

**ERASMUS UNIVERSITY ROTTERDAM**

**ERASMUS SCHOOL OF ECONOMICS**

**MSc Economics & Business**

**Master Specialization Economics of Management and Organization**

**Anonymous job applications in the Netherlands:  
A reconsideration of the The Hague experiment  
and a proposal for future experiments**



**Author:** J. Cuperus  
**Student number:** 347474  
**Thesis supervisor:** Dr. A. J. Dur  
**First reader:** Dr. S. V. Kapoor  
**Finish date:** August 12, 2017

## ABSTRACT

Anonymous Application Procedures (AAP) are a commonly used method to reduce discrimination in hiring workers. This topic became more and more prominent in the last years. In 2016, experiments about AAP have also been conducted in The Hague and Utrecht (The Netherlands). This study provides some additional research using the data from the experiment in The Hague. In contrast to the original study, I control for possible differences in the magnitude of hiring discrimination in jobs with different salary scales in different departments and with or without a managerial aspect. Furthermore, some power tests are performed in order to calculate the required sample size for a proposed new, large-scale RCT experiment in multiple municipalities in the Netherlands.

**Keywords:** Anonymously applying, AAP, hiring discrimination, salary scale, managerial function, power test, RCT experimental design.

---

<sup>1</sup> I would like to thank Robert Dur and Sacha Kapoor for their useful insights and comments on this research.

# TABLE OF CONTENTS

ABSTRACT .....	2
TABLE OF CONTENTS .....	3
LIST OF TABLES .....	6
LIST OF FIGURES .....	6
<b>1. Introduction.....</b>	<b>7</b>
1.1. Anonymously applying .....	7
1.2. Psychological cause of hiring discrimination.....	7
1.3. Earlier studies about hiring discrimination .....	8
1.4. The idea behind AAP as method to combat hiring discrimination .....	9
1.5. The main purpose of the study .....	10
1.6. The structure of the paper .....	10
<b>2. Some General Background.....</b>	<b>11</b>
2.1. Main findings of AAP experiments outside the Netherlands.....	11
2.1.1. Combating hiring discrimination through AAP .....	11
2.1.2. Critical notes on AAP .....	12
2.1.3. Differences in hiring discrimination through level of jobs.....	13
2.1.4. Postponed hiring discrimination .....	13
2.1.5. The methodologies for implementing AAP .....	14
2.1.6. Incentive firms for implementing AAP.....	14
2.2. Main findings of the AAP experiments in The Hague and Utrecht .....	15
2.2.1. Design and details about the AAP experiments .....	15
2.2.2. Request for implementing AAP in the Netherlands .....	17
2.2.3. Evaluation about AAP experiment.....	17
2.2.4. Results AAP experiment in The Hague .....	18

<b>3. Additional research about AAP experiment in The Hague .....</b>	<b>20</b>
3.1. Data description .....	20
3.1.1. Transformation of variables.....	21
3.1.2. Summary statistics.....	22
3.2. Analysis.....	25
3.2.1. Correlations .....	25
3.2.2. Regressions applying rate .....	26
3.2.2.1. Results regression ‘Bicultural Dutch candidates’.....	27
3.2.2.2. Results regression ‘Turkish, Moroccan, Surinamese or Antillean Dutch candidates’ .....	29
3.2.2.3. Results regression ‘Other Non-Western Dutch candidates’ .....	30
3.2.2.4. Results regression ‘Western Dutch candidates’ .....	32
3.3. Regressions total number of candidates .....	33
3.4. Summary findings.....	35
<b>4. Power tests .....</b>	<b>36</b>
4.1. Introduction power calculations .....	36
4.2. Theory power tests.....	37
4.2.1. Theory power test on single regressor .....	38
4.2.2. Theory power test on fit of a multiple linear regression .....	38
4.3. Results of the power tests.....	39
4.3.1. Power tests on single regressor .....	39
4.3.2. Power tests on the fit of the multiple linear regression.....	40
4.4. Required sample size calculation based on power calculations and previous research .....	41
<b>5. Proposal and design for experiment in the Netherlands on large scale.....</b>	<b>43</b>
5.1. Description proposal and design.....	43
5.2. Final remarks proposal and design.....	45

**6. Conclusions ..... 46**

REFERENCES..... 47

APPENDIX A ..... 50

APPENDIX B..... 50

# LIST OF TABLES

Table 1. Departments of the municipality the Hague .....16

Table 2. Summary statistic table ads .....22

Table 3. Independent sample t-test: difference between 2015 and 2016.....23

Table 4. Summary statistic table candidates .....24

Table 5. Correlations .....26

Table 6. Regressions on the percentage Bicultural Dutch .....28

Table 7. Regressions on the percentage of subcategory Turkish, Moroccan, Surinamese or Antillean Dutch .....30

Table 8. Regressions on the percentage Other Non-Western Dutch.....31

Table 9. Regressions on the percentage Western Dutch.....33

Table 10. Regression on total number of applicants .....34

Table 11. Overview relation between the probabilities  $\alpha$ ,  $\beta$  and power .....36

Table 12. Power of the single regressor AAP.....40

Table 13. Power of the fit and required sample size .....40

Table 14. Overview power and required sample size earlier AAP studies .....41

# LIST OF FIGURES

Figure 1. Graphical relation between type I and II error and power .....37

Figure 2. Power test for calculating the required sample size in proposed experiment.....42

Figure 3. Overview design experiment on total level .....44

Figure 4. Overview design experiment on municipality level .....45

# **1. Introduction**

## **1.1. Anonymously applying**

This decade anonymously applying for jobs has become a prominent and popular subject of discussion. However, it is still in the beginning stage. Only a few researchers from the USA, Sweden, France, The Netherlands and Germany studied specifically Anonymously Applying Procedures (AAP) to investigate hiring discrimination. Anonymously applying is applying for a job without the recruiter knowing personal information such as name, age, marital status, number of children, photo, origin and / or gender of the candidates. However, not all personal information needs to be excluded in most cases. The specific kind of information that will be excluded from the application forms, and therefore will not be available for the recruiters in the decision-making process, depends on the kind of discrimination that needs to be combated.

## **1.2. Psychological cause of hiring discrimination**

Before investigating hiring discrimination, it is worth stating that it is desirable for the employer to select the most qualified candidate for the job. Even with the intention to do so, employers do not always hire the best man for the job, due to both explicit and implicit biases. Devine, Plant, Amodio and Harmon-Jones (2002) argued that 'cognitive' explicit bias is consciousness and therefore controllable. Implicit bias is unconsciousness, has an ingrained nature and is therefore much more complex than explicit bias. A person might not be able to think of all different persons as equals, but society requires us to show respect to every single individual. Anonymous Application Procedures (AAP) are a method that reduces hiring discrimination by eliminating the implicit bias when determining which candidate gets an interview invitation. According to Kraus, Rinne and Zimmerman (2012), AAP imply that candidates are invited for an interview based on qualifications, and not based on other characteristics of candidates, such as looks, age, gender or ethnical origin. AAP lead, according to Aslund and Skans (2007), to more transparency, objectivity and equal chances during the recruitment phase and is an important step forward towards a decision-making process without discrimination. When someone has been disadvantaged by hiring discrimination, this can have an enormous impact on personal lives, since it reduces the probability of having a paid job and the probability to move up to higher positions in an organization. Through the literature, two methods are used in order to get application forms fully anonymized before handing them over to the recruiter(s). The first method is letting the candidates fill in a standardized form, which will be handed over to the recruiter afterwards. The second method is letting someone check the original resume and motivation letter and deleting sensitive information that could lead to hiring discrimination before handing the application forms over to the recruiter. These two methodologies of anonymization will be discussed later in more detail.

### 1.3. Earlier studies about hiring discrimination

Discrimination exists in a large range of consumer markets. Yinger (1998) researched discrimination through ethnicity in the search for housing or buying a car in Chicago. Yinger (1998) found that black and Hispanic households face significantly more discrimination than other races when closing a deal in purchasing a house or car. Despite anti-discrimination campaigns and the potential for hefty fines<sup>2</sup>, labor market discrimination remains a major issue. The focus of this paper is ethnic or racial discrimination in classical recruitment procedures. For example, Pager, Bonikowski, and Western (2009) argued that African Americans in New York were half as likely as whites to receive a job interview invitation after applying. McGinnity, Nelson, Lunn and Quinn (2009) studied hiring discrimination in Ireland and drew a similar conclusion: the callback rate for Irish applicants was twice as high as candidates of African, Asian, or German origin. In the Netherlands, the Social and Cultural Planning Office investigated hiring discrimination, where the focus was in particular on the magnitude and the nature of hiring discrimination. Andriessen, van der Ent, van der Linden and Dekker (2015) investigated whether employers are more interested in hiring native Dutch candidates than candidates with a Moroccan or Surinamese Hindustan background with similar qualifications. Fictitious candidates, who only differed in ethnical background, responded on the same, real ad. The name was the only signal for recruiters to 'recognize' the ethnical background of the candidates. Andriessen, van der Ent, van der Linden and Dekker (2015) concluded that native Dutch candidates had a fifty percent higher callback rate than Surinamese Hindustan Dutch candidates with the same qualifications, and an eighty percent higher callback rate than Moroccan-Dutch candidates with the same qualifications. No significant differences in callback rates were found through gender. According to Andriessen, van der Ent, van der Linden and Dekker (2015), the reason for hiring discrimination in ethnical groups is image formation of minorities. Pérez, Fortuna and Alegría (2008) stated that the link between discrimination and image formation of minorities can be explained by sociocultural differences (psychological mechanism). Aigner and Cain (1977) stated that this relation can be explained by the image formation that foreigners have a lower productivity and are therefore a liability (economical mechanism). The psychological mechanism is investigated by Andriessen, van der Ent, van der Linden and Dekker (2015) by adding a paragraph in the resume of the (fictitious) Moroccan or Surinamese Hindustan Dutch candidate that showed clearly (1) expressed involvement in the Dutch society (volunteer work) and (2) identification as a Dutch person. Andriessen, van der Ent, van der Linden and Dekker (2015) showed that the discrimination of Surinamese Hindustan Dutch candidates disappeared when including the additional paragraph to the resume, what implies that the callback rate does not significantly differ between native Dutch candidates and Surinamese Hindustan Dutch candidates who do show additional involvement in society and identification as Dutchmen-or women. For Moroccan Dutch candidates hiring discrimination is still present, even after (partly) overcoming sociocultural differences. The economical mechanism is investigated by adding (1) two years of extra work experience, (2) successfully completed relevant courses and (3) a paragraph about motivation and dedication to the (fictitious) Moroccan or Surinamese Hindustan Dutch candidates'

---

<sup>2</sup> After all, race discrimination is against the law. Ying (1998) referenced to the *Civil Rights Act of 1964: Title VII (Equal Employment Opportunities)*. It prohibits employment discrimination based on race, color, religion, sex, and national origin.



resumes. The result was that the callback rate of Moroccan Dutch candidates with the abovementioned adjustments in their resumes is not significantly different than native Dutch candidates without the adjustments in their resumes. For Surinamese Hindustan Dutch candidates the hiring discrimination is still present, even after partly overcoming the low productivity preconception. In conclusion, employers would like to have more certainty about the productivity of future Moroccan Dutch workers before hiring them and image formation about sociocultural differences plays a role in the hiring process for the Surinamese Hindustan Dutch candidates.

Another Dutch research about hiring discrimination was conducted by Blommaert, Coenders and van Tubergen (2014), who have studied the discrimination of Arabic-named applicants in different phases of an online recruiting procedure. Blommaert, Coenders and van Tubergen (2014) provided strong evidence of discrimination in the first phase, where the recruiter had to decide whether viewing candidate's complete resume after seeing a characteristic preview or not. Dutch-named candidates are fifty percent more likely to be viewed than Arabic-named candidates. However, Arabic-named candidates did not get significantly less positive reaction than Dutch-named candidates when the complete resume was viewed (phase two).

#### **1.4. The idea behind AAP as method to combat hiring discrimination**

The field of discrimination in hiring decisions is nowadays an upcoming, fast-growing scientific topic. The concept of anonymously applying is however not that new: last decade lots of single blind, double blind research or 'blindfold' experiments have been conducted (Goldin & Rouse, 2000; Blank, 1991; Siles, Hanson & Lindon, 1994). All these studies do have one thing in common: the result of the decision-making is often different when only the skills and competence are included in the evaluation and no other characteristics can influence the evaluation. Reducing bias in hiring is easiest achievable in the initial stages of the hiring process, since no personal contact is required in this stage which makes excluding bias-sensitive information easier. This can be useful, since a brief glance at an applicant's name, gender or age can be enough to form a stereotypical image of the person without meeting the candidate, with possible excluding from the next stage in the hiring process as a result. These biases affect most often people with a migrant background, women (with children) and older workers (Rinne, 2014). AAP are one of the measures that can be used to address labor market discrimination and promote an inclusive labor market. Discrimination in hiring decisions is an undesirable bias in terms of recruiting, since it is in the employers best interest to hire the most qualified employee for the job, independently from candidates' characteristics like gender, age, race, religion, ethnicity, sexual orientation or disabilities. A possible explanation for this bias is elucidated by the theory of taste based discrimination. Becker (1957) modeled discrimination as a personal prejudice or taste against associating with a particular group. Taste based discrimination is acting like there is a non-pecuniary cost of associating with a particular group, which can be reduced by AAP and may lead to hiring a more qualified worker. A few researchers (Finkelstein, 1966; Phelps, 1972 and Tomaskovic-Devey & Skaggs, 1999) provided evidence that the theory of statistical discrimination holds in practice. Statistical discrimination can also be reduced by AAP, but it is not certain that AAP lead to hiring more qualified workers, since important, but sensitive, information could not be used in the

hiring procedure. The main assumption of the statistical discrimination theory is that employers are less willing to provide women and minorities a job or paying women and minorities less since their productivity is lower on average. The term statistical refers to the fact that stereotypes are based on the discriminated group's average behavior. Fang and Moro (2011) also supported the statement that racial or gender inequality does even exist and persist between demographic groups when economical agents are rational and not prejudiced.

### **1.5. The main purpose of the study**

In this paper, based on the research about anonymously applying in The Hague, the difference in hiring discrimination between job types will be determined. In the investigation of the relation between AAP and applying behavior through minorities, I control for possible differences in the magnitude of hiring discrimination in jobs with different salary scales, in different departments and with or without a managerial aspect. I expect that AAP have a smaller effect on applying behavior in job ads with a higher salary scale that concerns a managerial function. These further investigations are important, since it makes clear which minority group in what situations faces most hiring discrimination and what departments / specific job vacancies characteristics need the most attention in the attempt to reduce the hiring discrimination. This research shows that hiring discrimination is present in the municipality The Hague, in particular for the Turkish, Moroccan, Surinamese or Antillean Dutch migrants. On average, the higher the salary scale, the less discrimination the candidates face. If the job has a managerial component, the hiring discrimination is also lower, but only significantly lower for other Non-Western migrants. The changes in composition of jobs offers through time, salary scale, department and whether a job had a managerial aspect, but also other factors could partly explain the change of the percentage applying candidates of a minority group in The Hague. The total impact can therefore not only be assigned to AAP (which is a year dummy in the The Hague study).

