

Between Meritocracy and Ethnic Discrimination: The Gender Difference.

Mahmood Arai, Moa Bursell and Lena Nekby ^{*†}

April 18, 2008

Abstract

Using a two stage correspondence test methodology, this study tests employer priors against job-applicants with Arabic names compared to job-applicants with Swedish names. In the first stage, employers are sent CVs of equal observable quality. Thereafter, in the second stage, the CVs with Arabic names are given an advantage of, on average, two more years of relevant work experience. This setup allows us to test the strength of unfavorable priors against job-applicants with Arabic names and to what degree these priors are revised, on average, when resumes are enhanced. Results indicate no significant differences in call-backs for female applicants when CVs with Arabic names are enhanced. The call-back gap for men however remains large and significant despite a positive adjustment of CVs with Arabic names. This implies that negative priors against male job applicants with Arabic names are not revised by an increase in observable merits.

JEL Classification: J15; J16; J71

Keywords: Correspondence Testing, Ethnic Discrimination, Biased Testing, Gender

*Mahmood Arai: mahmood.arai@ne.su.se, Department of Economics, Stockholm University and the Stockholm University Linnaeus Center for Integration Studies (SULCIS). Moa Bursell: moa.bursell@sociology.su.se, Department of Sociology, Stockholm University and SULCIS. Lena Nekby: lena.nekby@ne.su.se, Department of Economics, Stockholm University, SULCIS and IZA.

†The authors are grateful for comments from Magnus Bygren, Carl le Grand, Lena Schröder, Heidi Stoeckl, Peter Skogman Thoursie and Lars Vahtrik as well as seminar participants at the Department of Economics, Stockholm University and the Aage Sørensen Memorial Conference for Graduate Students, Harvard University (2008). The authors also wish to thank Lars Deullar for excellent research assistance. Financial support from the Swedish Council for Working Life and Social Research (FAS) and the Swedish Research Council (VR) is gratefully acknowledged.

1 Introduction

Numerous studies on racial and ethnic discrimination have been conducted using the correspondence-testing methodology which measures differences in employer responses (call-backs) for observationally equivalent fictive resumes (CVs) sent to advertised job vacancies.¹ These studies consistently find fairly large levels of unequal treatment for racial/ethnic minorities. Employers, however, base their decisions not only on the written information available on the resumes (CVs) but also on their priors concerning unobservable characteristics that are assumed to be correlated to ethnicity as identified by observed names.²

Employer priors about job candidates reflect employer preferences for or against different groups in society as well as estimates about group productivity that are unobserved on CVs but correlated with group affiliation. Observed call-back gaps between groups can therefore be seen as a measure of negative priors against the disfavored group, denoted in the literature as taste-based discrimination or statistical discrimination.

A potential drawback of previous correspondence testing studies is that the call-back gaps reported in these studies do not reveal the intensity of employer priors. If one group is only marginally favored it implies that employers, all else equal, prefer applicants from the favored group. An applicant from the disfavored group would only suffer from discrimination when there is an applicant from the favored group applying for the same job with equivalent observable merits. This is very different from the case where there are large negative priors against the disfavored group. In this case the disfavored group would suffer from discrimination also in a situation characterized by

¹Jowell and Prescott-Clarke (1970); Firth (1981); Riach and Rich (1991); Bertrand and Mullainathan (2004). For a Survey see Riach and Rich (2002). For Swedish evidence see Carlsson and Rooth (2007); Bursell (2007).

²For a critical discussion of audit studies and correspondence testing see Heckman and Siegelman (1993) Heckman (1998). See Fryer and Levitt (2004) for discussion of names as signal of individuals background characteristics.

observationally superior CVs in comparison to the favored group.

Two previous studies report that job-applicants with Arabic names have significantly lower chances of being called to job interviews in comparison to job applicants with typically Swedish names (Bursell (2007) and Carlsson and Rooth (2007)).³ This call-back gap can be due to a slight preference among Swedish employers for persons with Swedish names in comparison to those with Arabic names or to strong priors about the Arabic-named population. This study adds to this literature by estimating the strength of negative priors against applicants with Arabic names.

Employer priors against job applicants with Arabic names are examined using a two stage correspondence test. In the first stage of the experiment, employers are sent CVs of equal observable quality. Thereafter, in the second stage of the experiment, the CV with an Arabic name is given an advantage of on average two more years of relevant work experience. This setup allows us to examine the strength of negative priors against applicants with Arabic names and to what degree these priors are revised when resumes are enhanced. Employer priors are tested separately for female and male applicant pairs to allow for differences in priors attributable to job characteristics in a market characterized by considerable occupational gender segregation. In addition employer priors may be based on group stereotypes that differ by gender.

