Getting Grey Hairs in the Labour Market. An Alternative Experiment on Age Discrimination

Abstract

This study presents a new field experimental approach for measuring age discrimination in hiring. In addition to the classical approach in which candidates' ages are randomly assigned within pairs of fictitious résumés that are sent to real vacancies, we randomly assign between these pairs the activities undertaken by the older candidates during their additional post-educational years. When applying this design to the case of Belgium, we find that age discrimination depends fundamentally on the older candidates' career patterns. Older age only robustly affects call-back if the older candidate was employed in an out-of-field job during his or her extra post-educational years.

Keywords: Age discrimination; Design of experiments; Field experiments; Difference in Post-Educational Years Problem; Ageing; Hiring discrimination.

PsycINFO Classification codes: 3620; 3660.

JEL Classification codes: C90; C93; J14; J71.

1   Introduction

Population ageing is one of the most pressing challenges facing the western world. In 2050, for every person over the age of 65 there will only be 2.7 individuals of working age in the United States and 1.9 in Western Europe, compared to 5.4 and 4.2, respectively, in 2000.[1] This ageing places substantial pressure on public finances with respect to pensions and medical care. It is a widespread belief that the only viable way to address the challenge of ageing, without compromising living standards, is to encourage older workers to remain in the labour force (Lahey, 2008; OECD, 2006; Riach & Rich, 2010). During the past decade, many countries have carried out reforms in pursuit of this policy. Sonnet, Olsen, and Manfredi (2014) reviewed these reforms and concluded that they have improved the labour market situation of older workers. Nevertheless, activity rates among older age groups are relatively low, especially in Europe. In the EU-15, the employment rate in 2013 for the 50 to 64 age group was 61.2%, compared to 77.0% for the 25 to 49 age group. In Belgium, the country of analysis in this study, these numbers are even more divergent: 54.0% versus 79.9%, respectively.[2] It is important to determine the cause of these remaining differences so that further effective policy actions can be designed. Theoretically, there are three explanations for these gaps: (i) differences in employee productivity; (ii) differences in employee preferences and behaviour; and (iii) differences in employer preferences and behaviour (Kautonen, Hatak, Kibler, & Wainwright, 2015; Lahey, 2008; OECD, 2006). While policy discussions have traditionally focussed only on supply-side factors, policy attention has recently shifted more to the third of these channels, better known as discrimination (Sonnet et al., 2014). In this study, we focus on age discrimination in the first decision in the hiring process – the employer's decision (not) to invite a candidate to a job interview. This is the moment at which age discrimination is most likely to occur, as it is then least likely to be detected and, therefore, least costly for employers (Bendick, Brown, & Wall, 1999; Cédiey, Foroni, & Garner, 2008).

---

[1] Source: United Nations, World Population Prospects: The 2015 Revision ("Old-age dependency ratio 65+/(15–64) by major area, region and country"; Estimates 1950–2015; Medium fertility variant prospections 2015–2100). Western Europe comprises Austria, Belgium, France, Germany, Luxembourg, the Netherlands and Switzerland.

[2] Source: Eurostat, Labour Force Survey ("Employment rate by age groups").

Hiring discrimination against older candidates can be expected on the basis of the theories of taste-based discrimination (Becker, 1957) and statistical discrimination (Arrow, 1973). The former suggests that employers, customers and co-workers may experience a disutility when interacting with older workers. The latter suggests that employers may judge individual older workers on group characteristics rather than purely on their individual merits. Adverse group characteristics of the elderly, at least in the perception of employers, might be: (i) showing less energy, motivation, creativity, flexibility and/or adaptability; (ii) being more difficult to supervise; (iii) having poorer health; (iv) facing obsolescence of their human capital; and (v) having higher salary aspirations (Albert, Escot, & Fernández-Cornejo, 2011; Bendick, Brown, & Wall, 1996; Lahey, 2008; Riach & Rich, 2010).

Over the last two decades, scholars have attempted to measure age discrimination in the labour market. To this end, 11 studies applied the gold standard, namely correspondence experiments, to identify unequal treatment in the labour market. In this type of experiment, pairs of fictitious job applications are sent in response to real job advertisements. Each pair of applications is virtually the same except for the age of the applicants, which is the ground of discrimination on which we focus. By monitoring the subsequent positive reactions from employers (or "call-back"), we can identify unequal treatment arising from this characteristic and give the unequal treatment a causal interpretation. Using this experimental setting, high levels of age discrimination have been found in Australia, England, France, Spain, Sweden and the United States (Ahmed, Andersson, & Hammarstedt, 2012; Albert et al., 2011; Bendick et al., 1996; Bendick et al., 1999; Gringart & Helmes, 2001; Lahey, 2008; Neumark, Burn, & Button, 2015; Riach & Rich, 2006a, 2007, 2010; Tinsley, 2012). However, the (classical) application of the correspondence experimentation framework by most of the contributions cited here is problematic, because of a complication that in the present study we call the Difference in Post-Educational Years Problem. An older person inevitably has, in comparison with a younger person with the same educational background, a higher number of post-educational years at the moment of her/his application. Except for Neumark et al. (2015), who conducted their experiment simultaneously with ours, the correspondence studies mentioned above "filled in" these additional years in the résumé of the older candidates with a particular activity (additional in-field employment, out-of-field employment or inactivity), or let employers fill them in themselves (by only

3

mentioning professional experience for the most recent career years). The particular listed or perceived activity undertaken by the older candidates may, however, send out an additional positive or negative signal to an employer, and might thereby bias the discrimination measures downwards or upwards, respectively. In other words, the experimental design applied most in the earlier research makes it difficult to determine whether unequal treatment is attributable to age discrimination or to discrimination based on differences in human capital or career gaps.

In the present study, we deal with the Difference in Post-Educational Years Problem by proposing an extended correspondence experiment. In this extended design, we combine the classical within-pair randomisation of the age of fictitious job candidates with a between-pair randomisation of the activity undertaken by the older candidates during their additional post-educational years. With this procedure, we avoid focusing on only one type of extra post-educational experience of the older worker, as previous studies have done. By taking this approach, we obtain an upper and a lower bound for the amount of age discrimination.

We report on the application of this framework to measure age discrimination in the Belgian labour market. During six months, we sent out pairs of fictitious job applications (from both women and men) to a balanced number of vacancies for the lower skilled occupations of production operator, administrative clerk and waiter, and for the higher skilled occupations of laboratory analyst, management assistant, and sales representative. For each vacancy, the younger age of 38 or 44 was disclosed in one of the applications, and the older age of 44 or 50 was disclosed in the other. Furthermore, we randomly assigned to the older member of the pair one out of three activities performed during her/his 6 or 12 extra post-educational years: in-field employment, out-of-field employment or inactivity. As a consequence, this setting allowed us to measure unequal treatment based on age in three realistic situations.

The present study contributes to the existing literature for a number of additional reasons. First, while most former contributions have focussed on estimating age discrimination in one or two, mostly lower skilled, occupations, we test vacancies in six occupations at two different qualification levels and in three different areas. This design is not only preferable in terms of the generalisability of the research results, but it also allows us to investigate whether age discrimination

is heterogeneous by qualification level and/or by occupation characteristics. Second, we use both pairs of female and pairs of male applicants, which also allows us to inspect gender heterogeneity in age discrimination. Third, in our setting the difference in age within a pair is 6 or 12 years. On the one hand, these age differences are substantially smaller than the age differences adopted in the literature (on average 20 years),[3] which, in our opinion, yields a more realistic experimental setting. We argue that the unequal treatment of relatively close age groups is a stronger signal of age discrimination as such, as it is less likely that other factors (e.g., public policy) can have a different influence on the productivity-wage ratio of applicants of different ages. On the other hand, this variation in age difference allows us to look into whether unequal treatment varies linearly with age difference.

In the next section, we provide the reader with a systematic literature review of the earlier correspondence studies on age discrimination, with a focus on how these contributions dealt with the Difference in Post-Educational Years Problem. In Section 3, we explain how we take this problem into account in our experimental design. Section 4 reports our measures of age discrimination based on an econometric analysis of the experimentally gathered data. A final section concludes.

2 Related Literature

2.1 Evidence on Age Discrimination from Correspondence Experiments

Over the past few decades, economists have conducted various correspondence experiments to measure discrimination in the labour market. In this type of experiment, pairs of fictitious job applications are sent in response to real job advertisements. These applications differ only in the characteristic that is to be tested. This characteristic is randomly assigned within each pair of applicants. By

---

[3] Own calculation based on columns (5) and (6) of Table 1.

monitoring subsequent call-back, unequal treatment in first hiring decisions based on this characteristic can be identified.