### **1.6. The structure of the paper**

The structure of this paper is as followed. In the next chapter, a general overview of the earlier findings about the impact of anonymously applying on applying behavior and hiring decision making will be given. First, earlier research outside The Netherlands will be discussed, followed by an overview of results of the experiment in The Hague and Utrecht. In chapter 3, the data description and the analysis of the additional research about the anonymously applying study in the Hague is presented. In chapter 4, several power tests are performed in order to calculate the minimal required sample size for the proposed design for a large-scaled experiment in the Netherlands, which can be found in chapter 5. In chapter 6, this research will be concluded.

## **2. Some General Background**

### **2.1. Main findings of AAP experiments outside the Netherlands**

#### **2.1.1. Combating hiring discrimination through AAP**

Hiring discrimination may occur based on many characteristics of employees, for example gender, age, race, religion, ethnicity, sexual orientation or disabilities. In this paper, the main concern is ethnical / race discriminations. Bertrand and Mullainathan (2004) studied the effects of race on callback rate for an interview. Bertrand and Mullainathan (2004) sent similar resumes to newspapers in Boston and Chicago in reaction to help-wanted ads, one with a typically African-American (black) name, like Lakisha Washington or Jamal Jones, and one with a typically white name, like Emily Walsh or Greg Baker. Bertrand and Mullainathan (2004) found that resumes with typical black names were less likely to get a call for an interview, whether those fake resumes contained higher qualifications or not. Bertrand and Mullainathan (2004) concluded that African-Americans received different returns to resume quality than white Americans in the sense that African-Americans may face a lower incentives to invest in higher skills than white Americans. The results did not vary across different occupations, industries and employer size, where the level of required competence and pay may differ.

Kaas and Manger (2012) studied ethnic discrimination in Germany's labor market with a correspondence test as well, which implies two identical applications for a student traineeship with only different names. Kaas and Manger (2012) found that candidates with a Turkish name are much less likely to be invited for an interview than candidates with a German name with further identical resumes. These results are also consistent with the findings of Drydakis and Vlassis (2007), who analyzed the labor market opportunities of Albanians in Greece, and Carlsson and Rooth (2006) who analyzed those difference between Sweden and 'Middle Eastern' Sweden. Kaas and Manger (2012) also claimed that the anonymous applying effect is even 75 percent larger in small-and medium-sized firms in Germany. The effect is however significant, but not large in big firms in Germany. Kaas and Manger (2012) explained this finding by stating that the sample contained mostly well-educated students and two-or third generation Turkish German persons, where the foreign influence could be minimal. This is in line with Rinne (2014), who stated that anonymously job applying has the potential to reduce discrimination significantly only when discrimination is high.

Behaghel, Crépon and Le Barbanchon (2012) described an experiment in France to determine possible discrimination in gender and racial group in recruiting processes. The experiment involved about 1.000 firms in eight local labor markets and lasted about 10 months. Behaghel, Crépon and Le Barbanchon (2012) provided evidence that the callback rates of migrants were lower with AAP than with the standard procedure. However, Behaghel, Crépon and Le Barbanchon (2012) argued that this (unexpected) effect could be caused by the fact that application documents were not fully anonymous. Bicultural candidates could be identified based on the address of their school or language skills, which may have biased the result. The firms who were willing to participate to the anonymous recruitment program, gave more often migrants a chance than other firms in the

standard procedure condition for reputation issues. The most plausible explanation for this (unexpected) effect is that anonymously applying takes away the possibility to discriminate positively. Behaghel, Crépon and Le Barbanchon (2012) also referred to the John Henry effect, which implies that participating recruiters in the control group try hard to behave like the recruiters in the treatment group. Some candidates, in particular from minority groups, could be encouraged to apply, since disadvantaged groups could think that their chances for the job increase with anonymously applying. Without anonymously applying, minorities could choose not to apply because of the fear to be discriminated against. According to Hoogendoorn and van Praag (2012), another positive side effect of diversity in the workplace is the increasing productivity, which is always in line with the policy of the firm.

### **2.1.2. Critical notes on AAP**

Aslund and Skans (2007) noted that also highly relevant parts of an application may carry information about gender, ethnicity or any type of sensitive information. In the attempt to reduce hiring discrimination, also these relevant parts are left out in the documents before handing them over to the recruiter during AAP. Even a prestigious foreigner, who graduate with high grades and credentials from a good college and university at his / her own country does not get credits for that in the Anonymously Applying Procedures. Context-specific information may be interpreted disadvantageously when the origin of an applicant is crucial for understanding the full picture. Also, according to Rinne (2014), structural differences in skill or qualifications between migrants and autochthones could be present. The opportunities for education are not equal between the native Dutch and minorities, which could lead to different callback rates for minority candidates, even without knowing the origin of the candidate. Anonymous applying focused within the hiring process on skills and qualifications of the candidate. Unfortunately, if other types of discrimination in society lead to differences in the competence of an applicant, anonymously applying cannot solve this problem. AAP are unfortunately not able to tackle the underlying problems of educational difference and the associated differences in work experience.

Krause, Rinne and Zimmerman (2012a) mentioned that no reform or law exist what makes anonymously job applying obligated for firms, and claimed that a reform or law should not be necessary since most firms are willing to hire the best man for the job. Rinne (2014) also argued that decision-making only based on skills and qualifications, should lead to hiring the most capable and productive worker, which is in line with any firm policy. However, if a mindset or certain appearance is important for a particular firm, the screening filter for recruiting for these firms is not going to chance. Mentionable is the fact that according to Kaas and Manger (2012), some firms in Germany have taken AAP over as a standard for their screening process in hiring decision due to a good experience or an image issue.

### **2.1.3. Differences in hiring discrimination through level of jobs**

The degree of hiring discrimination differ through the level of jobs. For low-level jobs, the market is often oversaturated, which implies that too many employees are available with enough credentials to perform well. In this type of market discriminating does occur quite often. For high-level jobs, the market is often unsaturated. Kaas en Manger (2012) stated that competition for hiring high-skilled employees is intense, since discriminating firms cannot survive the “war for talents” and will be driven out of the market. Therefore, according to Kaas and Manger (2012), hiring discrimination is higher for low-level jobs than in high-level jobs. The main difference in applying behavior for low-level jobs and high-level jobs is that for high-level jobs more specific knowledge, skills and capabilities are required. This can be indicated by having labor market experience, high educational level and less blank years in the employment history. Also completing certification degrees, acquiring some foreign language skills or awarding some honors skew the chances in your favor. Kuhn and Shen (2013) studied gender- and age discrimination in internet job advertisements in China. They also found that the higher the education level, responsibility and competence is required for a job, the less hiring discrimination candidates face.

Krause, Rinne and Zimmermann (2012b) investigated anonymous job applications of Ph.D. economists in the academic job market and concluded that age and ethnicity did not significantly influence the invitation probability for a job interview. Taken into account that these ads were for high level jobs, it is reasonable that the anonymously applying effect is small and are therefore not significant. Furthermore, Krause, Rinne and Zimmermann (2012b) found evidence for gender discrimination in the sense that the probability that women are invited for a job interview is higher than for men with standard applying procedures , while the AAP group did not show this difference. This can be explained by the fact that the ratio male-female is quite high in the academically job market, which makes hiring women more likely.

### **2.1.4. Postponed hiring discrimination**

While, amongst others, Bertrand and Mullainathan (2004) studied the effect of race on the interview offer probability, Kang and Banaji (2006) studied the effects of race on the job offer probability. The main difference between the interview offer probability (callback rate) and job offer probability in terms of anonymously applying is simply the fact that the anonymously applying can influence the callback rate, but job offer probability is based on a job interview, which does not ensure anonymously for the candidate anymore. In the study of Kang and Banaji (2006), white colleges students watched recorded parts of interview of black and white candidates with similar credentials for the job. Kang and Banaji (2006) concluded that these white college students choose significantly more often the white candidate over the black candidates for the job. More interesting, Kang and Banaji (2006) claimed that this bias also affect the interviews indirect by making interactions awkward and leading the candidate to perform worse at the job applying. Aslund and Skans (2007) used standardized anonymously application forms in order to detect the difference in gender and minorities in

both interview offer probability and job offer probability as a result of AAP. The experiments were held in Goteborg (second to the largest city of Sweden) from 2004 till 2006. Aslund and Skans (2007) found that the probability of being interviewed and also the probability to get hired did converge between men and women when the standardized anonymously application forms were used. Furthermore, the racial minority experienced a higher call rate for the job interview when the application was anonymous, but not significantly more migrants got hired in the end. The job offer probability was not significantly higher during AAP. This is exactly what Rinne (2014) meant when he stated that anonymously applying may simply postpone discrimination in the hiring process. In the phase of job interviews, ensuring anonymity is no longer feasible, so hiring discrimination could simply occur in a later phase of the hiring process.

#### **2.1.5. The methodologies for implementing AAP**

In the introduction, two main methodologies to implement AAP are mentioned. The first method is using a standardized application form and the second method is checking resume and motivation letter and manually deleting the sensitive parts that could lead to hiring discrimination. Aslund and Skans (2007) and Krause, Rinne and Zimmermann (2012a) used standardized anonymously application forms and Moha and Konings (2016)<sup>3</sup> used the method where sensitive parts were manually removed from the application. According to Krause, Rinne and Zimmermann (2012a), a standardized application form makes it easier for recruiters to review and compare the applications. The candidates can be narrowed down faster by focusing on some important criteria which can be found at the same location on each application form what saves time. Moha and Konings (2016) argued that the more standardized the application forms are, the less candidates can sell their self and present their skills and capabilities in their own and unique way. The way candidates present themselves do say something valuable about them. The usage of standardized application forms is a very efficient method, but only when an optimal form is developed. Suboptimal forms can be costly, time- and labor consuming, and error-sensitive, but this can also be the case for manually withdrawing sensitive information.

#### **2.1.6. Incentive firms for implementing AAP**

In line with several researchers above, Hausman (2012) also concluded that anonymously applying decreases discrimination and helps firms with hiring more productive workers. However, Hausman (2012) argued that this conclusion is counterintuitive in the sense that firms need an incentive to do so. Therefore Hausman (2012) proposed an incentive: reducing liability insurance premiums for recruiters who hire new employees anonymously. This policy would not only reduce discrimination in hiring through gender or racial minorities, but also religion, sexuality, weight, size or even attractiveness, without the law to protect the candidates directly. In imitation of Kang and Banaji (2006), Hausman (2012) suggested that firms should

---

<sup>3</sup> Publication of the investigation in municipality The Hague, discussed in paragraph 2.2:  
[http://west.storage.regiogrid.nl/sitemanagerdata/uploads/GemeenteDenHaagPilotAnoniemSolliciterenEindrapportage13\\_VoorPublicatie.pdf](http://west.storage.regiogrid.nl/sitemanagerdata/uploads/GemeenteDenHaagPilotAnoniemSolliciterenEindrapportage13_VoorPublicatie.pdf).

anonymize resumes and eliminate interviews, since most of the hiring discrimination is not intentional and interviewing is a poor measure for predicting job performance. Employers are often conservative though and claim that interviews are necessary for good hiring decision making. However, a few dated and often ignored psychological researchers showed that in-person and telephone interviews have a much worse prediction performance than statistical evaluation (Sarbin, 1943; Gaertner, 1990). Bertrand and Mullainathan (2004) claimed that also implicit bias contaminates selection interviews, which implies that recruiters are more likely to select candidates with corresponding characteristics. Behaghel, Crépon and Le Barbanchon (2012) and Kang and Banaji (2006) supported this statement. Behaghel, Crépon and Le Barbanchon (2012) found evidence that recruiters were likely to select candidates with the same gender within standardized applying procedures and this effect disappeared when sensitive parts were in fact excluded from the applications. Kang and Banaji (2006) concluded that white college students choose significantly more often the white candidate than an ethnical minority candidate for the job. Downsides to this proposal are the facts that the incentive will provide the largest payoff for the most discriminated firms (adverse selection) and liability insurance is the most attractive instrument for firms with the most diverse workplace.

To sum up, in most studies AAP lead to an increase of the interview invitation probabilities of subordinated groups. Anonymously applying could have the side effect of delaying the hiring discrimination effect to a later stage of the process, which could explain why the job offer probabilities of minorities is often not significantly higher for bicultural Dutch applicants when AAP are implemented and the call back rate is in fact significantly higher. The effect of AAP is less for high-level job ads.

## **2.2. Main findings of the AAP experiments in the Hague and Utrecht**

### **2.2.1. Design and details about the AAP experiments**

As mentioned in paragraph 2.1., two experiments have been conducted in The Netherlands about hiring discrimination by implementing AAP, one in The Hague and one in Utrecht. In both towns, various signals about increasing discrimination, in particular on grounds of origin and religion, were circulating. Both experiments were performed from 2015 to 2016 and the documentation was published at the end of 2016 by Moha and Konings (2016). The duration of the experiment in Utrecht was four months, from January 1, 2016 till May 1, 2016 and the duration of the experiment in The Hague was six months from January 1, 2016 till June 30, 2016. The data of 2016 were compared with the job application data from a year earlier: January 1, 2015 till May 1, 2015 in Utrecht and January 1, 2015 till June 30, 2015 in The Hague. By comparing year-to-year, seasonal influences on the job applications and candidates are excluded.

The biggest and most important difference between the experiment in Utrecht and The Hague is the magnitude. In Utrecht only eight job ads in salary scales 10 till 14 were published in 2016 and only 215 candidates applied. The sample in the experiment in Utrecht is therefore unfortunately too small to state anything meaningful. The experiment in The Hague contained 36 job ads and 1.576 candidates in 2015 and 57

job ads and 1.880 candidates in 2016, which is more suitable for a few analyses. With this data, however, no causal effects can be estimated, only correlation research can be done, since it does not concern a fully randomized research with control group (RCT). Another difference is the fact that in municipality The Hague resumes and motivation letters were anonymized manually and in Utrecht a standardized anonymously applying format was used. Krause, Rinne and Zimmermann (2012a) claimed that the use of a standardized applying format is less labor- and time intensive than manually remove sensitive parts of a resume and / or motivation letter and is at least even practical in use. However, by using the standardized anonymously application form, the personal aspects of a job application are (partly) gone, which could also provide useful information for the employer or recruiter. In a sense, it can be stated that through anonymously applying with a standardized form the possibility to sell yourself as applicant is (partly) gone. Unfortunately, since the experiment in Utrecht did not yield any meaningful results, a comparison analysis between the two methods cannot be performed. Lastly, it must be stated that the municipality of both towns clearly indicated that the key point of implementation of AAP is awareness of the problem, not an intentional quantitative effect per se.

Both experiments focused in particular on the ethnic origin of the candidates and both experiments only included job ads for the mid-high salary scales 10 till 14, which implies a gross salary between 2.446 and 6.265 euros per month in 2015 and between 2.519 and 6.453 euros per month in 2016<sup>4</sup>. Among other things, the ethnic diversity within the organization is lower in high-level jobs, which is the main reason why this study only focused on job application in the salary scale 10 to 14. Furthermore, in the first half of 2015 and 2016, eight different departments of the municipality The Hague did place vacancies in order to hire some new workers. The eight departments can be found in table 1 below.