Results from the first stage correspondence test indicate large and significant differences in call-backs for applicants with Arabic names despite observationally equivalent CVs and regardless of gender. In the second stage of the correspondence test, when applications with Arabic names are enhanced,

³A large part of immigration to Sweden, since the Second World War, has been and continues to be from other Nordic countries. Since the mid-1970s, non-European immigration has increased and now constitutes approximately 50 percent of total immigration. Today, approximately 12 percent of the population in Sweden is foreign born. Employment levels for the foreign born are approximately 6 percentage points lower than for the native born population. Employment gaps to natives are however largest for those stemming from countries in Africa, Asia and the Middle East (Schröder (2007)).

the call-back gap disappears for women. The call-back gap is however, unaltered for male applicants implying that negative priors against male job applicants with Arabic names are not revised by an increase in observable merits.

The remainder of the paper is as follows. Section 2 introduces the theoretical framework while Section 3 briefly discusses employer priors with respect to gender and ethnicity. Section 4 describes the experiment design. Results are presented in Section 5. Finally, the paper is concluded in Section 6.

2 Theoretical Framework

In this section we provide an illustrative model of how employers respond to job applications with equal observable merits and names signalling different group affiliation. The model presented also describes the compensatory condition, that is to say, an augmentation of observable merits necessary to overcome employer priors against one group.

Employer Priors

Assume that employers $j = 1, 2, \dots, K$ evaluate posted resumes (CVs) of individuals $i = 1, 2, \dots, N$ and assign a score value V_{ji} based on a vector of observed individual merits supplied to the employer j , \mathbf{X}_{ji} , the group affiliation, $G_{ji} \in \{A, S\}$ disclosed by the name and a vector of unobserved productivity-related group characteristics \mathbf{Z}_{ji} , characterized by the following $k = 1, 2, 3, 4$ moments $\boldsymbol{\mu}_{kj}^G = E_j[\mathbf{Z}_G^k]$.⁴ The employers' evaluation is defined as: $V_{ji} : (\mathbf{X}_{ji}, G_{ji}, \boldsymbol{\mu}_{kj}^G) \rightarrow \mathbb{N}$.

An employer's evaluation V_{ji} is assumed to be non-decreasing in \mathbf{X}_{ji} on an interval $\mathbf{X}_{ji} \leq \tilde{\mathbf{X}}_{ji}$, and non-increasing on $\mathbf{X}_{ji} > \tilde{\mathbf{X}}_{ji}$ where $\tilde{\mathbf{X}}_{ji}$ is the level

⁴The reader can think of group affiliation as A being equal to Arabic backgrounds as signalled by Arabic names and S being equal to Swedish backgrounds as signalled by Swedish names.

of qualification where the candidate is perceived as overqualified. We assume that employers rank candidates according to the V_{ji} . The top candidates are then called to a job-interview. For every employer observing equal merits and thus equal average observed group merits: $\bar{\mathbf{X}}_{jS} = \bar{\mathbf{X}}_{jA} = \bar{\mathbf{X}}$, we define the employer's evaluation gap as the difference in the employers average evaluation of groups S and A as follows:

$$\Delta\bar{V}_j = \bar{V}_j(\bar{\mathbf{X}}, S, \boldsymbol{\mu}_{kj}^S) - \bar{V}_j(\bar{\mathbf{X}}, A, \boldsymbol{\mu}_{kj}^A) \quad (1)$$

where $\Delta\bar{V}_j$ is the measure of employer j 's unequal treatment of candidates according to their group affiliation. For simplicity we assume that S -candidates are chosen when $\Delta\bar{V}_j = 0$. A non-negative $\Delta\bar{V}_j$ represents discriminatory regimes corresponding to the following two cases:

$$\text{I: } E_j[\mathbf{Z}_S^k | \bar{\mathbf{X}}] = E_j[\mathbf{Z}_A^k | \bar{\mathbf{X}}] = \bar{\boldsymbol{\mu}}_{kj} \quad \text{and} \quad V_j(\bar{\mathbf{X}}, S, \bar{\boldsymbol{\mu}}_{kj}) - V_j(\bar{\mathbf{X}}, A, \bar{\boldsymbol{\mu}}_{kj}) \geq 0 \quad (2)$$

and/or

$$\text{II: } V_j(\bar{\mathbf{X}}, S = \nu, E_j[\mathbf{Z}_S^k | \bar{\mathbf{X}}]) \geq V_j(\bar{\mathbf{X}}, A = \nu, E_j[\mathbf{Z}_A^k | \bar{\mathbf{X}}]) \quad (3)$$

