



# **Do State Laws Protecting Older Workers from Discrimination Laws Reduce Age Discrimination in Hiring? Experimental (and Nonexperimental) Evidence**

David Neumark, Ian Burn, Patrick Button, and Nanneh Chehras



# **Do State Laws Protecting Older Workers from Discrimination Laws Reduce Age Discrimination in Hiring? Experimental (and Nonexperimental) Evidence**

**David Neumark**

University of California–Irvine

**Ian Burn**

University of California–Irvine

**Patrick Button**

Tulane University

**Nanneh Chehras**

University of California–Irvine

September 2016

Michigan Retirement Research Center  
University of Michigan  
P.O. Box 1248  
Ann Arbor, MI 48104  
[www.mrrc.isr.umich.edu](http://www.mrrc.isr.umich.edu)  
(734) 615-0422

## **Acknowledgements**

The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium through the University of Michigan Retirement Research Center Award RRC08098401. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof.

## **Regents of the University of Michigan**

Michael J. Behm, Grand Blanc; Mark J. Bernstein, Ann Arbor; Laurence B. Deitch, Bloomfield Hills; Shauna Ryder Diggs, Grosse Pointe; Denise Ilitch, Bingham Farms; Andrea Fischer Newman, Ann Arbor; Andrew C. Richner, Grosse Pointe Park; Katherine E. White, Ann Arbor; Mark S. Schlissel, *ex officio*

# **Do State Laws Protecting Older Workers from Discrimination Laws Reduce Age Discrimination in Hiring? Experimental (and Nonexperimental) Evidence**

## **Abstract**

We provide evidence from a field experiment — a correspondence study — on age discrimination in hiring for retail sales jobs. We collect experimental data in all 50 states and then relate measured age discrimination — the difference in callback rates between old and young applicants — to variation across states in antidiscrimination laws offering protections to older workers that are stronger than the federal age and disability discrimination laws. We do a similar analysis for nonexperimental data on differences across states in hiring rates of older versus younger workers. The experimental evidence points consistently to evidence of hiring discrimination against older men and more so against older women. However, the evidence on the relationship between hiring discrimination against older workers and state variation in age and disability discrimination laws is not so clear; at a minimum, there is not a compelling case that stronger state protections reduce hiring discrimination against older workers. In contrast, the non-experimental evidence suggests that stronger disability discrimination protections increase the relative hiring of older workers.

## **Citation**

Neumark, David, Ian Burn, Patrick Button, and Nanneh Chehras. 2016. “Do State Laws Protecting Older Workers from Discrimination Laws Reduce Age Discrimination in Hiring? Experimental (and Nonexperimental) Evidence.” Ann Arbor, MI. University of Michigan Retirement Research Center (MRRC) Working Paper, WP 2016-349.  
<http://www.mrrc.isr.umich.edu/publications/papers/pdf/wp349.pdf>

## **Authors’ acknowledgements**

We received generous support from the Alfred P. Sloan Foundation and the Social Security Administration, through a grant to the Michigan Retirement Research Center. The views expressed are our own, and not those of the Foundation or of the Social Security Administration. This study was approved by UC Irvine’s Institutional Review Board (HS#2013-9942)

## **Introduction**

Age discrimination may make it difficult for policymakers to encourage increased employment of older workers, to help address population aging. Indeed, policymakers may want to consider whether supply-side reforms that increase incentives to work longer should be complemented by stronger laws protecting older workers from discrimination in the labor market. To that end, this study does two things. First, it significantly builds upon a large-scale field experiment to measure age discrimination in hiring. Second, it studies whether stronger laws protecting older workers from discrimination in some U.S. states reduce hiring discrimination against older workers. Because many seniors transition to part-time or shorter-term “partial retirement” or “bridge jobs” at the end of their careers (Cahill et al., 2006; Johnson, 2014), or return to work after a period of retirement (Maestas, 2010), new hiring of older workers is likely to be essential to significant lengthening of work lives.

