ORIGINAL PAPER

The impact of attractiveness on job opportunities in Italy: a gender field experiment

Giovanni Busetta[1] · Fabio Fiorillo[2] · Giulio Palomba[2]

Abstract

This paper assesses the impact of being attractive and not being native on the gender gap in the opportunity of obtaining a job in Italy. To do so, we propose a field experiment that consists in sending 9680 fictitious curricula vitae to real firms looking for employees. We estimate an Heckit model in order to consider different response from firms and then to calculate the probability to receive a callback. We show that gender gap in opportunity of receiving a callback is a very important issue and this gap is affected by interaction with both attractiveness and not being italian natives, especially for the most qualified jobs.

1 Introduction

In this paper we aim to shed light on the profile of discrimination in the Italian labour market. In particular, we deal with gender discrimination, focusing on how it is shaped by not being Italian or being a physically attractive candidate. For this purpose, we carry out an empirical analysis using a database created ad hoc by sending fictitious curricula vitae (hereafter CVs) to real job openings. In particular, we have sent the same resume with the same skills several times to all the job postings displayed online between September 2011 and august 2012, changing the attached photo or attaching no photo at all; we also sent CVs of Italian candidates with photo

✉ Fabio Fiorillo
    f.fiorillo@univpm.it

1   Department of Economics, Business, Environmental Sciences and Quantitative Methods, Università di Messina, Via T. Cannizzaro, 278, 98122 Messina, Italy

2   Department of Economic and Social Sciences (DISES), Università Politecnica delle Marche, Piazzale Martelli 8, 60121 Ancona, Italy

 Springer

in order to calculate the impact of attractiveness for Italian candidates. In addition, our sample includes CVs of Italian and immigrants candidates without photos. This feature allows us to determine the impact of being immigrants compared to being Italian, regardless of the impact of attractiveness. It is worth to note that the scale of our experiment is larger than that of similar analyses.

The focus of our research is twofold: on the one hand, we consider the differences in the skills of the individuals who are employed and, on the other hand, we analyse the reasons why job recruiters perceive the various candidates differently. Consequently, we will try to answer the following questions: do employers discriminate against women and men more if they are not attractive or if they are foreigners? How does gender-based discrimination interact with that based on attractiveness or nationality in different kind of jobs?

In comparison to previous papers on discrimination in the labour market based on field experiments (for a complete review, see Rich 2014; Baert 2017; Neumark 2018), we build a large database which is unique in several respects. Indeed, we analyse the marginal impact of discrimination on the basis of the joint effect of gender and attractiveness and also that of gender and not being a native Italian. Moreover, while other articles basically analyse only the main characteristics of the job (e.g. hard work or front office), we also collected data on the types of work (managers, technicians, sales, etc.). In this way, we can study the impact of discriminatory variables such as gender, physical appearance and nationality on each type of work. This kind of investigation is uncommon in Italy (Patacchini et al. 2015). Actually, several scholars study gender inequality and gender discrimination in Italy, analysing the gender pay gap with non experimental tools (see, for example, Naticchioni and Ricci 2012; Mussida and Lucarelli 2014), while literature on this topic generally focuses on attractiveness and racial discrimination against the differential salaries (Campos-Soria and Ropero-Garcia 2016). In our paper, we identify the gender gap with the difference in callback rates, since we have sent fictitious CVs to companies that require work. This allows us to study the gender gap in opportunities and not in wages.

One criticism is that several analyses on the influence of attractiveness and nativeness on the creation of a gender gap during the hiring process consist of small samples of students who answered hypothetical questions about hiring decisions. Instead our analysis is based on a much larger sample of real job openings posted by actual employers. The underlying idea is thereby to evaluate whether attractiveness and nativeness interact with gender gap according to the different kind of job.

A second criticism revealed by most of the experimental studies concerns the impossibility for researchers to control for employee qualifications and skills. Conversely, the design of our experiment gives us complete control and observability over candidate backgrounds: since all the applications should fulfill employer requirements, for each job offer, we sent CVs of applicants who are identical in terms of education, work experience, language and computer skills, but they are associated to different names, nationality, sex and pictures (or lack thereof).

Our experiment has never been conducted before in Italy. We chose this country mainly because it is considered one of the main fashion countries in the world where physical appearance has always been considered very important. Moreover,

Schwab (2019) in the Global Competitiveness Report 2019 mentions that Italy is ranked 62th out of 141 countries, in terms of "Female participation in the labour force". It represents an improvement with respect to 2015 (93rd out of 144), but the issue of gender discrimination in Italy still exists. Therefore our investigation may be helpful in explaining the low participation of women in the Italian labour force market. Some contributions confirm that, historically, Italy represents a country in which beauty and attractiveness have always played a relevant role, as underlined by Gundle (1997, 1999): "Feminine beauty has been more discussed, appreciated, represented in art and associated with national, cultural identity in Italy than in any other country".

Finally, in Italy immigration represents a relative recent phenomenon which drove the country from an emigration to a new immigration nation (Bauer et al. 2000). According to ISTAT (2017), in 2016 immigrants were more than five million, reaching 8.3% of the Italian resident population. This makes Italy a perfect country to also study the impact of not being Italian native (from now on we will use the word "native") on the gender gap. In fact, our results show a huge gender gap in Italian labour market that increases when gender interacts with not being natives or with a lack of attractiveness, especially when the job requires a high qualification.

The paper is organized as follows: Sect. 2 reviews the main literature on the topic, while in Sect. 3 we describe the data used in the empirical analysis. Section 4 focuses on the statistical methodology we applied and Sect. 5 presents the main empirical results. Finally, Sect. 6 concludes.

2 Conceptual framework

Our paper is inspired by several contributions about occupational segregation and gender gap. According to Bergmann (1981), "Occupational segregation is the distribution of workers across and within occupations, based upon demographic characteristics, most often gender".

Dolado et al. (2004) and Meulders et al. (2010) claim that occupational segregation in the USA seems lower than in Europe. Within the European context, Dolado et al. (2004) and Campos-Soria and Ropero-Garcia (2016) find that northern countries almost always have been characterized by higher levels of occupational segregation than southern ones. During the past decade northern countries have undergone a desegregating process whereas Southern countries, especially Italy and Spain, have experienced an increase in segregation (Bettio and Verashchagina 2009). According to Simon (2012) and Campos-Soria and Ropero-Garcia (2016), segregation seems to explain a larger share of the gender wage gap, which appears to be substantial in the Italian labour market.

In our paper we follow the approach of gathering experimental data on gender differences in promotion opportunities proposed by Baert et al. (2016): "the research question determine the data to be obtained instead of the data determining the questions that can be asked". Baert (2018), underlined that studies based on self-reported information from employers suffers from at least two methodological limits. First of all, employer's attitude and behaviour may not reflect their actual hiring believes

(see also, Pager and Quillian 2005). Second, they could adapt their answers to the perception of what is socially considered desirable (Azmat and Petrongolo 2014). Baert et al. (2016) overcome the methodological problems coming from laboratory and field experiment using fictitious job applications in order to investigate gender gap in the opportunity to receive a callback, instead of investigating on wage gap.[1] Bygren et al. (2017) and Brandén et al. (2018) use the same methodology: the former contribution tests gender discrimination in Sweden labour market and investigates whether a fatherhood premium or a motherhood penalty exist, while the latter studies the trailing spouse phenomenon. The use of fictitious CVs does not allow researchers to investigate on actual wage offers, but only on the hiring opportunity throughout callback rates. Consequently, we set the following hypothesis.

Hypothesis 1 Gender gap in opportunities depends on the characteristics of workers, but also on characteristics of the industries.

The second part of the hypothesis is common to studies that does not use field experiment and investigates on wage differences. For example, Meyersson et al. (2001), Bayard et al. (2003) and Korkeämaki and Kyyrä (2006) suggest that gender gap could arise even if there is no strong wage difference within each job, because female dominated jobs are in lower paying industries than male dominated ones (see also, Campos-Soria and Ropero-Garcia 2016).

Focusing on the analysis of gender gap regardless to the impact of the most male and female dominated industries, would produce an underestimation of gender inequalities in labour market. For instance, economic theory on family migration for job reasons suggests that couples move for the sake of the man rather than the woman because investing in the man's career produce benefits for the couple, or the family, as a whole (Mincer 1978). This makes a greater level of segregation for jobs that imply a higher level of migration, and it depends on the characteristics of the job (Brandén et al. 2018). For the same reason fathers are usually advantaged over mothers in the labour market (see, Charles 2011; Bygren et al. 2017). Moreover, Baert et al. (2016) claim that women receive fewer callback for job interview if they apply for higher positions due to the general distaste of employers, coworkers and customers to collaborate with them. Åslund and Skans (2012) show that removing names from the applications reduce the difference between men and women in terms of callback rates and in terms of discrimination. This is the reason why our analysis of gender gap in the Italian labour market includes the different categories of jobs for which candidates can apply.

Following the pioneering paper of Dion et al. (1972), physically attractive people are perceived to be more sensitive, kind, modest and outgoing. According to Feingold (1992), a robust association for both men and women exists between physical attractiveness and numerous personality traits, such as social skills, mental health

---

[1] For instance, Baert et al. (2016) found that, in the Belgian labour market, women have a significantly lower hiring opportunity than men when they they apply for jobs that imply a promotion. On the contrary, the degree of discrimination seems not to depend on the level of authority of the job.

and intelligence. His main idea is that companies prefer attractive rather than unattractive people because attractive people are considered to be more competent. Hamermesh (2013) discussed the advantages of beautiful people in labour, loans and marriage markets, in sales and in happiness.

Several laboratory experiments have investigated the role of beauty in the labour market (see, for example, Cann et al. 1981; Hamermesh and Biddle 1994): man and attractive candidates are significantly preferred over women and their non attractive counterparts, even after evaluating specific skills. Moreover, the penalty for not being attractive is greater for women than men, and it is robust across occupations. Moreover, the advantages from attractiveness can be gender specific depending on job type (Heilman and Saruwatari 1979; Parrett 2015), years of career (Biddle and Hamermesh 1998). Such preference persists even if physically attractive workers do not appear to be more skillful than less attractive ones (Mobius and Rosenblat 2006).

