HAL
open science

Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications

Anthony Edo, Nicolas Jacquemet, Constantine Yannelis

To cite this version:

Anthony Edo, Nicolas Jacquemet, Constantine Yannelis. Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. Review of Economics of the Household, 2019, 17 (1), pp.349-376. 10.1007/s11150-017-9391-z . halshs-02042942

HAL Id: halshs-02042942

https://shs.hal.science/halshs-02042942v1

Submitted on 20 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Language Skills and Homophilous Hiring Discrimination: Evidence from Gender- and Racially-Differentiated Applications*

Anthony Edo† Nicolas Jacquemet‡ Constantine Yannelis §

March 2019

Abstract

This paper investigates the importance of ethnic homophily in the hiring discrimination process. Our evidence comes from a correspondence test performed in France in which we use three different kinds of ethnic identification: French sounding names, North African sounding names, and "foreign" sounding names with no clear ethnic association. Within the groups of men and women, we show that all non-French applicants are equally discriminated against when compared to French applicants. Moreover, we find direct evidence of ethnic homophily: recruiters with European names are more likely to call back French named applicants. These results show the importance of favoritism for in-group members. To test for the effect of information about applicant's skills, we also add a signal related to language ability in all resumes sent to half the job offers. The design allows to uniquely identify the effect of the language signal by gender. Although the signal inclusion significantly reduces the discrimination against non-French females, it is much weaker for male minorities.

Keywords: Correspondence testing, Gender discrimination, Racial discrimination, Ethnic homophily, Language skills.

JEL Classification: J15, J64, J71.

*Published in Review of Economics of the Household, vol. 17(1), pp. 349-376. The paper previously circulated as CES WP n°13-58). We gratefully acknowledge David Neumark for sharing his data and estimation programs with us. We also thank Francis Bloch, Nick Bloom, Emmanuel Duguet, Christelle Dumas, Raquel Fernandez, Stéphane Gauthier, James Heckman, Shelly Lundberg, Muriel Niederle, Phillip Oreopoulos, Paolo Pin, Chris Taber, Marie-Anne Valfort as well as participants to various conferences and seminars for their thoughtful comments on the paper. We are grateful to the CEPREMAP for financial support. Nicolas Jacquemet acknowledges the *Institut Universitaire de France* for its support. Constantine Yannelis thanks the Alexander S. Onassis foundation for generous support.

†CEPII. 113 rue de Grenelle, 75007 Paris, France. `anthony.edo@cepii.fr`.

‡Paris School of Economics and University Paris 1 Panthéon-Sorbonne. Centre d'Economie de la Sorbonne, 106 Bd de l'hopital, 75013 Paris. nicolas.jacquemet@univ-paris1.fr

§NYU Stern School of Business. 44 W. 4th Street, New York, NY 10012. `constantine.yannelis@stern.nyu.edu`.

1 Introduction

Homophily, the idea that people apply preferential treatment to similar individuals, has long-standing roots in both sociology (Lazarsfeld and Merton, 1954; Hamm, 2000; Mollica, Gray, and Treviño, 2003; Wimmer and Lewis, 2010) and psychology (Vigil and Venner, 2012). Although it has been successfully applied to marriage (Grossbard-Shechtman, 1993; Chiswick and Houseworth, 2011; Grossbard, Gimenez-Nadal, Molina, et al., 2014), friendship (Shrum, Cheek Jr, and MacD, 1988) and work relations (Lincoln and Miller, 1979; Ibarra, 1995), it has only recently been suggested as a source of discrimination (Stoll, Raphael, and Holzer, 2004; Giuliano, Levine, and Leonard, 2009). The aim of this paper is to test experimentally the extent of homophily in labor market discrimination, and to explore its mechanisms.

We test for homophily by comparing the disadvantage faced by clearly identified minorities, such as African-Americans in the US, to that faced by foreign applicants for whom no clear origin is identified by the employer. Our design follows Jacquemet and Yannelis (2012). We replicate the experiment in France, using North African applicants as a benchmark, and interact ethnicity and gender by using two sets of ethnically differentiated applications—one male, one female.[1] To test for the effect of enhanced information about language skills ability, we alter the resumes sent to half the job postings by adding language related extra activities or grades.

This paper has three main results. First, we find strong evidence of ethnic homophily in the French labor market—hiring discrimination is not directed toward specific ethnic groups, but rather arises due to a favoritism for in-group members.[2] Specifically, we find that non-French applicants—whether their specific minority group is identified or not—are equally disadvantaged as compared to French applicants. This result holds for the groups of men and women *and* whether or not a language skill signal is added in applications. Also, we directly observe a proxy for the ethnicity of recruiters and show that ethnic discrimination almost vanishes when the recruiters have a non-European origin.

Individuals can associate with similar individuals because of a preference for homogeneity through social or other conformity motives, or rather because they tend to share common characteristics that make communication, mutual trust and relationship formation easier (McPherson, Smith-Lovin, and Cook, 2001; Putnam, 2007). From an empirical point of view, one testable difference between the two explanations is that only the second one is sensitive to available information on applicants. Our second result is that discrimination is drastically reduced if a signal related to language skill abilities is included in the applications.

[1] We control for the potentially confounding effect of religion (as documented by Adida, Laitin, and Valfort, 2010) by selecting foreign names that are not perceived as being Muslim.

[2] We refer to homophily as workers from one group being favored over others from employers belonging to the same group. Some authors label this phenomenon "in group bias" (see, for example Currarini and Mengel, 2012), so as to distinguish unequal treatment behavior and self selection into groups. With some abuse of language, we only refer to homophily, since the two have similar empirical implications, at least in our setting.

Third, we find an asymmetric effect of the signal across gender groups. Contrary to the group of women, enhanced information does not change the level of hiring discrimination within the group of men. One interpretation is that information has nothing to do with males being discriminated against. Another one is that the lack of information driving ethnic homophily within the group of men is related to other productive characteristics than language skills. We also show that our results are robust to the heterogeneity of employer's and job's characteristics, as well as to potential bias that may arise due to differences in the distribution of the unobservable characteristics across the French and the non-French applicants (Heckman and Siegelman, 1993; Neumark, 2012).

The rest of this paper is organized as follows. The next section summarizes the existing literature on hiring discrimination, ethnic homophily and how to measure statistical discrimination. Section 3 describes our experimental design and its implementation in Paris and its suburbs. In Section 4, we provide the main results of the experiment. Section 5 shows that our results are robust to the bias induced by ethnic differences in the distribution of unobservables (Heckman and Siegelman, 1993). Section 6 concludes.

2 Ethnic Homophily as a source of discrimination

Several studies have shown the prevalence of discrimination against minority applicants, with discrimination ratios ranging from 1.3 to 1.7. This is the case in many countries with very diverse immigration and colonial histories—United States (Bertrand and Mullainathan, 2004), Canada (Oreopoulos, 2011), Australia (Booth, Leigh, and Varganova, 2012), Sweden (Carlsson and Rooth, 2007), France (Duguet, Leandri, L'Horty, and Petit, 2010) as well as in online markets (Doleac and Stein, 2013); see Riach and Rich (2002) for a complete survey of existing results. Recent evidence also shows that such results vary little according to the specific minority used to test for discrimination. In Canada, Oreopoulos (2011); Dechief and Oreopoulos (2012) show that Asian, Indian and Greeks fare similarly poorly compared to native Canadians. Duguet, Du Parquet, L'Horty, and Petit (2015) study the success of male and female applications of Senegalese, Moroccan and Vietnamese origin in France. Although they focus on a high-skill occupation (computing with a Masters degree), they obtain discrimination ratios that roughly fall in the above mentioned range, and which are similar against most non-French applicants. Because these studies rely on correspondence tests, they have the well-known drawback that no "real" person is behind experimental applications. By implementing an audit study based on Latino and Black applications in New-York, Pager, Western, and Bonikowski (2009) however confirms the existence of discrimination and the fact that applicants from the majority ethnic group are always favored.

Such stability across countries and across minorities within countries suggests that a more encompassing mechanism may underlie observed discrimination. The sociological notion of homophily points to such a mechanism, based on the principle that "birds of a feather flock together". This literature (reviewed in McPherson, Smith-Lovin, and Cook, 2001) not only shows that homo-

geneity is a driving force of social networks formation, but also that ethnicity is certainly the most influential factor of this process. While homophily is a well recognized phenomenon in the economics literature on network formation, the application to the effect of ethnicity on labor market outcomes only recently emerged. Based on observational data from four cities in the U.S., Stoll, Raphael, and Holzer (2004) show that black employers tend to hire more black applicants, not only because they receive more black applications but also because they hire a greater proportion of blacks who apply. Using data from a large U.S. retail firm, the study by Giuliano, Levine, and Leonard (2011) finds an own-race bias in manager-employee relationships in regards to quits, dismissals and promotion. Jacquemet and Yannelis (2012) study ethnic homophily based on a correspondence test in Chicago including foreign names with no clear ethnic identification. They show that the discrimination rates against these applicants are the same as the one experienced by African-American applicants in all fields under study—accounting, nursing and programming. As discrimination is directed against all members of the non-majority ethnic group, rather than specific minorities, they show the important role played by racial homophily in shaping hiring discrimination.[3]

Two main reasons have been raised to explain homophilous behavior in social relationships.[4] First, Currarini, Jackson, and Pin (2009) develop a model of network formation in which homophily results from an intrinsic individual preference for similar individuals. Applied to discrimination, this mechanism is very similar to the taste-based model introduced by Becker (1971) and its extension to nepotism by Goldberg (1982)—see Salamanca and Feld (2016) for a model combining the two. Second, belief-based explanations have been proposed for homophily. Such a pattern in social behavior indicates that beliefs and information on who the others are; as well as compliance to what the others are believed to expect plays a greater role than preferences.[5] Since

[3]The present paper goes beyond the study by Jacquemet and Yannelis (2012) in four important ways. First, our experiment is implemented in France instead of Chicago. Second, we interact ethnic discrimination with gender discrimination by adding a male equivalent to each female applicant. Third, we consider three different levels of job occupations (in the same sector) to assess the sensitivity of our results to the kinds of skills required. Fourth, we directly observe a proxy for the ethnicity of recruiters, allowing us to identify homophily through recruiters favoring applicants of the same ethnicity. Our results generalize Jacquemet and Yannelis (2012) to the French context as well as to both males and females, and to all occupational categories under study. We take this as robust evidence that homophily has explanatory power on observed discrimination in hiring. Our main treatment of interest tries to disentangle the reasons behind this behavior.