Case I is when S is preferred to A at equal observed and unobserved merits. This is the case commonly denoted as "taste based discrimination (Becker (1957)) implying that the employer always prefers S to A at equal wage costs. Case II instead corresponds to statistical discrimination when based on true statistics or error discrimination when based on prejudices about non-observed productivity-related group characteristics.⁵ These beliefs concerning group-related productivity characteristics can pertain to mean differentials in productivity, differences in variances between groups,

⁵For statistical discrimination see Arrow (1972), Arrow (1973), Phelps (1972) and Aigner and Cain (1977). For error discrimination see England (1992).

differences in symmetry (where the bulk of the group is placed in the productivity distribution) or differences in the existence/prevalence of outliers (i.e. persons not representative of the group).

Allowing for variation in $\Delta\bar{V}_j$ across employers, individual job applicants will face different chances of retaining a call back when facing different employers. Employers priors are given by the following vector:

$\Delta\bar{\mathbf{V}} = (\Delta\bar{V}_1, \Delta\bar{V}_2, \dots, \Delta\bar{V}_J)$. Then n_S is the number of non-negative elements of $\Delta\bar{\mathbf{V}}$. This means that n_S employers would call an S candidate when choosing between S and A . Thus $n_A = K - n_S$ will be the number of negative elements of $\Delta\bar{\mathbf{V}}$, in other words, the number of employers who would call an A candidate. The groups relative job-market chances can then be given by the relative call-back rate $R = \frac{n_S}{n_A}$.

Previous results from Sweden indicate call back ratios in favor of those with typically Swedish names (S); $R > 1$. This might correspond to a case when S is preferred to A at equal observed and unobserved merits that is taste-based discrimination or to statistical/error discrimination in favor of S applicants based on the assumption that the unobserved productivity related group characteristics of S are more valuable than those of A .

Advantages in Observables

Given that employers priors are characterized as above, we can define the *compensatory condition* for an employer with a preference for S candidates, $\Delta\bar{V}_j \geq 0$, as:

$$\Delta\bar{V}_j = \bar{V}_j(\bar{\mathbf{X}}, S, E_j[\mathbf{Z}_S^k|\bar{\mathbf{X}}]) - \bar{V}_j(\bar{\mathbf{X}} + \delta\mathbf{X}_j, A, E_j[\mathbf{Z}_A^k|\bar{\mathbf{X}} + \delta\mathbf{X}_j]) = -1. \quad (4)$$

where $\delta\mathbf{X}_j$ is an augmentation in observed merits for the employer j . The difference in observed merits $\delta\mathbf{X}_j$ as defined above should be seen only as an

augmentation of the observed merits of a group of candidates so that the groups are still perceived by employers to be competing in the same segment of the labor market. This is essential as job applications that deviate too much in terms of enhanced merits may be seen as distinctly overqualified for the position in question. It is important that both the enhanced and the regular application are perceived by employers as typical potential job candidates. This means that augmenting the merits of the disfavored group of job candidates with an amount $\delta\mathbf{X}$ will decrease the call-back gap depending on the fraction of employers holding priors $\Delta\bar{V}_j$ that can be compensated with the augmented merit. Any observed decrease in the call-back gap after augmentation of CVs therefore discloses to what degree negative employer priors can be re-evaluated. No significant change in call-back rates after augmented merits for the disfavored group, implies that employers evaluate candidates from S and A groups so differently that the augmented merits does not satisfy the *compensatory condition* (4).

Note that augmenting merits is one way to compensate for a lower call-back rate. Another method would be to increase the number of applications sent by the disfavored group in order to increase the probability of meeting employers with lower tastes for discrimination or less negative priors concerning unobserved group productivity, i.e., to employers with $\Delta\bar{V}_j \leq 0$. Disfavored applicants would have to increase the number of applications by the amount $\frac{n_S - n_A}{n_A}$ to overcome the call-back gap, under the assumption that the chances of being called to an interview is a linear increasing function of the number of applications.