Correspondence experiments focus on a very specific interaction between employers and employees: by means of a correspondence test, the researcher simply measures call-back rates for first interviews. The research results of such a test do not allow the researcher to make any statement about discrimination in the later stages of the selection process, let alone in promotion opportunities and wages. However, Bertrand and Mullainathan (2004) argue that it would be expected that a lower number of interviews would lead to a reduced number of job offers and to lower earnings. Moreover, since job interviews are costly, firms invite candidates to an interview only if these candidates have a reasonably high chance of getting the job. In addition, Lahey (2008) argues that employers who wish to discriminate against certain groups of workers without being sued would prefer to do this in the first step of the hiring stage, since it is more difficult for individuals to determine why they failed to be invited to a job interview than it is to determine why they did not get the job after an interview or why they did not get promoted or were fired after being hired. This reasoning is consistent with the international evidence that shows (i) a relatively low number of job losses among older workers compared to other age categories, on the one hand; and (ii) a relatively low probability of older workers leaving unemployment for work within a given period of time (Tinsley, 2012), on the other hand. Moreover, Bendick et al. (1999) and Cédiey et al. (2008) reported (based on their field experiments comprising all stages of the hiring process) that, at the moment of their study, about 75% of all age discrimination in the United States and about 85% of all ethnic discrimination in France, respectively, occurred in the first stage of the recruitment process.

Correspondence testing is the gold standard for identifying unequal treatment in the labour market, as it allows, in theory, discrimination to be disentangled from supply-side determinants of labour market outcomes (Eriksson & Rooth, 2014; Kroft, Lange, & Notowidigdo, 2013; Neumark, Forthcoming; Pager, 2007; Riach & Rich, 2002). This is the case when, in practice, the experiment is designed in such a way that it strictly controls all productivity-related characteristics of the fictitious applicants, so that the only distinguishing feature between the pair members is the characteristic that is to be tested. In its application to race and sex discrimination, unequal treatment is easily isolated by randomly assigning names to the pair

members. In the case of age discrimination, strictly equating candidates except for their age is trickier, as will become clear in the following subsection.

In Table 1, we systematically review the previous correspondence studies on age discrimination that are known to us.[4] As indicated in column (1) and column (2), these studies were published between 1996 and 2015 and reported experiments in Australia, England, France, Spain, Sweden and the United States. In columns (3) to (7) important features of these experiments are summarised: the occupation(s) tested, the gender of the applicants, the age combinations used and the sample size. Columns (8) and (9) focus on the level of unequal treatment estimated. In column (9), the overall measure of discrimination is presented for each study. This measure is the positive call-back ratio. It is calculated by dividing the percentage of the younger candidates' applications receiving a positive call-back (whose ages are in column (5)) by the corresponding percentage for the older candidates (whose ages are in column (6)). The definition of positive call-back used in the studies is mentioned in column (8). For instance, the positive call-back ratio of about 3.2 found by Ahmed et al. (2012) indicates that the 31-year-old candidates in their experiment received on average 3.2 times more job interview invitations than their 46-year-old counterparts, when applying for a position as a sales assistant or a waiter in Sweden.

<Table 1 about here>

We can identify three important patterns from an examination of Table 1. First, in all the countries tested, age discrimination is present and is relatively high in magnitude. All the positive call-back ratios presented are statistically significantly different from 1, at the 5% significance level. Moreover, the levels of age discrimination are higher than the levels of discrimination based on ethnicity, gender or sexual orientation, as estimated by other field experiments in the same countries during the same period (Bertrand & Mullainathan, 2004; Riach & Rich, 2006b; Wood, Hales, Purdon, Sejersen, & Hayllar, 2009).

Second, age discrimination is present even for those who are relatively young. For instance, Albert et al. (2011) reported that Spanish firms show a substantial drop

---

[4] In addition, Farber et al. (Forthcoming) conducted a large-scale experiment on multiple factors determining call-backs to job applications by the unemployed, among which age. They found that job candidates aged 50 and older are significantly less likely to receive positive call-back compared to younger candidates.

in interest for 38-year-old candidates compared to candidates aged 24 or 28.

Third, the outlined studies differ substantially in the features of their experiments. The age of the younger candidates ranges from 24 to 45 and the age of their older counterparts from 38 to 62. As mentioned in Section 1, these earlier contributions have focussed on estimating age discrimination in specific, mostly lower skilled, occupations. Administrative (clerk) jobs and vending jobs (sales assistants and waiters) dominate the positions tested, except in the study by Lahey (2008), who studied a variety of entry-level jobs (e.g., teachers, sales persons, managers, craftsmen and operators). Remarkably, the positive call-back ratios are more to the detriment of older candidates in experiments studying vending occupations than in experiments on administrative occupations. We come back to this observation later when we present our own research findings. Further, no clear patterns emerge in the discrimination ratio when looking at the region (country or continent) of the experiment, the gender of the applicants or the particular age combinations used.

2.2 Difference in Post-Educational Years Problem

To obtain clean estimates of discrimination, correspondence experiments require that the two fictitious candidates in each pair sent to employers are equal except for the characteristic that is to be tested. However, in the case of age as a ground of discrimination, equating the past careers of candidates with the same education but a different age is essentially impossible. We term this problem the "Difference in Post-Educational Years Problem".

This problem is caused by the fact that the older candidate in each pair will have a higher number of post-educational years at the moment of her/his application. As a result, it is only feasible to equate part of the career of the two members of the pair, over a period of the length of the younger one's career. The remaining part may result in more active years for the older candidate and, therefore, more (potentially in-field) experience. This additional amount of experience might yield a positive signal to employers and, therefore, raise the hiring chances, *ceteris paribus*. On the other hand, more experience may translate into higher salary aspirations so that the effect on hiring chances may just as well go in the opposite direction. Alternatively,

the older candidate's additional years may result in more inactive years, filled by, for example, unemployment or at-home caring tasks. This might give a negative signal to employers and, thus, lower the hiring chances, *ceteris paribus*.

As a result, it is not possible, by means of a classical correspondence test in which the post-educational years of the older candidate are filled in a particular way, as with all contributions listed in Table 1 except for Neumark et al. (2015), to disentangle the unequal treatment caused by this additional trajectory from any discrimination based on age. As mentioned in the previous section, in these studies, the younger and older candidates within the pairs sent to employers differed in age by, on average, 20 years. It is unlikely that the activities of job candidates during such a long period have no effect on employers' hiring decisions. This problem was mentioned in many of the studies, and the authors often provided the reader with an argument for their particular choice for (not) filling in the surplus years. However, except for Neumark et al. (2015), it was never tackled fundamentally, either by adapting the correspondence testing framework or by any other means. The approaches towards the Difference in Post-Educational Years Problem are summarised in column (10) of Table 1.

A first approach, adopted by Ahmed et al. (2012), Bendick et al. (1996), Bendick et al. (1999) and Gringart and Helmes (2001), is to fill the additional post-educational years in the résumé of the older candidates with inactivity or out-of-field employment, i.e. employment unrelated to the vacancy for which the candidate applies. For instance, the additional years were ascribed to military employment in Ahmed et al. (2012) and to military employment or teaching in Bendick et al. (1999). When the authors chose to fill in the additional years with inactivity, older candidates mentioned, for example, that they had been out of the labour force while bringing up children. Strictly speaking, these solutions yielded younger and older candidates with the same educational attainment as well as the same relevant (i.e. in-field) work experience. However, as mentioned above, both the choice to work in another field and the choice to be out of the work force might be perceived by employers as a negative signal. Employment in another field may be perceived as a signal of less interest in the occupation applied for (or lower specific human capital), while inactivity may be perceived as indicative of a lack of career progression or a lack of professional aspiration (or lower general and specific human capital). Following this reasoning, the measure of discrimination against older workers found

by this approach might be biased upwards.

A second approach used in the literature is to fill in the extra post-educational years of older candidates with in-field employment, as is done by Albert et al. (2011) and Riach and Rich (2006a, 2007, 2010). These authors argue that one should accept that, in reality, the experience component of (the human capital of) a candidate does vary between age groups and, therefore, they allow this dimension to vary within the experiment. On the other hand, Riach and Rich (2010) acknowledge that age discrimination in this setting should be seen as a lower bound ("a very significant level of prejudice") for age discrimination in reality. In other words, they accept that their measures of unequal treatment are, in fact, a combination of age discrimination (expected to be to the detriment of older candidates) and a profit-maximising response to differential human capital (in favour of older candidates).