[4]These two explanations refer to what sociologist call *choice homophily* (Kossinets and Watts, 2009). A third one, formalized for instance by Jackson and Rogers (2007); Bramoullé and Rogers (2009) in the context of network formation, explains homophilous behavior by a higher probability of meeting an in-group partner—this is called *induced homophily*, because homophily results from the structural opportunities to interact. Because all applications are sent to each employer, correspondence testing is ill-equipped to provide reliable measures of such a phenomenon. It does not mean it cannot contribute to the observed differences between applicants—it will be the case, for instance, if employers' decisions are correlated to the relative composition of the overall application pool, and the pool is not exogenous to the employers' characteristics. In the empirical part, we will include covariates related to the employer location to provide some control over this channel.

[5]This result is supported by McPherson, Smith-Lovin, and Cook (2001); Putnam (2007); Kets and Sandroni (2014) who suggest that individuals may cooperate with similar others for ease of communication, mutual trust and closed cultural features that smooth the coordination of activity. Ramachandran and Rauh (2013) moreover

information drives discriminatory behavior in this case, this mechanism echoes the statistical view of discrimination introduced by Arrow (1973); Phelps (1972).

The key difference between preference and belief based mechanisms is the role played by information. If beliefs drive the decision to discriminate, then observed characteristics, such as origin or gender, are used to infer unknown characteristics, such as language skill. This results in unequal treatment of individuals endowed with the same observable characteristics. Belief-based discrimination is thus sensitive to the extent of information available. In the context of hiring discrimination experiments, Oreopoulos (2011); Dechief and Oreopoulos (2012) are the first to systematically manipulate the content of the applications in order to investigate the extent of statistical discrimination. These experiments use a very rich set of resume characteristics as treatment variables such as language fluency, multinational firm experience, active education from highly selective schools and extracurricular activities. None of them affects the ethnic discrimination observed in Canada regardless of gender. The authors also conducted debriefing interviews with recruiters, from which language skills is a major concern (although mentioning fluency on the resume does not affect the level of ethnic discrimination).

Building on this evidence, we experimentally vary the language skills displayed on applications. As employers learn more about the productivity of workers, they will rely less on noisy signals that are correlated with productivity, such as race. Altonji and Pierret (2001), for instance, use the arrival of information associated to increased experience in the same firm to test for statistical discrimination in wage setting. We develop a similar strategy, which compares discrimination in cases where a prospective employer has more or less information about an individual's underlying ability. Our treatment variable is an additional resume experience (detailed below) related to the practice of French language. We add this information to all applications sent to half the job offers. If perceived language skills are correlated with both ethnic background and productivity, prospective employers that statistically discriminate should rely less on ethnic background to screen applicants. Hence, the signal premium should be higher for minority applicants implying a lower level of discrimination.

show that discrimination induced by such a mechanism can persist through coordination failures even once biased norms and beliefs about others have disappeared. Laboratory experiments evidence reported in Habyarimana, Humphreys, Posner, and Weinstein (2007) tend to support this belief-based explanation. The experiments take place in the neighborhood of Kampala, the Uganda's capital, and use as a treatment variable the ethnic composition of the group of people playing together. The evidence relies on standard social preferences games and shows that *(i)* higher homogeneity inside groups is associated with higher voluntary contributions to public goods while *(ii)* there is no difference in the level of gift chosen in a dictator game. It is only when the tribe of both players is known that people treat others from their own tribe more favorably. The preferred interpretation of the authors is the existence of a norm within ethnic groups.

3 Design of the Empirical Analysis

The empirical analysis relies on three main ingredients: the identity of fictitious applicants; the content of the applications and the job offers to which we apply on this basis.

3.1 Treatment Variable I: Fictitious Applicants

We rely on six fictitious identities. To test for gender discrimination, three of them are male and three are female. We also use three different origin. Our benchmark is the level of discrimination experienced by North African sounding names as compared to French sounding names. We choose North African names as our benchmark minority group since it is the largest minority group in the French population. Thus, it is both the most widely studied in correspondence studies run in France, and also the closest French equivalent to African-American used in US studies. We add a third group of names *(i)* perceived as foreign but *(ii)* with no clear ethnic origin, so that they are unidentifiable by employers.

The choice of the names is based on a preliminary survey that asks respondents to indicate to what origin, if any, they associate the names. We gather further control information by also asking respondents their guess for a gender and a religion associated to the name.[6] The survey mixes a sample of 32 names with clearly identifiable French and North African names as well as a set of names intended to be unidentifiable. The sample of names has been created based on public sources of information and the study by Duguet, Leandri, L'Horty, and Petit (2010). The survey has been conducted in various area in Paris using 150 employees mainly from the public sector (nurse, nursery nurse and temps) and 150 college students *(i.e.* a total of 300 respondents).

The results of the survey for the six names we consequently selected for the study are displayed in Table 1. The results confirm very high rates of correct guesses in terms of origins and gender for both French and North African sounding names, suggesting a very high reliability of the treatment variable. Second, the share of respondents who are unable to associate any particular origin to the last two names (leaving the field blank or putting a question mark) is always higher than two thirds. This means both that these names are largely perceived as non-French, and at the same time that no specific immigrant group is associated to these names. In spite of this high uncertainty about the names' origins, each one is associated with one particular gender for a large majority of respondents. These two names will be used in the study as the male and female treatment variable used to test ethnic homophily, *i.e.* discrimination against applicants for the only reason that they do not belong to the majority ethnic group.

One potential challenge to such an interpretation is that their treatment effects could potentially be driven by strong and consistent beliefs from a minority of employers. Inspection of the

[6] The question about religion was not included in the first wave of the survey. We ran a second wave, with again 300 respondents, including only this question. We decided not to ask the origin again to this second sample in order to avoid that respondents mechanically relate religion to origin.

Table 1: Perceived Origins, Gender and Religion for the Six Identities used in the Study

Names	Origin and Sex Guesses			Other Perceived Origin				Perceived Religion	
1. French	Correct	M	F						
LECLERC Pascal	99%	97%	1%	Unknown	1%	-		Christian	96%
ROUSSET Sandrine	97%	1%	98%	Unknown	2%	-		Christian	95%
2. North African	Correct	M	F						
BENBALIT Rachid	94%	96%	3%	Unknown	2%	Israeli	2%	Muslim	94%
BENOUNIS Samira	92%	1%	99%	Israeli	3%	Unknown	3%	Muslim	77%
3. Foreign	Unknown	M	F						
ALDEGI Jatrix	83%	73%	10%	East Eur.	5%	South Eur.	3%	Christian	67%
HADAV Alissa	70%	1%	83%	North Afr.	9%	Israeli	6%	Jewish	55%

Notes: The table reports the perceived origins, gender and religion of names for the six identities used in this study. The first two variables have been collected on a sample of 300 respondents; the religion question has been asked in the a second wave with 300 other respondents. The first column shows the share of correct guesses for the two first sets of names (French and North African) and the share of respondents who do not identify the origin of Foreign names. The percentage of respondents who perceived the name as male or female is indicated in columns M (male) and F (female). In addition to the prevalent guess, the *middle* columns display the most frequent answers among residuals respondents. Finally, the *right-hand* side of the table provides the most likely perceived religion of those names.

middle column of Table 1 discards such an interpretation. Among the 17% of respondents answering the origin question about ALDEGI Jatrix, and the 30% answering it for HADAV Alissa, the guesses are dispersed in terms of both geographic location and religion.[7] Interestingly, both names are associated by a majority of respondents to different religions: Christian for Jatrix, Jewish for Alissa. These are also different from the perceived religion of the North African names, which are unsurprisingly identified as Muslims, while French names are widely identified as Christian.[8]

3.2 Treatment Variable II: Content of the Applications

To maximize the job offers collected during the study, we select a dynamic sector which is less sensitive than others to economic fluctuations hence offering a large quantity of vacancies. We focus on accounting jobs, in particular vacancies advertising openings for accounting assistant,

[7] The most frequent origins are Eastern Europe (5%) and Southern Europe (3%) for Jatrix Aldegi; and North African (9%) and Israeli (6%) for Alissa Hadav.

[8] Among the residual respondents about the religion question, Jatrix is identified as Jewish by 16% of the sample and as Muslim by 9% of the sample; Alissa is perceived as either Christian (26%) or Muslim (14%).

accounting secretary and accountants. These jobs exhibit varying shares of women, since females amount to 66% of people employed in accounting jobs and 84% of people in assistant and secretary jobs according to the French labor force survey, 2010 wave.