<b>Department code</b>	<b>Department name (Dutch)</b>	<b>Department name (English)</b>
BSD	Bestuursdienst	Administrative Service
SZW	Sociale Zaken en Werkgelegenheidsprojecten	Social Affairs and Employment
DPZ	Dienst Publiekzaken	Public Affairs Service
OCW	Onderwijs, Cultuur en Welzijn	Education, Culture and Welfare
DSO	Dienst Steden Ontwikkeling	City Development Services
IDC	Intern Diensten Centrum	Internal Services Centre
TUO	Tijdelijke UitvoeringsOrganisatie	Temporary Implementation Organization
DSB	Dienst Staatsbeheer	State Administration Service

Table 1. Departments of the municipality The Hague.

The departments are included in my analysis, since the application rate of ethnical groups can also be different through the departments. When the distribution of the applications over the departments changed in 2016 compared to 2015, this could explain some changes in the percentage bicultural Dutch which is not due to AAP. Anonymously applying could make applying for disadvantaged groups more attractive. The number and composition of candidates were investigated, since applying anonymously may affect the willingness to apply

<sup>4</sup> <https://www.rijksoverheid.nl/documenten/publicaties/2015/09/10/salaristabel-per-1-januari-2016>



for the candidates. Rinne (2014) claimed that due to anonymously applying the number and diversity of candidates is likely to increase.

Prior to the pilot, in consultation with the municipalities The Hague and Utrecht, guidelines were established regarding the aspects of the letter and the CV which had to be made anonymous. The following items were deleted from all documents: last name, first name, email address, place of birth, nationality, photograph of the candidate and in some resumes the language skills. Furthermore, in this study, bicultural Dutch are divided into three categories: Turkish, Moroccan, Surinamese or Antillean immigrants (TMSA), other non-Western immigrants (ONW), which have origins in any of the countries in the continents of Africa, Latin America and Asia, excluding Indonesia and Japan, and lastly, Western ethnic minorities (origin North America, Europe, Indonesia or Japan). This classification is created, since in these groups the signals of exclusion in the labor market are overall the greatest.

### **2.2.2. Request for implementing AAP in the Netherlands**

The reason for encouraging and monitoring local pilots with anonymously applying in The Netherlands is the request by Verweij (politician, member of the PvdA: Dutch Labour Party) and Van Weyenberg (politician, member of D66: Dutch Democratic Party)<sup>5</sup> in the House of Representatives in 2015. Verweij and Van Weyenberg proposed that AAP should be operational with interim evaluations to reduce hiring discrimination for minorities in the Netherlands. The fully (translated) proposition can be found in Appendix A. The scale of this experiment is unique: never before an experiment had this magnitude in the Netherlands. Moha and Konings (2016) therefore claim that due to the large size and the design of the study, it is possible to determine the relation between AAP, interview offer probability and job offer probability<sup>6</sup> of bicultural candidates.

### **2.2.3. Evaluation about AAP experiment**

To find out the rationale for the candidates to apply, all candidates were asked to fill in a questionnaire. The questionnaire is taken by 17.5 percent (329 of the 1.880) of the candidates and also by a substantial percentage recruiters in The Hague to see how much impact anonymously applying had on the consideration to apply for the job or hiring decision. The main concerns of anonymously applying were the increasing amount of administrative work and the lack of opportunity to discriminate positively. However, most recruiters thought the effects of anonymous job applications in the hiring procedure would be nihil. A quarter of the recruiters thought that ethnicity played no role in hiring procedure and three quarter of the recruiters

---

<sup>5</sup> <https://www.parlementairemonitor.nl/9353000/1/i9vvij5epmi1ey0/vizicbprvut6>

<sup>6</sup> However, in the study of Moha and Konings (2016), the results of the job offer probabilities were not discussed, since most of the vacancies were not filled yet.

thought that anonymously applying did not lead to an increase in bicultural candidates over all jobs. Interesting is the fact that half of the candidates claim that they did not know about the introduction of AAP in 2016, approximately the same percentage through all ethnic groups and native Dutch candidates. One out of five candidates thought that ethnicity had no effect on the hire decision at all, since competences and experience are more important indicators. Candidates also argued that the selection committee would be professional enough to make a decision on characteristics that do matter. Bicultural Dutch suspected more preferential treatment for the recruiters' 'own kind of people'. Native Dutch candidates thought more often that important information is lost when applications are anonymized than bicultural Dutch candidates. The questionnaire also showed that the implementation of AAP has an influence on the image of a municipality. Candidates considered the municipality therefore more objective and fairer. In addition, candidates thought that the city government of The Hague showed that equivalence is important through introducing AAP. This increased the attractiveness of the municipality under bicultural Dutch. It can be stated that the pilot anonymous job applications had positive impact on the recruitment of bicultural Dutch employees.

As already mentioned, AAP is only applicable in the first phase of the recruiting process. Some recruiters argued that due to anonymously applying, they could not be included certain sensitive information in a positive way in the first phase screening process. Mentionable is the fact that the municipality The Hague continued the intervention, because of the positive intermediate results. Candidates and recruiters are also predominantly positive about the pilot. However, the important question is whether the intermediate results in terms of hiring discrimination should still be positive, have the same magnitude and are significant after controlling for job characteristics.

#### **2.2.4. Results AAP experiment in The Hague**

The main finding in the experiment in The Hague was that the introduction of anonymously applying has led to an increase in the number of applying bicultural candidates for jobs at the municipality The Hague (31 percent in 2016 versus 24 percent in 2015), in particular candidates with a Turkish, Moroccan, Surinamese or Antillean (TMSA) background (18 percent in 2016 versus 12 percent in 2015). This indicates that bicultural Dutch are more inclined to apply during AAP than during the standard procedure. For reference, in Utrecht no significant increase in bicultural candidates in 2016 compared to 2015 was visible. In The Hague the probability that bicultural candidates were invited for a first job interview was larger in 2016 (introduction AAP) than in 2015 (standard applying procedure), but only significantly larger for candidates with a Turkish, Moroccan, Surinamese or Antillean (TMSA) background. Risk assessments of selectors seem to play a role here.

Also more bicultural candidates (again, especially the TMSA group) were called back for a job interview. The percentage bicultural candidates invited for a job interview was 13 percent in 2015 (control year) and increased with 11 percent to 24 percent in 2016 (AAP). These numbers imply that the probability for

bicultural candidates to get a job interview increased since the introduction of AAP. However, the call rate is still lower than for the native Dutch. It seems therefore that anonymizing job applications is not sufficient enough to close the gap between native Dutch and bicultural Dutch in terms of hiring discrimination. Furthermore, more bicultural candidates (again, especially the TMSA group) were hired in 2016 compared to 2015 (36 percent versus 11 percent). Note that this increase cannot be attributed to AAP anymore, since the job interview and other follow-up meetings are no longer anonymously.

The differences in callback rate between native Dutch and the minorities looks bigger when taking the relative job interview probability into account, which means the bicultural Dutch job probability in comparison with the native Dutch job probability. Native Dutch candidates have 2.15 times more chance to get invited for a job interview than bicultural Dutch. This factor decreased to 1.49 in 2016, but implies that even when AAP are implemented, the native Dutch employee has more chance to get an interview than a bicultural Dutch. The job offer probability was also higher for bicultural Dutch in 2016 than in 2015.

A side note is that when the results of the experiment anonymously applying in The Hague had to be evaluated, not all jobs were given to one of the final candidates yet. In other word, the second round was not completed for all job applications. This implies that the findings regarding hiring candidates (job offer probabilities) were based on a more limited number of cases (39 instead of 57). Since only the first round of the applying procedure is relevant for this research, for the analysis on the job interview invitation rate, this is not problematic for this study. However, for the probability to get hired no significant differences could be found between AAP and the standard application procedures. AAP increased the probability for minorities to get an interview, but did not increase significant the probability to get hired, taken into account that relative more employees from disadvantaged groups has survived the first cut. A possible explanation could be the small amount of available job applications or the occurrence of delayed hiring discrimination (see paragraph 2.1.4).

### 3. Additional research about AAP experiment in The Hague

#### 3.1. Data description

As discussed above, commissioned by the municipality The Hague, an experiment is designed about anonymous job applications in order to examine whether there is a link between applying behavior of ethnic groups and the introduction of AAP in The Netherlands (Moha and Konings, 2016). In response to the results of Andriessen, van der Ent, van der Linden and Dekker (2015), who concluded that bicultural candidates have a lower probability to get a job interview invitation, the pilot '*anonymously applying*' was executed a year later. The study of Andriessen, van der Ent, van der Linden and Dekker (2015) was also commissioned by the municipality The Hague. In the Netherlands, no large-scale pilots have been conducted where the effects can be identified and thoroughly investigated. In previous small-scale pilots in the Netherlands (Alphen a/d Rijn and Nijmegen), no clear results could be found. For the study, the municipality The Hague did ask the external firm Motivaction to design the experiment and evaluate the process carefully. All candidates who applied on external vacancies in mid-job salary scales (10 till 14) were offered anonymity during the period January 1 and June 30, 2016 to the job holders. The results are compared with the same period in 2015. The main questions of this research are what AAP do with the recruiting (people who are persuaded to apply), selection (who are invited for a job interview) and what AAP do to the internal attitude towards the topic diversity. Among others, application number, salary scale, number of candidates per application and number of candidates through ethnic groups of all published vacancies in salary scale 10 till 14 for The Hague's municipality in the first six months of 2015 and 2016 can be found in Appendix B. In the municipality The Hague, 36 applications were published in 2015 and 57 applications in 2016<sup>7</sup>.

Moha and Konings (2016) collected data about job description, salary scale and origin of the applicants, but Moha and Konings (2016) argued that the number of vacancies seems too small to split the data for more detailed outcomes. This is not entirely correct and it seems valuable to do so. Splitting the results to other criteria, like gender, age, tenure and education was not possible, since unfortunately neither Motivaction nor the municipality The Hague was willing to provide the more detailed data for my study. Another important factor to take into account is the difference in access to jobs for candidates where certain qualifications and credentials are required. Krause, Rinne and Zimmermann (2012a) argued that unequal access to jobs across ethnic groups has an important implication for labor market outcomes of employees. In imitation of this statement, the idea of extending the research of the experiment of The Hague was born. In addition to the study, the salary scale, the department and the degree of leadership skill a job required is added to the regression to get closer to the true effect of anonymously applying on the hiring decisions of employers. Ideally, the anonymously applying effects on ethnical candidates in applying behavior, interview offer probability and the job offer probability should be investigated. Unfortunately, the documentation of Moha

---

<sup>7</sup> In the appendix of Moha and Konings (2016), only 56 applications in 2016 can be found. However, the numbers in their conclusions are the same as in my calculations, what suggest that no one has applied for the missing job application (yet).

and Konings (2016) does not contain raw data for the interview offer probability and the job offer probability analysis, since that phase was not finished at the publication date of Moha and Konings (2016).

In this analysis, some important correlations will be discussed. Furthermore, the regressions will be performed on top-level (all bicultural Dutch candidates), but also on the three subcategories (Turkish, Moroccan, Surinamese or Antillean immigrants (TMSA), Other non-Western immigrants (ONW) and Western immigrants (W)) to detect differences between those ethnic groups in applying behavior during AAP and AAP compared to the reference year. Two regressions will be performed: the first one is with the applying rate as dependent variable, because I am interested in the correlation between AAP, a few job characteristics and the applying behavior of minorities. The second regression is with the total number of applicants as dependent variable, since it is also interesting how much the results of the first regression are driven by an effect on the total number of applications. The linear regression analyzes will be discussed separately and conclusions will be based on both statistical significance and economically significance. Since the sample is not that large ( $N = 92$ ), results might be insignificant, while there is in fact an effect. One speaks of statistical significance if the estimate is quite precise and economically significant if the effect is important and large enough. For example, a coefficient that is almost equal to zero can be statistically significant, but never economically significant. A large coefficient is quite often economically significant, but not always statistically significant. When a coefficient is economically significant, there will be primarily looked at the magnitude and sign of the coefficient. All conclusions are drawn under the 'ceteris paribus' assumption. Lastly, a short overview of the most important findings will be given in paragraph 3.4..

### **3.1.1. Transformation of variables**

In this paragraph the variables in the regression will be discussed. The dependent variable is the percentage bicultural Dutch of all the applying candidates. This is calculated by simply summing the Turkish, Moroccan, Surinamese or Antillean immigrants (TMSA), other non-Western immigrants (ONW) and Western immigrants (W) per application and divided by the total number of candidates. The three subcategories, the TMSA-, ONW- and W-Dutch candidates, are also dependent variables. The measure for anonymously applying is a dummy variable that equals to 0 if the vacancy was published in 2015 and equals to 1 if the vacancy was published in 2016. Since this is a time-dummy, the coefficients of this variable does not have to be necessarily fully indicating the effects of anonymous applications, but any positive or negative shift can also be a result of other influences that developed that year. According to Moha and Konings (2016), the likelihood that substantial changes were present within one year is quite small, since no large changes in personnel composition and firm structure did occur. It is, though, an odd statement when taking into account that the number of vacancies increase in 2016 with more than fifty percent compared to 2015. This indicates that it is likely that some change did occur. However, this is the best proxy for anonymously applying there is available. Furthermore, the salary scale variable will be transformed to four dummy variables (salary scales 11 to 14, salary scale 10 is the reference category), to see the combined relation of anonymously applying and salary

scale of the job ads with the percentage candidates of a minority group. Also, another dummy is created that is equal to 0 if the application does not contain an obvious leadership / management component and is equal to 1 when the application does. A ratio variable should have left this too much to interpretation. Lastly, seven dummies are created for the departments, using department DSB as reference group<sup>8</sup> to see the difference in applying behavior through departments.

### 3.1.2. Summary statistics

In order to get an overview of the data in Appendix B, table 2 below shows the summery statistics of the characteristics of the ads, which will be discussed. First of all, the number of published applications in 2016 is more than 50 percent higher than in 2015. The average job salary scale is not significantly different over the years (table 3,  $p = .391$ ), but the distribution of the salary scales is different between 2015 and 2016. In 2015 less job ads with salary scale 12 and more job ads with salary scale 10 and 11 were published. Furthermore, the percentage managerial jobs is higher in 2015, but according to table 3 not significantly higher ( $p = .316$ ). The distribution of applications through department is shifted: there is an increase of vacancies in departments BSD, DPZ and DSB and a decrease in departments OCW, IDC and TUO. Table 3 also indicates that the percentage bicultural Dutch candidates increased in 2016, in particular due to the TSMA group. The percentage of Western candidates did however decrease.

