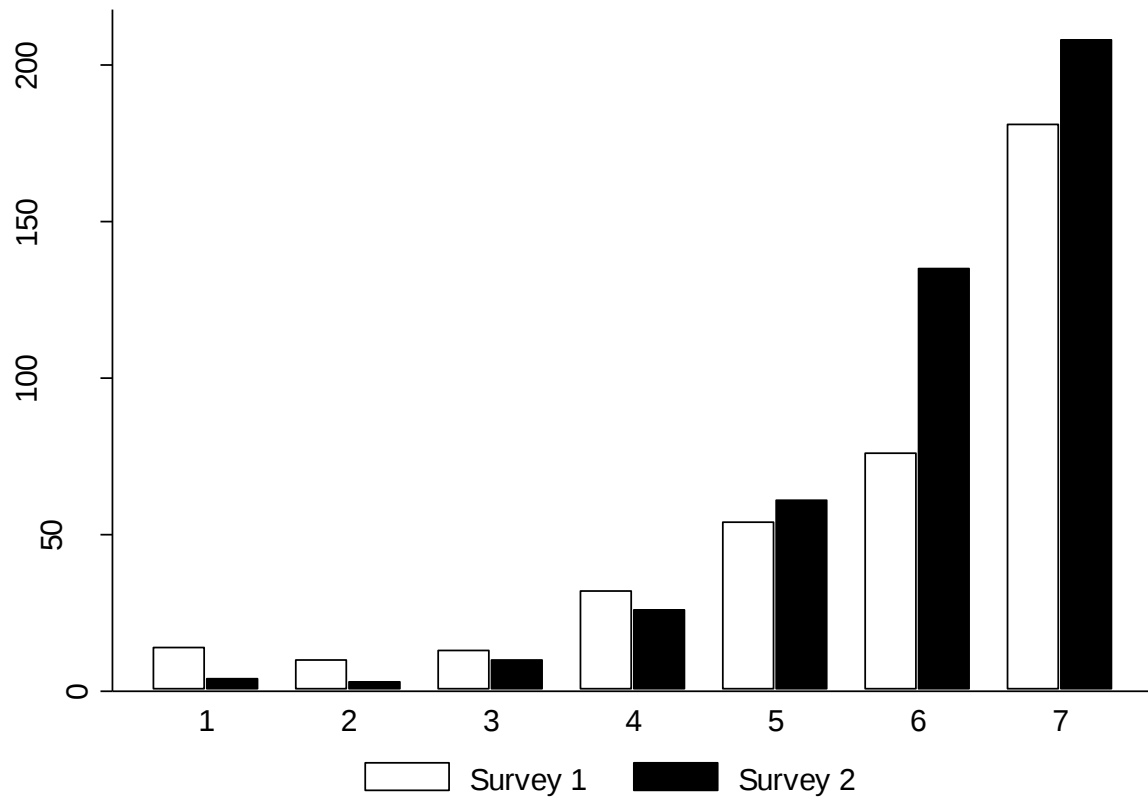
Figure A2. Fraction job advertisements with requirements on language proficiency, by occupation.



Notes: $N=1,311,091$. See also the notes below Figure 1.

A2. Additional results from employer surveys

Appendix Figure A3. The importance of good knowledge of Swedish at the workplace.



Notes: The employers in Survey 1 (N=382) and 2 (N=437) were asked “How important is good knowledge of Swedish at your workplace?” which they answered on a Likert type scale from 1 to 7. See also the notes to Figure 5.

A3. Description of the different language levels

The texts with the first language level are at a low level with many and clear second-language features/errors. The texts are characterized by simplicity and by many (on average 60) and clear second-language features. There is an overuse of frequent verbs and expressions. The texts have no major variation in sentence construction and do not handle inversion. They maintain basic grammatical structures such as inflection of verbs and marking of the plural, but there are also examples of uninflected or omitted verbs. Prepositions are also sometimes omitted.