We construct six resumes from actual ones accessible online and alter them to create a distinct set that would not be associated with their owners. Resumes are further arranged so that they differ in terms of content and form (layout, style and typeface) to limit the risk of detection. These small differences between applications could induce uncontrolled systematic differences in the perceived quality of resumes by the employer. To orthogonalize the differences in the content of the application and the applicant's identity, a random rotation system is implemented across names and resumes from one job offer to the other.

To limit the noise in the relative success of applicants introduced by the randomization, we constrained (resume specific) characteristics of the applicants to be similar according to the following dimensions. Applicants were born in 1988 (aged 23) and have never repeated a year. They all are single. The nationality of all applicants is French, as is signaled by standard French application practices—non-nationals explicitly mention their nationality on resumes, notably because their administrative conditions of hiring are very different. The postal addresses are also application specific (they do not change as the applicant name changes) and have been chosen within the same area of the South of Paris. Applicants have a BTS (*Brevet de Technicien Supérieur* - BTS) in accounting and management obtained in 2009.[9] The educational background of the six fictitious applicants is also similar, and consistent with their educational attainment. Applicants are not employed when they apply to show their immediate availability. Their work experience ranges from 18 to 22 months composed of two or three different spells, whose job titles and corresponding job descriptions are real but slightly altered and randomly assigned to the six resumes. Moreover, a motivation letter is created for each application, whose form is consistent with the associated resume. In order to be realistic and representative, the motivation letters are also created from multiple examples accessible online. Last, we account for gender differences between our applicants by adding clear and implicit indicators of gender through spelling in resumes and/or motivation letters. For instance, we added gender links such as "assistant" for male and "assistante" for female, as well as some abbreviation like "M." and "Mlle".

All of the above features are common to all applications. The content of the applications is also used as a treatment variable, through additional skill signals related to language ability. In the applications sent to half the job postings, we add one of each of the six following sentences on all resumes sent to the employer:

- Tutorial for pupils with difficulties in lecture and redaction – *Work Experience Category*
- French Tutoring—*Work Experience Category*

[9] According to the International Standard Classification of Education (ISCED), this level of education corresponds to the first stage of tertiary education. This educational attainment is the most requested for accounting jobs.

- Member of a reading club – *Leisure Category*
- Participation in Scrabble and Crossword competition—*Leisure Category*
- Writer of an inter-college newspaper—*Leisure Category*
- Reward for a well-known competition on French language skills in 2003 (rank: 54/8,500)—*Leisure Category*

The signals are neither application- nor identity- specific: we match them to applications and identity using an additional systematic rotation scheme. All signals are strongly related to language skills—two of them refer to past experience in French teaching, the others are related to past leisure or extra activities centered around use of the French language.

Although our applicants are French citizens and were enrolled in French schools since their first age, there are strong reasons to expect recruiters to be concerned about their French abilities. First, the existing literature documents that second and third generation immigrants have distinctive ways of speaking (Trimaille, 2004). This is also supported by Vallet and Caille (1996) who shows that, in 1989, foreign pupils and children of immigrants overall achieved lower scores than other pupils in a French assessment test. Second, many studies show a decline in the level of language of French pupils over the last two decades. For instance, Daussin, Keskpaik, and Rocher (2011) report that the average number of misspelling in a dictation administered to 7th grade pupils rose from 10.7 in 1987 to 14.7 in 2007 Additionally, the share of pupils with more than 15 mistakes jumping from 26% to 46%. During the same period, the share of pupils at 6th grade with reading difficulties raised from 20.9% to 31.3% in suburbs with high immigrant shares ("*zones d'éducation prioritaire*"), as compared to 14.9% and 19% in the general population. Overall, heterogeneity in French speaking abilities remains even among French citizens enrolled in French schools, and correlates with the ethnic origin. It is thus reasonable to assume that some recruiters might expect minority applicants to have more difficulties in speaking, writing or communicating with their managers, customers or co-workers.

3.3 Experiment Implementation

The experiment was conducted between September 2011 and February 2012 in Paris and its suburbs. We responded daily to job advertisements provided by the major employment agency *Pôle Emploi*. Other sources of job vacancies, such as `APEC.fr` and `cadreemploi.fr`, were also used to complete our sample. No unsolicited applications were sent. Jobs listings are included in the study according to the following criteria: full or part-time job, short or long-term contracts (excluding temporary jobs) and located in the Ile-de-France (Paris and its suburbs). Offers were not included if they required a Master's degree, some specific work experience, or more than two years of experience.

We sent six applications (one for each name in Table 1) to each job offer. They are sent sequentially and in random order, usually between 10 AM and 2 PM on the same day as it appeared on the website. Resumes are addressed to employers in word format via direct email.[10] For each fictitious identity, an email address and a telephone number (including an automatic answering service) were registered at large internet providers (*i.e.* Yahoo, Laposte, Hotmail or Gmail) and at a phone company (SFR). All of this information is applicant-specific and moves across resumes along with the identity of the applicant. Using both email and phone individual accounts, we record whether a call back is elicited within two weeks after sending the application. Our outcome variable is the number of callbacks elicited by each applicant. In order to minimize the effect of the study on the labor market, any invitation is promptly declined by email.

Because almost all job advertisements are collected through *Pôle Emploi*, we gather several additional control variables. First, the advertisements are standardized and provide a multitude of information on the number of employees in the firm, its location, the term of job contract, as well as the gender and name to whom we send the applications. Second, the agency works almost only with employers themselves, rather than specialized intermediaries such as interim or recruitment agencies. This allows us to consider the name provided as the (likely) recruiter's one, from which we build an origin and gender variable.

4 Main Results

Overall, we responded to 504 job offers and sent 3024 resumes. Table 2 provides the overall success rate of each applicant, by gender (in row) and origin (in column). Three points are immediately clear from the callback results.

First, applications with French names always elicit far more callbacks than non-French ones, with an overall rate of 17% against 10%. The overall discrimination ratio is equal to 1.7. Second, a decomposition of callback rates for each non-French origin indicates that North African (9.9%) and Foreign applicants (10.1%) are equally treated, and both disfavored relative to French ones (17.3%). The fact that Foreign applicants are discriminated against while their ethnic group is not identified by employers confirms the prevalence of ethnic homophily among employers—*i.e.,* a general mistrust against all members of the non-majority ethnic group. These two observations remain true conditional on the applicant's gender: the non-French names are equally discriminated as compared to the French name inside each gender group. Third, in terms of gender based discrimination, females are considerably more successful than males in getting through the initial job screening stage: the gap ranges from 19.6% vs 14.9% among French applications, to 12.7% vs 7.1% for North African sounding names and 12.1% vs 8.1% for Foreign sounding ones.

Moreover, we find some evidence of inter-sectionality (the interrelation between several sources of discrimination—see, for example, Crenshaw, 1991) as discrimination ratios are stronger among

[10] In a few cases, applications were sent by postal mail.

Table 2: Success Rates by Origin/Gender with and without the Inclusion of the Signal

	French Names			North African Names			Foreign Names		
	Control	Signal	Overall	Control	Signal	Overall	Control	Signal	Overall
Male	14.3%	15.5%	14.9%	6.0%	8.3%	7.1%	8.3%	7.9%	8.1%
Female	20.6%	18.7%	19.5%	10.7%	14.7%	12.7%	9.1%	15.1%	12.1%
Total	17.5%	17.1%	17.3%	8.3%	11.5%	9.9%	8.7%	11.5%	10.1%
Signal Effect on Racial Discrimination									
	H_1: French≠North African			H_1: French ≠ Foreign			H_1: North African ≠ Foreign		
Male	4.17***	3.48***	5.40***	2.73***	3.91***	4.63***	-1.61	0.28	-0.96
Female	4.25***	1.77*	4.29***	4.85***	1.57	4.57***	0.94	-0.23	-0.49
Total	5.35***	3.70***	6.46***	4.96***	3.58***	6.09***	-0.36	0.00	-0.24
Gender: t-tests on Mean Callback Differences (H_1: Male ≠ Female)									
	-2.63***	-1.30	-2.77***	-2.47***	-2.87***	-3.79***	-0.39***	-3.48***	-2.74***

Notes: ***, **, * denote significance at the 1%, 5%, 10% level. The *upper* part of the table displays the callback rates elicited by each applicant, by gender (in row) and ethnicity (in column). We provide the overall share of elicited callbacks, as well as its decomposition according to whether resumes included the signal or not. The total number of applications is 504 for each identity, and equal to 252 for the sub-samples with and without the signal. The *middle* part provides the statistics of student t-tests of equality in mean callbacks between origins. The *lower* part of the Table provides the statistics of student t-tests of equality in mean callbacks between gender of the same origin.

men (around 2) than women (1.6). The difference in the strength of ethnic discrimination for the group of men and the group of women is compatible with either of the two following interpretations: the underlying mechanisms through which discrimination operates are the same, but they are weaker against women; or the reasons underlying discrimination against men and women from the same origin are different. The decomposition of callback rates according to whether a signal is included on the resumes provides some empirical evidence supporting the second interpretation.