Field experiments confirms preferences for physical attractiveness. Busetta, Campolo and Panarello (2020) manipulates CV photos digitally in order to have both normal weight and obese applicants, finding evidence of discrimination against the latter group in the Italian labor market. A similar approach was taken previously by Rooth (2009). The analysis starts as an experiment on obesity but it ends up as an experiment on attractiveness and obesity. The results show that both men and women receive significantly lower callback rates in the event of obesity, but also that the results tend to be driven by obesity for women and by attractiveness for men. Ruffle and Shtudiner (2015) respond to advertised job openings in Israel, finding that CVs of women with no picture have a significantly higher callback rate than those of attractive or plain-looking women.[2] Lopez Bóo et al. (2013) find that attractive applicants are called more often than unattractive ones, but unlike Ruffle and Shtudiner (2015), they find stronger effects among male candidates. An Italian field experiment using fictitious CVs is conducted by Patacchini et al. (2015). They study the interaction between homosexuality and physical appearance and find a strong penalty for homosexuals and a beauty premium for females only but this premium is much lower when the "pretty" woman is skilled.

Baert (2018) investigated the impact of the public information about job candidates emerging by the profile picture on Facebook on hiring choices. He found a strong effect of attractiveness, which become even higher when applicants are highly educated and recruiters are female. From the above literature we can formulate the following hypotheses.

Hypothesis 2 Attractiveness matters, there is a beauty premium and a penalty for non attractive employees.

Hypothesis 3 Gender gap and attractiveness interact. The interaction is different among job types.

---

[2] This result is motivated by the female jealousy of attractive women and a negative perception of women (but not men) who include pictures of themselves on their CVs.

Hypothesis 4 The beauty premium and the differences among genders are not related with individual skills.

As we already mentioned, the discrimination carried out on non native applicants is crucial. One of the most important factors to define the success of immigrants integration is their participation to the labour market. In this regard, the situation in Italy does not seem to be favourable. Indeed, recent immigrants to Italy seem to struggle in the labour market, dealing with significantly lower wages (Venturini and Villosio 2008) and higher unemployment rates (ISTAT 2013). Moreover, they are usually forced to find employment in lower skilled occupations, not suitable to their education (Riva and Zanfrini 2013). Their income is about 20% lower than Italians' income, skills and jobs being equal, and this gap doubles to about 40% in terms of family disposable income (Saraceno et al. 2013) when we consider female candidates, which makes discrimination further increasing.

There are many papers similar to our analysis in terms of experimental design on racial discrimination. Bertrand and Mullainathan (2004), for instance, use different names distinctly associated with Whites and African Americans and find large racial differences in callback rates. Applicants with White sounding names need to send about 10 CVs to get one callback, whereas those with African American names must send approximately 15 resumes to receive one callback. This 50% gap in callback rates is statistically significant, and a White name yields as many more callbacks as an additional eight years of experience. Since applicants names are randomly assigned, this gap can only be attributed to name manipulation. Carlsson and Rooth (2007) adopt a similar methodology in studying the situation of Arabs in Sweden. They find large differences across occupations, with callback rates varying from 10% for computer professionals to over 100% in the case of shop sales and cleaning. Drydakis and Vlassis (2010) analyse the labour market opportunities of Albanians in Greece and conclude that Albanians face a 43% smaller chance of access to occupations, and also a significantly lower level of insurance coverage. Wood et al. (2009) conduct a correspondence test in Britain, finding that there are considerable gaps in callbacks between whites and several different ethnical groups. Oreopoulos (2011) analyses response to online job postings in Toronto to investigate why immigrants struggle in the labour market. He finds substantial discrimination against applicants with Indian, Pakistani, Chinese, and Greek names compared with English names. Busetta et al. (2018) form a database of ficious CVs of first- and second-generation immigrants and find that ethnic and gender discrimination in the Italian labour market is significant.

Hypothesis 5 Racial differences matters in terms of labour discrimination. Discrimination is stronger in high specialized jobs than in low ones.

It is worth to note that, previous papers find changes of the gender gap connected to the interaction either with attractiveness or with nativeness, but none

of them investigates on the impact of both attractiveness and nativeness on the gender gap at the same time as we do.

3 Experimental design

In literature, conventional labour force and household surveys collect data which can not be easily used to measure discrimination or to analyse its mechanisms. This is because they do not contain all the characteristics that employers observe when hiring, promoting or setting wages. The difficulty in using conventional data has led many scholars to use one of the two main field experiment techniques, namely audit testing (situation testing in the UK), and correspondence testing. In the correspondence testing (see, Baert 2017, for an overview of these experiments), one member of the minority and one of the majority group apply for the same jobs in order to check for discrimination. This is different from audit testing which does not involve individual testers. In this second case, pairs of written applications are sent to job openings, making the applications similar in all relevant aspects, but the one to be tested (Bursell 2007).

The first strand of the literature, such as Goldin and Rouse (2000) who examine the effect of blind auditioning on the hiring process for selecting musicians for symphony orchestras, measures the amount of gender discrimination. Many other contributions carry out an audit technique (see, among others, Bertrand and Mullainathan 2004; Carlsson and Rooth 2007; Drydakis and Vlassis 2010; Wood et al. 2009; Rooth 2009; Ruffle and Shtudiner 2015; Lopez Bóo et al. 2013). The weakness of this method lies in three main aspects. First, even if attempts are made to match auditors on several characteristics and to train them for several days, it is not always possible to eliminate all the differences between auditors (see, Heckman and Siegelman 1993; Heckman 1998). Second, such methods are not double-blind: as auditors know the purpose of the study, they could behave in such a way as to influence data either in favor or against the existence of discrimination (Turner et al. 1991). Third, audit methods are extremely expensive, therefore it is difficult to generate samples large enough to avoid significant differences in results between couples. Given these weaknesses, we assess the impact of gender, attractiveness and not being native on job opportunities, following Baert et al. (2016). Hence, we carry out a field experiment based on the correspondence technique by sending 9680 European format[3] CVs to 1210 firms looking for employees.

The Italian labour market represents an adequate context for our research, especially regarding the impact of physical attributes. Since the Italian habit is to send CVs without photos, this lack should not penalize applicants because it does not necessarily indicate that they are bad-looking. However, lately many websites suggest

---

[3] European format or "Europass" is an European Union (Directorate General for Education and Culture) initiative to increase transparency of qualification and mobility of citizens in Europe (Decision 2241/2004/EC, Article 1). It aims to make a person's skills and qualifications clearly understood throughout Europe (including the European Union, European Economic Area and EU candidate countries).

to include the photo, a request motivated by the use of a professional dress code. Thus, even in the presence of the picture there should not be any kind of penalty.

We sent all the resumes in the period between September 2011 and August 2012. At that time, the use of social networks was not yet widespread, so the probability that false profiles were discovered by potential employers was very low. We regularly scrutinized job postings on all the main online job service websites[4] offering positions in Italy. We chose only the websites that require no registration in order to prevent firms to detect that CVs in question were fictitious.

3.1 CVs sent to employers

As our goal is to obtain as many responses as possible from employers, we included in the CVs all the characteristics required by the advertised job postings so that the applicants would not be perceived as overqualified. Using this procedure, we sent eight CVs to each advertised job posting,[5] identical in every respect except name, surname, nationality and photo (or lack thereof). In this way, we intended to polarize the results focusing on the effect of attractiveness, nativeness and gender, regardless to other differences in candidate's profile. Thus, all applicants for the same vacancy have the same characteristics in terms of age (28 years old), education,[6] and amount of work experience. The only other difference in the CVs was font and font size as in Rooth (2009).

Consequently, each employer received 8 CVs from 4 females and 4 males,[7] of 28 years old each one, living in Rome, and with exactly the characteristics required for the job. We randomly matched first name, surname and photos. Four CVs contain photos of applicants representing attractive and unattractive Italian women, and attractive and unattractive Italian men respectively. The remaining four CVs have no photos attached and concern Italian, and North African women and men. These CVs do not include photos because we are not interested in evaluating the impact of

---

[4] Namely lavoratorio.it, Lavoro & Stage, Miojob, Lavorare.net, Page Personnel, Trovalavoro, Kijiji, Inique Agenzia, Archimede agenzia per il lavoro, Manpower divisione Horeca, Combinazioni s.r.l, Quanta agenzia per il lavoro, Humangest, Alma, Orienta agenzia per il lavoro, Varese centro per l'impiego, Adecco, Obiettivo lavoro, Temporary agenzia per il lavoro, Free work, Maw, Euro Interim, Mr Comunication, and Open Job.

[5] As a precautionary measure, in order not to let employers realize that they were receiving identical CVs, we staggered the dispatch of the CVs to the same firms over a few days. As each firm receives thousands of CVs, we are convinced that receiving eight CVs over few days should not make them suspicious, neither loosing too much time in relative terms. For the same reason, we used different names and addresses. All the addresses belong to the city of Rome in order not to make the scrutinizers perceive the candidates as different because of where they lived. Finally, we randomly chose the order of CVs sent to the same firm.

[6] To prevent the scrutinizers being influenced by the prestige of the school or the university in which the applicant had studied, we used institutions considered comparable. For the CVs with university degree, we used "La Sapienza" University, the largest university in Italy and located in Rome.

[7] In this context, the best experimental design would be to send applications with identical information to the same employer, except for the photo. As pointed out by Oreopoulos (2011), such a strategy would be impossible to implement without employers becoming suspicious. Therefore, we decided to associate a different name and address to each different photo (or no photo included).