A third alternative approach used by earlier correspondence studies is the one outlined by Lahey (2008) and Tinsley (2012). In their experiments, both the younger and the older candidates only mentioned the last part of their career trajectories in their written résumés. For this reason, the authors did not have to inform the employer about what the older candidate did during her/his extra post-educational years. In our opinion, this approach is cleaner than the former two. In addition, Lahey (2008) argues that her design is realistic. In a footnote in her study, she mentions that recruiters in the United States even indicated that ten-year career histories are – or at least were at the moment of her research – the gold standards for résumés. While this setting might, indeed, be realistic (for women in entry-level jobs) in the United States, this may not be the case (for males or females in other positions) in other countries. For instance, in Belgium, a sample of ten human resource managers declared that résumés most commonly comprise an exhaustive enumeration of the person's previous jobs.[5] In addition, by not mentioning the complete career of the job candidates, one actually lets employers fill in the gaps themselves. As a result, the setting in which age discrimination is measured is unclear. Moreover, the effect on hiring chances of limiting one's work history to the last five or ten years might be heterogeneous by the candidate's age, which might bias discrimination measures.

---

[5] These ten human resource managers were selected from the LinkedIn contacts of the first author of the present study and were informally contacted via telephone in September 2014.

By picking out just one of the activities that could be undertaken by older candidates during their extra post-educational years (or by not mentioning these activities), the previous correspondence studies on age discrimination presented unequal treatment in a particular situation, without knowing how far their measures were in reality from the upper or lower bound of discrimination. In addition, the different approaches adopted by the prior correspondence studies on age discrimination make their results very difficult to compare.

We are aware of only one study that tackled fundamentally the Difference in Post-Educational Years Problem. Simultaneously with our field experiment, Neumark et al. (2015) conducted a large-scale experiment on age discrimination in the United States. They took into account the Difference in Post-Educational Years Problem by alternating between older candidates with the same experience as younger applicants and older candidates having experience commensurate with their age. However, for the older candidates having the same experience as younger applicants, Neumark et al. (2015) truncated job histories at a particular year, yielding the same methodological limitation as the one met by Lahey (2008) and Tinsley (2012). Somewhat surprisingly, Neumark et al. (2015) cannot reject that age discrimination differs by the level of (explicitly mentioned) experience of older candidates.

In the present study, we propose extending the classical correspondence testing framework by the randomised assignment to the older candidates' résumés of one out of the three possibilities mentioned above for their activity in their additional post-educational years (in-field employment, out-of-field employment and inactivity). In this way, we are able to study age discrimination in three realistic situations. Moreover, the discrimination measures in each of these situations provide us with an indication of the importance of the Difference in Post-Educational Years Problem.

3 Experiment

We conducted our experiment between December 2014 and May 2015 in the labour market of Flanders, the Northern part of Belgium. Two applications were sent in response to each vacancy we selected from the database of the Public Employment Agency of Flanders, the region's major job search channel. For each occupation in the experiment, we constructed two comparable job application templates that differed only in details and layout. For each vacancy, we randomly assigned the younger person to one of these template types and the older to the other, and sent the resulting combinations in a randomised order to the employer. In addition, three particular age combinations, three different activities undertaken by the older candidates during their extra post-educational years and the candidates' gender (the same for both members of the pair) were randomly assigned between the pairs of templates. Thereafter, reactions from the employers were analysed to investigate unequal treatment in hiring based on age, in general, and by specific candidate and vacancy characteristics.

3.1 Selection of Vacancies

We sent pairs of fictitious job applications to real vacancies for six occupations in three areas (industry, administration and vending) and two classification levels (higher secondary (ISCED 3)[6] and lower tertiary education (ISCED 5 – Bachelor) as the functional level). These occupations were: production operator, administrative clerk, waiter, laboratory analyst, management assistant and sales representative.[7]

Testing occupations in multiple areas and at multiple qualification levels enabled us to avoid the danger inherent to many earlier correspondence experiments in which researchers selected one or only a few occupations with, potentially, an unrepresentatively high (or low) effect of the characteristic being tested. In addition, because of the particular occupations chosen, we were able to measure whether

---

[6] We employ the International Standard Classification of Education (ISCED) levels of education.

[7] The proportion of workers aged between 50 and 65 in these occupations in the Sixth Round of the European Social Survey was at least 10%.

there were differences in the age-related unequal treatment that arose from the educational attainment required by the vacancy or the area of the occupation. In general, theoretical and empirical studies show that candidates' higher education levels can act as a prejudice-reducing device (Baert, Cockx, Gheyle, & Vandamme, 2015; Wood et al., 2009). Thus, we might expect age discrimination to decrease with the educational attainment of candidates. However, some of the aforementioned characteristics of older workers that are perceived by employers (such as showing less creativity) might play a more important role in the assessment of the higher educated candidates, so that the interaction between age discrimination and education level might just as well go in the opposite direction. Concerning the area of the occupation, based on the model of taste-based discrimination mentioned earlier, one could expect more unequal treatment in occupations where personal contact with the employer, customers or co-workers is higher (Lahey, 2008; Riach & Rich, 2010). We took this aspect into consideration in our choice of occupations. As a consequence, the experimental design enables us to compare age discrimination between occupations where employer and customer contact is, on average, low (production operator and lab analyst) and occupations where employer contact (management assistant) or customer contact (waiter and sales representative) is particularly high.

3.2 Construction of Job Application Template Pairs

For each occupation, we constructed two types of templates (type A and type B), comprising a résumé and a motivation letter matching the general requirements for these kinds of jobs. To ensure that our applications were realistic and representative, example résumés and motivation letters from the Public Employment Agency of Flanders were calibrated for our purposes.

The type A and B templates differed in inessential peculiarities and layout, to avoid detection. It is important to note that these minimal differences between the type A and the type B job application templates could not bias the discrimination measures, as the younger and older ages were randomly assigned to these types. The same is true with respect to (heterogeneity in discrimination by) the variables that were randomly assigned between the pairs of applicants. Moreover, our

regression analysis presented in Section 4.2 will show that the small differences between the application templates did not yield different call-back outcomes for the type A and type B versions.

All fictitious applicants were born and living in Antwerp or Ghent, the largest cities of Flanders, with approximately 503,000 and 251,000 inhabitants, respectively. Applicants mentioned in their résumés that they were married and had two children (without disclosing their ages). All applicants had graduated from the same type of school, with comparable reputations. The candidates applying for a job as a production operator, administrative clerk or waiter had graduated from secondary education with a degree in mechanics or commerce. Those applying for a job as a lab analyst, management assistant or sales representative held a Bachelor's degree in chemistry, management assistance and commercial sciences, respectively. To prevent employers from detecting the experiment, a variety of common wordings was used for these degrees. The candidates with secondary education degrees had graduated at the age of 18, and those with Bachelor's degrees at the age of 21. These are the common ages in Belgium when an individual does not repeat any years.

The resumes provided recruiters with the complete career trajectory of the job candidates. All candidates were employed at the start of the experiment in in-field jobs that had started in 2006. In addition, they had worked in similar jobs immediately after graduation. This experience ensures that the employer perceives the candidates as being attached to the labour market and, more specifically, being attached to the sector of the vacancy. For the years in between, we randomly assigned to the older workers one out of three activities. In Section 3.4 we provide more details as to the exact procedure we followed to do this.

In addition, we added to all applications the following features: a typically Flemish-sounding first name and surname, a Belgian nationality, a day and month of birth, an address with an existing street name but a non-existent house number in a middle-class neighbourhood, a telephone number and an email address from a major provider, adequate Dutch, English and French language skills, comparable computer skills, one sporting activity and a driver's licence. The résumé and motivation letter templates are available upon request.

14

3.3 Randomised Assignment of Older Age within Pairs

The younger and older ages were alternately assigned to the type A and the type B application templates. This was done just by adding a different year of birth to these templates. As mentioned in the previous subsection, the day and month of birth were fixed at the application template type level. In addition, the ages of the candidates were also signalled by the summary of their past careers, on which we elaborate in the next subsection.

3.4 Randomised Assignment of Particular Age Combinations, Career Trajectory and Gender between Pairs

In order to measure age discrimination at multiple (older) ages, on the one hand, and for varying differences in age between the younger and older candidates, on the other hand, we alternately assigned three age combinations to the pairs of applicants. More concretely, we tested all (pairwise) combinations of candidates aged 38, 44 and 50 years. This yielded a variation in age difference between the pairs of fictitious candidates (6 versus 12 years). Our *a priori* expectation was that a double dose of older age would result in a more adverse effect on employment opportunities.