One simple way of augmenting the observable merits of a job applicant is to increase the candidate's historical rate of success in obtaining work. Given that employers have no negative priors about A candidates, that is to say if $\Delta\bar{V}_j \leq 0$, a candidate in the disfavored group with an A -name that signals more years of experience than an S -candidate must be regarded

by employers as a more successful candidate (as long as the higher years of experience are in a comparable occupation). For the case of employers with priors against A , $\Delta\bar{V}_j > 0$, observing an A -candidate with superior merits may lead the employer to deduce that the candidate belongs to the upper part of the overall merits distribution and, depending on the nature of priors, this information may be enough to compensate for the previous group-difference in priors. Employers may react to higher observable merits but also to what these merits indicate about unobservable characteristics. Note that an additional year of experience for an A candidate is a stronger signal of ability than an additional year of experience for a S candidate when candidates from group A face difficulties in obtaining employment.⁶

3 Employer Priors, Ethnicity and Gender

The theoretical model introduced in the previous section discusses employer priors about different groups of job applicants in a general framework. To be more specific, the focus in this study concerns employer priors about individuals sharing names with a common linguistic origin: job-applicants with Swedish or Arabic Names. The choice of Arabic names is motivated by results in previous studies indicating lower call-back probabilities for job applicants with Arabic names in Sweden (Bursell (2007) and Carlsson and Rooth (2007)). Employers are assumed to use names as signals of ethnicity and base their employment decisions on both observed merits and priors concerning unobserved characteristics that are related to ethnicity. Negative priors about the unobserved productivity characteristics of applicants with Arabic names then translate to ethnic differences in call-backs.

A gender perspective is necessary as ethnically based employer priors may differ for female and male applicants. There are two potential reasons for this.

⁶See Meyer (1991) and Arai et al. (2001).

First of all, the labor market is segregated with respect to gender. As many jobs are typically female or male, employer evaluations of job applicants may differ by gender due to the nature of the job being applied to. The second issue concerns gender differences in unequal treatment. Fershtman and Gneezy (2001) report results from name-based experiments in Israel indicating that observed discrimination was toward male subjects and primarily practiced by males. Ahmed (2004) reports similar results for Sweden.⁷

Studies within social psychology show that stereotypes about a group are often closely correlated to the stereotypes about the men belonging to that group while the stereotypes about women from the same group may differ greatly from the group stereotype. Eagly and Kite (1987) empirically examined this hypothesis for 28 nationalities finding that national stereotypes are largely in line with the male stereotypes of that nation while the stereotype about women from the same nation may differ greatly from the male national stereotype. This is especially true when large gender differences in equality are included in the national stereotype. In such cases women are stereotyped according to general female stereotypes rather than specific national stereotypes.

For these reasons, the experiment carried out in this study allows for differences in employer priors by the gender of applicants. As described in detail in the next section, CVs in our field experiment are sent in female or male pairs and ethnic differences in call-backs examined separately for male and female applicants.

⁷This pattern is in line with previous studies indicating that immigrant women are less disfavored than immigrant men in the Swedish labor market; see le Grand and Szulkin (2002), Carlsson and Rooth (2007) and Arai and Skogman Thoursie (2008).

4 Experiment Design

The first stage of the experiment, a traditional correspondence test, was conducted between March 2006 and October 2007. Job applications were sent to job openings in the Stockholm metropolitan area advertised on Sweden's main internet-based employment site ("Platsbanken"). To ensure an authentic look, applications were designed in line with already existing applications that actual job applicants had up-loaded on the Platsbanken job-applicant-pool. We also consulted specialists within each occupation to review and critique our applications.

The applications were created as follows. When a job opening in one of the five occupations was announced on Platsbanken, two applications were constructed each consisting of a personal letter of introduction (cover letter) and a CV. Initially CVs were constructed to match the job requirements specified in the job announcement. Age, schooling and experience levels were therefore determined by the job announcement and set to be equal between any given pair of applications.⁸ Thereafter, the actual names of educational institutions and previous employers matching the levels set initially and of equal quality were randomly assigned to each CV. Two CVs sent to any give employer are therefore of equal observable quality but are not identical.⁹ Cover letters were formed based on random assignment of pre-written modules and were randomly matched with a CV to ensure not only variation in the applications, but personal letters with random design. Various addresses were then randomly assigned to each of the two job applications. Finally, before being sent to employers, ethnicity was randomly assigned each pair of job applications such that one application had a Swedish sounding name and

⁸Note that the CVs were never assigned lower education or experience levels than required by job advertisements. To increase the probability of call-backs, some CVs pairs were assigned higher levels of experience than required but only by at most one year of higher experience in order to avoid the risk of being perceived by employers as overqualified.

⁹Due to the random assignment of actual names of schools and employers, any subjective quality differences between applications should be eliminated over time.

the other a Arabic sounding name. As each pair of applications were also randomly assigned the same gender, the applicant pairs consisted of applications with a typical male Arabic name and a typical Swedish male name or likewise female names. The names used in the experiment are listed in Table 1. These names are easily recognized as Arabic or Swedish names.