Table 1 Ranking of candidate's photos

id photo	1	2	3	4	5	6	7	8	9	10
Females scores	858	767	752	695	651	632	598	572	543	489
Males scores	926	842	820	772	751	712	702	651	632	618
id photo	11	12	13	14	15	16	17	18	19	20
Females scores	452	432	421	382	364	332	310	296	274	223
Males scores	600	576	534	515	472	437	382	320	285	247

attractiveness on immigrants opportunities to be called for a job interview; in other words, we aim to use them as control variables.

The design of our experiment gives us a complete control over candidates' backgrounds. This is in line with, as previously illustrated by Ruffle and Shtudiner (2015) and Lopez Bóo et al. (2013). Indeed, our applicants are very similar in every respect for each kind of job offer as long as the differences only concern gender, nativeness, photo and attractiveness. This methodology ensures that perceived productivity characteristics on the supply side are held constant.

The degree of differential treatment can be noticed from the number of callbacks for a job interview. Differences in response rates between candidates can only be due to different pictures or lack thereof. Responses are considered callbacks if the employers invite an applicant to a job interview. In order to minimise the inconvenience of an interview, we promptly declined any invitation via email, addressing the reason of a previous acceptance of another position.

3.2 Ranking photos by attractiveness

In choosing the pictures to be included in the resumes, we selected 20 male and 20 female photographs from the internet, modified in order to make them unrecognizable. All faces used in the photos are caucasian smiling individuals, thus eliminating racial preferences. One hundred students (50 women and 50 men) from the University of Messina were invited to evaluate the CV photos using a score between 1 and 10. We then calculated the total score obtained by each photo, and identified unattractive/attractive males and females as those obtaining the lowest/highest scores (see Table 1). In order to test the robustness of the ranking, we calculated how many students assigned the maximum to the most attractive male and female photos. The inverse procedure was followed for the less attractive photos.[8] Overall, the most attractive woman has received a score of 858, while the least attractive 247, the most attractive man received 926 points and the least attractive 223. Since a large majority of students (over 85%) agreed on who is the most and least attractive among women and men, we assume that our rankings are not influenced by any subjectivity.

---

[8] A different strategy would be to send to each firm a multiple of eight CVs. This strategy would have the advantage of increasing the size of the sample population, but it has a great disadvantage: a larger number of photos would have generated a subjective beauty ranking among each employer's responses without having the possibility to control for it.

Table 2 Names, surnames and photos chosen for the profiles

	Name	Surname	Gender	Photo	Nationality
1	Giulia	Rossi	Female	Attractive	Italian
2	Chiara	Mancini	Female	Unattractive	Italian
3	Francesca	Ricci	Female	No photo	Italian
4	Luca	Proietti	Male	Attractive	Italian
5	Francesco	Bianchi	Male	Unattractive	Italian
6	Marco	Russo	Male	No photo	Italian
7	Fatima	Benkeran	Female	No photo	Moroccan
8	Mohammed	Elalawe	Male	No photo	Moroccan

3.3 Assigning attributes to candidates

Unlike the procedure applied by Rooth (2009), once we made the association between names, surnames, address and photos, we kept it for the entire experiment. In practice, we maintain consistency in the application, since we can not exclude that the same vacancy may be present in different job offers on the web. Accordingly, whatever is the distortion produced, it will affect each job offer in the same way.[9]

In order to minimize the effects of differences in names influencing our results, we chose the most common first names and surnames in Rome, home to the eight fictitious applicants[10] (see Table 2).

As regards names, surnames and addresses of the fictitious foreign candidates, we apply the same procedure. Since we aim to assess the impact of nationality on the probability of obtaining a job interview, we assign to foreign candidates the most common non white nationality of immigrants in Italy, which is Moroccan. We made this choice also because Moroccan are recognizable as not being Italian, which make them potentially subjects to higher levels of discrimination based on physical appearance or on religious reason. The foreign applicants have the most common surnames in Morocco, namely Elalawe and Benkeran; the chosen names are Mohammed for males and Fatima for females. We created eight email accounts,

---

[9] The web search engines do not make any selection of the candidates, since they are just as a repository for the CVs. Moreover, in several job postings, there is no information on the single firm offering vacancies because the web search engine hides this information.

[10] We chose the six most popular surnames in Rome: Rossi (2644 families), Mancini (1676), Proietti (1399), Ricci (1369), Russo (1116), and Bianchi (887). We excluded De Angelis (1437), De Santis (1294) and Conti (982) since their surnames could be perceived as noble. In terms of first names we adopted the same procedure both for females and males. The most common first names in Rome for males are Luca, Marco and Francesco and for females they are Giulia, Chiara and Francesca. We excluded Andrea because this name in Italy can be used either for males or females and we did not wish to generate any kind of confusion.

one for each fictitious individual, to collect employer responses and we included this email address in the CV as contact information.[11]

Basically, we customize the CVs sent for responding to each job posting. In practice, we merge each association among name, surname, address and photo, with the personal characteristics in terms of education, work experience, language and computer abilities which completely fulfill the skills required by the job. This design strategy should prevent any matching problems by making each candidate comparable with the other applicants.

4 Data and methodology

4.1 The data

Table 3 summarizes the characteristics of job openings in our dataset. The variable names in parentheses are those that will be used in our analysis. The target population is equally divided into CVs which include photos of an attractive (variable A) and an unattractive (U) person, CVs including no photo of an Italian and a foreigner (F), and CVs for men (M) and for women (W).

Sending CVs with no photos allows us to consider as a benchmark the Italian individuals with no information on their attractiveness and to control for discrimination based on not being Italian. Moreover, our design strategy of sending fictitious CVs that exactly meet the firms' requirements should exclude any difference in the rate of response, due to matching problems among candidates.

Being aware that beauty might be relevant and contribute to worker productivity in some of the fields, we divide job positions into front and back office tasks. Thus, we classify all job openings according to whether the position involves face-to-face contact with the public. In particular, we define as front office jobs (FO) those which either explicitly state that the job requires face-to-face contact with people, or where such contact could be unequivocally inferred from the job advertisement. Otherwise, the job is classified as back office. For instance, we include jobs belonging to fields like sales and customer service in the first category. By contrast, in the back office category we put jobs like accounts management, budgeting, industrial engineering, and computer programming. As for front and back office jobs, we define as hard work (HW) jobs for which physical strength is explicitly required, or those for which it may be unequivocally inferred. Otherwise, they are classified as jobs which do not imply hard work.

In our analysis we also consider that the job offers may a high school diploma (High), a university degree (Grad) or may not require any qualification. In terms of functions offered, we follow the International Standard Classification of Occupations (ISCO) and the target population of job offers is divided into the

---

[11] Almost all the job postings require the inclusion of an email address and/or telephone number in the application form. We decided to include only the email address in the applications in order to make the collection of callbacks easier.

Table 3 Summary statistics

	CVs sent		Callbacks	
	9680	–	2711	28.006%
Candidate characteristics				
Picture				
Attractive Italian (A)	2420	25.000%	1208	49.917%
Unattractive Italian (U)	2420	25.000%	327	13.512%
Foreigner (F)	2420	25.000%	257	10.620%
No photo Italian	2420	25.000%	919	37.975%
Gender				
Women (W)	4840	50.000%	1311	27.087%
Men (M)	4840	50.000%	1400	28.926%
Job characteristics				
Public				
Front office (FO)	3832	39.587%	1021	26.644%
Back office	5848	60.413%	1690	28.899%
Strength				
Hard work (HW)	1304	13.471%	501	38.420%
Soft work	8376	86.529%	2210	26.385%
Qualification				
Graduation (Grad)	2728	28.182%	628	23.021%
High school (High)	4288	44.298%	1188	27.705%
No qualification	2664	27.521%	895	33.596%
Function offered				
Managers	544	5.620%	87	15.993%
Professionals	880	9.091%	229	26.023%
Technicians	3888	40.165%	1030	26.492%
Clerical jobs	1344	13.884%	426	31.696%
Sales workers	1440	14.876%	360	25.000%
Services workers	440	4.545%	207	47.045%
Skilled and craft workers machine and plant operators	752	7.769%	256	34.043%
Elementary occupation	392	4.050%	145	36.990%

following types of work: managers, professionals, technical jobs, clerical jobs, commercials, skilled and elementary occupations.

Table 3 also reports the distribution of callback rates. Focusing on attractiveness, in our sample the attractive Italian people have a callback rate of almost 50%, while unattractive applicants and Italians with no photo reach 13.512% and 37.975% respectively. Discrimination on not being Italian also appears to be significant, if we consider the markedly lower callback rates of 10.620% associated to foreign candidates. Furthermore, men get 28.926% of callbacks, while women 27.087%.

The callback rate for back office jobs is about 29% and that for front office jobs is 26.644%; 26.385% is the callback rate for jobs involving hard work, and 38.420% for those not entailing hard work. In terms of qualifications required, we get the highest callback rate of 33.596% for jobs that do not require any qualification, while jobs for graduate candidates and jobs for high school diploma candidates reach 27.705% and about 23% respectively (see Table 3). In terms of the ISCO classification, we have approximately 47% for services, about 37% for elementary occupations, 26.492% for wire workers, about 34% for craftsmen and skilled workers, and definitely lower callback rates for managers, and scientific and intellectual professions (around 16% and 26%).

Table 4 shows the correlations between job characteristics and classification. The matrix highlights that graduate jobs are negatively correlated to sales, front office and hard work jobs, while in case of executive and specialized jobs they are strongly positively correlated. Obviously, vacancies requiring high school qualifications are strongly negatively correlated to those requiring university degrees. Front office jobs are highly positively correlated to sales staff, and finally hard work is strongly positively correlated to service work, workmen and to unskilled work, and negatively correlated to technical jobs.

4.2 The model

As we mentioned in Sect. 3, we sent 9680 CVs to 1210 advertised job openings. Since we received 2711 callbacks that correspond approximately to the rate of 30%, our proposed analysis aims to inquire which are the principal attributes that affect the probability of obtaining a job interview. The dependent variable is the dichotomic variable `RESP`), which represents the employers responses; it is equal to 1 if the employer emailed the applicant to invite him/her for an interview, and 0 if the email was not sent. Since 312 firms did not reply at any CV, we must also consider a distinction between responding and non responding employers in our analysis. From the statistical point of view, the possibility that firms can not reply implies a problem of selection bias which can arise from censoring. In order to address this problem, we use the so-called Heckit model (Heckman 1979); this method was introduced to correct the selection bias occurring in nonrandomly selected samples and provides consistent estimates which eliminate the specification error in the case of censored data.