We chose 38 as the youngest age, as from this age onwards the probability of a woman becoming pregnant becomes fairly low. Statistics from the Flemish Agency for Care and Health show that in Flanders the age-specific fertility rate peaks at the age of 29. Only one out of twenty mothers is 38 or older when she gives birth. Before this age, it would be difficult, when considering female candidates, to distinguish between unequal treatment based on age and unequal treatment based on potential maternity leave in the future.

In addition, we chose 50 as the oldest age, as employed individuals at this age are not yet eligible for any old age labour market programmes. This is important, because if the employer would get a wage subsidy for hiring candidates of one of

15

our fictitious groups,[8] it would be impossible to disentangle age discrimination from profit-maximising unequal treatment based on older candidates' eligibility for these subsidies. In addition, at the age of 50, job candidates are not too close to their retirement age. In Belgium, the official retirement age is 65 – it will become 66 in 2025 and 67 in 2030. However, the widespread use of early retirement regimes means that the effective age of labour market exit was 59.6 for men and 58.7 for women in Belgium in 2013 (OECD, 2013). Candidates who are substantially older than 50 may, therefore, be treated unfavourably by employers, just because of the profit-maximising reflection that the employer's return on hiring investment will be lower. At the age of 50, however, the obsolescence of human capital should not yet be a rational worry (for most occupations), especially since our fictitious candidates were all working in similar jobs to the one for which they were applying. The descriptive evidence of Vandenberghe, Waltenberg, & Rigo (2013) even suggests that, at least in Belgium, the average pay-productivity gap becomes positive only from the age of 56. If these arguments did not hold, one could expect that unequal treatment based on age would be more to the detriment of applicants aged 50 (when compared to those aged 44) than it is to applicants aged 44 (when compared to those aged 38), *ceteris paribus*. However, our regression analysis presented in Section 4.2 will show that this is not the case. From a policy perspective, measuring age discrimination at the age of 50 (and 44) is relevant: given the aforementioned low effective age of labour market exit in Belgium, extending work lives beyond 50 is crucial.

Next, to deal with the Difference in Post-Educational Years Problem, we alternated in the pairs sent to employers between the three relevant activities that could have been undertaken by the older candidate during their extra years. More concretely, the potential activities were those used individually by the former contributions listed in Table 1: in-field employment (by analogy with, e.g., Albert et al. (2011) and Riach and Rich (2010)), out-of-field employment (by analogy with, e.g., Ahmed et al. (2012) and Bendick et al. (1999)) and inactivity (by analogy with, e.g.,

---

[8] In Belgium there are two relevant subsidies for hiring older workers. Firstly, wage subsidies within the ACTIVA program are government funded payments that cover part of the wage costs when hiring a new employee who is older than 45 and who comes from unemployment. Secondly, employers have to pay less social security contributions for all their workers aged 58 and above (Albanese & Cockx, 2015). Neither of these policies thus plays a role yet in our setting as our fictitious candidates were not unemployed at the moment of their application and as they were younger than 58.

Bendick et al. (1996) and Gringart and Helmes (2001)). Figure 1 schematises the resulting career trajectories of our experimental groups. As mentioned in Section 3.2, all candidates had worked in jobs within the same occupation as the one mentioned in the vacancy immediately after graduation ("in-field job 1") and were employed in other similar jobs at the start of the experiment ("in-field job 2"). For the younger candidate in the pair, the first job was immediately followed by the second. The older candidate had undertaken one of three different activities during the additional 6 or 12 years falling in between her/his comparable periods in in-field job 1 and in-field job 2. In the first third of the cases, the older candidate had been employed 6 or 12 years longer in in-field job 1. In the second third of the cases, she/he had been employed in an out-of-field job during this period. For the lower educated profiles, this had been as a member of maintenance staff, and for the higher educated profiles, this had been as a teacher at an organisation offering apprenticeship and entrepreneurship training. In the final third of the cases, the older candidate mentioned in her/his résumé that she/he had been inactive: out of the labour force while carrying out child-rearing tasks during this period. This inactivity might send different signals, depending on the gender of the applicants. More specifically, women might face less discrimination than men for a period of home-making tasks.

Finally, we alternated between female and male pairs of fictitious candidates. Our *a priori* belief was that if there were any gender differences in age discrimination, these would be to the detriment of older women. This belief was formed because various studies have argued that females age sooner than males in terms of (perceived) employability. This relates to women's perceived attractiveness being more age-dependent and to their observed earlier retirement and (thereby) lower employment rates at older ages (Gringart & Helmes, 2001). Empirically, Gringart and Helmes (2001) and Neumark et al. (2015) found that older females were, compared to their younger counterparts, indeed relatively more disadvantaged than older males, albeit in a statistically insignificant way.

3.5 Classification of Call-Back

We sent the 36 resulting combinations of the two job application templates, three

particular age combinations, three career patterns for the older candidates and two genders in an alternating order to the selected job postings, with a one-day delay in between. To avoid detection, we only applied to each employer with one pair of applications. By reference to the relative supply of vacancies in the occupations mentioned in Section 3.1, we decided to send out 72 pairs of applications to each of the occupations of administrative clerk, waiter, laboratory analyst and sales representative, and 144 pairs to each of the occupations of production operator and management assistant, resulting in 1,152 job applications for 576 vacancies.

While the size of our sample is substantially lower than the size of the data gathered by Lahey (2008) and Neumark et al. (2015) in the United States, it is at least comparable to the other studies summarised in Table 1. Overall, five previous correspondence experiments on age discrimination tested more vacancies, while six previous experiments tested fewer vacancies. However, also the latter six experiments were able to detect statistically significantly unequal treatment based on age at the level of their total dataset and, most of them, at the level of some subdatasets. For instance, Ahmed et al. (2012) tested 466 vacancies with two applicants and could reject unequal treatment at the 0.1% significance level (i) within their total sample; (ii) within their subsample of 263 vacancies for restaurant workers; (iii) within their subsample of 284 vacancies in the Nonmetropolitan area; (iv) within their subsample of 321 vacancies for full-time positions; and (v) within their subsample of 180 vacancies for temporary positions.

Reactions from the employers were received via telephone voicemail and email. To minimise inconvenience for the employers, we immediately terminated the recruitment procedure if we received a positive reaction. All call-backs received more than 30 days after the date of the application were discarded. However, this turned out to be an unnecessary restriction, as we hardly received any (positive) responses after 30 days.

As mentioned in Table 1, a small majority of the previous correspondence studies on age discrimination defined positive call-back as receiving some sort of positive reaction. This might be an invitation to a job interview, the receipt of an alternative job proposal, a request to provide more information or a request to contact the recruiter. The other studies defined positive call-back in a stricter sense to mean being invited for an interview for the job for which the application was

made. In the present study, we will present the measures of unequal treatment based on age separately for these two definitions.

4   Results

In this section, we present the empirical insights, based on a statistical examination of the experimentally gathered data. First, we report in general on the positive call-back rates for the younger and older fictitious candidates, classified according to the activity undertaken by the older worker during her/his extra post-educational years, and classified by vacancy and candidate characteristics. Second, we discuss a regression analysis that allows us to control for vacancy fixed effects and to measure the independent effect of relevant variables interacted with the older age of the candidate.

4.1   Discrimination Ratios

Table 2 and Table 3 present the positive call-back rates for the younger and older candidates within the pairs of fictitious job applications we sent out. In Table 2 (Table 3) we follow the strict (broad) definition of positive call-back. Statistics with regards to all the tested vacancies together are presented in Panel A of these tables. In total, the younger candidates were invited for a job interview (received any positive reaction) in 8.0% (16.5%) of their applications, while their older counterparts were invited (received a positive reaction) in only 4.9% (11.8%) of the cases. Column (3) of Table 2 and Table 3 shows that the positive call-back ratio in the strict sense is about 1.6, while the positive call-back ratio in the broad sense is 1.4. These numbers indicate that the younger candidate within the pairs had a 64.3% higher chance of being invited for a job interview and a 39.7% higher chance of receiving some positive reaction. Both statistics are significantly different from 1 at the 1% significance level.