The second level is an intermediate level with clear second-language features. The texts are relatively linguistically advanced with variation in sentence structure and some precision and variation in word choice. The texts are adapted to the genre, but with many (on average about 40) and clear second-language features. In relation to the highest level, many expressions have been simplified. The word order is varied with simpler grammatical structures than the highest level, the writer can handle inversion, but not with more advanced grammatical structures. The texts have several deviations from the highest level in determination and congruence and deviating prepositional use. These texts contain the same type of errors as level 3, but many more.

The texts at level 3 are at an advanced level with some second-language features/errors. The texts are linguistically advanced, well adapted to the genre and mainly correct, but with some (on average about 16) second-language features that are common even at advanced levels such as deviant prepositional use, deviations in determination and congruence, incorrect word order in complex sentence construction (e.g., pointed clauses, sentence adverbial placement in clauses) and simplifications of idiomatic expressions. They have good variety and precision both grammatically and lexically. The second-language features are generally few.

The texts at level 4 are at an advanced level. They are authentic personal letters that are linguistically advanced (grammatically and lexically), well adapted to the genre and correct.

Starting from the baseline cover letter written in error-free Swedish with language proficiency level 4 (shown in Appendix Figure A5), Appendix Table A1 shows in detail the second-language features introduced into the cover letter as we gradually decrease the level of language proficiency from 4 to 3, from 3 to 2, and from 2 to 1.

Appendix Table A1. Introduction of common second–language features into the cover letter.

Introduced second-language features when the level of language proficiency decreases from		
4 to 3	3 to 2	2 to 1
(i) Simplifications of idiomatic expressions (förenkling av idiomatiska uttryck, ordförråd) (3 changes): Delaktig i → med; Vidareutveckla detta → utveckla; Jag är mån om → jag är mån (ii) Incorrect word order (Ordföljd, ej inversion) (2): Innan jag kom till Sverige arbetade jag → innan jag kom till Sverige jag arbetade; Oavsett min arbetsplats så är jag → oavsett min arbetsplats jag är (iii) Deviation in determination and congruence (bestämmdhet och kongruens) (4): Ständigt → ständig Ett positivt sätt → Ett positivt sätt; Ett representativt sätt → Ett representativt sätt; Mitt senaste jobb → mitt senaste jobbet; Våra många vänner → våra många vännerna (iv) Verb form: Förändras → förändrar (v) Deviant prepositional use (prepositionsavvikelse) (3): Stor vana av → stor vana på; I alla typer av jobb → på alla typer av jobb; Erfarenhet av → erfarenhet på; Van vid → van på	(i) Simpler grammatical structure (grammatisk förenkling)(5): Om jag ska berätta lite om mig själv så heter jag → Jag heter; Vad jag skulle vilja jobba med → vad jag vill; Vilket → som; Våra många vänner → många vänner; Utvecklades → utvecklade (ii) Simplification of idiomatic expressions (förenkling av idiomatiska uttryck) (14): förändras → ändras; Jobba med → jobba; Jag är mån om att → jag vill; det utannonserade jobbet → jobbet; samt → och; delaktig → med; vidareutveckla detta → utveckla; i samband med → på; goda kunskaper → bra kunskaper; liknande → samma även → också; ta ansvar för att → ta ansvar att; hög arbetstakt → arbeta mycket (iii) Incorrect word order (ordföljd, ej inversion) (4): arbetade jag → jag arbetade; så är jag → jag är; Som person är jag → som person jag är (iv) Verb form: Ändras → ändrar (v) Deviation in determination and congruence (bestämmdhet och kongruens) (6): Ständigt (förändras) → som ständig (ändras); Höga ambitioner → hög ambitioner; Ett positivt sätt → ett positivt sätt; Ett representativt sätt → ett representativt Mitt senaste jobb → mitt senaste jobbet; Min erfarenhet → min erfarenheter (vi) Deviant prepositional use (prepositionsavvikelse) (5): På ett positivt sätt → i en positivt sätt; På Scandic hotell → i Scandic; I alla typ av jobb → på alla typ av jobb; Vana av → vana på; Erfarenhet av → erfarenhet på (vii) Genus (1): Ett sätt → en sätt	(i) Simplifications of idiomatic expressions ((förenkling av idiomatiska uttryck, ordförråd)): Jag fick bra kunskaper om mat och dryck → jag lära mycket om; Oavsett min arbetsplats → på alla mina jobbet; Att jobba → jobba; har lärt mig att ta ansvar → ta ansvar; ständigt → alltid; Min erfarenhet har gett mig stor vana → jag har mycket erfarenheter på; Representativ → trevlig; På jobbet → på jobba; Samma arbetsuppgifter → samma arbetsuppgifterna; Alla typer av jobb → alla typer jobb; van vid en hög arbetstakt → tycker om att jobb hela tiden vid en intervju → på intervju (ii) Incorrect word order (ordföljd, ej inversion): Sedan tidigare har jag → tidigare jag har; Då tar jag → då jag tar