		2015	2016
Applications	# (number)	36	56
	% (percentage)	100%	100%
Job with managerial aspect	# (number)	16	19
	% (percentage)	44,4%	33,9%
Job with salary scale 10	# (number)	11	17
	% (percentage)	30,6%	40,4%
Job with salary scale 11	# (number)	10	27
	% (percentage)	27,8%	48,2%
Job with salary scale 12	# (number)	11	8
	% (percentage)	30,6%	14,3%
Job with salary scale 13	# (number)	4	1
	% (percentage)	11,1%	1,8%
Job with salary scale 14	# (number)	0	3
	% (percentage)	0,0%	5,4%
Job in Department BSD	# (number)	0	8

<sup>8</sup> DSB is a large department, twenty out of ninety-two vacancies are published by this departement, and therefore a good reference group.

	% (percentage)	0,0%	14,3%
	# (number)	8	12
Job in Department SZW	% (percentage)	22,2%	21,4%
	# (number)	1	4
Job in Department DPZ	% (percentage)	2,8%	7,1%
	# (number)	7	5
Job in Department OCW	% (percentage)	19,4%	8,9%
	# (number)	4	8
Job in Department DSO	% (percentage)	11,1%	14,3%
	# (number)	10	3
Job in Department IDC	% (percentage)	27,8%	5,4%
	# (number)	2	0
Job in Department TUO	% (percentage)	5,6%	0,0%
	# (number)	4	16
Job in Department DSB	% (percentage)	11,1%	28,6%

Table 2. Summary statistic table ads.

Variable	t-statistic	N	df	Significance ( p –value)	95% Confident Interval of the difference	
Salary Scale	-,863	92	90	,391	-,62	,24
% Leadership jobs	-1,009	92	90	,316	-,31	,10
Number of candidates	-1,306	92	90	,195	-25,83	5,34
% Bicultural candidates	2,512**	92	90	,014	1,60	13,68
% TMSA candidates	1,957**	92	90	,045	,10	13,07
% ONW candidates	,604	92	90	,548	-3,26	6,11
% W candidates	-2,771***	92	90	,007	-4,20	-,69

Table 3. Independent sample t-test: difference between 2015 and 2016.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

In table 4, the summary statistics of the candidates are presented. Nineteen percent more candidates applied for jobs in the municipality The Hague in 2016 compared to the previous year, but the average number of candidates per vacancy is lower in 2016 compared to 2015. However, this decrease is according to table 3 not significant lower ( $p = .316$ ) and among other things, this insignificant decrease is due to the fact that the number of vacancies increased over time. Furthermore, the standard deviation of the total candidates in 2016 is larger, which implies that the range in candidates through the applications were larger than in 2015. The differences in applying behavior between popular ads and less popular ads were much better visible in 2016. The TMSA-, ONW-, and W-candidates are part of the group bicultural candidates. In other words, in this study

the total number of bicultural candidates is simply the sum of the three Dutch minority groups. The number of applying bicultural candidates increased during AAP heavily (55 percent), which is 36 percent more than average. This increase can be attributed to a large extent to TMSA candidates, since the number of TMSA candidates increased with more than 75 percent in one year. The total number of ONW candidates did not change much and therefore the average ONW-candidate per application decreased. The number of the Western Dutch candidates and the average per application however did increase. Lastly, the regressions are based on the percentage bicultural Dutch that applied for a specific job. So in relative terms, the percentage TMSA candidates did increase ( $p = .014$ ), the percentage ONW candidates did not significantly differ ( $p = .548$ ) and the percentage Western candidates decreased ( $p = .007$ ) in 2016 compared to previous year.

		2015	2016
Total candidates	# (total)	1.576	1.880
	# (average per application)	43,78	33,54
	$\sigma$ (standard deviation)	30,30	40,28
	% (percentage)	100%	100%
Bicultural candidates	# (total)	380	588
	# (average per application)	10,56	10,50
	$\sigma$ (standard deviation)	9,11	12,43
	% (percentage)	25,97%	32,58%
TMSA candidates	# (total)	187	330
	# (average per application)	5,19	5,89
	$\sigma$ (standard deviation)	5,61	6,73
	% (percentage)	11,94%	19,58%
ONW candidates	# (total)	92	98
	# (average per application)	2,56	1,75
	$\sigma$ (standard deviation)	2,31	3,11



	% (percentage)	7,89%	9,31%
W candidates	# (total)	101	160
	# (average per application)	2,81	2,86
	$\sigma$ (standard deviation)	2,51	3,77
	% (percentage)	6,13%	3,69%

Table 4. Summary statistic table candidates

## 3.2. Analysis

### 3.2.1. Correlations

The correlations between the percentage bicultural applying candidates (or which subcategory whatsoever), AAP, salary scale and leadership components<sup>9</sup> can reveal some important insights before coefficients in a regression can be interpreted well. Also, an assumption in the pilot in The Hague is that since all external job applications are included in the study and there is a short period of time (1 year) between reference period (first half of 2015) and the trial period (first half of 2016), it is not likely that major changes have occurred in the type of jobs. Through correlational research, this assumption will be investigated in more detail. Only the most valuable findings are discussed. First, table 5 claims that significant more applications with salary scale 10 and 11 and less applications with salary scale 12 and in lesser extent salary scale 13 were published in 2016 than in 2015. Kaas and Manger (2012) concluded that in low-level jobs discrimination in hiring is higher than in high-level jobs, which could imply that the impact of anonymously applying is in fact smaller than in the regression, since the negative shift in the average salary level of the applications. However, the independent sample t-test showed that the average salary scale did not significantly differ between 2015 and 2016 (table 3). Furthermore, the higher the salary scale and the level of the job, the less candidates apply for the job. Noticeable is the finding that the percentage TMSA Dutch candidates decreases when the job requires higher level of education, experience or a higher degree of leadership capacities. Salary scale and degree of leadership component in the job correlate positively, what implies that jobs with a managerial aspect require more skills, fall into a higher salary scale and therefore maintain a higher payment than jobs without a leadership aspect on average. The minority group *other non-Western Dutch* candidates apply less on jobs with a managerial aspect. The correlation between AAP and the percentage bicultural Dutch and TMSA Dutch candidates is positive and between the year dummy and Western Dutch candidates negative, which implies that relative more TMSA Dutch candidates and less Western Dutch candidates were applying in 2016 compared to 2015.

<sup>9</sup> The correlations with the department dummies are left out in table 5, since the correlations with any other variable do not contain meaningful information.

		Year	Salary scale	SS 10	SS 11	SS 12	SS 13	SS 14	Leadership
<b>Year</b>	coefficient	1	-,091	-,002	,203*	-,196*	-,201*	,147	-,106
	<i>sign.</i>		,391	,984	,052	,061	,055	,161	,316
<b>Salary Scale</b>	coefficient	-,091	1	-,730***	-,089	,452***	,451***	,228***	,138
	<i>sign.</i>	,391		<,001	,400	<,001	<,001	<,001	,189
<b>SS 10</b>	coefficient	-,002	-,730***	1	-,543***	-,337***	-,159	-,121	,114
	<i>sign.</i>	,984	<,001		<,001	,001	,131	,249	,278
<b>SS 11</b>	coefficient	,203*	-,089	-,543***	1	-,418***	-,197*	-,151	-,323***
	<i>sign.</i>	,052	,400	<,001		<,001	,060	,152	,002
<b>SS 12</b>	coefficient	-,196*	,452***	-,337***	-,418***	1	-,122	-,094	,098
	<i>sign.</i>	,061	<,001	,001	<,001		,245	,374	,353
<b>SS 13</b>	coefficient	-,201*	,451***	-,159	-,197*	-,122	1	-,044	,207**
	<i>sign.</i>	,055	<,001	,131	,060	,245		,677	,048
<b>SS 14</b>	coefficient	,147	,228***	-,121	-,151	-,094	-,044	1	,108
	<i>sign.</i>	,161	<,001	,249	,152	,374	,677		,304
<b>Leadership</b>	coefficient	-,106	,138	,114	-,323***	,098	,207**	,108	1
	<i>sign.</i>	,316	,189	,278	,002	,353	,048	,304	
<b>Total candidates</b>	coefficient	-,136	-,189*	,165	-,022	-,095	-,023	-,121	-,040
	<i>sign.</i>	,195	,071	,115	,836	,365	,826	,250	,703
<b>Percent Bicultural</b>	coefficient	,202*	-,144	-,061	,217**	-,020	-,204*	-,135	-,170
	<i>sign.</i>	,053	,172	,564	,038	,848	,052	,200	,105
<b>Percent TMSA</b>	coefficient	,256**	-,212**	,024	,218**	-,136	-,184*	-,118	-,013
	<i>sign.</i>	,014	,043	,819	,037	,195	,078	,263	,901
<b>Percent ONW</b>	coefficient	,063	,020	-,135	,127	,083	-,112	-,048	-,232**
	<i>sign.</i>	,548	,852	,199	,228	,430	,290	,650	,026
<b>Percent W</b>	coefficient	-,280***	,134	,036	-,257**	,176	,153	,020	,003
	<i>sign.</i>	,007	,203	,731	,013	,094	,146	,846	,978

Table 5. Correlations.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

### 3.2.2. Regressions applying rate

In the bicultural Dutch and the three subcategories TMSA-, ONW- and Western Dutch, the percentage candidates of bicultural group will be regressed on the year-dummy variable, which is the indicator of AAP (model 1), adding the salary scale dummy variables to the model (model 2), adding separately the leadership-dummy to the model (model 3), adding the variables salary scale and leadership together to the model (model 4) and finally adding the departments to the model (model 5). In all five models, the main focus is to see whether and how the coefficient of anonymously applying changes when adding the salary scales, leadership and department dummies. The regression formulas per model for the four ethnic groups can be found below and the regressions results are categorized by group.

$$\text{MODEL 1: Applying rate \{ethnic group\}_i = \beta_0 + \beta_1 * \text{YEAR}_i + \varepsilon_i$$

$$\text{MODEL 2: Applying rate \{ethnic group\}_i = \beta_0 + \beta_1 * \text{YEAR}_i + \beta_2 * \text{SS11}_i + \beta_3 * \text{SS12}_i + \beta_4 * \text{SS13}_i + \beta_5 * \text{SS14}_i + \varepsilon_i$$

$$\text{MODEL 3: Applying rate \{ethnic group\}_i = \beta_0 + \beta_1 * \text{YEAR}_i + \beta_2 * \text{LEADERSHIP}_i + \varepsilon_i$$

$$\text{MODEL 4: Applying rate \{ethnic group\}_i = \beta_0 + \beta_1 * \text{YEAR}_i + \beta_2 * \text{LEADERSHIP}_i + \beta_3 * \text{SS11}_i + \beta_4 * \text{SS12}_i + \beta_5 * \text{SS13}_i + \beta_6 * \text{SS14}_i + \varepsilon_i$$

$$\text{MODEL 5: Applying rate \{ethnic group\}_i = \beta_0 + \beta_1 * \text{YEAR}_i + \beta_2 * \text{LEADERSHIP}_i + \beta_3 * \text{SS11}_i + \beta_4 * \text{SS12}_i + \beta_5 * \text{SS13}_i + \beta_6 * \text{SS14}_i + \beta_7 * D_{BSD}_i + \beta_8 * D_{SZW}_i + \beta_9 * D_{DPZ}_i + \beta_{10} * D_{OCW}_i + \beta_{11} * D_{DSO}_i + \beta_{12} * D_{IDC}_i + \beta_{13} * D_{TVO}_i + \varepsilon_i$$

### 3.2.2.1. Results regression 'Bicultural Dutch candidates'

The regressions on the bicultural Dutch candidates can be found in table 6. The coefficients of all five regressions are given and the standard error are shown in parentheses. The first model indicates that twenty-six percent of the candidates were bicultural Dutch in 2015. The percentage applying bicultural Dutch increased with 6.6 percent to 32.6 percent in one year. The constant is statistical ( $p < .05$ ) and economically significant and the variable *year* is marginal statistically significant ( $.05 < p < .10$ ). All variables are economically significant. In model 2, the salary scale dummies are added. In this regression, no variable is statistical significant. This happens quite often when the sample is small and a substantial amount of variables are included in the regression. However, all variables in this regression are economically significant. The percentage bicultural Dutch candidates that applied for jobs in salary scale 10 in 2015 was 25.1 percent. This percentage increased with 5.7 percent in 2016 and increased around 5 and 2 percent if it concerned a job respectively in salary scale 11 and 12 and decreased around 10 and 12 percent when a job vacancy had salary scale 13 and 14 respectively. For example, this model predicts that the percentage of bicultural candidates in the municipality The Hague in salary scale 11 in 2016 equals to 35.8 percent ( $25.1 + 5.7 + 5$ ). The third model predicts the percentage bicultural Dutch candidates based on AAP in 2016 and whether a job contains a leadership aspect. The constant is statistically significant ( $p < .001$ ), the variable *year* is marginal significant ( $p = .075$ ) and the variable *leadership* is not statistically significant ( $p = .149$ ). Twenty-eight percent of the candidates who applied to a job without a leadership aspect in 2015 had a bicultural origin. This percentage increased in 2016 with six percent compared to 2015 and no statistical conclusion can be drawn about leadership. All variables are economically significant, so the probability that a bicultural Dutch candidate would apply for a managerial job in comparison with a job without that kind of responsibilities was 5 percent lower. The final model, model 4, contains the year, leadership and all salary scale dummy variables. Again, no variables are statistical significant, but all variables

are economically significant. The percentage bicultural candidates in 2015 for a job in salary scale 10 without managerial responsibilities is 26 percent. This increases were respectively around 4 and 2 percent in case of a job with respectively salary scale 11 and 12 and decreases with 9 and 12 percent when it concerns a job in respectively salary scale 13 and 14. Note that this is approximately the same as in model 2. The percentage bicultural Dutch candidates increase with 5.5 percent in 2016, which is 1 percent lower than in model 1. The salary scales and the variable *leadership* explain this deviation. When the job contains a supervision aspect, the percentage is 2.5 percent lower than if it does, which is 2.5 percent less decline than model 2 indicates. The salary scales explain about the half of the decrease regarding to leadership. This is due to the significant positive correlation between the salary scales and leadership variable. By adding the departments in model 5, it can be stated that bicultural Dutch apply less in almost all other departments compared to the DSB department. Model 5 does not add much in terms of influencing the coefficients in the previous models. In particular, the department dummies do not partly explain the impact of AAP on the applying rate of ethical minorities.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
(Constant)	25,97*** (2,64)	25,08*** (3,64)	28,16*** (3,03)	26,24*** (4,05)	30,06*** (6,53)
Year	6,618* (3,38)	5,66 (3,53)	6,10* (3,38)	5,58 (3,55)	5,46 (3,96)
Leadership	-	-	-4,95 (3,40)	-2,38 (3,60)	-3,03 (3,90)
SS 11	-	5,00 (3,92)	-	4,36 (4,06)	2,53 (4,37)
SS 12	-	1,89 (4,67)	-	1,90 (4,69)	0,63 (4,86)
SS 13	-	-9,81 (7,69)	-	-9,04 (7,81)	-10,92 (7,91)
SS 14	-	-12,48 (9,56)	-	-11,97 (9,62)	-8,54 (10,40)
DEP BSD	-	-	-	-	-7,95 (7,05)
DEP SZW	-	-	-	-	-4,18 (5,29)
DEP DPZ	-	-	-	-	-13,64 (8,24)
DEP OCW	-	-	-	-	0,27 (6,18)
DEP DSO	-	-	-	-	3,53 (5,99)