Table 3 rearranges the data according to the discrimination intensity at the employer level, broken down by gender. For both men and women, the majority (80%) of employers treat all applicants of different ethnic origins equally, either rejecting or calling back all applicants. This share is similar to what is generally observed in correspondence studies. The observed discrimi-

Table 3: Discrimination Intensity faced by Applicants, by Gender

		One Origin Favored			One Origin Discriminated		
	Equal Treatment	French	North African	Foreign	French	North African	Foreign
Male	435	39	5	8	3	8	6
	86.3%	7.7%	1.0%	1.6%	0.6%	1.6%	1.2%
Female	415	44	9	9	8	8	11
	82.3%	8.7%	1.8%	1.8%	1.6%	1.6%	2.2%

Notes. The Table reports, by gender, how employers treated applications: equal treatment (callbacks or no answer for all three origins), or favoritism towards one origin (middle part) – in the form of a call back for only the French name, the North African name or the Foreign name and no answer for the others – or discrimination against only one origin (right-hand side) – in the sense that all names elicits a callback except for the French name, the North African name or the Foreign name. The first row refers to the number of employers, the second one to the share among all employers that received applications.

nation thus derives from the behavior of the remaining 20%. Consistent with our interpretation of homophilous hiring discrimination, applicants with French names are far more likely to be the only applicant called back. Discrimination against one origin is equally common for French and non-French named female applicants. Male applicants with French names see slightly lower levels of discrimination, although the magnitude of this effect is smaller than male applicants with French names receiving favorable treatment. The identification of discriminatory behavior relies on the observed unequal treatment of applicants with names indicating differing ethnic origins. As a result, Table 3 suggests that our experimental results are mostly driven by applicants with French names facing favorable treatment as compared to both non-French ones.

4.1 Differences according to Language Skills Signals

We now turn to the effect of additional information in the applications. Table 2 also provides an overview of the comparison between the control applications (no additional skill related to language ability on the resume) and the treatment group. Focusing first on the control group, we observe that all trends discussed previously are reinforced: both non-French origins are discriminated against when compared to the French applicant. This result is true for both males and females, while females on average elicit more positive answers than their male counterparts.

The effect of the signal on French applications is negligible: while the inclusion of a signal generates slightly more callbacks for male applicants, it induces a moderate decrease for female ones (both are statistically not significant). These two slight variations together are enough to absorb the statistical significance of the difference between control applications—the t-test statistic of the difference between French males and females on applications once the signal is included

Table 4: Callback Rates Decomposition According to Job and Firm Characteristics

	Type of Job			Firm Size		Job Contract		Ethnic Diversity	
	Assistant (208)	Secretary (116)	Account (180)	Small (259)	Large (115)	Short (166)	Long (283)	High (395)	Low (104)
Pascal	16.8%	11.2%	15.0%	12.5%	14.8%	17.0%	13.7%	13.4%	19.2%
Rachid	6.3%	6.9%	8.3%	5.5%	8.7%	9.3%	5.9%	6.3%	9.6%
Jatrix	7.7%	6.0%	10.0%	5.8%	7.8%	9.9%	7.1%	6.3%	13.5%
Sandrine	19.2%	22.4%	18.3%	20.1%	18.3%	20.3%	19.3%	18.2%	25.0%
Samira	13.0%	16.4%	10.0%	10.6%	11.3%	14.3%	11.8%	10.9%	18.3%
Alissa	12.5%	11.2%	12.2%	10.6%	11.3%	12.6%	11.8%	10.9%	16.4%

Notes. The table distinguishes callback rates elicited by the six applicants based on type of job, firm size (large firms are defined as firms having more than 50 employees), type of job contract (long-term contracts have an indefinite-term while short-term contract are temporary) and location of the firm (based on the level of ethnic diversity at the city level). For each column, the total number of job offers to which we responded is in brackets.

becomes -1.30 (with a corresponding p-value equal to 0.20).

Among non-French applicants, the inclusion of a signal drastically improves the success rates, with similar magnitude among both non-French origins. The callback rate raises from 8.3% to 11.5% on average for North African sounding names and from 8.7% to 11.5% for Foreign sounding ones. However, for each of these two ethnic categories, we notice an important asymmetry between male and female applicants. The change in callback rates is much stronger for female applications than for male ones. Among women, the rise is both economically and statistically significant: thanks to the inclusion of a signal, the callback rate raises from 10.7% to 14.7% for North African applications, from 9.1% to 15.1% for foreign ones. Conditional on the signal, the difference with French female applicants remains significant only for females of North African origin. For foreign female applicants, the callback rates are statistically the same.

4.2 Decomposition according to Job and Firm Characteristics

Due to the design of the experiment, the randomization of applications applies at the job level on the basis of the openings during the experiment. The aforementioned discrimination patterns may thus be subject to composition effects at both the job's and firm's characteristics levels. Beyond the type of occupation, the characteristics of the job that we observe from the offers are the type of job contract offered (either long-term or short-term), the number of employees in the

firm and its location. We use the location variable to classify jobs' location according to the share of immigrants in the neighborhood.[11]

Table 4 disaggregates callbacks according to these characteristics. The first row of the Table indicates the distribution of the sample according to each dimension. Composition effects are actually far from being negligible: as the experiment does not randomize these characteristics, the sample is not balanced in most dimensions. However, from comparison of callback rates among applicants conditional on each characteristics, the patterns described previously do not seem to be sensitive to these decompositions. First, hiring discrimination is equally directed toward non-French names. The degree of ethnic discrimination is moreover approximately the same across characteristics. Second, conditional upon ethnic origin, callback rates of female applicants are always higher than those of males. However, the intensity of this kind of discrimination is not similar across characteristics. Although the intensity of gender discrimination does not vary by job contract or the ethnic composition of job location, the intensity of gender discrimination differs across types of job.[12] The gap in callback rates between male and female applications is the highest in secretary occupations (and at its lowest in accounting jobs). This is in line with the fact that female workers are overrepresented in secretarial jobs. Although this certainly drives part of the observed favoritism for women, it should be noted that females remain favored (although to a lesser extent) on all other types of jobs.

4.3 Decomposition according to Recruiters' Identity

In order to investigate the extent of gender and ethnic homophily on the labor market between the recruiters and applicants, we divide the callback rates by using the identity of the potential recruiters (gender and origin) in Table 5. The employer's characteristics are always hard to measure in correspondence testing studies, because the experimenter does not observe who is in charge of screening and selecting the applications. We rely on proxies of the employers' characteristics by using the observed identity of the person to whom we send the applications. Although it is likely that this name refers to the person in charge, this can also be the secretary of the human resource department or any other intermediary inside the organization. Such data are subject to measurement error, so that the effect of employer's identity on discrimination has to be interpreted with caution. Based on this identity, we first construct a measure of the gender of the employer. We also construct a variable reflecting the perceived origin of the employer deduced from their

[11] We use the 2008 wave of the French census (providing exhaustive information on the population living on the territory) to compute the share of immigrants in the population of each city. This is linked with the location variable based on the zip code. Ethnic diversity is assumed to be low (high) if the immigrant share is lower (higher) than 20%.

[12] Table 4 also shows a slight tendency towards a reduced discrimination from big firms against female minorities. The difference however remains significant. This tendency is however strongly reinforced when we classify firms according to whether they have more than 500 employees (which results in 40 observations out of the 504 employers in the sample): these very big firms do not discriminate against female minorities. The results are available from the authors upon request.

Table 5: Callback Rates Decomposition According to Recruiters' Identity

	Gender		Origin		
	Male (203)	Female (246)	French (280)	Non-French Eur. (101)	Non-European (63)
Pascal	11.3%	17.5%	15.0%	15.8%	11.1%
Rachid	4.0%	9.8%	6.4%	8.9%	6.4%
Jatrix	4.4%	10.2%	6.4%	8.9%	9.5%
Sandrine	13.3%	26.8%	21.1%	20.1%	19.1%
Samira	9.9%	15.0%	12.1%	9.9%	19.1%
Alissa	8.9%	15.9%	12.9%	11.9%	12.7%

Notes. The table distinguishes callback rates elicited by the six applicants based the identity (gender and origin) of potential recruiters. For each column, the total number of job offers to which we responded is in brackets.

names. We divide the origin of employers into two categories: French, Non-French European and non-European.[13] By design, such covariates are endogenous, since they result from hiring decisions by firms in the past. The aim of the decomposition is to substantiate our interpretation of the data based on correlations between recruiters' identity and applicants characteristics.

The first two columns of Table 5 indicate that female recruiters have a systematically higher callback rate.[14] They also appear to be mostly responsible for the discrimination in favor of female applicants among French-sounding ones. Among male recruiters, both French applications are treated the same way. In order to test this potential favoritism (*i.e.,* gender homophily among females), we implement two statistical tests of equality in callback rates between both male and

[13]The gender is deduced from first names and/or abbreviations before the last names like "M." and "Mme". Also, an employer is considered to be *(i)* French if his/her first and last names sound French, *(ii)* non-French European if his/her last name sound German, Spanish, Italian or Portuguese and *(iii)* non-European if his/her last name appears non-European. Most recruiters with non-European sounding names tend to be originated from North-African and Asian countries. The decomposition of callbacks by employer origin has to be interpreted with caution inasmuch the sample size of non-European sounding names of employers is low, around 15 % of all job offers. The crucial issues to assess the extent of the bias induced by the use of these measures are *(i)* the relative size of the noise captured when our correspondent is not the actual recruiter; *(ii)* whether this noise is likely to be correlated with callbacks. On the second issue, one important likely driving force is the size of the firm, as this may jointly affect the likelihood that collecting the application and making the final decision are separated, and the tendency to discriminate. We do not find such a pattern once the relationships between employer's identity and callback rates are interacted with firm size.

[14]This result is partly explained by the fact that female recruiters are overrepresented in firms which tend to exhibit higher callbacks. Female recruiters mainly work in large firms (70%) located in Paris (56%) and offering short-term contracts (60%).

female French applicants according to the gender of the employer. They confirm our conjecture: while callback rates between French applications are not significantly different when employers are male, they are different at a 1% level when employers are female.