<Table 2 about here>

<Table 3 about here>

19

This overall finding of hiring discrimination against older candidates is in line with the results of the correspondence experiments reviewed in Section 2. Our interview invitation ratio of 1.6 lies in between the comparable ratios found by Bendick et al. (1999) and Lahey (2008) for the United States, on the one hand, and those presented by Ahmed et al. (2012) and Riach and Rich (2010) for Sweden and England on the other hand. Our positive reaction ratio is higher than the ones presented by Gringart and Helmes (2001) and Lahey (2008) but lower than the ones found by the other studies among which Neumark et al. (2015). Baert et al. (2015) recently provided evidence for a negative correlation between discrimination and labour market tightness. Therefore, the fact that the overall level of age discrimination we measure is lower than the corresponding level in other European countries may be partly explained by the fact that labour market tightness is higher in Belgium than in other European countries.[9] Furthermore, our results corroborate the evidence from other countries in the sense that the levels of age discrimination found are higher than the levels of discrimination based on ethnicity, gender or sexual orientation that have been estimated for Belgium (Baert et al., 2015; Baert, De Pauw, & Deschacht, 2016).

Next, we break the positive call-back ratios down according to the activity of the older candidate within the pair during her/his additional post-educational years. Panel B of Table 2 and Table 3 indicates that equal treatment can be rejected (at the 5% significance level) for the subsamples in which the older candidates had years of out-of-field employment. The probability of being invited to a job interview decreases by 65.0%,[10] and the probability of some positive reaction decreases by 44.8%, for the older fictitious candidates who revealed out-of-field employment during their additional years, when compared to their younger counterparts. The punishment for older age in combination with inactivity is somewhat lower: 41.2% with respect to interview invitation, and 25.8% with respect to some positive reaction. In addition, this punishment is only statistically significant at the 10% level. Finally, we cannot reject equal treatment with respect to the pairs in which the older

---

[9] The labour market tightness in a country can be captured by its job vacancy rate, i.e. the number of vacancies as a proportion of the sum of this number of vacancies and the number of occupied jobs. In the second quarter of 2015, the job vacancy rate was 2.7% in Belgium, while it was 1.7% in the EU-28. In no European country was this statistic higher than in Belgium (source: Eurostat, Job vacancy rate). In contrast, during the last years, the job vacancy rate in the United States has been higher than the one in Europe (OECD, 2014).

[10] $65.0\% = 1 - (1/2.857)$.

candidate had more (in-field) experience. With regards to the probability of being invited to a job interview, the positive call-back ratio is even (insignificantly) smaller than 1. Apparently, the older candidates in our experiment could compensate for their older age by additional in-field experience.

Our results underline that the Difference in Post-Educational Years Problem, which we aim to take into account in this study, is an important problem. Discrimination measures vary fundamentally according to the way the older candidates are said to have spent their extra years in experiments on age discrimination. Our extended correspondence experiment gives us an indication of the lower (zero) and upper bound (the ratios in cases where the older candidates had out-of-field experience) of age discrimination in the tested occupations in Flanders. If we had picked just one out of these three potential situations, as the previous correspondence studies on age discrimination did, the experiment would only have allowed us to present one arbitrary value out of this band, without knowing how far from the upper or lower bound we were.

As a consequence, our results contrast with those reported by Neumark et al. (2015), who found, except for the occupation of janitor, no heterogeneity in discrimination by older candidates' level of work experience (equal to the younger candidates in their experiment versus commensurate with these older candidates' age). However, these differing results may be explained by the use of a differing experimental approach. While we explicitly filled in the extra post-educational years of the older candidates with extra in-field experience, out-of-field experience or inactivity, Neumark et al. (2015) truncated the job histories for the older candidates with non-commensurate experience at a particular year. As a consequence, employers might have filled in the unmentioned years with relevant work experience, yielding similar call-back rates for the older candidates by level of (explicitly mentioned) experience.

In Panels C, D and E of Table 2 and Table 3, we break down the data according to the tested occupation, the particular age of the candidates within the pair and their gender, respectively. First, the discrimination measures with regards to the classification by occupation are the highest for the occupations of management assistant (statistically significant at the 5% level for both outcomes) and waiter (statistically significant at the 5% level when using the strict sense definition of

21

positive call-back). This finding might be explained by the fact that these occupations might have requirements that relate to the stereotypes of older workers mentioned in the Introduction. For instance, the requirement of using up-to-date IT tools in the occupation of management assistant and the flexible working schemes in the catering industry, might not match with the lower flexibility and adaptability of older workers (in the perception of employers). In addition, this finding might be explained by employer and customer discrimination in the sense of the taste-based discrimination model of Becker (1957). On the one hand, personal cooperation with the employer is probably most important for a management assistant. Employers looking for a management assistant might have a taste for younger individuals in their direct environment. On the other hand, the high discrimination ratios in the occupation of waiter might suggest that customer discrimination partly underlies the age discrimination, as customer contact is on average higher in this occupation than in the others. This observation squares with the aforementioned pattern in Table 1 that the discrimination found in the prior contributions testing vending occupations was, on average, higher than the discrimination found in the studies testing other occupations. However, an important *caveat* in this respect is that we might have lacked of statistical power to reject unequal treatment in some of the other occupations for which we only tested 72 vacancies.

Second, our overall finding of age discrimination in terms of job interview invitations (any positive reactions) is driven by the comparison of the 38-year-old to the 44-year-old (44-year-old to the 50-year-old). Somewhat surprisingly, the discrimination ratios are the lowest for the comparison of the 38-year-old to the 50-year-old. Third, based on the ratios presented in Panel E, our overall ratios are driven by the male rather than by the female pairs. The latter finding contrasts to the results of Neumark et al. (2015): as aforementioned, in their setting, age discrimination was driven by unequal treatment of young versus older *female* candidates. We come back to the significance of the discussed dimensions of heterogeneity in age discrimination when we present our regression results in the next subsection.

4.2 Controlling for Vacancy Fixed Effects

As the younger and older ages are assigned randomly within our pairs of

applications, regressing positive call-back (in the strict or the broad sense) on an indicator of being the older candidate within the pair leads to exactly the same empirical conclusion as that based on Panel A of Table 2 and Table 3. In addition, as we randomly assigned these variables between pairs, regressions, including interactions between older age and (i) the activity the older candidate undertook during her/his additional post-educational years, (ii) the particular age combination of the candidates and (iii) the gender of the candidates, should lead to the same empirical pattern as the one in Panel B, Panel D and Panel E of Table 2 and Table 3, at least for a sample size approaching infinity. However, our sample size is finite. Thus, some of these variables, which were randomly assigned between pairs, may happen to correlate with observable and unobservable vacancy characteristics. As these characteristics may affect the hiring chances of our fictitious candidates, not controlling for them might yield biased measures of the heterogeneity of age discrimination by the older candidates' past careers, the particular age combinations and the gender of the candidates used. Therefore, we further explore the experimentally gathered data by means of a regression analysis controlling for vacancy fixed effects.

Table 4 and Table 5 present our regression results. We regress the outcome of positive call-back in the strict sense (Table 4) and in the broad sense (Table 5) on various sets of key and control variables by means of a linear probability model with vacancy fixed effects. Variables that are included in interaction with the indicator of older age are not included without this interaction, as they are all constant at the vacancy level and are, therefore, controlled by our fixed effects estimations. In what follows, we first focus on the results outlined in Table 4. Afterwards, we compare these results with those in Table 5.

<Table 4 about here>

<Table 5 about here>

In regression model (1), we only include an indicator of older age as an explanatory variable. From this, we find that older age decreases the probability of a job interview invitation by 3.1 percentage points. Obviously, this outcome equals the difference between the overall positive call-back rates in a strict sense among the younger and older candidates mentioned in Section 4.1 and Table 2, i.e. the difference between 8.0% and 4.9%. The ratio of this 3.1 percentage point effect and

23

the baseline interview rate of 4.9% for the older candidates is very close to the call-back ratio of 1.643 mentioned in Table 2 (and discussed in the previous subsection).

Next, in model (2), we interact the older age with indicators for the two least beneficial activities undertaken by the older pair member during her/his additional years (with additional in-field employment as a reference category). We observe, in line with the findings reported in Table 2, that for the reference category of older workers with additional in-field employment we cannot reject unequal treatment based on age. Secondly, we find a significantly negative effect on the probability of getting an invitation for a job interview of the interaction of the older age with additional out-of-field employment. The probability of positive call-back in the strict sense is 7.8 percentage points lower for older candidates with extra out-of-field employment compared to older candidates with extra in-field employment. Thirdly, again, older workers who filled their additional post-educational years with inactivity hold a middle position. On the one hand, the interaction of the older age with inactive years is only weakly significant. On the other hand, an F-test performed to assess the statistical difference between the interactions of the older age with either additional out-of-field employment or inactivity does not allow us to reject the hypothesis that they are equal.