DEP IDC	-	-	-	-	-1,51 (6,23)
DEP TUO	-	-	-	-	-20.48* (12.29)
<b>R<sup>2</sup></b>	,041	,114	,063	,119	,209
<b>N</b>	92	92	92	92	92

Table 6. Regressions on the percentage Bicultural Dutch.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

### 3.2.2.2. Results regression 'Turkish, Moroccan, Surinamese or Antillean Dutch candidates'

The regressions on the Turkish, Moroccan, Surinamese or Antillean Dutch candidates can be found in table 7. The coefficients are given and the standard error are shown in parentheses. In model 1 the constant and the variable *year* are statistically significant ( $p < .001$  and  $p = .014$  respectively) and economically significant. The average percentage applying TMSA Dutch candidates equals to 11.9 percent in 2015 and increased with 7.6 percent in 2016, which is a substantial increase. In model 2 only the variables *year* ( $p = .048$ ) is statistical significant. The percentage TMSA Dutch candidates increased from 2015 to 2016 with 6.4 percent to almost 20 percent. All variables are economically significant. If the salary scale of a job vacancy is 11, the percentage applying TMSA Dutch candidates is 2.5 percent higher than a job with salary scale 10, but if the salary scale of a job vacancy is 12, 13 and 14 respectively, the percentage applying TMSA candidates is 3, 9 and 12 percent lower than jobs with salary scale 10. This percentage increased another 6.5 percent when the application is published in 2016. Model 3 does not add much to this analysis, since the leadership coefficient was neither statistically ( $p = .892$ ) nor economically significant ( $\beta = 0.42$ ). Model 4 shows that the variable *year* is significant ( $p = .045$ ). The percentage TMSA Dutch applying candidates increased with 6.5 percent in 2016, which is 1 percent lower than in model 1. The dummy salary scale variables caused that effect. All variables are economically significant. The percentage TMSA Dutch applying candidates increased with 3.5 percent if it concerns a job in salary scale 11 in comparison with salary scale 10. That percentage decreases with 3, 10 and 13 respectively when it concerns a job with salary scale 12, 13 and 14 respectively. The coefficients of salary scale 11, 13 and 14 are in this model around one percent higher than in model 2. If the job requires managerial skills, this percent increased with 3 percent. Model 3 indicated that leadership had no effect on the TMSA Dutch, but including the salary scale in model 4, it does. Adding the department, we can state that the DSB department has the highest applying rate of the TMSA Dutch candidates. Six out of the seven dummies are (marginal) significant, but do not add much in terms of (partly) explaining the impact of AAP on the applying rate of ethical minorities.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
(Constant)	11,94*** (2,37)	13,24*** (3,29)	11,76*** (2,75)	11,71*** (3,64)	21,00*** (5,78)
Year	7,64** (3,04)	6,41** (3,19)	7,83** (3,08)	6,50** (4,19)	6,15* (3,50)
Leadership	-	-	,421 (3,09)	3,17 (3,24)	0,61 (3,50)
SS 11	-	2,55 (3,54)	-	3,41 (3,65)	1,81 (3,86)
SS 12	-	-3,23 (4,22)	-	-3,24 (4,22)	-3,06 (4,30)
SS 13	-	-9,13 (6,94)	-	-10,15 (7,03)	-11,67* (7,00)
SS 14	-	-12,40 (8,63)	-	-13,08 (8,66)	-5,75 (9,20)
DEP BSD	-	-	-	-	-17,63*** (6,23)
DEP SZW	-	-	-	-	-8,39* (4,68)
DEP DPZ	-	-	-	-	-13,31* (7,29)
DEP OCW	-	-	-	-	-6,51* (5,46)
DEP DSO	-	-	-	-	-10,40* (5,29)
DEP IDC	-	-	-	-	-6,69 (5,52)
DEP TUO	-	-	-	-	-19,49* (10,87)
R <sup>2</sup>	,066	,139	,066	,140	,254
N	92	92	92	92	92

Table 7. Regressions on the percentage of subcategory Turkish, Moroccan, Surinamese or Antillean Dutch.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

### 3.2.2.3. Results regression 'Other Non-Western Dutch candidates'

The regressions on the Other Non-Western Dutch candidates can be found in table 8. The coefficients are given, and the standard error are shown in parentheses. In model 1, the constant is statistically and economically significant, which implies that in 2015 only 8 percent of the applying candidates were ONW Dutch. The variable year is however not statistically significant ( $p=0.548$ ) and also not economically significant ( $\beta = 1.42$ ). In model 2, no variable is statistically significant, and only the variables *SS 11*, *SS 12* and *SS 13* are economically significant. The percentage applying ONW Dutch candidates in 2015 for jobs in salary scale 10 is around 6 percent, jobs in salary scale 11 around 9.5 percent and jobs in salary scale 12 around 10 percent. In

model 3 the constant and the variable leadership are statistically significant ( $p < .001$  and  $p = .031$  respectively), and economically significant, but unfortunately the year dummy is not statistically significant and also not significant in economically terms for this ethnical group. The percentage ONW Dutch candidates for jobs in 2015 without a leadership component was 10.2 percent. When the job did contain a supervision aspect, this percentage decreased with 5.1 percent to 5.1 percent (50 percent decrease). In model 4 only the variable leadership is marginal significant ( $p = .075$ ). The percentage applying ONW Dutch decreases with 4.5 percent when it concerns a job with managerial responsibilities, which is 0.5 percent less decline than in model 3. This 0.5 percent might be explained by the salary scale. The variables year, *SS13* and *SS14* are not economically significant. The constant, the variables leadership, *SS11* and *SS12* are in fact economically significant, which implies that the percentage applying ONW Dutch candidates in 2015 for non-leadership jobs in salary scale 10 is 8 percent. This percentage is equal to 10.5 percent or 12.3 percent when it concerns a non-leadership job in salary scale 11 or 12 respectively. This percentage is higher than in model 2, which is due to the variable leadership. And, as stated above, when the job has a leadership component, this percentage decreased with 4.5 percent, which is a smaller degree than in model 3, due to the salary scale dummies. Model 5 shows that ONW Dutch did apply in particular in the DSO, OCW and BSD departments. Since the year dummy is not statistical and economical significant in each model, the department dummies can therefore not partly explain the impact of AAP on the applying rate of ethical minorities.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
(Constant)	7,89*** (1,84)	5,80** (2,59)	10,17*** (2,08)	7,99** (2,83)	4,43 (4,50)
Year	1,42 (2,36)	1,17 (2,51)	0,89 (2,32)	1,04 (2,48)	0,75 (2,72)
Leadership	-	-	-5,13** (2,34)	-4,53* (2,52)	-3,19 (2,68)
SS 11	-	3,79 (2,79)	-	2,56 (2,84)	2,41 (3,00)
SS 12	-	4,25 (3,32)	-	4,26 (3,28)	3,22 (3,35)
SS 13	-	-2,37 (5,47)	-	-0,91 (5,46)	-0,83 (5,44)
SS 14	-	-1,09 (6,80)	-	-0,11 (6,73)	-2,62 (7,16)
DEP BSD	-	-	-	-	7,11 (4,85)
DEP SZW	-	-	-	-	2,19 (3,64)
DEP DPZ	-	-	-	-	-0,20 (5,67)
DEP OCW	-	-	-	-	5,84 (4,25)

DEP DSO	-	-	-	-	11,51*** (4,12)
DEP IDC	-	-	-	-	2,16 (4,29)
DEP TUO	-	-	-	-	-3,22 (8,46)
R <sup>2</sup>	,004	,044	,055	,079	,200
N	92	92	92	92	92

Table 8. Regressions on the percentage Other Non-Western Dutch.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

#### 3.2.2.4. Results regression 'Western Dutch candidates'

The regressions on the Western Dutch candidates can be found in table 9. The coefficients are given, and the standard error are shown in parentheses. In the first model, the constant and the variable year were statistically ( $p < .001$  and  $p = .007$  respectively) and economically significant. The percentage applying Western Dutch candidates in 2015 was 6.1 percent. This percentage decreased with 2.5 percent to 3.6 percent in 2016. Note that this was the only ethnical group where in 2016 the percentage applying candidates is in fact lower than in 2015. Model 2 indicates that the constant and year are statistically significant ( $p = .007$  and  $p = .042$ ) again, but no salary scale dummy variable is statistically significant. The percentage applying Western Dutch candidates in 2015 on jobs applications that fall in salary scale 14 was 6 percent. This percentage decreased with 1.9 to 4.1 percent in 2016. Only *SS13* is economically significant, so the percentage decreased around 1.5 percent comparing to jobs in salary scale 10. In model 3, again, the variable *year* is statistically and economically significant, but the variable *leadership* is not statistically significant ( $p = .792$ ) and not economically significant (magnitude coefficient too small), which is why no conclusions could be drawn about leadership based on model 3. In model 4, the constant and the variable *year* are statistically ( $p = .004$  and  $p = .038$  respectively) and economically significant. In addition to those variables, variables *SS11* and *SS13* are also economically significant. The percentage applying Western Dutch candidates in 2015 in jobs in salary scale 10 without a leadership component was 6.5 percent. This percentage decreased with 1.5 percent if it concerned a job with salary scale 11 and increased with 2 percent if it concerned a job with salary scale 12. The percentage applying Western Dutch candidates decreased with 2 percent in 2016. Model 5 does not add much to the previous models. The department dummies are almost all not statistical nor economical significant for the Western Dutch, so no further conclusions can be drawn.



Variable	Model 1	Model 2	Model 3	Model 4	Model 5
(Constant)	6,14*** (,69)	6,05*** (,96)	6,24*** (,80)	6,54*** (1,06)	4,63*** (1,74)
Year	-2,45*** (,88)	-1,93** (,93)	-2,47*** (,89)	-1,96** (,93)	-1,44 (1,06)
Leadership	-	-	-,24 (,90)	-1,02 (,95)	-0,45 (1,04)
SS 11	-	-1,33 (1,03)	-	-1,61 (1,07)	-1,70 (1,17)
SS 12	-	,87 (1,23)	-	,88 (1,23)	0,47 (1,30)
SS 13	-	1,70 (2,03)	-	2,03 (2,05)	1,58 (2,11)
SS 14	-	1,00 (2,52)	-	1,22 (2,53)	-0,16 (2,78)
DEP BSD	-	-	-	-	2,58 (1,88)
DEP SZW	-	-	-	-	2,02 (1,41)
DEP DPZ	-	-	-	-	-0,14 (2,20)
DEP OCW	-	-	-	-	0,94 (1,65)
DEP DSO	-	-	-	-	2,42 (1,60)
DEP IDC	-	-	-	-	3,02* (1,67)
DEP TUO	-	-	-	-	2,23 (3,28)
<b>R<sup>2</sup></b>	,079	,131	,079	,142	,203
<b>N</b>	92	92	92	92	92

Table 9. Regressions on the percentage Western Dutch.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

### 3.3. Regression total number of candidates

In this paragraph, a regression on the total number of applicants will be performed. In the previous regressions, I took the relative change of applying behavior into account in order to correct for possible absolute differences in total number of applicants. I described in paragraph 3.1.1. that I transformed the absolute data (number of ethnical applicants) to relative data (percentage ethnical applicants of total number of applicants) to see what happened with the proportion of the applying bicultural Dutch when AAP was introduced in 2016. It is, however, also good to investigate how much of the results of AAP are driven by an effect on the total number of applications. The independent variables are similar to model 5.

$$\begin{aligned}
 \text{MODEL 6: Total Number of Applicants}_i = & \beta_0 + \beta_1 * \text{YEAR}_i + \beta_2 * \text{LEADERSHIP}_i + \beta_3 * \text{SS11}_i + \\
 & \beta_4 * \text{SS12}_i + \beta_5 * \text{SS13}_i + \beta_6 * \text{SS14}_i + \beta_7 * D_{\text{BSD}_i} + \beta_8 * D_{\text{SZW}_i} + \beta_9 * D_{\text{DPZ}_i} + \beta_{10} * \\
 & D_{\text{OCW}_i} + \beta_{11} * D_{\text{DSO}_i} + \beta_{12} * D_{\text{IDC}_i} + \beta_{13} * D_{\text{TUO}_i} + \varepsilon_i
 \end{aligned}$$

The regressions on the total number of candidates can be found in table 10. The coefficients are given, and the standard error are shown in parentheses. The average number of applicants per ad is 44 in 2015 for a job without a leadership component in salary scale 10 in the department DSB. This average number is 20 lower in 2016 (AAP), which is almost 50% less than in 2015. This result is statistical significant. Much more applications were published in 2016 however (36 vs 56), so this explains the downsizing of the average number of applicants partly, taking into account that the total number of applicants in 2015 was 1.576 versus 1.880 in 2016. Moreover, whether a job requires an applicant with leadership skills or not, this does not affect the average number of applicants for a vacancy. The following rule apply for the salary scale: the higher the salary scale, the less applicants per ad on average. The only abnormal observation is that the average number of applicants in salary scale 13 is higher than in salary scale 12. Compared to BSD, in almost every other department, more people did apply. Departments SZW and DPZ were quite popular, but nothing compared with department BSD which has on average even 40 and 30 more applicants per ad than the popular departments SZW and DPZ respectively.

Variable	Model 6
(Constant)	43,80*** (14,11)
Year	-19,78** (8,55)
Leadership	-1,97 (8,42)
SS 11	-7,07 (9,43)
SS 12	-21,14** (10,50)
SS 13	-12,69 (17,09)
SS 14	-62,87 (22,46)
DEP BSD	67,50*** (15,23)
DEP SZW	25,49** (11,43)
DEP DPZ	36,24** (17,80)

DEP OCW	16,31 (13,35)
DEP DSO	-7,36 (12,93)
DEP IDC	-0,17 (13,47)
DEP TUO	-0,26 (26,56)
<b>R<sup>2</sup></b>	,294
<b>N</b>	92

Table 10. Regression on total number of applicants.

\*\*\*, \*\*, \* Significance at 1, 5 and 10 percent level respectively.

### 3.4. Summary findings

In sum, significantly more TMSA-Dutch candidates and significantly less Western Dutch candidates applied for jobs in the municipality The Hague in 2016 compared to 2015. For the ONW-Dutch candidates no statistically and economically significant result in terms of applying behavior can be found. In terms of salary scales, no salary scale is statistically significant in any minority group and the variable *leadership* is only statistically significant in the ONW-group. Around 5 percent less ONW-Dutch candidates were applying for jobs where managerial skills were required. When concluding in terms of salary scales on economically significant ground, especially the impact of salary scales is huge for the application behavior of the TMSA group, approximately eight times larger than for the Western group and five times larger than for the ONW-group. Lastly, around 15 percent (1 percent decrease of 6.6 percent) of the impact of anonymously applying on the percentage applying ethnic minorities can be explained by the salary scales and the variable *leadership*. The department dummies do not add much in terms of partly explaining the impact of AAP on the applying rate of ethnic minorities. The regressions on the total number of applicants shows that the average number of candidates is 50% lower in 2016 due to an increase of ads. Furthermore, the higher the salary scale, the less applicants per ad on average. Lastly, some departments are much more popular than other departments.