The remaining three columns show that the degree of ethnic discrimination for both gender groups depends on the employers' origin. Ethnic discrimination against foreign applicants are entirely driven by the subsample of European recruiters. When the recruiters have a non-European sounding name, the level of ethnic discrimination for both gender groups is not statistically significant (recruiters whose first name and last name sound North-African play a key role to generate this result, as they represent about two thirds of the non-European recruiters). This result highlights the important role played by the favoritism for in-group members (*i.e.*, ethnic homophily) in shaping the hiring discrimination.

5 Robustness Tests

This section investigates the robustness of our results to the heterogeneity of employer's and job's characteristics, as well as to potential bias that may arise due to differences in the distribution of the unobservable characteristics across the French and the non-French applicants (Heckman and Siegelman, 1993; Heckman, 1998; Neumark, 2012). In appendix, we develop a simple theoretical model of employer decision making that allows the variance of the unobserved productivity to differs across ethnic groups, based on Neumark (2012). We show*(i)* how Probit estimates of the effect of race on callback rates can provide a bias measure of discrimination and *(ii)* how it is possible to correct for the bias.

5.1 Main Regression Results

In order to recover estimates of discrimination that are not biased by the difference in the variances of unobservables, we need identifying variables which impact the probability of hiring, regardless of ethnic origin. We exploit three identifying variables. We follow Neumark (2012) and use academic honors received by applicants.[15] We also use the type of jobs (assistant, secretary or account) and the type of job contract (short or long-term) as they tend to affect hiring (see Table 4) and their effects do not vary with ethnicity. These three variables can thus be used to define identifying restrictions because they vary from one applicant to another (either within or between job offers) and affect their success.

We first estimate the effect of ethnic background by pooling the data from both genders. Table 6 presents the estimation results from three specifications of the model, which gradually adds in the control variables (used as identifying variables in the heteroskedastic Probit estimation) to assess the sensitivity of our results to this choice. The top panel presents estimates derived from the

[15] At the end of high-school, students have to take a national exam called "baccalauréat". According to the test scores, students can have academic honors (no honors, relatively good, good and very good honors).

Table 6: Conditional Estimates of Homophily

	Specification 1			Specification 2			Specification 3		
	Control (1)	Signal (2)	Overall (3)	Control (4)	Signal (5)	Overall (6)	Control (7)	Signal (8)	Overall (9)
				A. Basic Probit Model					
North Afr.	-0.084***	-0.053***	-0.069***	-0.084***	-0.053***	-0.069***	-0.084***	-0.053***	-0.069***
	(-5.46)	(-3.73)	(-6.58)	(-5.45)	(-3.74)	(-6.58)	(-5.48)	(-3.74)	(-6.59)
Foreign	-0.078***	-0.054***	-0.067***	-0.078***	-0.054***	-0.067***	-0.078***	-0.054***	-0.067***
	(-4.95)	(-3.66)	(-6.14)	(-4.95)	(-3.66)	(-6.13)	(-4.99)	(-3.65)	(-6.15)
Male	-0.038***	-0.057***	-0.048***	-0.038***	-0.057***	-0.048***	-0.038***	-0.057***	-0.048***
	(-2.59)	(-3.52)	(-4.35)	(-2.61)	(-3.51)	(-4.36)	(-2.60)	(-3.50)	(-4.34)
Log-lik.	-518.4	-519.0	-1111.8	-518.0	-577.7	-1110.0	-516.8	-577.7	-1109.2
				B. Heteroscedastic Probit Model					
North Afr.	-0.093**	-0.056***	-0.080***	-0.093**	-0.054***	-0.082***	-0.094***	-0.054***	-0.083***
	(-2.52)	(-2.81)	(-4.53)	(-2.39)	(-2.77)	(-4.70)	(-2.90)	(-3.43)	(-4.56)
Foreign	-0.099***	-0.069***	-0.074***	-0.098***	-0.066***	-0.074***	-0.097***	-0.066***	-0.076***
	(-4.63)	(-3.38)	(-4.32)	(-4.49)	(-3.44)	(-4.47)	(-4.81)	(-3.43)	(-4.40)
Male	-0.030*	-0.056***	-0.047***	-0.030	-0.057***	-0.047***	-0.032**	-0.057***	-0.046***
	(-1.66)	(-3.30)	(-3.80)	(-1.61)	(-3.43)	(-3.85)	(-2.08)	(-3.43)	(-3.77)
Log-lik.	-577.2	-578.1	-1111.3	-516.6	-576.7	-1109.3	-515.5	-576.7	-1108.4
Identifying Variables	Type of Job			Type of Job, Job Contract			Type of Job, Job Contract, Honors		
N	1,512	1,512	3,024	1,512	1,512	3,024	1,512	1,512	3,024

Notes. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level. The dependent variable in each specification is an indicator of whether or not an application received a callback from an employer. T-statistics are indicated in parentheses below the point estimate. The table reports the marginal effects of ethnicity and gender on the likelihood of eliciting a callback from both homoscedastic and heteroscedastic Probit models. Panel A presents the homoscedastic Probit model while Panel B presents the heteroscedastic Probit model. Identifying variables included are noted in the bottom panel. For each regression, we provide the value of the log-likelihood. Standard errors are clustered at the firm level.

homoskedastic Probit model, whereas the bottom panel presents results from a heteroskedastic Probit model. In each model, the effect of ethnic background is captured by two dummies. The first row presents the coefficient on a dummy variable which is equal to one if an individual has a North African name. The second row presents the coefficient on a dummy variable which is equal to one if an individual has a foreign-sounding name that is unidentifiable to French employers. The results are split between applications which include a signal of ability in the French language, and those which do not. Overall, the sign and magnitudes of our results are robust to the inclusion of different sets of identifying variables. The marginal effects from the heteroskedastic specification are always very close to those from the Probit estimates, indicating that differences in the distribution of the unobservables across the French and the non-French applicants do not play a major role in our experiment.

Reassuringly, both models reaffirm the experimental results found in the descriptive statistics. Applicants with both North African and unidentifiable Foreign names face a significantly reduced probability of receiving a callback in comparison to applicants with French names. Moreover, the coefficient on having a North African name is not significantly different from the coefficient on a Foreign name, suggesting that all individuals with non-French names face discrimination in the French labor market. This result indicates that discrimination is directed against all non-majority ethnic groups, and highlights the importance of homophily in hiring discrimination.

Moreover, Table 6 provides evidence of statistical discrimination. In the full sample, the effect of a non-French name on callbacks pools the effects of statistical and taste based discrimination. Statistical discrimination arises from a (real or perceived) gap in the mean or variance of characteristics of ethnic groups, as employers use ethnicity as a proxy for unobservable characteristics. The extent of this phenomenon is measured by contrasting the estimates from columns 1, 4, and 7—that restrict the sample to resumes that did not include a signal of proficiency in the French language—while columns 2, 5 and 8—that restrict the sample to resumes that include a signal of strong abilities in the French language. Both the ordinary Probit and the heteroskedastic Probit models indicate that the inclusion of the language signal reduces discrimination faced by applicants with non-French names, indicating the employers do engage in statistical discrimination.[16]

5.2 Asymmetric Effects by Gender

We now turn to conditional estimation of gender-specific discrimination, pooling the two non-French origins as a treatment dummy variable. Table 7 disaggregates the estimates of the effect of having an ethnic name on the probability of hiring for male, female and all applications. We also disaggregate the sample according to whether or not the signal was included in the resumes. Panel

[16]However, the effect of the signal on callbacks for the non-French applicants is not significant with a t-test statistic (p-value) equal to -1.50 (0.13). As explained below, the effect of the signal on callbacks is mainly driven by the sub-sample of female identities.

Table 7: Conditional Estimates of Statistical Discrimination, by Gender

	Male Applicants			Female Applicants			All Applicants		
	Control	Signal	Overall	Control	Signal	Overall	Control	Signal	Overall
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				A. Basic Probit Model					
Non-French	-0.065***	-0.067***	-0.066***	-0.098***	-0.037*	-0.069***	-0.082***	-0.054***	-0.068***
	(-3.95)	(-4.21)	(-5.76)	(-5.21)	(-1.86)	(-5.04)	(-5.76)	(-4.16)	(-7.07)
Male	-	-	-	-	-	-	-0.038***	-0.057***	-0.048***
							(-2.59)	(-3.50)	(-4.34)
Log-lik.	-230.4	-243.8	-480.9	-283.8	-330.7	-625.2	-516.9	-577.7	-1109.2
				B. Heteroscedastic Probit Model					
Non-French	-0.071***	-0.074***	-0.071***	-0.106***	-0.036	-0.073***	-0.085***	-0.056***	-0.073***
	(-3.69)	(-3.64)	(-4.75)	(-4.12)	(-1.20)	(-4.56)	(-4.19)	(-3.43)	(-6.54)
Male	-	-	-	-	-	-	-0.038***	-0.058***	-0.049***
							(-2.56)	(-3.63)	(-4.35)
Log-lik.	-230.4	-243.7	-480.9	-283.7	-330.6	-625.1	-516.8	-576.9	-1108.8
N	756	756	1,512	756	756	1,512	1,512	1,512	3,024

Notes. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level. T-statistics are indicated in parentheses below the point estimate. The dependent variable in each specification is an indicator of whether or not an application received a callback from an employer. The table reports the marginal effects of ethnicity and gender on the likelihood of eliciting a callback from both homoscedastic and heteroscedastic Probit models. Panel A presents the homoscedastic Probit model while Panel B presents the heteroscedastic Probit model. All specifications include type of job, type of job contract and academic honors. For each regression, we provide the value of the log-likelihood. Standard errors are clustered at the firm level.