Notes: The cover letter with the highest language proficiency level (4), which is similar to what could be expected from a native, is found in Appendix Figure A5. The columns list the modifications made to the cover letter as the level of language proficiency is decreased from 4 to 3, from 3 to 2, and from 2 to 1.

A4. The design of the resumes

Appendix Figure A4 shows the cover letter part, translated from Swedish, of the standardized resume template with the text not in bold being the same across occupations (the original text in Swedish is shown in Appendix Figure A5). The bold parts were replaced by other text depending on random variables and context when the complete resumes which we sent were generated. The parts in bold in square brackets varies across occupations to allow for occupation-specific job descriptions and job experience. The text replacing the parts in bold in curly brackets is either determined by the value of random variables (name and country) or the context (immigration year depends on the year we applied for the job, i.e., 2020 or 2021; age is determined by the required length of education and work experience, which in turn is determined by occupation).

Appendix Figure A4. The structure of the cover letter part of a resume.

Dear X,

The advertised job (**[name of job from job advertisement]**) sounds very interesting and fits exactly what I would like to work with in the future.

If I'm going to tell you a little about myself, my name is **{name}** and I'm **{age}** years old. I came to Sweden in **{immigration year}** from **{country}**. I already have experience working as **[job]**. My last job was as **[type of job]** at **[previous firm]**. My job tasks were **[job description 1]**. My employer had high ambitions, which meant that I developed a lot. I gained good knowledge of the job and have learned to take responsibility for further developing my knowledge. Before I came to Sweden, I worked with similar job tasks in **{country}**.

My experience has given me a great knowledge in dealing with customers in a positive and representative way. Regardless of my workplace, I am keen to do a good job where I am. I see a challenge in all types of jobs.

As a person, I am social, cooperative, careful and used to a high work rate. I am calm and in control of what I have to do. Together with my partner, I like to spend my free time exercising and spending time with our many friends.

I will be happy to come and introduce myself in more detail during an interview. Then I will of course bring my grades and references from previous jobs.

Best regards,
{name}

Notes: Own translation from Swedish. See the text in Section 3.1 for the explanations of the parts in bold in square and curly brackets.

Appendix Figure A5. Original cover letter in Swedish.

Hej,

Det utannonserade jobbet (**[name of job from job advertisement]**) på Platsbanken låter mycket intressant och passar precis vad jag skulle vilja jobba med framöver.

Om jag ska berätta lite om mig själv så heter jag **{name}** och är **{age}** år. Jag kom till Sverige **{immigration year}** från **{country}**. Sedan tidigare har jag erfarenhet av att jobba som **[job]**. Mitt senaste jobb var **[type of job]** på **[previous firm]**. Mina arbetsuppgifter var **[job description 1]**. Min arbetsgivare hade höga ambitioner, vilket gjorde att jag utvecklades mycket. Jag fick goda kunskaper om jobbet och har lärt mig att ta ansvar för att vidareutveckla mina kunskaper. Innan jag kom till Sverige arbetade jag med liknande arbetsuppgifter i **{country}**.