Five types of occupations were targeted; computer specialists, drivers, business economists, senior high school teachers and assistant nurses. The positions as computer specialist, business economist and high school teacher all require four to five years of tertiary education and can therefore be seen as qualified positions. The positions as assistant nurses require secondary education with a medical profile while job announcements for drivers do not usually require any formal education, only valid drivers licenses. However as it is common in Sweden to have completed high school, applications to drivers were assigned high school degrees.

The second stage of the experiment was carried out between May and October 2007 with the same basic set-up as above with one major difference. The CVs with an Arabic sounding name were assigned higher levels of relevant previous work experience than the CVs with typically Swedish names. As in the first stage of testing, both CVs were initially constructed to match the experience requirements of the job announcement, thereafter the CVs with Arabic names were randomly assigned one to three years of extra relevant work experience. On average, the CV with the Arabic sounding name was therefore enhanced with two extra years of work experience. As age and experience are correlated and in order to produce credible CVs, age was also adjusted for enhanced CVs with Arabic names as follows. If one year of extra experience was assigned to the CV with Arabic names, no adjustment in age was made. When experience was adjusted with two years, age was adjusted upwards by one year and with three years of additional experience, age

was adjusted upward by two years.¹⁰ The maximum age difference between the two groups of job candidates in the second stage of the experiment is therefore two years. A limit of three years of higher experience was imposed in order to avoid considerable age differences between applicants and the possibility of one candidate being perceived as over-qualified, both of which would hinder the general comparability of applicant pairs. Note that the experiment design implies that only one stimulus, relevant work experience, is implemented when CVs with Arabic names are enhanced by only one year of experience.

The applications included an e-mail address and a cell phone number. When employers contacted the job applicants through e-mail or by the voice mail on cell phones a positive call-back was registered and the job interview offer was politely declined.

Descriptive statistics shown in Table 2 show small but significant age differences by gender in each stage of the experiment. Note that by design there are no ethnic differences in age by gender in the first stage of the experiment. There are however differences in the distribution of jobs applied to by gender and by experiment stage.¹¹ Any comparison of ethnic discrimination across experiment designs therefore need to adjust for differences in occupations applied to as well as possible time effects. In Stage 2 of the experiment when CVs with Arabic names are enhanced with more labor market experience, the mean age of female applicants with Arabic names is 26.5, significantly higher than the mean age for female applicants with Swedish names, 25.5. The mean age for male applicants with Arabic names is 25.8 which is also significantly higher than the mean age for male applicants with Swedish names, 24.8.

¹⁰The adjustment in age implies a slightly lower mean age in stage one of the experiment (24.7 years) than in stage two (25.6).

¹¹Note that by design there are no within gender ethnic differences in occupational distribution.

5 Results

In the first stage of the experiment 566 (283 CV pairs) observationally equivalent CVs (192 females and 374 males) were sent to jobs within the five chosen occupational groups (computer specialists, high-school teachers, nurses, economists and drivers). As seen in Table 3, a call-back gap between applicants with Arabic and Swedish names is found for both men and women. The relative call-back rate is 2.0 for females and 1.8 for men and the difference between genders is not statistically significant. As there is substantial variation in relative call-back rates across occupations these overall figures are not very informative. The smallest relative call-back rate, for example, is observed among teachers, both for male and female applicants and the largest ethnic call-back difference for male drivers (around 2.4). Notice that relative call-back rates by occupation are also sensitive to a relatively small number of observations within occupations.

In the second stage of the experiment, 584 CVs (292 CV pairs, 260 female CVs and 324 male CVs) were sent to employers where applications with Arabic names were enhanced by higher levels of experience. Results show that the call-back gaps changed dramatically. The relative call-back rate for women decreases to 1.2 and is no longer statistically significant, while the relative call-back rate for men increases to 2.7. The low relative call-back rate for females is observed in all occupations except for drivers. As noted above, the the distribution of jobs applied to by occupation differs across the two stages of the experiment. A difference in the relative weights of occupations may therefore generate differences in the overall relative call-back rates. To achieve relative call-back rates that are comparable to those in the first stage of the experiment, we re-weighted the relative call-back rates using share of occupations as weights. The corresponding relative call-back rates changed from 1.2 to 1.1 for females and from 2.7 to 2.4 for males.

In short, we observe a significant difference across experimental setups in female-male relative-call-back gaps for applicants with Swedish and Arabic names. Enhancing the CVs with Arabic names with on average two years of experience seems to increase the call-back probability for female applicants with Arabic names but does not improve the call-back probability for male applicants with Arabic names.