A presents the estimation results from the homoskedastic Probit model, while Panel B presents the estimation results from the heteroskedastic Probit. The sign and significance of the coefficients of interest are consistent with the descriptive statistics and the results of Table 6. The results indicate evidence of a strong discrimination against minorities.

As well as differential levels of discrimination by gender, we also find differential reactions to the language signal by gender. Including the language signal drastically reduces discrimination for women. For men, there is some evidence of a reduction in discrimination however the results are

Table 8: Conditional Estimates of Statistical Discrimination, by Gender and Type of Job

	Assistant		Secretary		Account	
	Control (1)	Signal (2)	Control (3)	Signal (4)	Control (5)	Signal (6)
			A. Male Applicants			
Non-French	-0.076***	-0.098***	-0.058*	-0.022	-0.059**	-0.052*
	(-2.84)	(-3.75)	(-1.71)	(-0.90)	(-2.24)	(-1.95)
Log-lik.	-88.8	-107.2	-66.9	-24.7	-72.7	-103.6
			B. Female Applicants			
Non-French	-0.097***	-0.030	-0.119***	-0.039	-0.079***	-0.050*
	(-2.92)	(-0.96)	(-3.22)	(-0.77)	(-3.05)	(-1.92)
Log-lik.	-116.8	-139.8	-93.4	-57.0	-69.9	-127.3
N	300	324	192	156	264	276

Notes. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level. T-statistics are indicated in parentheses below the point estimate. The dependent variable in each specification is an indicator of whether or not an application received a callback from an employer. The table reports the marginal effects of ethnicity on the likelihood of eliciting a callback from homoscedastic Probit model. Panel A presents the results for male applicants, while Panel B focuses on female applicants. All specifications include type of job contract and academic honors. For each regression, we provide the value of the log-likelihood. Standard errors are clustered at the firm level.

more ambiguous. The ordinary Probit models and the heteroskedastic Probit models indicate that the language signal has a much larger effect for female applicants than for male applicants. For women, Table 7 shows that when a language signal is included the marginal effect of non-French names reduces by 0.061 points. The difference between the two estimates is significant at the 1% level. However for men the inclusion of a language signal does not result in any significant change in the observed level of discrimination.

The three occupations we consider in the study are likely to require different skills, especially concerning language abilities. We thus further disaggregate the results provided in Table 7 across the three types of job (assistant accountant, secretary accountant and accountant) to which applications were sent. Table 8 provides the estimated effect of being non-French on the likelihood to receive a callback by gender, according to whether or not a language signal is included in applica-

tions.[17] For assistant jobs, the estimates are similar to those presented in Table 7: the inclusion of a signal in resumes related to language skills removes ethnic discrimination in the group of women but does not affect discrimination against male minorities. However, our estimates indicate that language signal mitigates ethnic discrimination for both gender groups when focusing on secretarial jobs only. This result is consistent with the fact that language skills (e.g. ease of communication, writing and speaking) are relatively more valuable in secretarial jobs where these skills are key. Finally, we find that the signal is not sufficient to remove ethnic discrimination when applicants apply to accountant jobs, even for the group of women.

5.3 Heterogeneity

Table 9 assesses the robustness of our results to composition effects in the pool of job offers. Results are broken down according to two measures of neighborhood diversity as well as the gender and ethnic origin of the employer. Focusing first on diversity, Columns 1 and 2 break down the results by the percentage of foreign-born individuals in the locality. The sample is restricted to localities where more than 20% of the population is foreign-born in Column 1, while Column 2 restricts the sample to localities where less than 20% of the population is foreign-born. This split does not give rise to any noticeable difference: significant evidence of discrimination against applicants with non-French names is found in both areas with similar magnitude. This either suggests that discrimination does not vary with the share of immigrants around the job offer location, or that the share of immigrants is not a good proxy to capture the level of diversity in cities. To further investigate this dimension, Columns 3 and 4 present the results broken down by whether or not the employer is located in the city of Paris or in surrounding suburbs—as it is a well-known phenomenon that in Paris ethnic minorities tend to live outside the city proper in the *banlieue*, or suburbs. The results indicate that applicants with non-French names are discriminated against significantly in both the city of Paris and the suburbs. However the coefficient on non-French names indicates a greater disadvantage faced by non-French applicants in the city of Paris as opposed to the suburbs. This results is consistent with the findings of Jacquemet and Yannelis (2012) who find that in Chicago, where ethnic minorities tend to live in the city center, firms in the suburbs discriminate much more against African-American applicants than firms in the suburbs. In both cases, discrimination is greater in areas where a greater portion of the population hails from the majority group. In both Chicago and Paris, employers thus favor individuals from their own ethnic group according to location specific data, which is consistent with homophily.

The right-hand side of Table 9 also provides direct evidence of both ethnic and gender homophily. Columns 5 and 6 present results broken down by the ethnic origin of the recruiter, whether or not the recruiter has a European or non-European name. We only find significant

[17] Decomposing the effect of ethnic origin on elicited callbacks across types of job implies that we cannot use this latter dimension as an identifying variable when implementing our heteroskedastic probit regressions. In Table 8, we thus only run simple probit estimation.

Table 9: Conditional Estimates of Heterogeneity of Discrimination

	Firm Location				Recruiter Identity			
	Diverse	Homogenous	Paris	Suburb	European	Non-Eur.	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Basic Probit Model				
Non-French	-0.067***	-0.073***	-0.084***	-0.059***	-0.079***	-0.031	-0.057***	-0.095***
	(-6.15)	(-3.62)	(-4.75)	(-5.19)	(-6.97)	(-1.04)	(-3.27)	(-5.75)
Male	-0.046***	-0.060*	-0.039**	-0.054***	-0.053***	-0.075*	-0.021	-0.109***
	(-3.84)	(-2.18)	(-2.02)	(-4.06)	(-4.13)	(-2.26)	(-0.58)	(-2.83)
Log-likelihood	-798.4	-274.6	-424.4	-640.3	-826.9	-139.7	-336.7	-607.7
				B. Heteroscedastic Probit Model				
Non-French	-0.071***	-0.071*	-0.093***	-0.066***	-0.087***	-0.035	-0.062***	-0.095***
	(-5.26)	(-1.83)	(-4.28)	(-4.81)	(-6.43)	(-0.94)	(-3.51)	(-5.58)
Male	-0.047***	-0.059**	-0.041**	-0.055***	-0.539***	-0.075**	-0.032	-0.108***
	(-3.69)	(-2.03)	(-2.06)	(-3.97)	(-4.13)	(-2.25)	(-0.43)	(-2.70)
Log-likelihood	-798.3	-274.6	-424.2	-639.9	-826.5	-139.7	-147.0	-256.1
N	2,370	624	1,116	1,878	2,286	378	1,205	1,459

Notes. ***, **, * denote statistical significance from zero at the 1%, 5%, 10% significance level. T-statistics are indicated in parentheses below the point estimate. The dependent variable in each specification is an indicator of whether or not an application received a callback from an employer. The criteria by which each specification is restricted is given in the first row of the table. The table reports the marginal effects of ethnicity and gender on the likelihood of eliciting a callback from both homoscedastic and heteroscedastic Probit models. Panel A presents the homoscedastic Probit model while Panel B presents the heteroscedastic Probit model. All specifications include type of job, type of job contract and academic honors. Columns (7) and (8) include dummy variables for the origin of a recruiter and an interaction between the origin of the recruiter and the gender of the applicant. For each regression, we provide the value of the log-likelihood. Standard errors are clustered at the firm level.

discrimination against non-French applicants occurring for the set of recruiters with European names. While applicants with non-French names are also less likely to receive callbacks when the recruiter has a non-European name, the point estimate is much smaller than in the case of European recruiters and moreover is insignificant. We interpret this as direct evidence of homophilous discrimination: recruiters are more likely to call back applicants of a similar ethnic

origin.[18] Furthermore, Columns 7 and 8 also suggest that homophilous discrimination operates on the basis of gender as well as ethnic origin. Once the sample is split according to the gender of the recruiter, we find that female recruiters are significantly less likely to call back males, indicating that female recruiters prefer female applicants. This result is strongly driven by French female recruiters calling back more female applicants with French names. In particular, results from additional regressions show that once the sample is further split according to the origin of the applicant,gender homophily does not operate between French recruiters and non-French applicants: non-French males are always significantly discriminated against whatever the recruiter's gender. This confirms descriptive statistics evidence that ethnic homophily tends to overcome gender homophily.

6 Conclusion

This study makes two contributions. First, our experimental results indicate that 'non-French' applications are equally treated, and significantly disfavored relative to 'French' ones. At the initial job-screening stage, ethnic discrimination thus operates against all non-majority members. This result holds for the group of men and the group of women. Moreover, this discrimination is driven by recruiters calling back applicants from similar ethnic backgrounds. This highlights the important role played by ethnic homophily in shaping hiring discrimination. The results thus indicate that hiring discrimination may be faced by a large number of minority groups in a society, rather than being limited to some specific targeted groups.