Min erfarenhet har gett mig en stor vana av att bemöta kunder på ett positivt och representativt sätt. Oavsett min arbetsplats så är jag mån om att göra ett bra jobb där jag är. Jag ser en utmaning i alla typer av jobb.

Som person är jag social, samarbetsvillig, noggrann och van vid en hög arbetstakt. Jag är lugn och har kontroll på det jag ska göra. Tillsammans med min sambo ägnar jag gärna fritiden åt att motionera och umgås med våra många vänner.

Jag kommer gärna och presenterar mig närmare vid en intervju. Då tar jag givetvis med mina betyg och referenser från tidigare jobb.

Vänlig hälsning
{name}

Notes: See the text in Section 3.1 for the explanations of the parts in bold in square and curly brackets.

The CV part of the standardized resume template is a skeleton empty of content which determine the structure of the CV, i.e., where in the document to put the information determined by the random variables and the context. Appendix Figure A6 shows an example for accounting clerks of how a complete CV looks after the values of the random variables (e.g., name) and the context (e.g., occupation and city) is known.

Appendix Figure A6. Example of the CV part of a resume sent to an accounting clerk job.

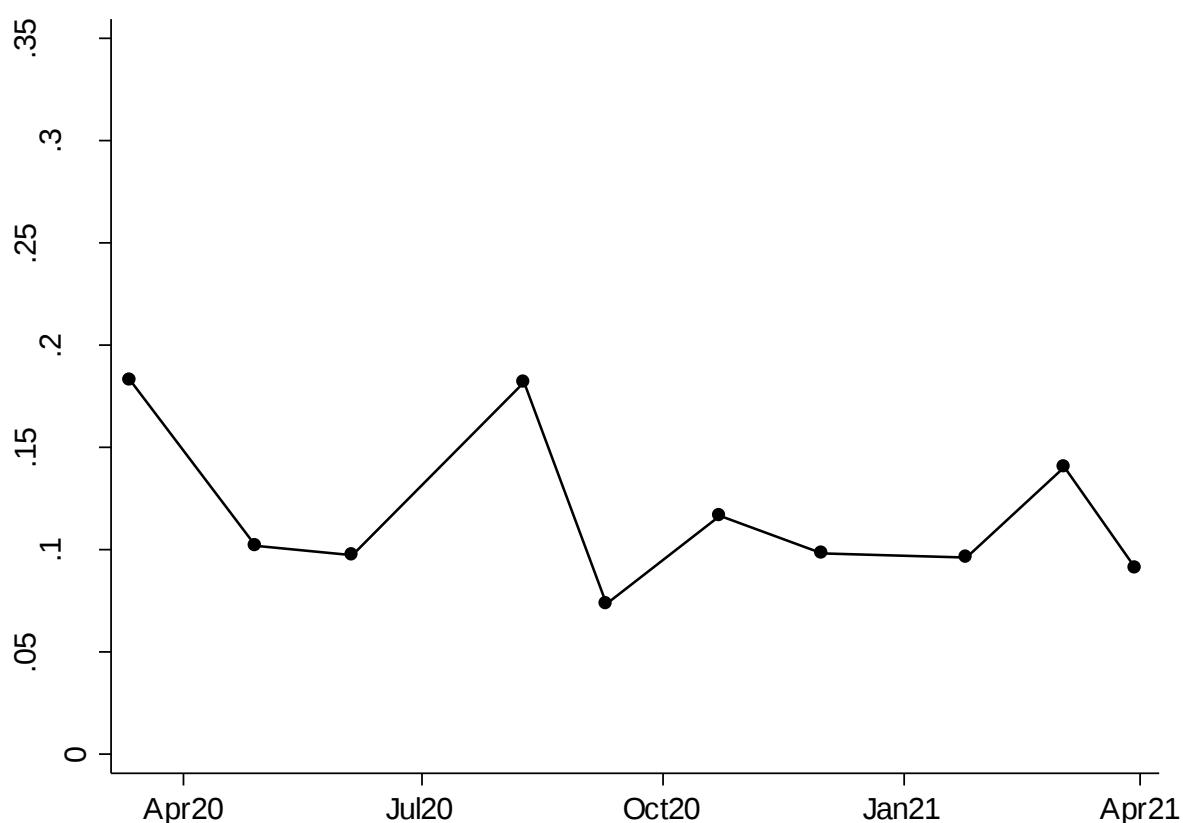
List of qualifications	
<i>Personal information:</i>	
Heidi Müller Slättåkragränd XX, 12572 Älvsjö 072-835 26 XX heidimuller716@XX.com Year of birth: 1995	
<i>Work experience</i>	
2019-02 –	MediaMarkt. Make credit assessments, accountant ledgers, invoice management, etc.
<i>Education</i>	
2002 – 2011	Education in economics (high school), Berlin (Germany)
2011 – 2014	Business Administration (Bachelor's Degree), Freie Universität Berlin, Berlin
<i>Language</i>	
German Swedish (SFI training) English	Mother tongue
<i>Computer knowledge</i>	
Word, Excel, email and the Internet.	
<i>Driving license</i>	
Passenger car	