## 4. Power tests

### 4.1. Introduction power calculations

In order to make sure that conclusions are not wrongly rejected (indicated insignificant results are then in fact significant results), ideally the power of the statistical test must be high to draw conclusions. In The Hague, but to a larger extent in Utrecht, the sample is not large enough to investigate this popular and prominent issue and draw significant conclusions instead of economical significant conclusions. In these studies the power was not high enough to analyze the AAP experiment in more detail in order to get significant results when an effect might be present. One of the important aspects of planning a study is the calculation of the minimal required sample size that is needed to get a statistically significant effect if there is in fact exists an effect. It is obviously neither practical nor feasible to study the whole population in a study. Therefore, in order to propose a good design for a large RCT experiment in the Netherlands with the minimal required magnitude to avoid unnecessary waste of money, but with enough statistically power to state something meaningful, power tests are performed.

	<b>H<sub>0</sub> is true</b>	<b>H<sub>0</sub> is false</b>
<b>Reject H<sub>0</sub></b>	Type I Error = $\alpha$	Correct Rejection (= power: $1 - \beta$ )
<b>Fail to reject H<sub>0</sub></b>	Correct Acceptation	Type II Error = $\beta$

Table 11. Overview relation between the probabilities  $\alpha$ ,  $\beta$  and power.

The power of a test is the probability that a test correctly rejects the null hypothesis ( $H_0$ ) when in fact the alternative hypothesis ( $H_a$ ) is true. Mathematically, this is represented as  $Pr(\text{reject } H_0 | H_a \text{ is true})$ . When a null hypothesis is not rejected, it is not straightforward that the alternative hypothesis is true. Table 11 gives quickly an overview of the relation between the type I error, type II error and the power of a statistical test. Type I error ( $\alpha$ ) and Type II error ( $\beta$ ) are both probabilities and ideally are (set) as low as possible. Taking table 11 into account and based on figure 1, there is a certain coherence between  $\alpha$ ,  $\beta$  and the power. It can be distracted from figure 1 that when the power ( $1 - \beta$ ) increases, for example, caused by a shift of the alternative hypothesis, the type II error ( $\beta$ ) decreases and the type I error ( $\alpha$ ) increases when effect size (ES) and sample size (N) do not change (*ceteris paribus*).

In this paper, two types of power tests will be performed. First, I perform power tests on a single regressor. In particular, we are interested in the power of the *year* variable (AAP). Second, power calculations are done for the fit (all the explanatory variables) in order to determine the minimal required sample for a proposed new, large-scale experiment in multiple municipalities in the Netherlands. Power analysis can be used to determine the minimal sample size (N) needed in order to be 'reasonably' sure to find an effect of a given magnitude where the type I error ( $\alpha$ ) and type II error ( $\beta$ ), both probabilities, can have any value between zero and one. Power analysis can also be used to calculate the minimum effect size that is likely to be detected in an

experiment using a(n in advanced) given sample size. The power of a test depends always on (1) the statistical significance criteria (the value of  $\alpha$  in order to reject a null hypothesis), (2) the effect size (ES) and (3) the sample size N (Kirby, Gebski and Keech, 2002). In the next paragraph, the theory behind the power calculations will be explained in more detail.

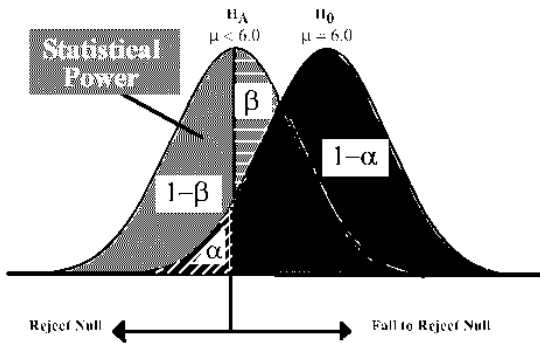


Figure 1. Graphical relation between type I and II error and power.

Source: <http://www.psychology.emory.edu>

## 4.2. Theory power tests

The power can be calculated for all mathematical measures that include a hypothesis test, for example correlations, t-test, ANOVA and multiple linear regression. Cohen (1992) argued that a different calculation is required for the effect size per mathematical approach in order to calculate an accurate power. The effect size for multiple linear regression is called Cohen's  $f^2$ . The underlying assumption is the fact that the amount of bias depends on the bias of the underlying measurement of explained variance, such as  $R^2$  (Steiger, 2004).

Cohen's  $f^2$  can be calculated as follows:

$$f^2 = \frac{\rho^2}{1 - \rho^2} = \frac{SS_{reg}}{SS_{res}} = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y_i - \hat{y})^2} = \frac{R_{adj}^2}{1 - R_{adj}^2}$$

$\rho^2$  is the squared multiple correlation between the dependent variable and the independent variables and SS is the sum of squares of the residuals or the regression. Multiple correlation indicates how accurately a variable can be predicted using a linear function. Under some assumptions, for example that the linear function intersect the origin and some nonlinearities, the squared multiple correlation equals the coefficient of determination ( $R^2$ ), which is a widely used measure to present how much variance of the dependent variable can be explained by a multiple regression. Unfortunately, these assumptions rarely hold. According to Gatsonis and Sampson (1989), the multiple correlation between the independent variables ( $X_1, X_2, \dots, X_N$ ) and the dependent variable (Y) can be calculated with the formula below:

$$\rho_{YX}^2 = \frac{\Sigma'_{YX} \Sigma_X^{-1} \Sigma_{YX}}{\sigma_Y^2}$$

The required matrices above can be extracted from the covariance matrix of the multivariate normal distribution of  $(Y, X_1, X_2, \dots, X_N)$ :

$$\begin{pmatrix} \sigma_Y^2 & \Sigma'_{YX} \\ \Sigma_{YX} & \Sigma_X \end{pmatrix}$$

$\rho^2$  is however approximately equal to  $R_{adj}^2$ , the adjusted  $R^2$ , which equals approximately to  $R^2$  when the number of observations is large and the number of independent variables is small, which can be deduced from the formula below:

$$R_{adj}^2 = \frac{(1 - R^2)(N - 1)}{N - k - 1} \quad \text{with } N = \text{sample size and } k = \text{number of predictors.}$$

#### 4.2.1. Theory power test on single regressor

For the power tests on a single regressor, the (1) effect size (the size of the effect relative to the standard error:  $\frac{D}{se(D)}$ ) and (2) the type I error ( $\alpha$ ) are required. The numerator of the effect size  $D$  is equal to  $\hat{\beta} - \beta_0$ , which equals to  $\hat{\beta}$ . I assume an  $\alpha$  of 0.05 and will perform a one-tailed test, because it is expected that the introduction of AAP will have a positive influence on the applying rate of the bicultural Dutch candidates. According to Dupont and Plummer (1998), the formula for the power of an one-sided hypothesis test<sup>10</sup> for a single regressor is as followed:

$$POWER_{single\ regressor} = \Pr\left(t_{df} > -\frac{D}{se[D]} + t_{df, \alpha}\right) \quad \text{with } df = N - 1$$

#### 4.2.2. Theory power test on the fit of a multiple linear regression

In order to calculate the power for the fit of any multiple regression, the number of data points ( $N$ ), the number of independent variables ( $k$ ), the type I error ( $\alpha$ ) and the  $R^2$  are required. The  $R$ -squared is an input here, because the  $R$ -squared is by definition the combined correlations between all variable included in the regression corrected for biases. It is intuitive that a regression with high correlated variables needs a smaller effect size in order to get the same power to provide insight in the required sample size.

First, the effect size can be calculated with the  $R^2$ :

$$f^2 = \frac{R^2}{1 - R^2}$$

---

<sup>10</sup> Note that the calculation for a two-sided test would be as followed:

$$POWER_{single\ regressor} = \Pr\left(t_{df} \leq -\frac{D}{se[D]} - t_{df, \frac{\alpha}{2}}\right) + \Pr\left(t_{df} > -\frac{D}{se[D]} + t_{df, \frac{\alpha}{2}}\right) \quad \text{with } df = N - 1$$

The degrees of freedom of the regression of the F-test equals the number of predictors ( $df_{reg} = k$ ) and the degrees of freedom of the residuals of the F-test equals the number of data point minus the number of predictors minus one ( $df_{res} = N - k - 1$ ). If a random variable  $X$  has a F-distribution with parameters  $df_1$  and  $df_2$ , then  $X \sim F(df_1, df_2)$ . The critical F-value depends then on the  $\alpha$ ,  $df_1$  and  $df_2$ . In this particular case the  $\alpha$ ,  $df_{reg}$  and  $df_{res}$  are required to calculate the critical F-value.

The cumulative density function (cdf) of the F-distribution is:

$$F(x; df_{reg}, df_{res}) = I_{\frac{df_{reg} * x}{df_{reg} * x + df_{res}}} \left( \frac{df_{reg}}{2}, \frac{df_{res}}{2} \right) = \frac{B(x; df_{reg}, df_{res})}{B(df_{reg}, df_{res})} = \frac{\int_0^1 t^{df_{reg}-1} (1-t)^{df_{res}-1} dt}{\int_0^x t^{df_{reg}-1} (1-t)^{df_{res}-1} dt}$$

Note that  $I$  is a regularized incomplete beta function. Furthermore, the critical F-value is needed in order to perform a noncentral F-test for determining the power in a regression with  $N$  data points,  $k$  predictors, a significance level of  $\alpha$  and a given  $R^2$ . The noncentrality parameter is  $\lambda$  and can be calculated by multiplying the effect size and the number of data points ( $\lambda = f^2 * N$ ).

The cumulative density function (cdf) of the noncentral F-test is:

$$F(x | df_{reg}, df_{res}, \lambda) = \sum_{j=0}^{\infty} \frac{(\frac{1}{2}\lambda)^j}{j!} e^{-\frac{\lambda}{2}} I\left(\frac{df_{reg} * x}{df_{reg} * x + df_{res}} | \frac{df_{reg}}{2} + j, \frac{df_{res}}{2}\right)$$

Again,  $I$  is a regularized incomplete beta function. Furthermore, note that when  $\lambda = 0$ , the noncentral F-distribution becomes the F-distribution. Using the critical value of  $F$ ,  $\lambda$ ,  $df_1$  and  $df_2$ , the type II error ( $\beta$ ) can be calculated and with an easy transformation the power as well.

### 4.3. Results of the power tests

In paragraph 4.2 the theory behind the two types of power test are explained. In this paragraph, I will show the outcome of the power calculations. Recall from chapter 3 that five regressions per minority group are performed. There are three minority groups where we also regress on all bicultural Dutch candidates, which can be further subdivided into the three minority subgroups: (1) Turkish, Moroccan, Surinamese or Antillean (TMSA) Dutch candidates, (2) Other Non-Western (ONW) Dutch candidates and (3) Western (W) Dutch candidates.

#### 4.3.1. Power tests on single regressor

In table 12, the effect size and corresponding power of the AAP variable can be found. Overall, the power is quite low, in particular for the ONW group in all models, where the power is around .10. This is due to the fact that for all five models of ONW Dutch candidates, the magnitude of the *year* variable is quite small and statistical insignificant. All the bicultural Dutch applicants together also do not show high power on the variable AAP. On the other hand, the power of the AAP regressor in model 1 and model 3 for TMSA and Western Dutch

applicants are even above .80, which is quite high. The other three models also do not have a substantial power for the AAP coefficient.

		Model 1	Model 2	Model 3	Model 4	Model 5
<b>Bicultural</b>	<i>Effect Size</i>	1,96	1,60	1,80	1,57	1,38
	<i>df</i>	91	91	91	91	91
	<i>Power</i>	,62	,48	,56	,46	,39
<b>TMSA</b>	<i>Effect Size</i>	2,51	2,01	2,54	1,55	1,76
	<i>df</i>	91	91	91	91	91
	<i>Power</i>	,80	,63	,81	,46	,54
<b>ONW</b>	<i>Effect Size</i>	,60	,47	,38	,42	,28
	<i>df</i>	91	91	91	91	91
	<i>Power</i>	,14	,12	,10	,11	,08
<b>Western</b>	<i>Effect Size</i>	-2,78	-2,08	-2,78	-2,11	-1,36
	<i>df</i>	91	91	91	91	91
	<i>Power</i>	,87	,66	,87	,67	,38

Table 12. Power of the single regressor AAP.

#### 4.3.2. Power tests on fit of the multiple linear regression

In table 13, the  $R^2$ , effect size and the associated power of the twenty regressions can be found. Also the required number of applications is calculated assuming the power equals .80, which is a value that most researchers assess as a standard for adequacy. The  $R^2$  fluctuates between .004 and .142, the effect size fluctuates between .034 and .124 and most important, the power fluctuates between .238 and .839. Noticeable is that the power of the ONW-group is considerably lower than the other two bicultural Dutch groups. Lastly, note that for each bicultural group the power of model 2 is smaller than the power of model 1, while the effect size and  $R^2$  are both larger in model 2 compared to model 1. This is due to the heavily increase of predictors (adding salary scale dummies) in the regression.

		Model 1	Model 2	Model 3	Model 4	Model 5
<b>Bicultural</b>	$R^2$	,041	,114	,063	,119	,209
	<i>Effect Size</i>	,076	,107	,087	,105	,127
	<i>Power</i>	,669	,574	,634	,521	,560
	<i>N (power = .80)</i>	122	136	127	147	137
<b>TMSA</b>	$R^2$	,066	,130	,066	,140	,254
	<i>Effect Size</i>	,097	,118	,090	,124	,156



	Power	,790	,634	,648	,600	,687
	N (power = .80)	95	123	124	129	112
<b>ONW</b>	R <sup>2</sup>	,004	,044	,055	,079	,200
	Effect Size	,034	,051	,063	,075	,121
	Power	,324	,238	,592	,347	,532
	N (power = .80)	273	291	137	207	144
<b>Western</b>	R <sup>2</sup>	,079	,131	,079	,142	,203
	Effect Size	,108	,118	,100	,120	,123
	Power	,839	,638	,705	,607	,542
	N (power = .80)	84	122	111	127	142

Table 13. Power of the fit and required sample size.

#### 4.4. Required sample size calculation based on power calculations and previous research

Cohen (1988) argued that a  $f^2$  value of .02, .15, and .35 represent small, medium, and large effect sizes respectively. Transforming this to the  $\rho^2$  values, through the formula  $f^2 = \frac{R^2}{1-R^2}$ , this implies that a  $\rho^2$  value of .02, .13 and .26 represent small, medium and large effect sizes. This implies that the effect sizes in the twenty regressions are mostly small-medium. Earlier research related to anonymously applying also studied AAP in order to state something meaningful in terms of power and associated variables in this particular field. The number of applying candidates, job applicants, number of independent variables in the regression, the R-squared and the effect size is collected. The overview can be found below in table 14.