Analytically, our Heckit model is

$$\begin{cases} s_i^* = \mathbf{Z}_i'\boldsymbol{\gamma} + \eta_i & i = 1, 2, \ldots, N \\ \text{RESP}_i = \mathbf{X}_i'\boldsymbol{\beta} + \varepsilon_i & \Leftrightarrow s_i^* > 0, \end{cases} \tag{1}$$

where N is the sample size. The first equation is the "selection equation", where s_i^* is an unobservable latent variable that is positive if the employer replies to the applicant, $\mathbf{Z}_i'$ is the $1 \times m$ containing all the characteristics that determine whether the reply is made or not, while $\boldsymbol{\gamma}$ is a vector of unknown parameters. The second equation is the "principal equation" that consists of the linear model of interest, where RESP_i is the dependent variable, $\mathbf{X}_i$ is an k-dimensional vector of exogenous variables, and $\boldsymbol{\beta}$ is a vector of unknown parameters. The potential sample

Table 4 Correlation between job characteristics and classification The sample size is $N = 9680$ CVs sent to employers

	Profess.	Techn.	Clerical	Sales	Services	Skilled	Element.	Front office	Hard work	Grad.	High school
Managers	−0.077	−0.200	−0.098	−0.102	−0.053	−0.060	−0.050	−0.190	−0.096	0.389	−0.218
Professionals		−0.259	−0.127	−0.132	−0.069	−0.078	−0.065	−0.109	−0.125	0.505	−0.282
Technicians			−0.329	−0.342	−0.179	−0.203	−0.168	0.016	−0.299	−0.026	0.294
Clerical				−0.168	−0.089	−0.099	−0.082	−0.110	−0.158	−0.081	0.089
Sales					−0.091	−0.104	−0.086	0.516	−0.002	−0.262	0.141
Services						0.881	−0.045	−0.177	0.425	−0.128	−0.163
Skilled							−0.051	−0.193	0.421	−0.147	−0.192
Elementary								−0.055	0.496	−0.129	−0.183
Front office									−0.082	−0.334	0.207
Hard work										−0.247	−0.352
Graduation											−0.559

selection bias in this equation has been corrected by inserting the inverse Mills ratio $\lambda_i = \phi(\mathbf{Z}_i\boldsymbol{\gamma})/\Phi(\mathbf{Z}_i\boldsymbol{\gamma})$ as additional regressor, where $\phi(\cdot)$ and $\Phi(\cdot)$ are respectively the density and the cumulative density of the standardised Gaussian. The explanatory variables in $\mathbf{X}_i$ could also be included in $\mathbf{Z}_i$ and viceversa.[12] Moreover, we assume that the random disturbances ε_i and η_i are i.i.d. and jointly distributed as a multinormal random variable with zero mean and a full covariance matrix. When the covariance between disturbances is non-zero, OLS estimation yields biased and inconsistent estimates of β (see, Heckman 1979).

In our proposed model, selection bias arises because the callback RESP_i is observed only when the ith CV is sent to those employers who reply at least one time. In this case we observe the variable $\text{REPLY}_i = 1$ as the approximation for the latent variable s_i^*; the model is estimated by maximum likelihood (ML) with robust standard errors, which jointly returns consistent and asyptotically normal estimates of both equations in the system (1).

Table 5 contains some descriptive statistics regarding the number of no reply on respect to ISCO classification and Italian regions.

There is no regularity among regions, while the percentage of no reply is significantly higher when firms are looking for a managerial position, probably because it is a more competitive sector. Therefore, our analysis consists of estimating three different heckit models based on three specifications for the vector $\mathbf{X}_i$. For all models the benchmark applicant is Italian and applies in Lazio (whose capital is Rome), sending a CV with no photo attached. Moreover, the selection equation about the probability of receiving an email with the invitation for an interview is always estimated via a probit model where the row vector of the explanatory variables is partitioned as

$$\mathbf{Z}_i' = \begin{bmatrix} \mathbf{p}_i' & \mathbf{q}_i' & \mathbf{jad}_i' & \mathbf{reg}_i' \end{bmatrix}, \tag{2}$$

where

1. $\mathbf{p}_i' = [\texttt{Clerical}_i\ \texttt{Elementary}_i\ \texttt{Managers}_i\ \texttt{Professionals}_i\ \texttt{Sales}_i\ \texttt{Services}_i\ \texttt{Skilled}_i]$ contains the job sectors based on ISCO classification. The reference sector is represented by the variable $\texttt{Technicians}_i$ which is excluded to avoid exact collinearity,
2. $\mathbf{q}_i' = \begin{bmatrix} \texttt{FO}_i & \texttt{HW}_i & \texttt{Grad}_i & \texttt{High}_i \end{bmatrix}$ is composed by the job characteristics (variables $\texttt{FO}_i$ for front office and $\texttt{HW}_i$ for hard work) and the education level (variables $\texttt{Grad}_i$ for graduation and $\texttt{High}_i$ for high school diploma) dummies,
3. $\mathbf{jad}_i'$ and $\mathbf{reg}_i'$ are sets of controls that include the job advertisement dummies (see, for instance, Riach and Rich 2002) and the regional dummies; the variables in the first partition take the value 1 if the vacancy is posted by a given job search

---

[12] It is well known that vectors $\mathbf{X}_i$ and $\mathbf{Z}_i$ can not have many variables in common, since λ_i may show an approximately linear dependency with respect to both of them, so the model will be under-identified for all practical purposes. Puhani (2000) shows that in this case the Heckit model suffers from some problems of collinearity. We adopt the most general solution which consists of dropping from the equation of RESP_i those variables that show good predicting power in the selection equation.

Table 5 Percentages of firms with no reply (CVs sent in parenthesis)

	Managers	Professionals	Technicians	Clerical	Sales	Services	Skilled	Elementary	TOT.
Aosta V.			37.5% (64)	25% (32)	– (24)	– (16)	33.33% (24)	50% (16)	27.27% (176)
Piedmont	62.5% (64)	50% (48)	35.14% (296)	22.22% (72)	– (72)	100% (8)	16.67% (48)	100% (8)	33.77% (616)
Lombardy	52.63% (152)	20% (160)	32.69% (832)	50% (192)	35.29% (136)	37.5% (64)	44.44% (72)	– (40)	35.44% (1648)
Trentino A.A.	50% (16)	– (32)	27.27% (88)	20% (40)	25% (32)	– (24)	25% (32)	– (16)	20.00% (280)
Veneto	28.57% (56)	36.84% (152)	20.45% (352)	22.22% (144)	20% (120)	33.33% (48)	22.22% (72)	16.67% (48)	24.19% (992)
Friuli V.G.		25% (32)	30.77% (104)	– (24)	– (24)	– (8)	– (8)		20.00% (200)
Ligura	– (16)		33.33% (48)	33.33% (24)	20% (40)	– (56)	33.33% (24)	– (16)	17.86% (224)
Emilia Rom.	– (16)	40% (40)	13.04% (184)	– (24)	20% (80)	100% (8)	50% (16)	42.86% (56)	22.64% (424)
Tuscany	50% (32)	27.78% (144)	28.89% (360)	19.05% (168)	12.5% (128)	60% (40)	25% (64)	33.33% (24)	26.67% (960)
Umbria		– (24)	11.11% (72)	– (40)	– (48)		– (32)	– (8)	3.57% (224)
Marche	36.36% (88)	28.57% (112)	28.24% (680)	25% (288)	28.95% (304)	30% (80)	18.75% (128)	– (40)	26.98% (1720)

Table 5 (continued)

	Managers	Professionals	Technicians	Clerical	Sales	Services	Skilled	Elementary	TOT.
Lazio	—(8)	66.67%(24)	25%(64)	14.29%(56)	—(24)			—(16)	20.83%(192)
Abruzzo	—(8)		60%(40)	50%(16)			—(16)	—(8)	36.36%(88)
Molise	100%(16)	100%(8)	66.67%(24)	—(16)	—(24)	—(8)	—(8)		38.46%(104)
Campania	—(16)		11.76%(136)	20%(40)	33.33%(48)	—(16)	—(72)	—(24)	11.36%(352)
Apulia	50%(16)	25%(32)	33.33%(96)	—(16)	12.5%(64)	—(8)	33.33%(24)	33.33%(24)	25.71%(280)
Basilicata	—(8)		—(32)	50%(16)	—(8)		33.33%(24)		18.18%(88)
Calabria	—(16)	20%(40)	17.86%(224)	12.5%(64)	13.33%(120)	—(40)	50%(48)	—(24)	16.67%(576)
Sicily	100%(16)	25%(32)	25%(160)	12.5%(64)	8.33%(96)	50%(16)	33.33%(24)	33.33%(24)	24.07%(432)
Sardinia			—(32)	—(8)	16.67%(48)		50%(16)		15.38%(104)
TOT.	42.65%(544)	29.09%(880)	27.16%(3888)	23.81%(1344)	18.33%(1440)	25.45%(440)	24.47%(752)	18.37%(392)	25.79%(9680)

engine and 0 otherwise, while the variables included in the second partition take the value 1 if the job vacancy comes from a given Italian region and 0 otherwise. Some preliminary estimations led us to exclude the job advertisement dummies in **jad**$_i$ because they are collinear with the explanatory variables contained in vectors $\mathbf{p}_i$ and $\mathbf{q}_i$.