In order to control for potential differences in call-back rates between occupations and over time, linear probability models on call-backs (defined as a zero/one variable equal to one if applicants are contacted by employers) are estimated separately by gender. Two models are estimated for each stage of the experiment, the first controlling only for differences in names between applications, the second controlling also for occupation applied to and common time effects. As applications are sent over a period of several months, controlling for common time effects via time dummies, defined according to the date of application submission, is necessary. In addition, standard errors are clustered by date of application.

Results, presented in Table 4, show that in the first stage of the correspondence testing when CVs are observationally equivalent, there are significantly lower call-backs from employers for applicants with Arabic names. An Arabic name on an application is associated with, on average, a 20-21 percentage point lower probability of contact from employers than an application with a Swedish name (model 1). Lower call-back rates for applications with Arabic names are found for both male and female applicants. Adding controls for occupation applied to and common time effects yield similar results, applicants with Arabic names are associated with 23-24 percentage points lower call-back probabilities from employers, regardless of gender (model 2). These estimates are in line with results from previous correspondence tests in Sweden indicating that employers have negative priors regarding the unobservable productivity characteristics of job applicants with Arabic names

and/or tastes for discrimination against persons with Arabic backgrounds.

Estimations from the second stage of the experiment, when CVs with Arabic names are randomly enhanced with one to three years of relevant work experience, yield results indicating that ethnic differences in call-back rates for female applicants are eliminated. The coefficient for female applications with Arabic names is small and no longer significant implying no differences in call-backs from employers between female applications with Arabic names and more qualified CVs and female applicants with Swedish names and standard applications. For male applicants, enhanced work experience on applications does not alter previously reported differences in call-back probabilities. On average, a CV with a male Arabic name is still associated with a 27 percentage point lower call-back probability than a male applicant with a Swedish name, despite observationally higher levels of relevant work experience (model 1).¹² The call-back increases significantly (from model 1) when occupation and time dummies are included in estimation to a 39 percentage point difference (model 2). The difference in call-backs for male Arabic applicants is however not significant across experimental stages indicating that employer responses to male applicants with Arabic names are similar in both stages of the experiment. In summary, these results imply that employers react to higher merits or to what higher merits signal about unobserved productivity characteristics for female applicants with Arabic names but not for male applicants with Arabic names.

Results from separate estimations on applicant pairs by level of enhanced experience indicate that differences in call-backs probabilities disappear for women already when only one year of extra experience is assigned to appli-

¹²We recognize that employers may be reacting to higher age levels on CVs with Arabic names rather than higher levels of experience. The purpose of the second stage experiment is however to evaluate the strength on negative priors on unobserved characteristics and how these priors are adjusted when merits are enhanced, regardless of whether the enhancement is due to experience or age.

cations with Arabic names. As age is not altered on the CVs with Arabic names when only one year of extra work experience is added to the CV, i.e., age is equal across CV pairs, these results stem solely from employer responses to higher experience. For men, significant ethnic differences in call-back probabilities exist for each level of enhanced experience.¹³

We also run regressions on data from the second stage of the experiment including dummies for each level of experience-gap relative to applicants with Swedish names, i.e., one, two or three years of more experience, as well as interactions between the experience dummies and the (Arabic) name dummy.¹⁴ The results indicate that the main effect on call-backs for Arabic names remains insignificant for females and negative and significant for males. There is no significant call-back difference between various experience-gap levels nor do we observe significant effects of interactions between experience-gap dummies and the Arabic name dummy.¹⁵ These results indicate that one more year of experience eliminates call-back gaps for females but that increased experience thereafter do not further reduce call-back differentials.¹⁶

Pooled estimation on data from both stages of the correspondence testing including interactions between test stage (first or second stage of the experiment), names (Arabic or Swedish) and gender (female or male) as well as each partial interaction yield results confirming those reported above. The difference-in-difference-in-difference estimate, i.e., the effect of an Arabic name for females with enhanced CVs is positive and significant. Other results, reported in Table 5 show that the direct effect of an Arabic name on an application is a significantly lower call-back probability while the direct effect of an enhanced CV is positive and (weakly) significant (model 2).

¹³Results available from authors by request.

¹⁴Results available from authors by request.

¹⁵F- and t-statistics show that the null hypothesis that these variables are individually or jointly equal to zero cannot be rejected.

¹⁶Note that significance levels may be affected by relatively small sample sizes within each level of experience.

Finally, the interaction effect between Arabic names and enhanced CVs as well as the interaction effect between gender and enhanced CVs are both insignificant.