Second, the study focuses on the underlying mechanisms behind homophily. We test for the role played by information about applicants' skills by including a signal in half the applications. The inclusion of a signal drastically undermines ethnic discrimination within the group of females, but it has a more ambiguous effect on discrimination against male minorities. Such an asymmetric impact indicates that the causes of discrimination are different across gender. The literature suggests a number of possible explanations for the differential effect of the language signal by gender. First, discrimination may be driven by the lack of information in the case of women, while information is not the key driver of discrimination towards men. If driven by information, it can be that employers discriminate men based on other characteristics than language skills—e.g. computing skills. In that sense, the informational distance between the majority group and minorities may differ for male and female minorities (Lundberg and Startz, 2004). Alternatively, the informational content of the language skill signal can go beyond language abilities *per se*. For instance, employers may view language ability as a signal of assimilation. If, at the same time, employers are concerned with e.g., women becoming pregnant—since immigrant women in

[18] European employers favoring French applicants is consistent with homophily. We observe reduced discrimination for North African candidates from non-European employers. This would be consistent with North African employers favoring North African candidates and other non-European employers discriminating against these applicants. Non-European employers are not all North African, so we are measuring a heterogeneous effect.

France have higher birth rates (Westhoff and Frejka, 2007)—then the signal could be interpreted in terms of a lower "risk" of fertility. Our study however minimizes the noise induced by differences between applicants by considering a very narrow range of ages. Future work could explicitly test this hypothesis by varying applicants' age and interacting this variable with the signal. Last, the signal might be seen as more credible for women if, for instance, the activities added to the resumes are perceived as more feminine. Similarly, the asymmetric effect could be induced by homophily in the accounting sectors in which women are highly represented. If employers prefer female applicants in the sectors studied, and give extra consideration to female applicants, the signal will only affect female applicants through implicit discrimination (as suggested by Dechief and Oreopoulos, 2012): employers may carefully read only the content of applications of females. Under this last interpretation, the channels through which discrimination affects men and women are identical. The perceived asymmetry would be only driven by employers preferring female applicants in the sectors studied, and the corresponding extra attention paid to female applicants.

These results thus open interesting avenues for future studies. First, a wider range of skills could be introduced as a signal treatment variable in order to document the whole scope of skills involved in belief based discrimination. Our results also show some heterogeneity in the effect of the signal according to the occupation considered, supporting the idea that the signal is used to infer applicants' skills that are unobserved to the employer. This occupation specific effect of the signal raise potential implications for a wider range of occupations, such as blue collar professions and male intensive positions which could be studied in future work.

The fact that language skills are identified as an important source of variation in callbacks has policy implications. The fact that employers are using language skills to discriminate amongst candidates should turn attention to the implementation of learning and certification programs designed to improve language abilities. Such programs should be devoted to provide and certify sufficient language abilities for all applicants. A label could therefore be transcribed in resumes so as to modify employer beliefs or perception on applicant language skills.

References

Adida, C., D. Laitin, and M. Valfort (2010): "Identifying barriers to Muslim integration in France," *Proceedings of the National Academy of Sciences*, 107(52), 22384–22390.

Altonji, J., and C. Pierret (2001): "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116(1), 313–350.

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

Becker, G. S. (1971): *The Economics of Discrimination*. University of Chicago Press, Chicago (IL).

Bertrand, M., and S. Mullainathan (2004): "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94(4), 991–1013.

Booth, A., A. Leigh, and E. Varganova (2012): "Does Ethnic Discrimination Vary Across Minority Groups? Evidence from a Field Experiment," *Oxford Bulletin of Economics and Statistics*, 74(4), 547–573.

Bramoullé, Y., and B. W. Rogers (2009): "Diversity and Popularity in Social Networks," *CIRPEE Discussion Paper*, 09(03).

Carlsson, M., and D.-O. Rooth (2007): "Evidence of ethnic discrimination in the Swedish labor market using experimental data," *Labour Economics*, 14(4), 716–729.

Chiswick, B. R., and C. Houseworth (2011): "Ethnic intermarriage among immigrants: Human capital and assortative mating," *Review of Economics of the Household*, 9(2), 149–180.

Crenshaw, K. (1991): "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color," *Stanford Law Review*, 43(6), 1241–1299.

Currarini, S., M. O. Jackson, and P. Pin (2009): "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, 77(4), 1003–1045.

Currarini, S., and F. Mengel (2012): "Identity, Homophily and In-Group Bias," *FEEM Working Paper*.

Daussin, J.-M., S. Keskpaik, and T. Rocher (2011): "L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années," *INSEE, France, Portrait Social*.

Dechief, D., and P. Oreopoulos (2012): "Why do some employers prefer to interview Matthew but not Samir? New evidence from Toronto, Montreal and Vancouver," *CLSRN Working Papers*, (2012-8).

Doleac, J., and L. Stein (2013): "The Visible Hand: Race and Online Market Outcomes," *Economic Journal*, 123(572), F469–F492.

Duguet, E., L. Du Parquet, Y. L'Horty, and P. Petit (2015): "First Order Stochastic Dominance and the Measurement of Hiring Discrimination: A Ranking Extension of Correspondence Testings with an Application to Gender and Origin," *Annals of Economics and Statistics*, Forthcoming.

Duguet, E., N. Leandri, Y. L'Horty, and P. Petit (2010): "Are young French jobseekers of ethnic immigrant origin discriminated against? A controlled experiment in the Paris area," *Annals of Economics and Statistics*, pp. 187–215.

Giuliano, L., D. Levine, and J. Leonard (2009): "Manager Race and the Race of New Hires," *Journal of Labor Economics*, 27(4), 589–631.

Giuliano, L., D. I. Levine, and J. Leonard (2011): "Racial Bias in the Manager-Employee Relationship: An Analysis of Quits, Dismissals, and Promotions at a Large Retail Firm," *Journal of Human Resources*, 46(1), 26–52.

Goldberg, M. S. (1982): "Discrimination, Nepotism, and Long-Run Wage Differentials," *Quarterly Journal of Economics*, 97(2), 307–319.

Grossbard, S. A., J. I. Gimenez-Nadal, J. A. Molina, et al. (2014): "Racial intermarriage and household production," *Review of Behavioral Economics*, 1(4), 295–347.

Grossbard-Shechtman, S. (1993): *On the economics of marriage: a theory of marriage labor and divorce*. Boulder Colorado/Oxford England Westview Press 1993.

Habyarimana, J. P., M. Humphreys, D. N. Posner, and J. Weinstein (2007): "Why Does Ethnic Diversity Undermine Public Goods Provision?," *American Political Science Review*, 101(4), 709–725.

Hamm, J. V. (2000): "Do birds of a feather flock together? The variable bases for African American, Asian American, and European American adolescents' selection of similar friends," *Developmental Psychology*, 36(2), 209–219.

Heckman, J., and P. Siegelman (1993): *Clear and Convincing Evidence: Measurement of Discrimination in America*chap. The Urban Institute Audit Studies: Their Methods and Findings, pp. 187–258. The Urban Institute Press, Washington, D.C.

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

Ibarra, H. (1995): "Race, opportunity, and diversity of social circles in managerial networks," *Academy of management journal*, 38(3), 673–703.

Jackson, M. O., and B. W. Rogers (2007): "Meeting Strangers and Friends of Friends: How Random Are Social Networks?," *American Economic Review*, 97(3), 890–915.

Jacquemet, N., and C. Yannelis (2012): "Indiscriminate Discrimination: A Correspondence Test for Ethnic Homophily in the Chicago Labor Market," *Labour Economics*, 19(6), 824–832.

Kets, W., and A. Sandroni (2014): "A belief-based theory of homophily," *mimeo*.

Kossinets, G., and D. J. Watts (2009): "Origins of Homophily in an Evolving Social Network," *American Journal of Sociology*, 115(2), 405–450.

Lazarsfeld, P., and R. K. Merton (1954): "Friendship as a Social Process: A Substantive and Methodological Analysis," in *Freedom and Control in Modern Society*, ed. by M. Berger, T. Abel, and C. Page, pp. 18–66. Van Nostrand, New-York (NJ).

Lincoln, J. R., and J. Miller (1979): "Work and friendship ties in organizations: A comparative analysis of relation networks," *Administrative science quarterly*, pp. 181–199.

Lundberg, S., and R. Startz (2004): "Information and Racial Exclusion," *Journal of Population Economics*, 20(3), 621–642.

McPherson, M., L. Smith-Lovin, and J. Cook (2001): "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 27, 415–444.

Mollica, K. A., B. Gray, and L. K. Treviño (2003): "Racial Homophily and Its Persistence in Newcomers' Social Networks," *Organization Science*, 14(2), 123–136.

Neumark, D. (2012): "Detecting Discrimination in Audit and Correspondence Studies," *Journal of Human Resources*, 47(4), 1128–1157.

Oreopoulos, P. (2011): "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Six Thousand Resumes," *American Economic Journal: Economic Policy*, 3(4), 148–171.

Pager, D., B. Western, and B. Bonikowski (2009): "Discrimination in a Low-Wage Labor Market: A Field Experiment," *American Sociological Review*, 74(5), 777–799.

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

Putnam, R. D. (2007): "E Pluribus Unum: Diversity and Community in the Twenty-first Century," *Scandinavian Political Studies*, 30(2), 137–174.

Ramachandran, R., and C. Rauh (2013): "Discrimination without Taste - How Discrimination can spillover and persist," *Working paper*.

Riach, P. A., and J. Rich (2002): "Field Experiments of Discrimination in the Market Place," *Economic Journal*, 112(483), F480–F518.

Salamanca, N., and J. Feld (2016): "A short note on discrimination and favoritism in the labor market," *The BE Journal of Theoretical Economics*.

Shrum, W., N. H. Cheek Jr, and S. MacD (1988): "Friendship in school: Gender and racial homophily," *Sociology of Education*, pp. 227–239.

Stoll, M. A., S. Raphael, and H. J. Holzer (2004): "Black job applicants and the hiring officer's race," *Industrial and Labor Relations Review*, 57(2), 267–287.