4.2.1 Preliminary analysis on the gender effect

In Table 3, the percentages concerning the callback rates for females and males are quite similar. In order to asses whether men and women have different chances of being hired or discriminated against, we estimate the probability to be called back as an exclusive function of the gender. We call this basic specification as the "Model 0" which represents the starting point of our empirical analysis. Its formal definition is

$$\textbf{Model}\,0 : \quad \mathbf{X}'_i = \begin{bmatrix} \mathtt{W}'_i & \lambda_i \end{bmatrix}, \tag{3}$$

where the explanatory variable $\mathtt{W}_i$ is a scalar dummy variable for women. The complementary dummy $\mathtt{M}_i$ is omitted to prevent exact collinearity.

The estimation results in Table 6 represent a raw evaluation of the gender effect because it does not take into accounts other available regressors. Focusing on the principal equation, at a first glance the negative and significant coefficient of $\mathtt{W}_i$ can be interpreted as a disadvantage due to being a woman compared to being a man on the job market. The use of the Heckit approach is justified by the statistical significance of estimated coefficient of the Mill's Ratio λ_i.

4.2.2 Models on attractiveness and nationality with gender effect

Given the task of estimating the impact of attractiveness and nativeness together with the gender effect on the probability to be called back, we propose three different Heckit specifications which test our five hypotheses. The first specification is the model that estimates the general effect of beauty on the probability to be called back. Its equation is

$$\textbf{Model}\,1 : \quad \mathbf{X}'_i = \begin{bmatrix} \mathbf{a}'_i & \lambda_i \end{bmatrix}, \tag{4}$$

where $\mathbf{a}'_i = \begin{bmatrix} 1 & \mathtt{A}_i & \mathtt{U}_i & \mathtt{F}_i \end{bmatrix}$ includes the constant and the scalar variables $\mathtt{A}_i$, $\mathtt{U}_i$ and $\mathtt{F}_i$ are dummy variables for the attributes already defined in Sect. 3. The corresponding parameters included in $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 & \beta_A & \beta_U & \beta_F & \beta_\lambda \end{bmatrix}'$ are crucial because they measure the difference between the applicant's feature of being attractive, unattractive or foreign and the reference applicant. As we claim before, β_λ is the coefficient associated to the inverse Mills ratio λ_i that indicates if some sample selection bias is operating.

The second specification aims to analyse the interactions between beauty, uglyness and not being Italian attributes and job characteristics. The regressors are

Table 6 Preliminary estimates (Model 0, dep. variable RESP, QML standard errors), $N = 9860$, censored 2496 (25.785%)

Variable	Coeff.	s.e.	Z-stat	p value
Const	0.130	0.007	19.253	0.000***
W	-0.028	0.006	-4.934	0.000***
λ				0.577*** (0.004)
σ				0.578
ρ				0.998
Log-likelihood				-8920.585
Akaike IC				17847.169
Schwarz IC				17867.808
Hannan–Quinn IC				17854.272
CM test				41.349*** [0.000]
LR test for regions				913.231*** [0.000]

Standard errors in parentheses, p values in brackets

IC information criteria, *CM* conditional moment test, *LR* likelihood ratio test

*Statistical significance at the 0.1 level

**Statistical significance at the 0.05 level

***Statistical significance at the 0.01 level

$$\textbf{Model}2: \quad \mathbf{X}'_i = \begin{bmatrix} \mathbf{a}'_i & (\mathbf{q}'_i \otimes \mathbf{a}'_i) & \lambda_i \end{bmatrix}, \tag{5}$$

where the partition $\mathbf{q}'_i \otimes \mathbf{a}'_i$ contains all the interactions between the attributes included in vector $\mathbf{a}_i$ and the components of vector $\mathbf{q}_i$. The symbol "$\otimes$" denotes the Kronecker product.

Finally, the third specification is based on the cross products

$$\textbf{Model}3: \quad \mathbf{X}'_i = \begin{bmatrix} (\mathbf{p}'_{i*} \otimes \mathbf{a}'_i) & \lambda_i \end{bmatrix}, \tag{6}$$

where $\mathbf{p}'_{i*} = \begin{bmatrix} \text{Clerical}_i \ \text{Elementary}_i \ \text{Managers}_i \ \text{Professionals}_i \ \text{SaleServ}_i \ \text{Skilled}_i \end{bmatrix}'$. This vector differs from $\mathbf{p}_i$ already defined in Eq. (2) because it contains the variable $\text{SaleServ}_i = \text{Sales}_i + \text{Services}_i$ that has been built because the use the two addends separately leads to exact collinearity. For the same reason, the vector $\mathbf{a}_i$ is not a partition of $\mathbf{X}_i$.

The model in system (1) does not take into account potential gender discriminations, therefore we split the regressors by defining $\mathbf{w}_i = \text{W}_i \mathbf{X}_i$ and $\mathbf{m}_i = \text{M}_i \mathbf{X}_i$, where W_i is a scalar dummy variable for women and M_i is a scalar dummy variable for men. The resulting model is the one we call "genderified model", whose the principal equation is

$$Pr(\text{RESP}_i = 1) = \mathbf{w}_i \boldsymbol{\beta}_w + \mathbf{m}_i \boldsymbol{\beta}_m + \varepsilon_i. \tag{7}$$

This specification allows us to carry out a battery of Wald tests on the null hypothesis of no gender discrimination given by the element by element equality

$$H_0 : \beta_{w,j} = \beta_{m,j}, \qquad \text{for all } j = 1, 2, \ldots, k, \tag{8}$$

where and k is the dimension of vectors $\boldsymbol{\beta}_w$ and $\boldsymbol{\beta}_m$. It is worth noting that the "genderified model" corresponds to separate estimates for women and for men, so we can consider the complete covariance matrix to perform the test. Therefore, in our analysis we estimate the heckit twice, without and with the genderification, for each of the three models. The genderified version includes the Wald test on the estimated parameters according to the null hypothesis in Eq. (8). This allows us to introduce the concepts of female/male "premium" and "penalty" in being attractive or native.

5 Results

The analysis we performed consists of three heckit models for the dependent variable `RESP`. Following Eq. (1), for each Model 1, 2 and 3, we estimate the heckit twice, without and with the genderification. All estimation results are accompanied by some measures to evaluate the goodness of each model specification. In particular, since $\beta_\lambda = \rho_{\varepsilon\eta}\sigma_\varepsilon$, we estimate values of the correlation between the two errors ($\rho_{\varepsilon\eta}$) and the scale factor given by the standard deviation of ε_i (σ_ε). Moreover, we provide the condition number, in order to control the degree of collinearity among a large number of regressors, and the Akaike, Bayesian and Hannan–Quinn information criteria (AIC, BIC and HQC) to determine which specification best fits the data. We also carry out a likelihood ratio test (LR) about the null hypotheses of no effects generated by the regional control variables. Moreover, we perform a conditional moment (CM) test for normality via the Outer Product Gradients (OPG) regression suggested by Davidson and MacKinnon (1993), since the Heckit estimator could suffer from inconsistency when the normality assumption fails (see, Pagan and Vella 1989, for details).

All the results are reported in Tables 7, 8 and 9 for the principal equation, while Table 10 contains the selection equation estimates. A probit model with the same dependent and regressors of the selection equation is also provided in order to calculate the McFadden R^2, the correct prediction percentages to assess the models' goodness of fit and a normality test on the model residuals. In our estimations some of the explanatory variables among those presented in Table 3 are dropped to avoid exact collinearity, while other variables are excluded to reach the maximum possible reduction of parameters, without losing of any relevant information.

From the statistical perspective the estimation results seem robust among the proposed models. The estimated coefficient $\hat{\beta}_\lambda$ lies around the value of 0.54 and all the related t-statistics highlight that a sample selection mechanism is indeed operating. This result, already obtained for the Model 0 (see Table 6), suggests that the Heckit model is always superior to OLS. In general, all the regression statistics indicate that there are not substantial differences among models. Independently of the model specification and of genderization, the estimated correlation $\hat{\rho}_{\eta,\varepsilon} = 0.998$. Moreover, the condition number is always less than the critical value 30 indicating that our estimates

Table 7 Model 1 (dep. variable `RESP`, QML standard errors), $N = 9860$, censored 2496 (25.785%)

Variable	All	Woman	Man	Wald test
`const`	0.225*** (0.009)	0.210*** (0.008)	0.210*** (0.009)	0.012 [0.912]
A	0.013*** (0.002)	0.019*** (0.003)	0.010*** (0.003)	6.069** [0.014]
U	−0.133*** (0.013)	−0.179*** (0.014)	−0.056*** (0.010)	67.917*** [0.000]
F	−0.165*** (0.012)	−0.133*** (0.015)	−0.164*** (0.014)	3.178* [0.075]
			Non genderized	Genderized
λ			0.555*** (0.004)	0.546*** (0.004)
σ			0.556	0.547
ρ			0.998	0.998
lcondition number			5.794	7.016
Log-likelihood			−8416.312	−8376.333
Akaike IC			16,842.623	16,770.666
Schwarz IC			16,877.021	16,832.582
Hannan-Quinn IC			16,854.461	16,791.973
CM test			0.801 [0.449]	0.530 [0.589]
LR test for regions			94.508*** [0.000]	94.103*** [0.000]

Standard errors in parentheses, *p* values in brackets

IC information criteria, *CM* conditional moment test, *LR* likelihood ratio test

*Statistical significance at the 0.1 level

**Statistical significance at the 0.05 level

***Statistical significance at the 0.01 level

do not suffer from any problems of collinearity[13] and the values of the loglikelihood and all the information criteria are about the same in models. The CM test rejects the null hypothesis only for the non genderized Model 2 and Model 3, making us confident that, at least in all the other cases, the selection problem was satisfactorily dealt with. Finally, the LR tests lead to the conclusion that the regional variables are suitable.

When we consider only the impact of gender on the callback rate of a candidate sending no photo CVs, no gender gap in opportunity appears. On the contrary, when we consider the impact of attractivess and not being Italian, such gap emerges. Attractiveness and not being native matters, but the difference between genders is also significant (Hypotheses 1, 2 and 5).