6 Conclusions

Using the correspondence testing methodology, this study analyzes to what degree employers revise negative priors observed against applicants with Arabic sounding names when CVs are enhanced with higher levels of previous work experience. In the first stage of the correspondence test, observationally equivalent CVs are sent to employers with only one difference, ethnic background as signalled by names. Applicants with Arabic sounding names are found to have significantly lower call-back rates regardless of gender. In the second stage of the experiment, CVs with Arabic sounding names are enhanced by on average two years of relevant work experience. Results for the second stage correspondence test show that significant differences in call-backs between female applicants are eliminated. No changes in call-back rates are however observed for men, despite enhanced CVs for the disfavored group (applicants with Arabic names). These results suggest that employers react to higher levels of merits and/or revise their negative priors concerning unobservable characteristics for female ethnic minorities but do not react to signals of higher previous employment success for male minorities.

These results contradict the widely held belief that women with foreign backgrounds suffer from both ethnic and gender discrimination in the labor market. Rather, the results reported here suggest that it is Arabic men that suffer most from discrimination as higher qualifications do not overcome the negative priors of employers concerning this group. These results are in line with a number of recent studies showing greater unequal treatment for male ethnic minorities than female ethnic minorities in Sweden.

Although more research is necessary to determine how generalizable these results are to other groups, occupations, and labor markets, reported results are compatible with studies within social psychology showing gender differences in stereotypes against different groups in society. Employers may have stronger negative priors against Arabic men than Arabic women simply because the negative stereotype about those with Arabic backgrounds is, to a large degree, a male stereotype. According to these studies, the stereotypes about women with Arabic backgrounds may largely be generated from traditional gender stereotypes that place women in domestic and nurturing roles. In short, these stereotypes suggest that an Arabic woman successful in the labor market may be perceived by employers as deviating from the stereotypic norms associated with Arabic woman. On the other hand, greater labor market experience may not alter the negative stereotypes associated with Arabic males. This implies that employers weigh the higher labor market experience of Arabic women as a signal of higher productivity, i.e., of an Arabic woman having overcome the traditional role ascribed to her, while little or no weight is attached to the higher labor market productivity of Arabic men.

One may object that the female applications in the first stage of the experiment also had high quality CVs implying that the pattern of a smaller call-back gap for women should be discernable already in stage one, which it is not. This objection however misses the point that weak negative priors may cause as large of a call-back difference as strong negative priors when merits are observationally equivalent.

In conclusion, the results in this study suggest that male and female members of an ethnic minority do not always face the same type of employer priors on unobserved characteristics. Employers in Sweden appear to have stronger negative priors concerning the unobservable characteristics of Arabic men or inflexible tastes for discrimination against Arabic men implying that individual investment in human capital enhancement may not alone counter

unequal treatment in the labor market and that other policy initiatives may be necessary to guarantee equal opportunity.

References

- Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30, 175–187.
- Arai, M., Billot, A. and Lanfranchi, J. (2001). Learning by helping: a bounded rationality model of mentoring. *Journal of Economic Behavior and Organization*, 45, 113–132.
- Arai, M. and Skogman Thoursie, P. (2008). Renouncing personal names: An empirical examination of surname change and earnings. Mimeo, Department of Economics, Stockholm University.
- Arrow, K. (1972). Some mathematical models of race discrimination. In A. Pascal (Ed.) *Racial Discrimination in Economic Life*. Lexington: Lexington Books.
- Arrow, K. (1973). The theory of discrimination. In O. Ashenfelter and A. Rees (Eds.) *Discrimination in Labor Markets*. Princeton University Press, Princeton.
- Becker, G. S. (1957). *The Economics of Discrimination*. Chicago, University of Chicago Press.
- Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? : A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94, 991–1013.
- 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*.
- Carlsson, M. and Rooth, D.-O. (2007). Evidence of ethnic discrimination in the swedish labor market using experimental data. *Labour Economics*, 14, 716–729.
- Eagly, A. H. and Kite, M. E. (1987). Are stereotypes of nationalities applied to both women and men? *Journal of Personality and Social Psychology*, 53, 451–462.
- England, P. (1992). *Comparable Worth, Theories and Evidence*. Aldine de Gruyter, New York.