Trimaille, C. (2004): "Pratiques langagiéres chez des adolescents d'origine maghrébine," *Hommes et Migrations*, pp. 66–73.

Vallet, L.-A., and J.-P. Caille (1996): "Niveau en français et en mathématiques des élèves étrangers ou issus de l'immigration," *Économie et statistique*, 293(1), 137–153.

Vigil, J. M., and K. Venner (2012): "Prejudicial behavior: More closely linked to homophilic peer preferences than to trait bigotry.," *Behavioral and Brain Sciences*, 35(6), 448–449.

Westhoff, C., and T. Frejka (2007): "Religiousness and Fertility among European Muslims," *Population and Development Review*, 33(4), 785–809.

Wimmer, A., and K. Lewis (2010): "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook," *American Journal of Sociology*, 116(2), 583–642.

Yatchew, A., and Z. Griliches (1985): "Specification Error in Probit Models," *Review of Economics and Statistics*, 67(1), 134–139.

Appendix

A Hiring Discrimination with heteroskedastic Unobserved Heterogeneity

The data generating process of a correspondence study stems from employers treatment of the content of the applications. We denote $P_i^*(J, X, Z)$ the productivity of an application i in a given position. This productivity depends on the job's characteristics (measured mainly through firm specific variables) J. At the individual level, productivity results from two components: a set of individual characteristics that are observable to both the econometrician and the employer, X_i; and a component which is unobservable to them both, Z_i. We assume that ethnicity, denoted $R_i = 0$ for non-minorities (*e.g.*, whites) and $R_i = 1$ for minorities (*e.g.*, North Africans), do not enter productivity directly.[19] This allows focus on discrimination, which in this framework might occur for two reasons. First, let the employer callback decision be described by the latent variable T_i^*, which determines the treatment applied to a given application (it will drive a dichotomous choice variable below, but it can be thought of as any continuous outcome, such as *e.g.*, wage offers). Taste-based discrimination implies that such treatment depends not only on the productivity of the applicant but also on unproductive observable characteristics such as ethnicity:[20]

$$T_i^* = T^*[P_i^*, R_i] = P_i^* + \gamma R_i = \delta J + \beta X_i + Z_i + \gamma R_i \,.$$

Second, the invitation decision relies on employer's perceived productivity of applicant i, to whom Z_i is not observed while R_i is:

$$\mathbb{E}[P_i^*(J, X, Z)|X, J, R_i] = \delta J + \beta X_i + \mathbb{E}(Z|R_i)\,,$$

which implies statistical discrimination as long as $\mathbb{E}(Z|R_i) \neq \mathbb{E}(Z_i)$ – the unobserved productivity of an applicant is thought of as being dependent on ethnicity. When each employer receives two applications, one from a non-minority applicant and another from a minority one, discrimination shows up in the average differential treatment reserved to applications,

$$\mathbb{E}[T^*|R = 0] - \mathbb{E}[T^*|R = 1] = \beta[\mathbb{E}(X|R = 0) - \mathbb{E}(X|R = 1)] + \mathbb{E}(Z|R = 0) - \mathbb{E}(Z|R = 1) + \gamma\,.$$

[19] To ease exposition, we focus on the two ethnicities case with $R = 1$ for minority applicants, although the framework easily generalizes to more ethnic origins or to other specifications of the characteristics describing discriminated sub-populations, such as gender. It is also worth mentioning that we gather all sources of unobserved heterogeneity in Z. One can add an *i.i.d.* error term to the functionals without changing the main conclusions.

[20] In the following, we will impose the linearity assumptions that will be used to estimate the model. Note that Z is not only unobserved to the econometrician but also to the employer in a correspondence test (as opposed, *e.g.* to an audit study, where employers might gather additional information by meeting experimental applicants). As a result, employer's decisions should be deterministic conditional on the observables. To ease exposition and save on notation, we assume that employers take their callback decisions based on a random draw in the relevant distribution of unobservables—which generates random variations in decisions across firms. As discussed in (Neumark, 2012, p.1135), an alternative way to arrive at a statistical model would be to assume random productivity differences across firms that are multiplicative in the observed productivity of a worker.

By construction, this difference in callback cannot be driven by employer's characteristic. We thus omit the term δJ below, implicitly including it in the deterministic part of the model, βX.

A.1 The Content of Correspondence Test Data

By way of construction, a correspondence test controls the observables X in such a way that they do not systematically differ across sub-groups: the difference in observed characteristics must balance over job applications, hence $\mathbb{E}(X|R=0) = \mathbb{E}(X|R=1)$.[21] The difference in treatment over all job applications thus arises due to two parameters: taste based discrimination, γ, and statistical discrimination which induces a gap in the (perceived or actual) means of the unobserved productive characteristics $\mathbb{E}(Z|R=0) - \mathbb{E}(Z|R=1)$. Note that the two parameters can only be identified together unless the study provides enough control to guarantee that the distribution means are exactly equal, so that only taste based discrimination occurs.[22] However, all components of this aggregate effect arise as the result of an unequal treatment based on unproductive observable characteristics. As such, they correspond to the economic definition of discriminatory behavior: in the remaining, we thus focus on the identification of the gross handicap experienced by minority applicants, μ—standing for the sum of these two parameters: $\mu = \mathbb{E}(Z|R=0) - \mathbb{E}(Z|R=1) + \gamma$.

As noted by Heckman and Siegelman (1993); Heckman (1998) the observational context of a correspondence study may lead to biased estimates if the variance of the error term is ethnicity-specific.[23] This is the case because the observed outcome is non-linearly related to the underlying discriminatory decisions. To see this explicitly, denote c the perceived quality threshold an applicant has to reach to be invited for an interview: the outcome variable of the experiment is $T = \mathbf{1}[T^* > c]$. Further assume that the unobserved productivity Z is $i.i.d$ normally distributed in each ethnic sub-group, but with a (perceived or actual) variance that varies across groups. To ease of exposition, we denote $\sigma_0^2 = Var(Z|R=0)$ and $\sigma_1^2 = Var(Z|R=1)$. The statistical model generating the observed outcome thus gives the following specifications for the probability of obtaining a callback:

$$\begin{aligned}
\mathbb{P}[T=1|R=1,X] &= 1 - \Phi[(c - \mathbb{E}(Z|R=1) - \beta X + \gamma)/\sigma_1] = \Phi[(\beta X + \mathbb{E}(Z|R=1) + \gamma - c)/\sigma_1] \\
\mathbb{P}[T=1|R=0,X] &= 1 - \Phi[(c - \mathbb{E}(Z|R=0) - \beta X)/\sigma_0] = \Phi[(\beta X + \mathbb{E}(Z|R=0) - c)/\sigma_0]\,,
\end{aligned}$$

where Φ denotes the standard normal distribution. The difference between these two expressions is the empirical source of identification for discrimination. However, even in the extreme case

[21] This is the case for two reasons. First, the resumes are calibrated in such a way that all observed productive characteristics are equally likely. There still remains differences from one resume to another, but systematic differences across ethnicities are ruled out through randomization of the names-resumes matching.

[22] This condition is more likely to be met as the set of observable characteristics is wider. Still, the relative share of taste-based and statistical discrimination in the aggregate effect of ethnicity remains a matter of interpretation and cannot be empirically tested even in this kind of framework.

[23] The general principle that heteroskedasticity causes the coefficient estimates in discrete choice models to be inconsistent draws back to Yatchew and Griliches (1985).

with neither statistical nor taste-based discrimination—*i.e.*, $\gamma = 0$ and $\mathbb{E}(Z|R) = \mathbb{E}(Z)$, so that $\mu = 0$—the difference in probabilities still depends on the comparison between σ_1 and σ_0. For instance, if the X have been chosen at the lower tail of the skills distribution, so that $\beta X < c$, and $\sigma_0 < \sigma_1$ then it has to be that $\Phi[(\beta X - c)/\sigma_1] < \Phi[(\beta X - c)/\sigma_0]$, $\forall X$: the experiment produces (spurious) evidence of discrimination against people from $R = 1$ origin. In the more general case with both discrimination and differences in the variance of unobservables, this framework highlights the identification problem faced by correspondence test data.

A.2 Unbiased Measures of Discrimination

The identification problem arises because the average treatment effect $\Delta = \overline{T}_{|R=0} - \overline{T}_{|R=1}$ provides an estimate of $\frac{\mathbb{E}(Z|R=0)-c}{\sigma_0} - \frac{\mathbb{E}(Z|R=1)+\gamma-c}{\sigma_1}$ (up to the functional form of the distribution), while what one seeks to measure is $\mu = \mathbb{E}(Z|R = 0) - \mathbb{E}(Z|R = 1) + \gamma$: the systematic disadvantage experienced by minority applicants due to their belonging to an observable sub-population. As a result, the confounding effect of the dispersion in the observables shows up in the comparison of callback probabilities, thus affecting the mean comparisons between outcomes from the study.

As noted by Neumark (2012), one can however use restrictions on β to restore identification. If the study provides enough variation on relevant applications characteristics, grouped in X in the model above, one can estimate β/σ_0 and β/σ_1 from the (heteroskedastic) Probit model derived in the previous section. Under the assumption that the effect of X on the callback is homogeneous across ethnicity (*i.e.,* the true value of β is the same), the ratio of the two point estimations identifies σ_0/σ_1, which in turn allows us to estimate μ after the usual normalization setting one of the variance terms equal to 1. Since only one such identifying regressor is needed to achieve identification, any additional productivity control provides the usual specification tests of an over-identified model.