We focus not only on age discrimination laws, but also on disability discrimination laws. As argued in Neumark, Song, and Button (forthcoming) and Stock and Beegle (2004), disability discrimination laws may be important in protecting older workers, in particular, from discrimination. Disabilities that can limit work and hence trigger protection by disability discrimination laws rise steeply with age, especially past age 50 or so (e.g., Rowe and Kahn, 1997). Correspondingly, employer expectations that a worker will develop a disability in the near future should also rise steeply with age. Indeed, disability discrimination laws may do more to protect many older workers than age discrimination laws. Many ailments associated with aging have become classified as disabilities (Sterns and Miklos, 1995). This can give some older workers an option of pursuing discrimination claims under either the Age Discrimination in Employment Act (ADEA), the Americans with Disabilities Act (ADA), or the corresponding state laws. The combined effect of potential coverage under both age and disability

discrimination laws may be to increase protections. For example, the ADA does more to limit defenses against discrimination claims.<sup>1</sup> A disability discrimination claim does require, of course, proving a disability, but as we shall see, this can be substantially easier under state disability discrimination laws than under the ADA.<sup>2</sup>

It may seem obvious that stronger discrimination protections for older or disabled workers will increase hiring of older workers. However, these laws may be ineffective at reducing or eliminating age discrimination in hiring. Enforcement relies in large part on potential rewards to plaintiffs' attorneys. In hiring cases, it is difficult to identify a class of affected workers, which inhibits class action suits and thus substantially limits awards. In addition, economic damages can be small in hiring cases because one employer's action may extend a worker's spell of unemployment only modestly. Terminations, in contrast, can entail substantial lost earnings and pension accruals. Moreover, it could be worse: if age discrimination laws fail to reduce discrimination in hiring, but make it harder to terminate older workers, these laws could actually deter hiring of older workers (Bloch, 1994; Lahey, 2008a; Posner, 1995).

To garner evidence on whether stronger age and disability discrimination laws increase hiring of older workers, we marry two efforts in this paper. First, we substantially extend a recent large-scale resume correspondence study (Neumark, Burn, and Button, 2015), from 12

---

<sup>1</sup> Unlike the ADEA, the ADA does not include an exception for bona fide occupational qualifications (BFOQs). BFOQ exceptions arise when age is strongly associated with other factors that pose legitimate business or safety concerns (e.g., Stock and Beegle, 2004; Posner, 1995; Starkman, 1992). Furthermore, age-related disabilities might be judged as amenable to "reasonable accommodation" by employers under disability discrimination laws, which usually require "reasonable accommodation" of the worker, making it much harder to justify an apparently discriminatory practice on the basis of business necessity (Gardner and Campanella, 1991).

<sup>2</sup> Under the ADA and similar state laws, plaintiffs need to prove that they have a condition that "...substantially limits one or more major life activities..." (42 U.S. Code §12102 (1)): This has proved difficult, leading plaintiffs to lose the vast majority of cases (Colker, 1999). Even with the definition of disability being broader now after the ADA Amendments Act of 2008 (ADAAA), proving coverage is not easy for many conditions, unlike coverage under the ADEA which is obvious.

cities in 11 states to all 50 states (although we do not cover all the occupations included in the previous study). The evidence from the resume correspondence study provides direct measures of discrimination in hiring. Second, we utilize information on state age discrimination laws that extend beyond the federal ADEA and state disability discrimination laws that extend beyond the ADA to study the relationships between these state laws and the direct measures of age discrimination in hiring from the field experiment.<sup>3</sup> Our focus is on discrimination against job applicants ages 64 to 66, who are at or near the age of retirement.

Finally, we also provide some parallel findings from nonexperimental evidence on the relationships between relative hiring of older workers and these state laws. In contrast to this nonexperimental evidence, the experimental evidence circumvents issues of differences across states in which types of older workers, or how many older workers, seek employment. While the experimental evidence relates state antidiscrimination laws to explicit measures of hiring discrimination, the nonexperimental evidence could be more relevant to the policy question of whether stronger antidiscrimination laws would lengthen work lives, although it may be less likely to be causal.