Notes: Own translation from Swedish.

A5. Data collection during Covid-19

Appendix Figure A7 documents that the callback rate is rather stable during the experiment. However, we trained our first research assistant in the beginning of the experiment and trained our second research assistant at the end of the summer 2020 (a replacement because the first left for a new job). When we trained the research assistants, we could not immediately apply for jobs in all 17 occupations. This led to temporary compositional changes in which occupations we applied for jobs in and explains the somewhat different (higher) callback rate in the beginning of the experiment and at the end of the summer 2020.

Appendix Figure A7. Callback rate over time in field experiment.



A6. Occupational groupings

Appendix Table A2. Occupational groupings.

	Occupations historically requiring less Swedish	Immigrant dense	Female dominated
Panel A: Low-skill			
Cleaners	x	x	
Customer service workers			x
Shop sales assistants			x
Waiters/waitresses		x	x
Warehouse staff workers	x	x	
Panel B: Medium-skill			
Accounting clerks			x
Chefs	x	x	
HR assistants			x
Mechanics	x	x	
Office assistants			x
Personal assistants		x	x
Preschool teachers		x	x
Truck drivers	x	x	
Panel C: High-skill			
Accountants			x
Business salespersons			
Computer professionals			
IT support technicians			

Notes: These are the groupings of occupations used in Table 4. The first column identifies five occupations which historically have required less communication in Swedish to perform its work tasks. The second and third columns list the occupations where the share of immigrants and females, respectively, are above the medians (15% and 52%).

A7. Additional analysis

Appendix Table A3. The effect of language proficiency on the callback rate, by stating a literacy skill requirement in the job advertisement (or not).

	(1)
Language x literacy requirements mentioned	.0207** (.0088)
Language x literacy requirements not mentioned	.0246*** (.0060)
Literacy requirements mentioned	.0470** (.0190)
p-value (test of equal Language coefficients)	.70

Notes: $N = 3,153$. The regression in the table includes interaction variables between the continuous language variable and dummies for whether requirements on literacy skills are mentioned in the job advertisement. Note that the language variable itself is omitted in the regressions. We use the approach described in Section 2.1 to identify whether requirements on literacy skills are mentioned in a job advertisement. Out of the 3,153 observations in the main sample, we could not identify 43 advertisements in the database at the Swedish Public Employment Service; in the regression, we include a missing indicator and its interaction with the language variable. The regression includes no other variables than those listed in the table. Standard errors in parentheses are robust.

*** $p < .01$, ** $p < .05$, * $p < .1$