From the estimated parameter $\hat{\beta}_A$ in Model 1 we observe that attractive people have a higher chance of being contacted by the employers. These results are in line with those obtained by Garner-Moyer (2010) for French labour market. He shows a major difference in callback rates between attractive (42%) and unattractive (16%) candidates during the first stages of the hiring process. As expected, our estimates suggest that there is a female premium in being attractive, because the probability

---

[13] This is confirmed also by a check based on the variance inflation factors (VIF). All the preliminary results are available upon request from the authors.

Table 8 Model 2 (dep. variable `RESP`, QML standard errors), $N = 9860$, censored 2496 (25.785%)

	All	Woman	Man	Wald test
const	0.208*** (0.008)	0.205*** (0.008)	0.211*** (0.008)	4.402** [0.036]
FO_A	0.060*** (0.011)	0.073*** (0.013)	0.060*** (0.012)	3.231* [0.072]
HW_A	0.001 (0.012)	−0.020 (0.015)	0.001 (0.013)	3.786* [0.052]
$Grad_A$	−0.015*** (0.004)	−0.011** (0.005)	−0.019*** (0.005)	1.980 [0.159]
$High_A$	0.002 (0.005)	0.018*** (0.007)	−0.013** (0.006)	15.325*** [0.000]
FO_U	−0.099*** (0.013)	−0.115*** (0.013)	−0.101*** (0.020)	0.399 [0.527]
HW_U	0.040*** (0.013)	0.022* (0.013)	0.094*** (0.024)	11.345*** [0.001]
$Grad_U$	−0.061*** (0.009)	−0.150*** (0.016)	−0.039*** (0.006)	37.921*** [0.000]
$High_U$	−0.101*** (0.014)	−0.101*** (0.017)	−0.087*** (0.020)	0.327 [0.568]
FO_F	−0.029** (0.014)	0.008 (0.015)	−0.082*** (0.016)	17.282*** [0.000]
HW_F	0.024* (0.013)	0.016 (0.013)	0.031** (0.014)	2.786* [0.095]
$Grad_F$	−0.132*** (0.016)	−0.138*** (0.020)	−0.121*** (0.022)	0.386 [0.535]
$High_F$	−0.120*** (0.014)	−0.097*** (0.019)	−0.130*** (0.017)	1.937 [0.164]

			Non genderized	Genderized
λ			0.540*** (0.003)	0.538*** (0.004)
σ			0.541	0.539
ρ			0.998	0.998
condition number			9.380	9.416
Log-likelihood			−8277.257	−8249.345
Akaike IC			16,582.515	16,552.691
Schwarz IC			16,678.829	16,738.440
Hannan–Quinn			16,615.660	16,616.613
CM test			7.678*** [0.001]	1.381 [0.251]
LR test for regions			94.508*** [0.000]	215.429*** [0.000]

Standard errors in parentheses, *p* values in brackets

IC information criteria, *CM* conditional moment test, *LR* likelihood ratio test

*Statistical significance at the 0.1 level

**Statistical significance at the 0.05 level

***Statistical significance at the 0.01 level

of attractive women to be called back is about the double of that of attractive men. This is in line with Hypothesis 2. In general, attractiveness and nativeness influence the gender gap. Since $\hat{\beta}_U < 0$ and $\hat{\beta}_F < 0$, there is a penalty for unattractive and non native people. Indeed, the Wald tests confirm that the lack of attractiveness produce an evident penalty for women, while being non native reduces the probability to receive a callback for men. These confern our Hypotheses 2 and 5.

Table 9 Model 3 (dep. variable `RESP`, QML standard errors), $N = 9860$, censored 2496 (25.785%)

	All	Woman	Man	Wald test
`const`	0.215*** (0.014)	0.200*** (0.012)	0.203*** (0.012)	0.634 [0.426]
`A`	0.007** (0.004)	0.011** (0.005)	0.005 (0.004)	0.974 [0.324]
`U`	−0.148*** (0.018)	−0.154*** (0.018)	−0.042*** (0.010)	37.463*** [0.000]
`F`	−0.153*** (0.018)	−0.086*** (0.021)	−0.105*** (0.020)	0.662 [0.416]
`Managers`	−0.262*** (0.030)	−0.252*** (0.029)	−0.254*** (0.029)	0.045 [0.832]
`Professionals`	−0.044 (0.028)	−0.048* (0.025)	−0.024 (0.028)	4.322** [0.038]
`Clerical`	0.116*** (0.028)	0.110*** (0.031)	0.088*** (0.031)	3.627 [0.057]
`SaleServ`	−0.019 (0.023)	−0.004 (0.022)	−0.052** (0.026)	4.126** [0.042]
`Skilled`	0.210*** (0.038)	0.195*** (0.033)	0.249*** (0.037)	4.468** [0.035]
`Elementary`	0.110*** (0.038)	0.091* (0.051)	0.139*** (0.037)	0.960 [0.327]
`Managers`$_A$	0.010 (0.010)	0.022 (0.016)	0.005 (0.011)	0.800 [0.371]
`Professionals`$_A$	−0.021*** (0.008)	−0.006 (0.009)	−0.023* (0.012)	1.394 [0.238]
`Clerical`$_A$	0.013 (0.010)	0.167* (0.088)	0.008 (0.010)	3.285* [0.070]
`SaleServ`$_A$	0.039*** (0.013)	0.020** (0.009)	0.056** (0.025)	2.018 [0.155]
`Skilled`$_A$	−0.103*** (0.037)	−0.072*** (0.027)	−0.078*** (0.028)	0.029 [0.866]
`Elementary`$_A$	−0.207*** (0.057)	−0.260*** (0.060)	−0.048* (0.028)	10.744*** [0.001]
`Managers`$_U$	0.130*** (0.020)	0.145*** (0.019)	0.042*** (0.012)	25.082*** [0.000]
`Professionals`$_U$	0.113*** (0.021)	0.116*** (0.024)	0.024* (0.014)	11.986*** [0.001]
	All	**Woman**	**Man**	**Wald test**
`Clerical`$_U$	−0.139*** (0.034)	−0.102*** (0.039)	−0.145*** (0.040)	1.373 [0.241]
`SaleServ`$_U$	0.018 (0.026)	0.050* (0.029)	0.006 (0.030)	1.329 [0.249]
`Skilled`$_U$	0.119*** (0.025)	0.095*** (0.027)	0.037 (0.030)	2.227 [0.136]
`Elementary`$_U$	0.160*** (0.022)	0.189*** (0.051)	0.051*** (0.018)	6.728*** [0.009]
`Managers`$_F$	0.122*** (0.022)	0.081*** (0.022)	0.096*** (0.021)	0.342 [0.559]
`Professionals`$_F$	0.024 (0.039)	0.049* (0.026)	0.035 (0.033)	0.150 [0.698]
`Clerical`$_F$	−0.129*** (0.034)	−0.108** (0.045)	−0.139*** (0.037)	0.656 [0.418]
`SaleServ`$_F$	0.044 (0.031)	0.067*** (0.022)	0.032 (0.031)	1.003 [0.317]
`Skilled`$_F$	−0.024 (0.062)	−0.038 (0.039)	0.018 (0.034)	1.274 [0.259]
`Elementary`$_F$	0.168*** (0.022)	0.130** (0.052)	0.105*** (0.024)	0.207 [0.649]
			non genderized	**genderized**
λ			0.544*** (0.005)	0.541*** (0.004)
σ			0.545	0.542
ρ			0.998	0.998
condition number			6.194	7.255
Log-likelihood			-8266.911	-8176.180
Akaike IC			16591.821	16466.360
Schwarz IC			16791.330	16858.498
Hannan–Quinn			16660.479	16601.309

Table 9 (continued)

CM test	50.573*** [0.000]	0.536 [0.584]
LR test for regions	212.656*** [0.000]	89.017*** [0.000]

Standard errors in parentheses, *p* values in brackets

IC information criteria, *CM* conditional moment test, *LR* likelihood ratio test

*Statistical significance at the 0.1 level

**Statistical significance at the 0.05 level

***Statistical significance at the 0.01 level

Once these results were obtained, we proceed to further disentangle such differences in gender, considering both job characteristics and types. (testing Hypothesis 1).

The estimates of Model 2 show that sending a CV without photo increases the probability to be called back for men. As expected, in the case of front office jobs, attractiveness rises this probability and there is a preference of employers towards women. It can be noticed here a remarkable female premium in being not native. In the case of hard works, attractiveness is not a relevant attribute, and the gender gap always operates in favour of men.

Surprisingly, in our estimates all the coefficients of graduation are negative. Moreover, a relevant female penalty of being unattractive is observed. When higher levels of education are considered, our estimates reveal a strong gender difference only in the case of attractive applicants. Having a high school diploma or a degree produces opposite effects on the probability of being called back. Such effect is positive for attractive women, but it is negative for attractive men.

Hence, our analysis is substantial in line with Hypothesis 4: when two CVs contain the same skills, gender differences remain and attractiveness explains such differences.

Overall, our results support the presence of racial discrimination in the Italian labour market. Except in the case of hard works, where no specialisation is required, being a foreigner generally reduces the callback probability. This is consistent with Hypothesis 5.

Finally, Model 3 focuses on the impact of beauty and nativeness on different types of job.

Regarding the technicians (benchmark), the estimated coefficient of U_i highlights only a female penalty in being unattractive, while for attractive and foreign candidates substantial preferences between males and females do not emerge. In this context, positive estimated coefficients are associated only to beauty.

Attractiveness is desirable for all applicants for technical and sales/services job positions, while it reduces the probability of receiving a callback in the case of professional, skilled and elementary activities. The Wald tests on gender differences highlight a female premium in being attractive for clerical jobs, and a female penalty in the case of elementary activities.

In general, being unattractive mostly increases the callback rate for female applicants, and the same features is disadvantageous for clerical jobs. The lack

of attractiveness is likely to produce a gender gap in favour of women when they apply for managerial, professional and elementary activities. All these results confirm Hypothesis 3 according to which the gender gap mainly depends on physical attributes.