- Fershtman, C. and Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics*, 116, 351–377.
- Firth, M. (1981). Racial discrimination in the british labour market. *Industrial and Labor Relations Review*, 314, 265–272.
- Fryer, R. G. and Levitt, S. D. (2004). Causes and consequences of distinctively black names. *Quarterly Journal of Economics*, 119, 767–805.
- Heckman, J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, 12, 101–116.
- Heckman, J. J. and Siegelman, P. (1993). The urban Institute audit studies: Their methods and findings. In M. Fix and R. Struyk (Eds.) *Clear and convincing evidence: Measurement of Discrimination in America*. Washington D.C.: Urban Institute Press.
- Jowell, R. and Prescott-Clarke, P. (1970). Racial discrimination and white-collar workers in britain. *Race and Class*, 11, 397–417.
- le Grand, C. and Szulkin, R. (2002). Permanent disadvantage or gradual integration: Explaining the immigrant-native earnings gap in sweden. *Labour*, 16, 37–64.
- Meyer, M. A. (1991). Learning from coarse information: Biased contests and career profiles meyer. *Review of Economic Studies*, 58, 15–41.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62, 659–661.
- Riach, P. A. and Rich, J. (1991). Testing for racial discrimination in the labour market. *Cambridge Journal of Economics*, 15, 239–256.
- Riach, P. A. and Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal*, 112, 480–518.
- Schröder, L. (2007). From problematic objects to resourceful subjects: An overview of immigrant-native labour market gaps from a policy perspective. *Swedish Economic Policy Review*, 14, 7–40.

Table 1: Names of applicants used in the experiment

First Name	Surname	First Name	Surname
Fateme	Ahmed	Kamal	Ahmadi
Nasrin	Hassan	Abdallah	Mohammed
Halima	Mohammadi	Islam	Hashemi
Aïcha	Abdallah	Abdelaziz	Hussein
Fatima	Ahmad	Abdelhakim	Hassan
Sara	Andersson	Jonas	Söderström
Marie	Björkvist	Erik	Östberg
Johanna	Gustafsson	Johan	Nyström
Karolina	Svensson	Mikael	Andersson
Malin	Wallin	Martin	Berggren

Table 2: Descriptive Statistics, by Gender. Standard errors in parantheses

	Equivalent CVs		Enhanced (Arabic Name) CVs	
	Female	Male	Female	Male
Arabic	0.5	0.5	0.5	0.5
Age	24.3 (2.93)	24.9 (2.76)	26.0 (2.49)	25.3 (2.56)
Call-back	0.26	0.32	0.32	0.29
Computer Specialist	0.16	0.24	0.21	0.18
Driver	0.07	0.27	0.10	0.39
Economist	0.21	0.20	0.36	0.21
High School Teacher	0.23	0.16	0.15	0.08
Assistant Nurse	0.33	0.13	0.18	0.15
No. of observations	192	374	260	324

Table 3: Call-backs, by Gender

	Equivalent CVs		Enhanced (Arabic Name) CVs	
	Females	Males	Females	Males
Both Invited	22	70	56	46
Only Arabic Name Invited	12	16	20	6
Only Swedish Name Invited	46	88	36	92
Neither invited	112	200	148	180
Relative Call-Back Rate	2.0	1.9	1.2	2.7
Number of applicants	192	374	260	324

Table 4: Call-back Probabilities for Arabic Named CVs in Comparison to Swedish Named CVs (Linear Probability Models).

Equivalent CV				
	Females	Males	Females	Males
Arabic	-0.208*	-0.199*	-0.233**	-0.239*
	(0.062)	(0.049)	(0.101)	(0.072)
Occupation	NO	NO	YES	YES
Date	NO	NO	YES	YES
N	192	374	192	374
Enhanced Arabic CV				
Arabic	-0.062	-0.265*	0.041	-0.388*
	(0.067)	(0.053)	(0.081)	(0.074)
Occupation	NO	NO	YES	YES
Date	NO	NO	YES	YES
N	260	324	260	324

Note: * and ** denote significance at the one and five percent level. Estimations control for four occupation dummies and 73 date of application dummies. Standard errors in parentheses are clustered by the date of application.

Table 5: Difference-in-Difference-in-Difference Estimates (Linear Probability Models).

	Model 1	Model 2
Female×Arabic×Enhanced	0.214*	0.277*
	(0.099)	(0.117)
Female	-0.047*	0.010
	(0.129)	(0.075)
Arabic	-0.197*	-0.210*
	(0.049)	(0.066)
Enhanced CV	0.071	0.658*
	(0.124)	(0.183)
Female × Arabic	-0.010	-0.072
	0.073	(0.089)
Female × Enhanced	-0.228	-0.112
	0.177	(0.101)
Arabic × Enhanced	-0.068	-0.062
	0.072	(0.091)
Occupation	NO	YES
Occupation × enhanced	NO	YES
Date	NO	YES
N	1150	1150

Note: * denote significance at the five percent level. Estimations control for four occupation dummies and 73 date of application dummies. Standard errors in parentheses are clustered by the date of application.