Onderzoek	N (sample)	K (predictors)	R <sup>2</sup>	Effect Size	Power	Required N (power = .80)
Behaghel, Crépon and Le Barbanchon (2011)	1260 candidates, 595 job applicants	9	.109	.100	1.000	168
Aslund and Skans (2007)	3529 candidates, 109 job applicants	6 (+ 31 fixed effects)	.30	.069	0.993	108
Kraus, Rinne, Zimmerman (2012a)	809, 10 job applicants	2 (+ 3 fixed effects)	.016	.015	.715	948
Kaas en Manger (2012)	1056 applicants	7	?	-	-	-
Kraus, Rinne and Zimmerman (2012b)	82 job applicants	12	?	-	-	-

Table 14. Power and required sample size earlier AAP studies.

In terms of power, the researches of Behaghel, Crépon and Le Barbanchon (2011) and Aslund and Skans (2007) are truly remarkable, since the power of the studies are almost equal to 1. The effect size of the study of Aslund and Skans (2007) equals .069 and the effect size of Behaghel, Crépon and Le Barbanchon (2011) is .100. This difference in effect size can be explained by the large number of observations, and the high R<sup>2</sup>

(3529 observations and  $R^2$  equals .30) of the study of Aslund and Skans (2007). The research of Kraus, Rinne, Zimmerman (2012a) had a smaller power than Behaghel, Crépon and Le Barbanchon (2011) and Aslund and Skans (2007), what can be explained by the small effect size. In imitation of the researches of Behaghel, Crépon and Le Barbanchon (2011) and Aslund and Skans (2007), the desired power will be set to .80. The magnitude of the effect size should be chosen based on 'worst case' of the abovementioned earlier studies, so the required sample size can set to ensures us to detect even a small, but meaningful effect size. As worst case scenario, the effect size will be assumed to be equal to the effect size found by Kraus, Rinne, Zimmerman (2012a), namely .015. The type I error ( $\alpha$ ) will be held at .05 and these test will be performed one-tailed, since anonymously applying has in general a positive impact on percentage bicultural Dutch candidates. In order to determine the required sample size, as a last step, the number of predictors has to be established. In figure 2, the relation between the required sample size and the power is displayed. Noticeable is the fact that more power implies an exponential increase in the required sample size. This means that a small increase in power makes an experiment much more expensive due to the exponential increasing required sample size. For a more certain conclusion, much more data is needed. Based on the design, which is explained and described in detail in the next chapter, forty predictors seem to be a good conservative estimation. Taking all abovementioned assumptions into account (one-tailed testing,  $ES_{\text{worst case}} = .015$ ;  $\alpha = .05$ ; power = .80; number of predictors = 40), 1841 job applications are required to meet these criteria. In the proposal, for facilitation purposes and just to be on the safe side, I will assume that 2000 job applications are required to conclude about an AAP experiment with sufficient certainty and power.

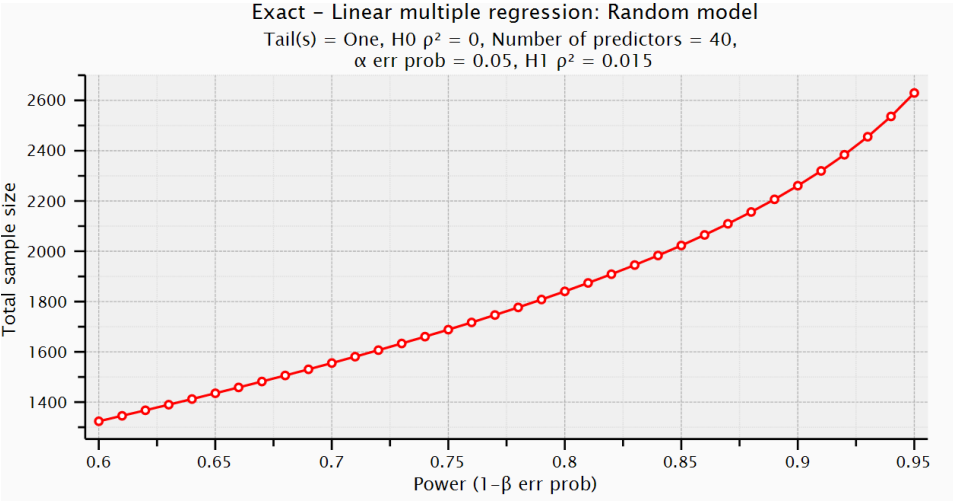


Figure 2. Power test for calculating the required sample size in proposed experiment.

## 5. Proposal and design for experiment in the Netherlands on large scale

In this chapter, a proposal will be presented in order to conduct an experiment on a large scale to investigate the effects of anonymously applications on the applying behavior of bicultural Dutch candidates. The feedback about the researches in The Hague and Utrecht was quite positive, with some critical remarks which overall implied that an extension of the current experiment(s) is needed to investigate anonymously applying on larger scale to draw more certain and detailed conclusions about the effect of AAP on different ethnic groups in the recruiting / hiring process. The studies in The Hague and Utrecht are experiments where conclusions can be drawn in terms of correlations, not in terms of causal effects of AAP. This is due to the fact that no control group and no randomization can be found in the design of the study in The Hague. My proposed design is a RCT (Randomized Controlled Trial), with an intervention group and a control group running parallel and random assignment of the experimental condition to a vacancy.

A short documentation<sup>11</sup> of the pilot in Den Hague is used by the Minister of Social Affairs and Employment, L.F. Asscher, as a reaction to a request of C. S. Marcouch to convince the House of Representatives to perform an anonymous job applications experiment on greater scale in the Netherlands. The vote was a close call, but the request was not accepted in December 2016. Therefore, in this section a new research proposal will be proposed, taking the power tests into account, in order to investigate the behavior of ethnic groups caused by the anonymously applying phenomenon. Also, the behavior of the recruiters will be investigated. In other words, the job offer probability will also be analyzed through the ethnical groups in a RCT design.

### 5.1. Description proposal and design

In this design, the same three ethnical groups<sup>12</sup> as the The Hague AAP study will be included in the analysis, since Moha and Konings (2016) argued that in these three groups the signals of exclusion in the labor market are overall the greatest. In imitation of the study in The Hague, the callback rate and the job offer probability through ethnical groups will be investigated. In particular, the analysis of the job offer probability did not lead to significant results in The Hague due to a low sample size.

In figure 3, a global overview of the design is given where the number of the minimal required job applications is shown in brackets. Based on the assumptions for this proposal, which are discussed and elaborated in paragraph 4.4., 2000 job applications are required when aiming for a power of .80. The experiment will preferably run for two years, from January 1, 2018 till December 31, 2019. In the study in The Hague, not enough vacancies were published in two periods of six months which were included in the research. In this design, the time span can be extended in order to attain the minimal required vacancies for analysis purposes to draw a meaningful, statistical conclusion in terms of causal effects. Also, seasonality effects could

---

<sup>11</sup> <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2016/12/05/kamerbrief-over-rijksmonitor-anoniem-solliciteren/kamerbrief-over-rijksmonitor-anoniem-solliciteren.pdf>

<sup>12</sup> Turkish, Moroccan, Surinamese or Antillean (TMSA) candidates, Other Non-Western (ONW) candidates and Western (W) candidates.

be investigated by collecting data on a timespan of at least two full years. The key improvement is the fact that the control group and the experimental group are running parallel in this design for the full time span. In the design of the AAP study in the Hague the AAP effect was measured by comparing the data of 2016 (AAP) with 2015 (control year). In that design, changes over time could bias the results, where in this proposal, the analysis contains more true hiring discrimination effect than a year dummy. Also, randomization is possible within this design. The applications will be randomly assigned into the control group or treatment group. Assuming that fifty applications per year will be published at the municipality of any city where this experiment will be performed, twenty cities should be willing to participate in this experiment for 2 years. A small overview in terms of needed job applications on municipality level will be given in figure 4.

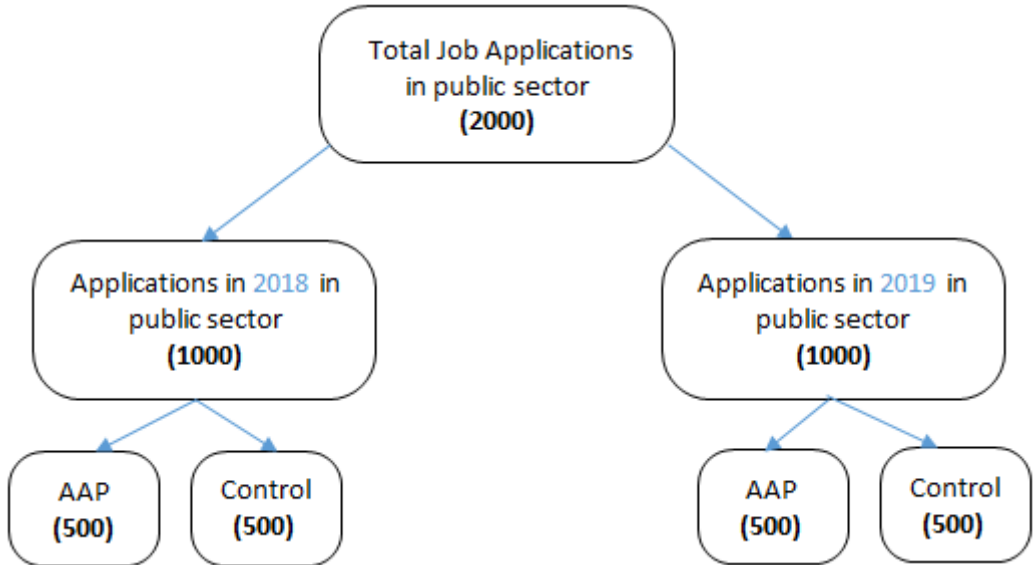


Figure 3. Overview design experiment on total level

Figure 3 shows that in this experiment, half of the job application will follow the traditional applying protocol and the other half of the job applications will follow AAP. In this study the focus will be on the effect of AAP in the public sector. Therefore, the anonymously applying experiment should get started at municipalities of twenty different cities. In this experiment hiring discrimination through ethnicity and gender will be investigated. Ideally, hiring discrimination through age would also be combated in this study, but that is not preferable. Experience is a very important part of a resume and should therefore not be left out in any resume, but unfortunately experience is a good indication of the candidates' age. In the choice making between standardized applying formats or remove sensitive parts manually, it is taken into account that standardized is less labor-and time intensive, but standardized applying forms exclude the possibility to present yourself in a way you like. In other words, the personal aspect of a job application is gone. Since the style of writing between men and women significantly differ (Rubin & Greene, 1992; Koppel, Argamon & Shimoni, 2002), the standardized applying forms are preferable in this design. In imitation of the research of the municipality The Hague, the focus should be on jobs in the salary scale range from 10 to 14. Also, a few other

fixed effects are included in the regression, such as hiring-, job- and personal characteristic fixed effects. As stated before, the applying rate of minorities that is caused by AAP will be investigated. After that, the call back rate will be explored and lastly the job offer rate will be analyzed in order to draw conclusions in terms of delayed hiring discrimination (paragraph 2.1.4.).

### 5.2. Final remarks proposal and design

It is worth noting that Aslund and Skans (2007) claimed that the different styles of applying / differences in standardizing the application document through countries could induce undesirable biases. Since no extensive research has been conducted over multiple cities in the Netherlands, it is wise to conduct a hiring discrimination research on large scale in the Netherland first. After evaluating, a possible expansion to other countries in Europe can be performed and taking abovementioned statement into account in that design. However, in this design, the difference in hiring discrimination through ethnical groups per town in the Netherlands will also be analyzed.

Further, minimizing the John Henry effect is hard, even in this RCT experiment. The John Henry effect occurs when participating recruiters of a control group know that the job application is part of the experiment and therefor select more of the discriminated candidates just for signaling to the rest of the world that they do not discriminate at all. In order to participate on this experiment the municipality or firm, among others, the whole selection process needs to be overhauled, which cannot be missed by any recruiter.

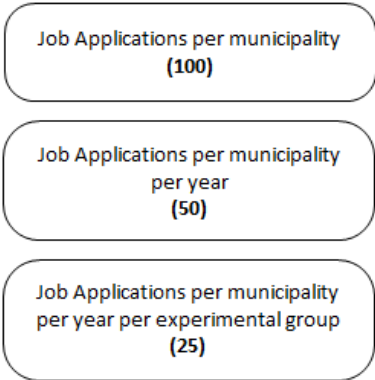


Figure 4. Overview design experiment on municipality level

In short, this proposal is designed in order to investigate hiring discrimination through ethnic groups and gender in a large-scale research in twenty cities in the Netherlands in a time frame of at least two years. The AAP and the ‘standard’ applying procedure will run parallel by randomizing the appearing job applications into these two groups. The same ethnical groups and salary scale restrictions as the study in The Hague are applied. The study will be having three phases: investigating the applying behavior, the call back probability and the job offer probability through ethnical groups and gender.

## 6. Conclusions

In this paper, some additional research about the anonymously applying experiment in The Hague has been performed in order to get more insight in the correlation between AAP, applying behavior of minorities and job characteristics. This study focused on differences in magnitude of hiring discrimination through jobs in different salary scales in different departments and whether with or without a managerial aspect. First, more TMSA-Dutch candidates and less Western Dutch candidates applied for jobs for the municipality The Hague in 2016 compared to 2015. Second, only ONW-Dutch candidates applied significantly less on jobs which required managerial skills (5 percent). Third, TMSA-Dutch candidates applied much less on high-level jobs than low-level jobs than the two other minority groups. Fourth, around fifteen percent of the impact of anonymously applying on the percentage applying disadvantaged groups can be explained by the salary scales and the variable leadership. Furthermore, the power of the twenty performed regressions are calculated. The result of the power tests is that the sample of the AAP study in The Hague is too small in order to get more detailed significant result whether AAP might influence the applying behavior of minorities in the Netherlands. Therefore, some additional power tests are performed in order to calculate the required sample size for a proposed, first large-scale RCT experiment in multiple cities in the Netherlands. It is useful to investigate anonymously applying in more detail to find a solution for the structural hiring discrimination problem. This proposal indicates that around 2000 job applications in twenty cities need to be analyzed in two years to state something meaningful about anonymously applying in causal terms.