In regression model (3), we extend the set of independent variables with variables indicating type B applications and applications that were the first of the pair to be sent to the employer. Next, in regression model (4), we add additional interactions between the indicator of older age and indicators for (i) higher educated candidates, (ii) vacancies for administrative occupations and vacancies for vending occupations (keeping vacancies in the industrial area as a reference category), (iii) candidates being 50 years old, (iv) candidates being 12 years older than their pair member and (v) candidates being female. As the extra variables included in models (3) and (4) are, because of the construction of the experiment, not correlated with the older age and career pattern dummies, their adoption does not change the coefficients for the main variables of interest.

In addition, the estimation results for model (3) show, as announced in Section 3.2, that the probability of receiving a job interview invitation does not vary across the template types. Moreover, it does not matter which application was sent as the first pair member. Concerning the interactions added in model (4), we find only a

24

weakly significant interaction effect with regards to the vending occupations (waiter and sales representative). The other interactions are not significant at all: we do not find a significant difference in age discrimination when candidates differ in age by 6 or 12 years or when they are male or female.

Finally, in model (5) we add indicators for vacancies offering part-time contracts and vacancies offered by an interim office in interaction with an indicator for being the older candidate within the pair. None of these interactions has a significant effect on the probability of job interview invitation. However, we are underpowered to formally test for heterogeneity in the penalty of older age by these vacancy characteristics as only 86 vacancies offered a part-time contract and only 64 vacancies were posted by an interim office. More importantly, the introduction of these variables hardly affects the regression coefficients of main interest.

Table 5 presents the corresponding results using the broad definition of positive call-back as an outcome variable. The most important difference when comparing column (5) of this table to the same column in Table 4 is that the weak significance of the interaction between the indicators of the older candidate and inactivity during her/his additional post-educational years disappears. Thus, for positive call-backs in the broad sense, we cannot reject that the treatment of older candidates is equal for those who were inactive during their additional years and those who gained additional in-field experience. However, we also cannot reject that the treatment of older candidates is equal for those who were inactive and those who were employed in an out-of-field job. A second difference between column (5) of Table 4 and the comparable column of Table 5 is that a positive outcome in the broad sense is significantly more likely for the candidate whose application is sent first. For this candidate, the chance of getting some positive reaction is 3.3 percentage points higher, *ceteris paribus*. As suggested by a reviewer of a former version of this article, the latter result can be read as a recommendation for future correspondence experiments to have less than 24 hours delay in between the candidates sent to the same employer. Job application via online tools like the database of the Public Employment Agency of Flanders may favour early submissions as many job seekers may apply very quickly when new vacancies appear online.

5 Conclusion

This study contributed to the empirical literature on age discrimination. We argued that former contributions to this literature reporting on classical correspondence experiments (the gold standard to identify unequal treatment in hiring) were not able to disentangle age discrimination from unequal treatment arising from the activities of older workers during their additional post-educational years (which may lead to additional human capital or career gaps).

In our approach to this "Difference in Post-Educational Years Problem", we proposed an extension of the correspondence testing framework. More concretely, we presented a setting in which the random assignment of a younger and an older age within pairs of fictitious job applications sent to real job openings is combined with the random assignment between these pairs of various realistic activities undertaken by the older worker during her/his extra post-educational years.

We applied this extended field experimental framework to the Belgian case. We sent out pairs of fictitious job applications in response to real vacancies in Flanders, the Northern part of Belgium, between December 2014 and May 2015. Alternately, in one application, the younger age of 38 or 44 was disclosed, and in the other, the older age of 44 or 50; the resulting variation in age difference (6 or 12 years) allowed us to look into whether age discrimination varied linearly with this difference – whether the penalty for being 12 years older is double that of being 6 years older.

We randomly assigned to the older candidate within the pair one out of the three activities she/he could most realistically have undertaken during her/his 6 or 12 years of additional labour market time: in-field employment, out-of-field employment and inactivity. We further contributed to the literature by measuring the heterogeneity in age discrimination by the skill level and the area of occupation and by the gender of the candidate. To this end, we sent out our pairs of applications to a balanced number of vacancies in six different occupations, and we randomly assigned the gender of the candidates between these pairs.

We found that, in line with the literature, the younger candidate within a pair of applicants had, on average, a 64.3% higher chance of being invited for a job interview than the older candidate, and a 39.7% higher chance of receiving some sort of

positive reaction. However, when we classified this invitation probability by the activity of the older candidate during her/his additional post-educational years, we found that our overall result of age discrimination was driven by older candidates who filled their additional post-educational years with out-of-field employment. On the other hand, none of our analyses provided evidence of discrimination against older aged individuals who compensated their older age by additional in-field experience. Finally, older candidates who filled their extra years with inactivity held a middle position. Thus, also empirically, the Difference in Post-Educational Years Problem turned out to be important.

We end by acknowledging some research limitations inherent to our research focus and experimental design. Firstly, we were not able to apply Neumark's (2012) correction for the fact that the variance of unobservable determinants of the call-back outcomes we measured may differ between younger and older candidates. To identify a heteroskedastic probit model allowing the error term to vary with older age, it is needed that the researcher observes job-relevant characteristics that affect the younger and older populations' propensities of call-back in the same way, as is the case in Neumark, Burn and Button (2016). This was not possible in our case given that we observed only variation in candidate characteristics by which age discrimination might be heterogeneous, at least from a theoretical point of view (particular age combinations, career pattern for the older candidate and gender of the candidates).

Secondly, our application of the proposed experimental approach to measuring age discrimination is limited in terms of scope. We measured unequal treatment based on age only for particular age combinations in relation to jobs posted in the database of the Public Employment Agency of Flanders. This limitation is less acute in our design than it is in earlier contributions, as we made the conscious choice to test vacancies in industry, administration and sales, and in lower skilled as well as higher skilled occupations. Moreover, because this limitation is expected to cause a similar shift in the punishment of all older candidates, irrespective of their past careers, it should not bias the conclusions with respect to the relative surplus arising from the various activities the older candidates undertook during their extra post-educational years. Nevertheless, differing labour market institutions and characteristics (e.g., with respect to labour market tightness) mean that we cannot simply generalise our results to other Western European countries, let alone to

other continental regions. In addition, it would be interesting to explore other dimensions of heterogeneity in age discrimination (for instance, the interaction between age and ethnicity). Therefore, we are in favour of large-scale, cross-country implementations of our experimental setting to measure age discrimination. The outcomes of these experiments might not only facilitate more accurate policies and interventions but may also make employers more aware of the responsibility they carry with respect to the labour market situation of older workers and, therefore, with respect to the problem of an ageing population.

References

Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2012). Does age matter for employability? A field experiment on ageism in the Swedish labour market. *Applied Economics Letters*, *19,* 403–406.

Albanese, A., Cockx, B. (2015). Permanent wage cost subsidies for older workers: An effective tool for increasing working time and postponing early retirement? *IZA Discussion Paper Series*, 8988.

Albert, A., Escot, L., & Fernández-Cornejo, J. A. (2011). A field experiment to study sex and age discrimination in the Madrid labour market. *International Journal of Human Resource Management*, *22*, 351–375.

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.) *Discrimination in labor markets*. Princeton, NJ: Princeton University Press.

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult? *ILR Review*, *68*, 467–500.

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

Becker, G. (1957). *The economics of discrimination*. Chicago, IL: University of Chicago Press.

Bendick, M., Brown, L. E., & Wall, K. (1996). Experimental discrimination against

older workers: An experimental study of hiring practices. *Journal of Aging & Social Policy*, *8*, 25–46.

Bendick, M., Brown, L. E., & Wall, K. (1999). No foot in the door: An experimental study of employment discrimination against older workers. *Journal of Aging & Social Policy*, *10*, 5–23.

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*, 991–1013.

Cédiey, E., Foroni, F., & Garner, H. (2008). Discrimination à l'embauche fondée sur l'origine à l'encontre des jeunes français(e)s peu qualifié(e)s. *Dares Premières Infos Premières Synthèses*, 06.3.

Eriksson, S., & Rooth, D.-O. (2014). Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. *American Economic Review*, *104*, 1014–1039.

Farber, H. S., Silverman, D., & von Wachter, T. (Forthcoming). Factors determining callbacks to job applications by the unemployed: An audit study. *Russel Sage Journal of the Social Sciences*.

Gringart, E., & Helmes, A. (2001). Age discrimination in hiring practices against older adults in Western Australia: The case of accounting assistants. *Australian Journal on Ageing*, *20*, 23–28.