### **Correspondence Study Evidence on Age Discrimination**

Experimental audit or correspondence (AC) studies of hiring are generally viewed as the most reliable means of inferring labor market discrimination (e.g., Fix and Struyk, 1993).<sup>4</sup>

---

<sup>3</sup> There is nonexperimental evidence that these state laws affect labor market outcomes for older or disabled worker. Neumark and Song (2013) find that the effects of increases in the Social Security Full Retirement Age on work and later retirement were larger in states with age discrimination laws that are stronger than the federal ADEA. Other analyses of state age and disability discrimination laws (a nonexhaustive list) include Neumark et al. (forthcoming), Jolls and Prescott (2005), Lahey (2008a), Stock and Beegle (2004), and Button (forthcoming). We do not review this evidence here; the reader is referred to those papers. We are aware of only two other papers that look at variation in experimental evidence on discrimination across jurisdictions with different anti-discrimination laws — Tilcsik's (2011) study of discrimination against gays, and Ameri et al.'s (2015) study of discrimination against the disabled.

<sup>4</sup> For discussions of why, see, for example, Bertrand and Mullainathan (2004) and Neumark (forthcoming). For critiques of this evidence, see Heckman and Siegelman (1993) and Heckman (1998).

While observational studies try to control for productivity differences between groups, AC studies create artificial job applicants in which there are intended to be no average differences by group, so that differences in outcomes likely reflect discrimination. Audit studies use actual applicants coached to act alike, and capture job offers, whereas correspondence studies create fake applicants (on paper or electronically) and capture “callbacks” for job interviews. Correspondence studies can collect far larger samples of job applications and outcomes, especially using the Internet. Because of the time costs of interviews, even large-scale, expensive audit studies typically have sample sizes only in the hundreds. Correspondence studies also avoid “experimenter effects” that can influence the behavior of the actual applicants used in audit studies (Heckman and Siegelman, 1993). For these reasons, we use a correspondence study in this paper.

There is past evidence on age discrimination in employment using correspondence study methods. The three main earlier studies, plus one recent study, point to substantial age discrimination in hiring for both men and women (Bendick, Jackson, and Romero, 1997; Bendick, Brown, and Wall, 1999; Riach and Rich, 2010; Lahey, 2008b; Farber, Silverman, and von Wachter, 2015). The recent Neumark et al. (2015) study was the first to focus on workers at or above the age of eligibility for Social Security benefits. Moreover, it addressed sources of bias in the estimates from these past studies that could be in either direction.

One issue is the practice of giving older and younger applicants similar labor market experience, consistent with the standard paradigm in correspondence studies. (One cannot, of course, match on the high experience of older applicants.) However, the absence of relevant experience commensurate with an older applicant’s age may be a negative signal,<sup>5</sup> and on real-

---

<sup>5</sup> Researchers are aware of this problem. Bendick et al. (1997) had both older and younger applicants report 10 years of similar experience on their resumes. However, they had the resumes for older

world resumes older applicants tend to report experience commensurate with their age. Neumark et al. (2015) addressed this question by using a variety of resume types for older workers, including some with experience commensurate with age, which we argued was more consistent with the central policy and legal questions regarding discrimination, and some with low experience matched to that of younger applicants, which hews more closely to the classic correspondence study paradigm. For one occupation (janitors), matching on low experience generated spurious evidence of discrimination against older male workers. However, for the retail sales occupations on which we focus in this paper, there was no such evidence.

Second, Heckman and Siegelman (1993) and Heckman (1998) have demonstrated that if the groups studied have different variances of unobservables, experimental estimates of discrimination can be biased in either direction (formally, it is unidentified) — the “Heckman critique.” This problem may be especially salient with respect to age, as the human capital model predicts greater dispersion in unobserved investments among older workers (Mincer, 1974; Heckman, Lochner, and Todd, 2006). If the average resume quality in the study is low compared to the distribution of resumes employers actually observe,<sup>6</sup> then the high variance group is more likely to exceed the threshold for hiring or a callback, creating a bias in favor of hiring older workers, and hence a bias against finding evidence of age discrimination. This problem was addressed by using a method developed in Neumark (2012), which is explained in more detail later in the paper. Neumark et al. (2015) found that the results for the sales occupations we study in the present paper are sensitive to correcting for this source of bias,

---

applicants indicate that they had been out of the labor force raising children (for the female executive secretary applications), or working as a high school teacher (for the male or mixed applications). Lahey (2008b) studies women, for whom she argues that time out of the labor force is less likely to be a negative signal. She then includes only 10-year job history for all applicants (in part based on conversations with three human resources professionals she cites who said 10-year histories were the “gold standard”). However, the older resumes in either study could convey a negative signal.