Not being native play a relevant role in our results, but relevant differences between genders do not emerge. This suggests that nativeness is a cause of discrimination, but it does not impact the gender gap, consistently with the Hypothesis 5.

6 Concluding remarks

In our analysis we carried out a field experiment based on real job online openings in Italy to ascertain whether gender, attractiveness and nativeness have an impact at the early stage of job search. The sample population that we analysed consists of 9680 CVs sent to firms looking for employers in response to advertised job postings displayed online in the period between September 2011 and August 2012. Comparing the response rates for different categories, we obtained the following results: attractive applicants are those who receive the highest number of positive answers; both unattractive and non Italian native candidates obtained lower callback rates. Attractiveness is quantitatively more important for women than for men. Attractive Italian women have higher callback rates than unattractive or non native women. This discrepancy is greater for women than for men. Generally, there is a female penalty in being unattractive, but this does not apply for all the examined kinds of jobs. It is worth noting that the lack of attractiveness increases the gender gap in favour of men when the interaction between gender and attractiveness is significant. On the other hand, attractiveness reduces this disparity and sometimes an attractive woman has more opportunity to be called back than an attractive man. Most responses to unattractive subjects involve low skilled jobs, which is a clear sign of occupational segregation. Unsurprisingly, attractiveness appears to be essential for front office and executive jobs, while the lack of attractiveness is likely to strongly reduce chances for clerical jobs too. This applies even more for women.

Our estimates suggest that attractive candidates should attach a photograph to their CVs when have the opportunity to do so because an image increases the likelihood of being called for an interview. The opposite is true for unattractive candidates. In other words, it seems that a woman aiming to find a good job in Italy has to be attractive, while an unattractive woman, even if she is highly qualified, has little chance of getting a highly skilled job, at least if she applies online and attaches a photo. These results lead us to conclude that attractiveness increases the gender gap in the labour market, especially in terms of occupational segregation.

We also found that racial discrimination appears to be substantial. However, in the case of women, racial discrimination is less prominent than discrimination based on physical features. Instead, our results highlight a male penalty in being non native because the estimated percentages of callbacks are generally very low, especially for "soft" or highly qualified jobs. Conversely, unattractive or

non native men have a higher probability of receiving a callback if they apply for a hard and poorly qualified work position. In general, non native women seem to be more likely to get a job interview in Italy than non native men; the exceptions are some job categories such as services, elementary and hard work for which men are preferred. On the other hand, non native men are taken into consideration only in the cases of hard or low skilled works.

In a future research it would be interesting to repeat our analysis using fictitious CVs with photos sent by candidates of different ethnicities. In this way it could be possible to estabilish whether nationalities interact with attractiveness in reducing or increasing gender discrimination, considering also the type of job the candidate is applying for.

Appendix

See Table 10.

Table 10 Selection equation (selection variable REPLY)

	Model 0	Model 1		Model 2		Model 3		Probit
	Non genderized	Non genderized	genderized	Non genderized	genderized	Non genderized	genderized	
const	0.587*** (0.059)	0.526 (0.020)	0.502*** (0.018)	0.509*** (0.017)	0.522*** (0.019)	0.523*** (0.028)	0.488*** (0.026)	0.604*** (0.057)
Managers	0.046 (0.128)	−0.011 (0.008)	−0.011 (0.007)	0.004 (0.007)	$3.5e-4$ (0.007)	−0.469*** (0.056)	−0.451*** (0.053)	−0.114* (0.066)
Professionals	−0.005 (0.099)	0.003 (0.007)	0.009 (0.006)	0.001 (0.006)	0.001 (0.006)	−0.103** (0.052)	−0.075 (0.047)	0.116** (0.059)
Clerical	−0.003 (0.036)	−0.033*** (0.006)	−0.031*** (0.006)	−0.011** (0.006)	0.020*** (0.006)	0.213*** (0.053)	0.176*** (0.059)	0.093** (0.045)
Sales	−0.094*** (0.019)	−0.003 (0.005)	−0.002 (0.005)	−0.001 (0.008)	0.011 (0.009)	0.003 (0.047)	−0.012 (0.041)	−0.064 (0.051)
Services	0.073 (0.374)	−0.019 (0.020)	−0.017 (0.020)	0.002 (0.012)	0.018 (0.024)	−0.100* (0.060)	−0.048 (0.044)	−0.078 (0.153)
Skilled	0.079 (0.341)	0.003 (0.016)	0.003 (0.016)	−0.004 (0.009)	0.009 (0.019)	0.383*** (0.082)	0.357*** (0.064)	0.212 (0.138)
Elementary	−0.039 (0.040)	0.002 (0.009)	0.012 (0.008)	−0.001 (0.008)	0.005 (0.012)	0.196*** (0.070)	0.228*** (0.066)	0.271*** (0.097)
FO	−0.058 (0.046)	0.018*** (0.005)	0.020*** (0.004)	0.085*** (0.017)	0.082*** (0.017)	0.028*** (0.006)	0.029*** (0.007)	0.385*** (0.039)
HW	−0.063* (0.032)	0.019** (0.009)	0.011 (0.008)	0.025 (0.022)	0.001 (0.023)	0.019* (0.012)	0.012 (0.009)	−0.194*** (0.072)
Grad	−0.068 (0.087)	0.000 (0.007)	−0.010 (0.006)	−0.048*** (0.008)	0.055*** (0.010)	0.001 (0.009)	−0.008 (0.008)	−0.312*** (0.060)
High	−0.093*** (0.032)	−0.005 (0.004)	−0.014*** (0.005)	−0.030*** (0.008)	0.042*** (0.009)	−0.009 (0.007)	−0.012* (0.007)	−0.153*** (0.048)
Piedmont	−0.004 (0.026)	0.003 (0.011)	0.005 (0.010)	0.008 (0.009)	0.017 (0.014)	0.018 (0.018)	0.013 (0.012)	0.090 (0.105)
AostaValley	0.042 (0.044)	−0.013* (0.007)	−0.010 (0.007)	−0.018** (0.008)	0.014* (0.008)	−0.017** (0.009)	−0.014** (0.007)	−0.106* (0.061)
Lombardy	0.075 (0.066)	−0.008 (0.005)	−0.006 (0.005)	−0.008 (0.005)	0.010* (0.005)	−0.009 (0.007)	−0.005 (0.005)	−0.144*** (0.047)
TrentinoAA	−0.053*** (0.019)	0.012 (0.009)	0.022** (0.009)	0.016* (0.008)	0.020* (0.010)	0.026* (0.016)	0.108 (0.164)	0.319*** (0.091)
Veneto	−0.009 (0.021)	0.005 (0.006)	0.009* (0.006)	0.004 (0.006)	0.009 (0.007)	0.007 (0.008)	0.022** (0.010)	0.132** (0.056)
FriuliVG	−0.057*** (0.020)	0.012 (0.011)	0.018* (0.010)	0.016 (0.010)	0.018* (0.011)	0.022 (0.018)	0.009 (0.010)	0.338*** (0.105)
Liguria	−0.009 (0.020)	0.008 (0.010)	0.019** (0.010)	0.015 (0.009)	0.047** (0.021)	0.029 (0.018)	0.034*** (0.013)	0.341*** (0.106)
EmiliaRomagna	−0.017 (0.017)	0.006 (0.008)	0.011 (0.007)	0.005 (0.008)	0.001 (0.008)	0.005 (0.011)	−0.001 (0.007)	0.154** (0.077)
Tuscany	−0.015 (0.015)	0.001 (0.006)	0.005 (0.005)	−0.001 (0.006)	0.005 (0.007)	0.003 (0.008)	0.003 (0.006)	0.038 (0.055)

Table 10 (continued)

	Model 0	Model 1 Non genderized	Model 1 genderized	Model 2 Non genderized	Model 2 genderized	Model 3 Non genderized	Model 3 genderized	Probit
Umbria	−0.039*** (0.015)	0.029*** (0.008)	0.052*** (0.009)	0.047*** (0.008)	0.079*** (0.030)	0.040*** (0.012)	0.312* (0.163)	1.232*** (0.162)
Marche	−0.041** (0.017)	0.017 (0.010)	0.021** (0.010)	0.020* (0.010)	0.015 (0.010)	0.020 (0.017)	0.010 (0.010)	0.173 (0.111)
Abruzzo	0.043 (0.269)	−0.019 (0.017)	−0.015 (0.016)	−0.014 (0.017)	9.8e − 6 (0.024)	−0.027 (0.017)	−0.019 (0.013)	−0.314** (0.147)
Molise	0.061 (0.096)	−0.005 (0.015)	−0.004 (0.014)	−0.006 (0.017)	0.009 (0.017)	0.000 (0.023)	−0.022 (0.015)	−0.265** (0.125)
Campania	−0.074*** (0.018)	0.010 (0.007)	0.023*** (0.008)	0.020** (0.008)	0.058*** (0.018)	0.020 (0.012)	0.036*** (0.012)	0.535*** (0.096)
Apulia	−0.003 (0.016)	−0.011 (0.009)	−0.007 (0.009)	−0.008 (0.011)	0.006 (0.013)	−0.017 (0.012)	−0.005 (0.010)	0.011 (0.088)
Basilicata	−0.069*** (0.024)	0.006 (0.013)	0.018 (0.013)	0.014 (0.012)	0.009 (0.012)	0.010 (0.020)	0.009 (0.013)	0.354** (0.167)
Calabria	−0.049*** (0.014)	0.002 (0.007)	0.011 (0.007)	0.000 (0.007)	0.018* (0.010)	0.009 (0.010)	0.010 (0.007)	0.333*** (0.071)
Sicily	−0.021 (0.015)	−0.021** (0.009)	−0.017** (0.008)	−0.020* (0.010)	0.024** (0.011)	−0.030*** (0.011)	−0.022** (0.009)	0.040 (0.073)
Sardinia	−0.083*** (0.019)	−0.005 (0.013)	0.002 (0.012)	−0.018 (0.027)	0.008 (0.041)	−0.027 (0.017)	−0.017 (0.013)	0.209 (0.153)

Probit regression statistics

	Model 0				
McFadden R²	0.051	Adjusted R²	0.045	Loglikelihood	−5242.926
Akaike IC	10,547.853	Schwarz IC	10,770.365	Hannan–Quinm IC	10,623.294
Correct predictions	7192 (74.298%)	Actual 0-predicted 0	16	Actual 1-predicted 1	7176
		Actual 0-predicted 1	2480	Actual 1-predicted 0	8
LR test	227.799 [0.000]				
Normality test	39.068 [0.000]				

Standard errors in parentheses, *p* values in brackets

IC information criteria, *CM* conditional moment test, *LR* likelihood ratio test

*Statistical significance at the 0.1 level

**Statistical significance at the 0.05 level

***Statistical significance at the 0.01 level

References

Åslund, O., & Skans, O. N. (2012). Do anonymous job application procedures level the playing field? *ILR Review*, *65*(1), 82–107.

Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, *30*(C), 32–40.

Baert, S. (2017). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. IZA discussion papers 10738, Institute of Labor Economics (IZA), April.

Baert, S. (2018). Facebook profile picture appearance affects recruiters' first hiring decisions. *New Media & Society*, *20*(3), 1220–1239.

Baert, S., Pauw, A.-S. D., & Deschacht, N. (2016). Do employer preferences contribute to sticky floors? *ILR Review*, *69*(3), 714–736.

Bauer, T., Lofstrom, M., & Zimmermann, K. F. (2000). Immigration policy, assimilation of immigrants, and natives? Sentiments towards immigrants: Evidence from 12 OECD countries. *Swedish Economic Policy Review*, *7*, 11–53.

Bayard, K., Hellerstein, J., Neumark, D., & Troske, K. (2003). New evidence on sex segregation and sex differences in wages from matched employee-employer data. *Journal of Labor Economics*, *21*(4), 887–922.

Bergmann, B. (1981). The economic risks of being a housewife. *The American Economic Review*, *71*, 81–86.

Bertrand, M., & 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.

Bettio, F. & Verashchagina, A. (2009). Gender segregation in the labour market: Root causes, implications and policy responses in the EU. Technical report, Luxembourg: European Commission. Expert Group on Gender and Employment (EGGE).

Biddle, J. E., & Hamermesh, D. S. (1998). Beauty, productivity and discrimination: Lawyers, looks and lucre. *Journal of Labor Economics*, *16*(1), 172–201.

Brandén, M., Bygren, M., & Gähler, M. (2018). Can the trailing spouse phenomenon be explained by employer recruitment choices? *Population, Space and Place*, *24*(6), e2141.

Bursell, M. (2007). What's in a name? A field experiment test for the existence of ethnic discrimination in the hiring process. *SULCIS Working Papers* (p. 7). Stockholm University, Linnaeus Center for Integration Studies—SULCIS.

Busetta G., Campolo M.G., Panarello D. (2020). Weight-based discrimination in the Italian Labor Market: an analysis of the interaction with gender and ethnicity. J Eco Inequ. https://doi.org/10.1007/s10888-020-09456-5.

Busetta, G., Campolo, M. G., & Panarello, D. (2018). Immigrants and Italian labor market: Statistical or test-based discrimination? *Genus*, *74*(4), 4.

Bygren, M., Erlandsson, A., & Gähler, M. (2017). Do employers prefer fathers? Evidence from a field experiment testing the gender by parenthood interaction effect on callbacks to job applications. *European Sociological Review*, *33*(3), 337–348.

Campos-Soria, J., & Ropero-Garcia, M. (2016). Occupational segregation and the female male wage differentials: Evidence for Spain. *Gender Issues*, *33*, 183–217.

Cann, A., Siegfried, W. D., & Pearce, L. (1981). Forced attention to specific applicant qualifications: Impact on physical attractiveness and sex of applicant biases. *Personnel Psychology*, *34*(1), 65–75.

Carlsson, M., & Rooth, D. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics*, *14*, 716–729.

Charles, M. (2011). A world of difference: International trends in women's economic status. *Annual Review of Sociology*, *37*, 355–371.

Davidson, R., & MacKinnon, J. D. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290.

Dolado, J., Felgueroso, F., & Jimeno, J. (2004). Where do women work? An analysis patterns of occupational segregation by gender. *Annales d'Économie et de Statistique*, 293–315.

Drydakis, N., & Vlassis, M. (2010). Ethnic discrimination in the Greek labour market: Occupational access, insurance coverage, and wage offers. *The Manchester School*, *78*(3), 201–218.

Feingold, A. (1992). Good looking people are not what we think. *Psychological Bulletin*, *111*, 304–341.

Garner-Moyer, H. (2010). L'hyper place de la beauté dans les critères de sélection en phase de recrutement. *Revue Internationale de Psychosociologie*.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of 'blind auditions on female musicians'. *The American Economic Review*, *90*(4), 715–741.

Gundle, S. (1997). Bellissima: Feminine beauty and the idea of Italy. Yale University Press **(ISBN: 9780300123876)**.

Gundle, S. (1999). Feminine beauty, national identity and political conflict in postwar Italy. *Contemporary European History*, *8*(3), 359–378.

Hamermesh, D. S. (2013). *Beauty pays. Why attractive people are more successful*. Princeton University Press **(ISBN 9780691158174)**.

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review*, *84*(5), 1174–1194.

Heckman, J. J. & Siegelman, P. (1993). The urban institute audit studies: Their methods and findings. In M. Fix & R. J. Struyk (Eds.), *Clear and convincing evidence: Measurement of discrimination in America* (pp. 187–258). The Urban Institute Press.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153–161.

Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, *12*(2), 101–116.

Heilman, M. E., & Saruwatari, L. R. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and nonmanagerial jobs. *Organizational Behavior & Human Performance*, *23*(3), 360–372.

ISTAT. (2013). *La popolazione straniera residente in Italia—Bilancio demografico. Anno 2012*. Statistiche, ISTAT

ISTAT. (2017). *Noi Italia: 100 statistiche per capire il Paese in cui vivamo*. Report, ISTAT.

Korkeämaki, V., & Kyyrä, T. (2006). A gender wage gap decomposition for matched employer? Employee data. *Labour Economics*, *13*, 611–638.

Lopez Bóo, L., Rossi, M., & Urzua, S. (2013). The labor market return to an attractive face: Evidence from a field experiment. *Economic Letters*, *118*(1), 170–172.

Meulders, D., Plasman, A., Rigo, R., & O'Dorchai, S. (2010). Horizontal and vertical segregation. Meta-analysis of gender and science research. Topic report, 7th RTD Framework Programme of the European Union, rTD-PP-L4-2007-1.

Meyersson, E. M., Petersen, T., & Snartland, V. (2001). Equal pay for equal work? Evidence from Sweden and a comparison with Norway and the US. *Scandinavian Journal of Economics*, *103*(4), 559–583.

Mincer, J. (1978). Family migration decision. *Journal of Political Economy*, *86*(5), 749–771.

Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, *96*(1), 222–235.

Mussida, C., & Lucarelli, C. (2014). Dynamics and performance of the Italian labour market. *Economia Politica, Journal of Analytical and Institutional Eco-nomics*, *1*, 33–54.

Naticchioni, P., & Ricci, A. (2012). Public sector reforms and the evolution of wage inequality: Evidence from Italy. *Economia Politica, Journal of Analytical and Institutional Economics*, *2*, 193–220.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799–866.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? A field experiment with thirteen-thousand resumes. *American Economic Journal: Economic Policy*.

Pagan, A., & Vella, F. (1989). Diagnostic tests for models based on individual data: A survey. *Journal of Applied Econometrics*, *4*, 29–59.

Pager, D., & Quillian, L. (2005). Walking the talk? What employers say versus what they do. *American Sociological Review*, *70*(3), 355–380.

Parrett, M. (2015). Beauty and the feast: Examining the effect of beauty on earnings using restaurant tipping data. *Journal of Economic Psychology*, *49*, 34–46.

Patacchini, E., Ragusa, G., & Zenou, Y. (2015). Unexplored dimensions of discrimination in Europe: Homosexuality and physical appearance. *Journal of Population Economics*, *28*(4), 1045–1073.

Puhani, P. A. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, *14*(1), 53–68.

Riach, P., & Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal*, *112*(483), F480–F518.

Rich, J. (2014). What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000. *IZA Discussion Paper No. 8584*, Institute for the Study of Labor, Bonn, Germany.

Riva, E., & Zanfrini, L. (2013). The labour market condition of immigrants in Italy: The case of Lombardy. *Revue Interventions Économiques*, *47*, 1–27.

Rooth, D. (2009). Obesity, attractiveness, and differential treatment in hiring. *Journal of Human Resources*, *44*(3), 710–735.

Ruffle, B. J., & Shtudiner, Z. (2015). Are good-looking people more employable? *Management Science*, *61*(8), 1760–1776.

Saraceno, C., Sartor, N., & Sciortino, G. (2013). *Stranieri e disuguali—Le disuguaglianze nei diritti e nelle condizioni di vita degli immigrati*. Il Mulino.

Schwab, K. (2019). *The global competitiveness report 2019*. Technical report, World Economic Forum, Switzerland

Simon, H. (2012). The gender gap in earnings: An international comparison with European matched employer-employee data. *Applied Economics*, 1985–1999.

Turner, M., Fix, M., & Struyk, R. (1991). Opportunities denied, opportunities diminished: Racial discrimination in hiring. *UI Report*, 91-9. Washington DC: The Urban Institute Press.

Venturini, A., & Villosio, C. (2008). Labour-market assimilation of foreign workers in Italy. *Oxford Review of Economic Policy*, *24*(3), 517–541.

Wood, M., Hales, J., Purdon, S., Sejersen, T., & Hayllar, O. (2009). A test for racial discrimination in recruitment practice in British Cities. Research Report 607, Department for Work and Pensions.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.