Kautonen, T., Hatak, I., Kibler, E., & Wainwright, T. (2015). The role of age-based self-image. *Journal of Economic Psychology*, *50*, 41–51.

Kroft, K., Lange, F., & Notowidigdo, M. J. (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *Quarterly Journal of Economics*, *128*, 1123–1167.

Lahey, J. N. (2008). Age, women, and hiring: An experimental study. *Journal of Human Resources*, *43*, 30–56.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, 47, 1128–1157.

Neumark, D. (Forthcoming). Experimental research on labor market discrimination. *Journal of Economic Literature*.

Neumark, D., Burn, I., & Button, P. (2015). Is it harder for older workers to find jobs? New and improved evidence from a field experiment. *NBER Working Paper Series*, 21669.

Neumark, D., Burn, I., & Button, P. (2016). Experimental age discrimination evidence and the Heckman critique. *American Economic Review*, *106*, 303–308.

OECD (2006). *Live longer, work longer*. Paris: OECD.

OECD (2013). *Pensions at a glance 2013: OECD and G20 indicators*. Paris: OECD.

OECD (2014). *OECD Employment Outlook 2014*. Paris: OECD.

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science*, *609*, 104–133.

Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, *112*, 480–518.

Riach, P. A., & Rich, J. (2006a). An experimental investigation of age discrimination in the French labour market. *IZA Discussion Paper Series*, 2522.

Riach, P. A., & Rich, J. (2006b). An experimental investigation of sexual discrimination in hiring in the English labor market. *B.E. Journal of Economic Analysis & Policy*, *5*, 1–22.

Riach, P. A., & Rich, J. (2007). An experimental investigation of age discrimination in the Spanish labour market. *IZA Discussion Paper Series*, 2654.

Riach, P. A., & Rich, J. (2010). An experimental investigation of age discrimination in the English labor market. *Annals of Economics and Statistics*, *99/100*, 169–185.

Sonnet, A., Olsen, H., & Manfredi, T. (2014). Towards more inclusive ageing and employment policies: The lessons from France, the Netherlands, Norway and Switzerland. *De Economist*, *162*, 315–339.

Tinsley, M. (2012). *Too much to lose: Understanding and supporting Britain's older workers*. London: Policy Exchange.

Vandenberghe, V., Waltenberg, F., & Rigo, M. (2013). Ageing and employability. Evidence from Belgian firm level data. *Journal of Productivity Analysis*, *40*, 111–136.

Wood, M., Hales, J., Purdon, S., Sejersen, T., & Hayllar, O. (2009). A test for racial

discrimination in recruitment practice in British cities. *DWP Research Reports*, 607.
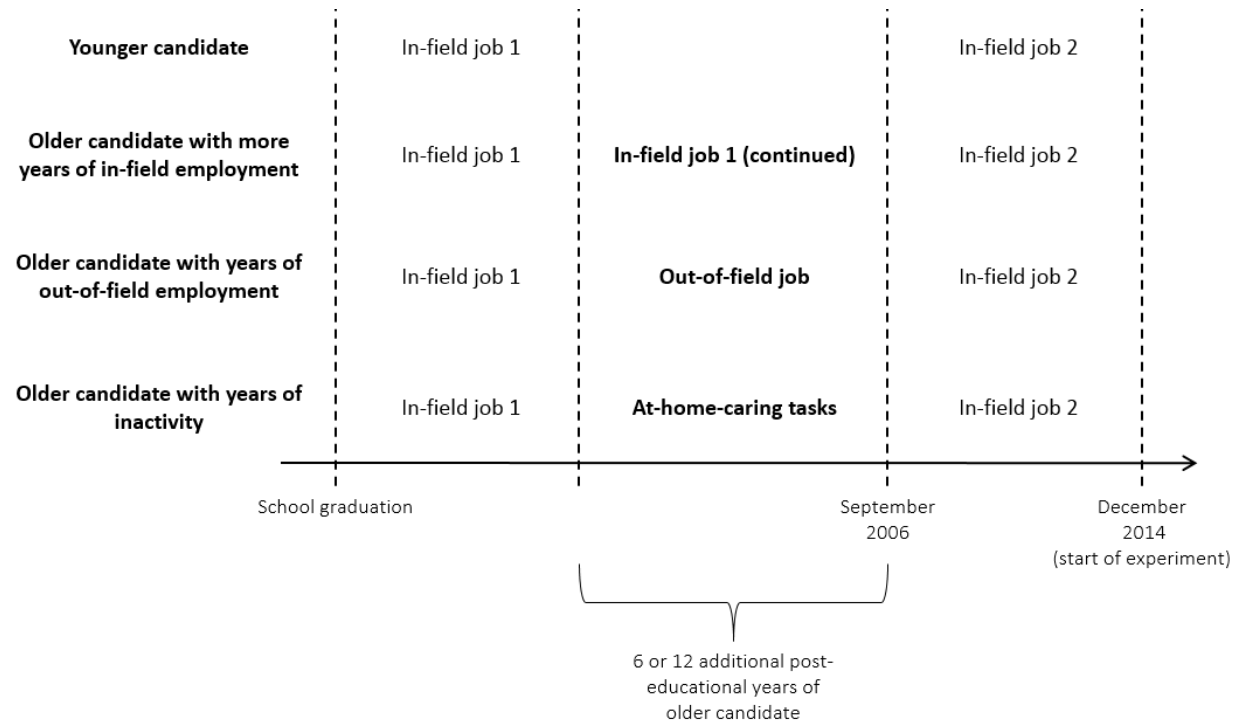
Figure 1. Experimental Groups

Table 1. Literature Review: Correspondence Experiments on Age Discrimination

(1) Study	(2) Country	(3) Tested occupation(s)	(4) Gender of applicants	(5) Age of younger applicants	(6) Age of older applicants	(7) # vacancies (# résumés / vacancy)	(8) Definition of positive call-back	(9) Positive call-back ratio	(10) Approach towards Difference in Post-Educational Years Problem
Ahmed et al. (2012)	Sweden	Seller and waiter	Male	31	46	466 (2)	Interview invitation	3.230	Years of out-of-field employment for the older candidate
Albert et al. (2011)	Spain	Clerk (in accountancy, administration or sales)	Female or male	24 or 28	38	1,062 (10)	Any positive reaction	1.694	More years of in-field employment for the older candidate
Bendick et al. (1996)	United States	Clerk (in administration, copy-writing or ICT)	Female or male[a]	32	57	775 (2)	Any positive reaction	1.467	Years of inactivity for the older candidate in the administration area and years of out-of-field employment for the older candidate in the other areas
Bendick et al. (1999)[b]	United States	Clerk (in administration or sales)	Female or male[c]	32	57	102 (2)	Interview invitation	1.135	Years of out-of-field employment for the older candidate
Gringart and Helmes (2001)	Australia	Clerk (in accountancy)	Female or male	32	57	452 (1 or 2)	Any positive reaction	1.238	Years of inactivity for the older candidate when using female pairs and years of out-of-field employment for the older candidate when using male pairs
Lahey (2008)	United States	A variety of entry-level jobs	Female	35 or 45	50, 55 or 62	8,002 (2)	Interview invitation / Any positive reaction	1.418 / 1.190	Work histories limited to the last ten years (for females applying for entry-level jobs)
Neumark et al. (2015)	United States	Administrative assistant, janitor, retail seller and security guard	Female or male	29, 30 or 31[d]	64, 65 or 66	13,412 (3)	Any positive reaction	1.537	Alternation between older candidates with the same experience as younger applicants and older candidates having experience commensurate with their age
Riach and Rich (2006a)	France	Waiter	Male	27	57	345 (2)	Any positive reaction	3.571	More years of in-field employment for the older candidate
Riach and Rich (2007)	Spain	Waiter	Male	27	57	340 (2)	Any positive reaction	3.500	More years of in-field employment for the older candidate
Riach and Rich (2010)	England	A variety of entry jobs, seller and waiter	Female or male[e]	21 or 27[f]	39 or 47	470 (2)	Interview invitation	1.883	More years of in-field employment for the older candidate
Tinsley (2012)	England	Clerk (in administration) and waiter	Female	24 or 25[g]	50 or 51	1,227 (2)	Any positive reaction	2.146	Work histories limited to the last five years

The positive call-back ratio is calculated by dividing the percentage of applications receiving a positive call-back (in the sense of column (78)) for the younger candidates by the corresponding percentage for the older candidates. The presented ratios are our own calculations, based on the mentioned studies – positive call-back ratios for different occupations or cities were weighted by their number of tested jobs. All these ratios are statistically significantly different from 1, at least at the 5% significance level. [a]Female pairs were sent to the administration and copy-writing jobs, and male pairs were sent to the copy-writing and ICT jobs. [b]This study is in fact an "audit experiment", as it comprises all stages of the hiring process (Pager, 2007; Riach & Rich, 2002). For reasons of comparability, the positive call-back ratio mentioned in this table for Bendick et al. (1999) only relates to the first stage of the tested recruitment process. Important to mention is that in this experiment, the older candidate always applied (shortly) before the younger one. [c]Female pairs were only sent to six job openings. [d]In addition, the authors sent a middle aged candidate (aged 49, 50 or 51) to each vacancy. [e]Female pairs were sent to the a variety of entry jobs and sales positions, and male pairs were sent to the positions as a waiter. [f]Candidates applying for entry jobs were 21 or 39 years old; the others were 27 or 47 years old. [g]Candidates applying for clerk positions were 24 or 50 years old; those applying for jobs as a waiter were 25 or 51 years old.