<sup>6</sup> As one example in a different context, Bertrand and Mullainathan (2004, p. 995) claim that they tried to avoid over-qualified applicants who employers might not bother trying to hire.

although they also noted that the evidence for these occupations might not be robust. Hence, we build the same bias correction into the present study.

Correspondence studies do not directly distinguish between taste discrimination and statistical discrimination. However, both are illegal under U.S. law.<sup>7,8</sup> Nonetheless, economists are interested in which model might explain discriminatory behavior, and the policy response may differ. Moreover, in applying these methods to older workers, there are many plausible channels of statistical discrimination. First, employers might expect older workers to have health problems, which could raise absenteeism, lower productivity, or pose accommodation costs. Second, employers might expect that older workers (our highest age range is 64-66) would be near retirement, and hence be less likely to want to invest in them. Third, an older applicant with experience commensurate to their age applying for the same job as a younger applicant might be viewed as less qualified or having less potential, because he or she has been at that job level for longer – i.e., has a slower “speed of success” (Tinkham, 2010). And fourth, employers may make assumptions about skill differences across cohorts – perhaps most important that older applicants have fewer computer skills.

We cannot definitively rule out a role for these explanations of the evidence – and as we noted above, it does not matter from a legal perspective. Nonetheless, Neumark et al. (2015) present a number of types of evidence suggesting that these potential sources of statistical discrimination do not play much of a role. Some of these are based on evidence external to the field experiment. For example, with respect to separations, younger workers are also likely to

---

<sup>7</sup> EEOC regulations state: “An employer may not base hiring decisions on stereotypes and assumptions about a person's race, color, religion, sex (including pregnancy), national origin, age (40 or older), disability or genetic information” (<http://www1.eeoc.gov/laws/practices/index.cfm?renderforprint=1>, viewed September 27, 2015).

<sup>8</sup> Customer discrimination is also illegal. For a number of cases showing failure of defenses based on customer preference, see <https://www1.eeoc.gov/eeoc/initiatives/e-race/caselist.cfm?renderforprint=1#customer> (viewed September 26, 2016).

leave a job; although this is for other jobs rather than retirement, the reason for turnover is irrelevant to the employer. In 2015:Q1 data from the Quarterly Workforce Indicators, the separation rate (relative to beginning-of-quarter employment) was 9.9 percent for workers ages 55-64, and 18.7 percent for workers ages 25-34 (our youngest age range is 29-31).<sup>9</sup> Other evidence comes from the study. For example, the study also used resumes with different kinds of “bridging” or partial retirement behavior, which we know is sometimes associated with declining health (Johnson, Kawachi, and Lewis, 2009; Johnson, 2014). Since employers should know this from past experience, if declining health is an issue, older applicants with “bridge resumes” should experience lower callback rates than other older applicants; but they do not.<sup>10</sup>

### **The Experimental Design**

The present study builds on the approach and findings from the prior study. The extension to all 50 states is critical for studying the effects of antidiscrimination laws. At the same time, the extensive resources required to extend to all 50 states necessitated omitting some of the occupations included in the previous study. In particular, we omit administrative assistant, security, and janitor jobs, and focus only on jobs in retail sales. A clear implication of this limitation is that the evidence must be regarded as a case study, which may not generalize to other low-skill jobs.<sup>11</sup> On the other hand, of the jobs included in Neumark et al. (2015), retail sales is the one for which both male and female applicants were submitted, so in the present study we obtain evidence on whether there are difference in the results for men and women. In addition, given the evidence from Neumark et al. (2015) that in retail sales there was no difference in measured age discrimination whether high-experience or low-experience resumes

---

<sup>9</sup> See <http://qwiexplorer.ces.census.gov/#x=0&g=0> (viewed August 11, 2016).

<sup>10</sup> The reader is referred to Neumark et al. (2015) for discussion of additional evidence on statistical discrimination, as well as other potential challenges to the validity of interpreted differences in outcomes by age as discrimination.

<sup>11</sup> These kinds of studies typically use a very limited number of jobs. For example, Farber et al. (2015) focuses only on age discrimination against women in administrative assistant jobs.

were used for older applicants, in this paper we use low-experience resumes that match those of younger applicants. This simplified the resume creation because a long work history did not have to be developed for the older applicants.