## References

- Aigner, D. J., & Cain, G. G. (1977). Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, 30(1), 175 – 187.
- Andriessen, I., Ent, van der, B., Linden, van der, M., & Dekker, G. (2015). Op afkomst afgewezen: Onderzoek naar discriminatie op de Haagse arbeidsmarkt. *Sociaal en Cultureel Planbureau*.
- Aslund, O., & Skans, O. N. (2007). Do Anonymous Job Application Procedures Level the Playing Field? *Industrial and Labor Relations Review*, 65(1), 362–372.
- Becker, G. S. (1957). Theory of Discrimination. *Chicago: University of Chicago Press*.
- Behaghel, L., Crépon, B., & Le Barbanchon, T. (2012). Do anonymous resumes make the battlefield more even? Evidence from a randomized field experiment. *Mimeo, Centre de Recherche en Economie et Statistique (Paris)*.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4), 991-1013.
- Blank, R. M. (1994). The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence. *The American Economic Review*, 81(5), 1041-1067.
- Blommaert, L., Counders, M., & Tubergen, van, F. (2014). Discrimination of Arabic-Named Applicants in the Netherlands: An Internet-Based Field Experiment Examining Different Phases in Online Recruitment Procedures. *Social Forces*, 92(3), 957-982.
- Carlsson, M., & Rooth, D. (2006). Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data. *Labour Economics*, 14, 716–729.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Devine, P. G., & Plant, E. A., Amodio, D. M., & Garmon-Jones, E. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, 82(5), 835-848.
- Drydakis, N., & Vlassis, M. G. (2007). Ethnic Discrimination in the Greek Labour Market: Occupational Access, Insurance Coverage, and Wage Offers. *Working Paper No. 0715, University of Crete, Department of Economics*.
- Dupont, W. D., Plummer, W. D. (1998). Power and Sample Size Calculations for Studies Involving Linear Regression. *Controlled Clinical Trials*, 19, 589 – 601.

- Fang, H., & Moro, A. (2011). Theories of Statistical Discrimination and Affirmative Action: A Survey, Jess Benhabib, Matthew Jackson and Alberto Bisin, eds: Handbook of Social Economics, Vol. 1A, Chapter 5, The Netherlands: North Holland, 2011, pp. 133-200.
- Finkelstein, M. O. (1966). The Application of Statistical Decision Theory to the Jury Discrimination Cases . *Harvard Law Review*, 80(2), 338 – 376.
- Gaertner, S. L. (1990). How Does Cooperation Reduce Intergroup Bias? *Journal of Personality and Social Psychology* 59(4), 692-704.
- Gatsonis, C., & Sampson, A. R. (1989). Multiple correlation: Exact power and sample size calculations. *Psychological Bulletin*, 106, 516-524.
- Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of ‘blind’ auditions on female musicians. *American Economic Review* 90(4), 715–741.
- Hausman, D. (2012). How Congress Could Reduce Job Discrimination by Promoting Anonymous Hiring. *Stanford Law Review*, 64, 1343 – 1370.
- Hoogendoorn, S. , Praag, van, M. (2012). Ethnic Diversity and Team Performance: A Field Experiment. *IZA Working Paper No. 6731*.
- Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany’s labour market: a field experiment. *German Economic Review*, 13(1), 1–20.
- Kang, J., Banaji, M. R. (2006). Fair Measures: A Behavioral Realist Revision of “Affirmative Action”. *California Law Review*, 94(4) ,1094-1139.
- Kirby, A., Gebski, V., & Keech, A. C. (2002). Determining the sample size in a clinical trial. *Medical Journal of Australia*, 177(5), 256–257.
- Koppel, M., Argamon, S., & Shimon, A. R. (2002). Automatically Categorizing Written Texts by Author Gender. *Lit Linguist Computing*, 17(4), 401-412.
- Krause, A., Rinne, U., & Zimmermann, K. F. (2012a). Anonymous job applications in Europe. *Journal of European Labor Studies*, 5(1), 1 – 20.
- Krause, A., Rinne, U., & Zimmermann, K. F. (2012b). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117, 441 – 444.
- Kuhn, P., & Shen, K. (2013). Gender discrimination in job ads: evidence from China. *The Quarterly Journal of Economics*, 128 (1), 287-336.
- McGinnity, F., Nelson, J., Lunn, P., & Quinn, E. (2009). Discrimination in Recruitment: Evidence from a Field Experiment. *Dublin: The Equality Authority/The Economic and Social Research Institute*.



- Moha, A. A., & Konings, D. (2016). Effecten van de pilot anoniem solliciteren. *Motivaction research and strategy*.
- Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a Low-Wage Labor Market: A Field Experiment. *American Sociological Review*, 74, 777–799.
- Pérez, D. J., Fortuna, L., & Alegría, M. (2008). Prevalence and correlates of everyday discrimination among U.S. Latinos. *Journal of Community Psychology*. 36(4), 421–433.
- Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review*, 62 (4), 659-661.
- Rinne, U. (2014). Anonymous job applications and hiring discrimination. *IZA World of Labor*.
- Rubin, D. L., & Greene, K. (1992). Gender-Typical Style in Written Language. *Research in the Teaching of English*, 26(1), 7-40.
- Sarbin, T. R. (1943). A Contribution to the Study of Actuarial and Individual Methods of Prediction. *American Journal of Sociology*, 48(5), 593-602.
- Siles, M., Hanson, S. D., & Lindon, J. R. (1994). Socio-Economics and the Probability of Loan Approval. *Review of Agricultural Economics*, 16 (3), 363-372.
- Steiger, J. H. 2004. Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods* 9(2), 164-182
- Tomaskovic-Devey, D., & Skaggs, S. (1999). An Establishment-Level Test of the Statistical Discrimination Hypothesis. *Work and occupation*, 26(4), 422 – 445.
- Yinger, J. (1998) Evidence on discrimination in consumer markets. *Journal of Economic Perspective*, 12(2), 23–40.

# Appendix

## Appendix A

Request Vermeij / Van Weyenberg about local experiments with anonymous job applications - Establishing the budget of the Ministry of Social Affairs and Employment (XV) for the year 2016.

Presented December 2, 2015

The house of Representatives,

heard the deliberations,

noting that everyone counts in full in the Netherlands and is an integral part of our society;

noting that unequal treatment, disadvantage or exclude people based on (personal) characteristics that are not relevant, stands in the way of the equality of people;

noting that employment discrimination is still too often experienced and does still occur;

whereas anonymous applying can reduce the bias in the application process in some cases;

whereas anonymous applying can be one of the means to combat labor market discrimination;

whereas on the basis of experiences elsewhere in Europe, where local initiatives have been launched, in the Netherlands those initiatives are still in the beginning phase;

requesting the government to promote and monitor the local experiments with anonymous job applications actively. The results and next steps thereafter need to forward to the summer recess of the House and thereby continuing the fight against discrimination,

and proceeds to the order of the day.

Vermeij

Van Weyenberg

## Appendix B

### Appendix B.1. Data per released application of the municipality The Hague in 2015.

Application Code	Application Name	Salary scale	Candidates	TMSA Candidates	Bicultural candidates-Western	Other non-Western (ONW)	% candidates TMSA+ONW	% candidates TMSA
SZW-2015-0196	Head of Customer Service	10	143	31	12	9	28%	22%
SZW-2015-0091	Communications adviser	10	130	16	2	7	18%	12%
SZW-2015-0024	Head Back Office team Intervention & Screening (TIS)	10	75	13	5	6	25%	17%

DSO-2015-0035	Tourism Policy Officer	11	76	8	4	9	22%	11%
DSB-2015-0111	Lawyer (civil law)	11	47	7	0	6	28%	15%
SZW-2015-0113	Head Bridges management	10	62	10	3	3	21%	16%
DSO-2015-0034	Creative Industries Policy Officer	11	66	8	5	2	15%	12%
OCW-2015-0148	Direction Secretary	12	69	7	1	3	14%	10%
OCW-2015-0176	Pediatrician	12	19	3	1	7	53%	16%
OCW-2015-0119	Project Manager Construction & Engineering	11	39	5	1	4	23%	13%
OCW-2015-0025	Functional manager	10	53	5	4	3	15%	9%
SZW-2015-0087	Senior communication consultant	11	46	4	4	4	17%	9%
SZW-2015-0147	Business Controller Planning & Control	12	37	3	4	5	22%	8%
DSB-2015-0084	Head enforcement teams	10	39	6	3	1	18%	15%
OCW-2015-0057	Department Manager Central Coordination	13	46	6	3	1	15%	13%
SZW-2015-0055	Policy Advisor	13	61	5	3	2	11%	8%
DPZ-2015-0029	Valuator WOZ	10	51	5	1	1	12%	10%
IDC-2015-0185	Senior IT Project Manager & Architecture	12	63	5	7	1	10%	8%
OCW-2015-0150	Head staff department Student Affairs	12	33	3	3	3	18%	9%
OCW-2015-0167	Doctor forensic medicine	12	13	2	0	4	46%	15%
TUO-2015-0089	Temporary communication consultant executive WMO	10	62	3	4	3	10%	5%

DSO-2015-0046	Team manager The Hague property Brigade	12	49	2	3	3	10%	4%
DSO-2015-0191	Project Manager project contracting work	10	19	4	1	1	26%	21%
IDC-2015-0052	Senior employment lawyer	12	30	3	1	2	17%	10%
IDC-2015-0157	Senior Service desk Advisor	10	41	5	3	0	12%	12%
IDC-2015-0187	Solution Architect Projects & Architecture	11	7	3	0	2	71%	43%
IDC-2015-0098	Purchasing Consultant A (senior)	12	28	0	2	4	14%	0%
IDC-2015-0184	IT Project Manager & Architecture	11	16	3	2	1	25%	19%
SZW-2015-0071	Experienced Head WMO (individual provisions)	10	32	4	3	0	13%	13%
DSB-2015-0207	Head Traffic Center and Port	12	21	2	1	1	14%	10%
IDC-2015-0044	senior communication consultant	11	23	3	1	0	13%	13%
IDC-2015-0186	Senior solution architect Projects & Architecture	12	7	2	1	1	43%	29%
IDC-2015-0083	Product Manager Projects and Architecture	13	20	0	2	2	10%	0%
DSB-2015-0140	Traffic Manager Central and Port Authority	11	27	1	0	0	4%	4%
IDC-2015-0099	Product Manager Knowledge Centre	13	8	0	1	0	0%	0%
TUO-2015-0088	Senior Communications Temporary executive organization WMO	11	18	0	1	0	0%	0%

## Appendix B.2. Data per released application of the municipality The Hague in 2016.

Application Code	Application Name	Salary scale	Candidates	TMSA candidates	Bicultural candidates-Western	Other non-Western (ONW)	% candidates TMSA+ONW	% candidates TMSA
BSD-2016-0032	Central confidant for inappropriate behavior	11	188	30	9	17	25%	16%
SZW-2016-028	Policy Officer, Work, Learn and Income	11	90	22	7	13	39%	24%
DPZ-2015-0373	Coördinator Reception staff	10	85	20	2	6	31%	24%
SZW-2016-0080	Policy Officer Poverty and WMO	11	81	19	5	3	27%	23%
BSD-2016-0103	Policy advisor International Affairs	11	208	18	19	18	17%	9%
OCW-2016-0078	Policy Officer Participation	11	66	17	4	6	35%	26%
OCW-2016-0228	Junior policy officer emancipation	10	50	16	3	3	38%	32%
DSB-2016-0133	Coordinator logistical planning, guidance and mediation	10	57	16	3	5	37%	28%
SZW-2016-0180	Management consultant WSP	11	49	11	1	2	27%	22%
OCW-2016-0200	Team Leader Student Affairs	10	72	11	0	4	21%	15%
DSB-2016-0219	Medior financial project controller	11	29	10	2	5	52%	34%
DSO-2016-0006	Employee real estate	10	43	10	4	3	30%	23%
BSD-2015-0353	Policy Officer and subsidy consultant	11	64	10	7	5	23%	16%
DPZ-2014-0209	Senior functional administrator	10	70	9	4	4	19%	13%

DSB-2016-0123	Jurist public space usage	11	29	8	0	3	38%	28%
DSB-2016-0021	Leader enforcement teams	10	46	8	1	3	24%	17%
IDC-2016-0087	Team Manager location management	12	57	8	2	5	23%	14%
DSO-2016-0082	Lawyer Land Administration	10	26	7	1	5	46%	27%
DSO-2016-0061	Lawyer Land Administration	11	22	6	1	5	50%	27%
DSB-2016-0001	Policy Officer	11	6	5	0	0	83%	83%
OCW-2016-0194	Analyst Student Affairs	10	22	5	1	3	36%	23%
SZW-2016-0170	Management Controller	11	20	5	1	1	30%	25%
SZW-2016-0135	Experienced principal for Debt Counselling	10	21	5	1	1	29%	24%
SZW-2016-0010	Quartermaster / contemplated Domain Manager Customer Contact, Quality and Learning	14	23	5	2	1	26%	22%
SZW-2016-0169	Medior functional specialist	10	27	5	1	1	22%	19%
DSO-2016-0144	Medior Plan Economist and Junior Plan Economist	11	7	4	0	0	57%	57%
DPZ-2015-0424	Management Controller	12	80	4	5	5	11%	5%
DSB-2016-0152	Experienced principal for Debt Counselling	11	7	3	0	1	57%	43%
BSD-2015-0427	Quartermaster / contemplated Domain Manager Customer Contact, Quality and Learning	11	20	3	0	5	40%	15%
DSB-2016-0188	Medior functional specialist	11	8	3	0	0	38%	38%

IDC-2016-0091	Medior plan economist and Junior Economist Plan	11	18	3	1	1	22%	17%
SZW-2016-0029	Head of Department & Events	12	19	3	1	0	16%	16%
IDC-2016-0157	Senior project design	10	25	3	0	0	12%	12%
DSB-2016-0045	Concern Advisor employment conditions / labor lawyer	12	4	2	0	0	50%	50%
DSB-2016-0136	Technical manager	10	4	2	0	0	50%	50%
DSB-2016-0147	Security and Safety Advisor	11	5	2	0	0	40%	40%
SZW-2016-0049	Senior policy Participation & Services debt counseling	10	19	2	1	1	16%	11%
DSB-2016-0035	Senior technical building administrator	13	35	2	1	1	9%	6%
DSB-2016-0187	Systems coordinator	11	2	1	0	0	50%	50%
BSD-2016-0175	Medior Project Manager	12	15	1	1	3	27%	7%
SZW-2016-0225	Junior Project Manager / Technical Manager	12	4	1	0	0	25%	25%
BSD-2016-0174	Head Employers Service Desk	12	16	1	0	2	19%	6%
DSB-2016-0046	Senior Principal / Project Manager	11	8	1	0	0	13%	13%
DPZ-2016-0215	Valuator WOZ	10	10	1	0	0	10%	10%
OCW-2016-0101	Epidemiological researcher with knowledge of big data	11	14	1	1	0	7%	7%
DSB-2016-0003	Senior traffic engineer Accessibility	11	15	1	0	0	7%	7%
DSO-2016-0017	Building Constructions Adviser	11	3	0	0	2	67%	0%

SZW-2016-0167	Architect Business Intelligence	11	2	0	0	1	50%	0%
DSO-2016-0083	Geo-ICT consultant	11	19	0	0	6	32%	0%
DSO-2016-0008	Specialist geographic information	10	18	0	0	5	28%	0%
DSB-2016-0127	Traffic engineering programmer traffic lights	10	6	0	1	1	17%	0%
BSD-2016-0186	Department Manager StrategieLab	14	15	0	1	2	13%	0%
DSO-2016-0108	Senior policy researcher	12	19	0	4	2	11%	0%
BSD-2016-0205	Lead Management Consulting and Strategy / Senior Advisor Public Affairs	14	2	0	0	0	0%	0%
DSB-2016-0149	Policy advisor	11	1	0	0	0	0%	0%
SZW-2016-0232	Senior functional specialist	11	7	0	0	0	0%	0%

---