Table 2. The Probability of Interview Invitation by Age: Descriptive Analysis

	(1) Interview rate younger candidate	(2) Interview rate older candidate	(3) Interview ratio: (1) / (2)
A. All vacancies [N = 576]	0.080	0.049	1.643*** [3.021]
B. Classification by the older worker's activity during her/his extra years			
In-field employment [N = 192]	0.047	0.057	0.818 [0.706]
Out-of-field employment [N = 192]	0.104	0.036	2.857*** [3.724]
Inactivity [N = 192]	0.089	0.052	1.700* [1.818]
C. Classification by the tested occupation			
Production operator [N = 144]	0.111	0.090	1.231 [0.774]
Administrative clerk [N = 72]	0.042	0.042	1.000 [0.000]
Waiter [N = 72]	0.167	0.069	2.400** [2.164]
Laboratory analyst [N = 72]	0.042	0.028	1.500 [1.000]
Management assistant [N = 144]	0.042	0.014	3.000** [2.021]
Sales representative [N = 72]	0.083	0.042	2.000* [1.757]
D. Classification by age of the candidates			
38 versus 44 [N = 192]	0.073	0.019	3.876** [2.336]
38 versus 50 [N = 192]	0.073	0.047	1.556 [1.293]
44 versus 50 [N = 192]	0.094	0.068	1.385* [1.675]
E. Classification by gender of the candidates			
Male [N = 288]	0.090	0.049	1.857*** [2.713]
Female [N = 288]	0.069	0.049	1.429 [1.503]

T-statistics, indicating whether the probability of getting an invitation for a job interview is the same for candidates from both age groups, are in brackets. Standard errors are corrected for clustering of the observations at the vacancy level. * (**) ((***)) indicates significance at the 10% (5%) ((1%)) significance level.

Table 3. The Probability of Any Positive Reaction by Age: Descriptive Analysis

	(1)	(2)	(3)
	Positive reaction rate younger candidate	Positive reaction rate older candidate	Positive reaction ratio: (1) / (2)
A. All vacancies [N = 576]	0.165	0.118	1.397*** [3.327]
B. Classification by the older worker's activity during her/his extra years			
In-field employment [N = 192]	0.135	0.125	1.083 [0.446]
Out-of-field employment [N = 192]	0.198	0.109	1.810*** [3.358]
Inactivity [N = 192]	0.161	0.120	1.348* [1.799]
C. Classification by the tested occupation			
Production operator [N = 144]	0.167	0.146	1.143 [0.726]
Administrative clerk [N = 72]	0.083	0.097	0.857 [0.445]
Waiter [N = 72]	0.222	0.125	1.778* [1.981]
Laboratory analyst [N = 72]	0.194	0.139	1.400 [1.424]
Management assistant [N = 144]	0.153	0.069	2.200*** [2.900]
Sales representative [N = 72]	0.181	0.153	1.182 [0.815]
D. Classification by the age of the candidates			
38 versus 44 [N = 192]	0.135	0.099	1.368* [1.706]
38 versus 50 [N = 192]	0.172	0.141	1.222 [1.096]
44 versus 50 [N = 192]	0.188	0.115	1.636*** [3.205]
E. Classification by the gender of the candidates			
Male [N = 288]	0.163	0.104	1.567*** [2.911]
Female [N = 288]	0.167	0.132	1.263* [1.774]

T-statistics, indicating whether the probability of getting any positive reaction is the same for candidates from both age groups, are in brackets. Standard errors are corrected for clustering of the observations at the vacancy level. * (***) indicates significance at the 10% (1%) significance level.

Table 4. The Probability of Interview Invitation by Age: Regression Analysis

	(1)	(2)	(3)	(4)	(5)
Older candidate	-0.031*** (0.010)	0.010 (0.018)	0.010 (0.018)	-0.000 (0.029)	0.001 (0.030)
Older candidate x Out-of-field employment		-0.078*** (0.025)	-0.078*** (0.025)	-0.078*** (0.025)	-0.078*** (0.025)
Older candidate x Years of inactivity		-0.047* (0.025)	-0.047* (0.025)	-0.047* (0.025)	-0.046* (0.025)
First application sent			0.014 (0.010)	0.014 (0.010)	0.013 (0.010)
Job application template type B			-0.011 (0.010)	-0.012 (0.012)	-0.012 (0.012)
Older candidate x Higher educated				0.008 (0.021)	0.008 (0.022)
Older candidate x Vacancy for administrative occupation				-0.003 (0.025)	-0.004 (0.026)
Older candidate x Vacancy for vending occupation				-0.052* (0.027)	-0.055* (0.028)
Older candidate x 50 years old				0.018 (0.023)	0.018 (0.023)
Older candidate x 12 years older				-0.009 (0.033)	-0.008 (0.033)
Older candidate x Female				0.021 (0.021)	0.020 (0.021)
Older candidate x Vacancy offering part-time contract					0.008 (0.031)
Older candidate x Vacancy offered by interim office					-0.013 (0.033)
Constant	0.080*** (0.007)	0.080*** (0.007)	0.068*** (0.010)	0.067*** (0.011)	0.067*** (0.011)
Vacancy fixed effects	Yes	Yes	Yes	Yes	Yes
F-test for the equality of the coefficients for "Older candidate x Out-of-field employment" and "Older candidate x Years of inactivity" (p-value)		0.211	0.212	0.215	0.215
Observations	1,152	1,152	1,152	1,152	1,152

The presented results are linear probability model estimates with standard errors in parentheses. The dependent variable is being invited to a job interview. * (**) ((***)) indicates significance at the 10% (5%) ((1%)) significance level.

Table 5. The Probability of Any Positive Reaction by Age: Regression Analysis

	(1)	(2)	(3)	(4)	(5)
Older candidate	-0.047*** (0.014)	-0.010 (0.024)	-0.010 (0.024)	0.012 (0.040)	0.010 (0.041)
Older candidate x Out-of-field employment		-0.078** (0.034)	-0.078** (0.034)	-0.078** (0.034)	-0.079** (0.034)
Older candidate x Years of inactivity		-0.031 (0.034)	-0.031 (0.034)	-0.031 (0.034)	-0.032 (0.034)
First application sent			0.033** (0.014)	0.033** (0.014)	0.034** (0.014)
Job application template type B			-0.019 (0.014)	-0.022 (0.016)	-0.022 (0.016)
Older candidate x Higher educated				-0.029 (0.029)	-0.030 (0.030)
Older candidate x Vacancy for administrative occupation				-0.009 (0.034)	-0.004 (0.035)
Older candidate x Vacancy for vending occupation				-0.025 (0.037)	-0.019 (0.038)
Older candidate x 50 years old				-0.012 (0.032)	-0.012 (0.032)
Older candidate x 12 years older				-0.014 (0.045)	-0.015 (0.045)
Older candidate x Female				0.024 (0.028)	0.025 (0.028)
Older candidate x Vacancy offering part-time contract					-0.020 (0.042)
Older candidate x Vacancy offered by interim office					0.024 (0.046)
Constant	0.165*** (0.010)	0.165*** (0.010)	0.139*** (0.014)	0.138*** (0.015)	0.137*** (0.015)
Vacancy fixed effects	Yes	Yes	Yes	Yes	Yes
F-test for the equality of the coefficients for "Older candidate x Out-of-field employment" and "Older candidate x Years of inactivity" (p-value)		0.174	0.168	0.170	0.173
Observations	1,152	1,152	1,152	1,152	1,152

The presented results are linear probability model estimates with standard errors in parentheses. The dependent variable is getting any positive reaction. ** (***) indicates significance at the 5% (1%) significance level.