### *Basic Analysis Framework*

The core analysis uses probit models for callbacks ( $C$ ) as a function of dummy variables for age ( $S$  for older/senior) and observables (from the resumes)  $X$ . The latent variable model (for  $C^*$ ) is

$$(1) \quad C_i^* = \alpha + \gamma S_i + X_i \delta + \varepsilon_i.$$

In this basic model, the null hypothesis of no discrimination implies that  $\gamma = 0$  (for older workers). We always estimate the model for men and women separately.

Here we outline the solution proposed in Neumark (2012) to address the Heckman critique; the original paper provides details. To see the intuition behind the solution, recall that in a probit model, all that is identified is the ratio of the coefficient in the latent variable model to the standard deviation of the unobservable. If we are willing to assume that  $\delta$  in equation (1) is the same for younger and older applicants, then we can identify the ratio or the standard deviation of the unobservables, denoted  $\sigma_S/\sigma_Y$ , from the ratios of probit coefficients older (senior) and younger applicants. Thus, information from a correspondence study on how variation in observable qualifications is related to callback outcomes can be informative about the relative variance of the unobservables, and this, in turn, identifies an unbiased estimate of the effect of discrimination.

The parameters are estimated using a heteroskedastic probit model with variance differing between younger and older applicants, and requires that at least one element of  $\delta$  is

equal for younger and older workers.<sup>12</sup> With data on multiple productivity-related characteristics in  $X$ , there is an overidentifying restriction that the younger/older ratios of coefficients on any element of this vector are equal (to the same  $\sigma_S^{\Pi}/\sigma_Y^{\Pi}$ ). The method also requires that at some of the applicant characteristics in  $X$  affect the callback probability (since if all the effects are zero we cannot learn about  $\sigma_S^{\Pi}/\sigma_Y^{\Pi}$  from these coefficient estimates). AC studies typically do not try to include variables that shift the callback probability, but instead create one “type” of applicant for which there is only random variation in characteristics that are not intended to affect outcomes. However, we build this information into the study design, through assignment to some resumes of random elements of a vector of skills and other characteristics that should increase the callback probability.

### *Resume Creation*<sup>13</sup>

The core of a correspondence study is the bank of resumes created for the artificial job applicants, since these resumes constitute the study data. Our over-arching strategy was to use empirical evidence whenever possible in making decisions about creating the resumes, to minimize decisions that might limit the external or “comparison” validity of the results. In many cases, this empirical evidence came from a large sample of publicly available resumes we downloaded from a popular national job-hunting website. We downloaded a sample of more than 25,000 resumes, which we then scraped for a variety of types of information that we use in our resume design decisions. In addition, we used public-use data to inform other issues in designing the resumes.

---

<sup>12</sup> Thus, we could have begun by writing equation (1) with different coefficients on  $X$  for young and old workers.

<sup>13</sup> Many additional details are provided in the on-line appendix to Neumark et al. (2015), although with some differences because that paper presents a more complex study with additional occupations, additional resume types, etc. We do not do anything in the current paper that extends beyond what was done in Neumark et al. (2015), but in some cases what we do is more limited.

## Basic Parameters

Past studies have tended to use workers near age 30 as the young group, and workers near age 50 as the older group. We include a similar age range for young workers (29-31), but compare results to older workers near the retirement age (64-66), who are the focus of policy efforts to respond to population aging. We convey age on the resumes, via high school graduation year, which is common on the actual resumes we examined. Given these age ranges, we chose common names (by sex) for the corresponding cohorts based on data from the Social Security Administration. To focus on age, we chose first and last names that were most likely to signal that the applicant was Caucasian. In response to each job ad, we send out a quadruplet of resumes consisting of a young and old male applicant and a young and old female applicant.

Neumark et al. (2015) used the resume database to document that there are older applicants in retail sales, which is consistent with data from the Current Population Survey (CPS) Tenure Supplement showing a sizable representation of low-tenure older workers in the occupations that make up retail sales (retail salespersons and cashiers in the Census occupational classification). Furthermore, the presence of older resumes on the resume posting website suggests that older workers do use on-line resources such as we use in this study to apply for jobs. That paper also showed that retail sales capture appreciable shares of new hiring of older workers (and, of course, higher shares for the types of low-skill jobs that could plausibly be candidates for the study), and are in the upper tier in terms of the proportions of older people hired.<sup>14</sup>

As noted above, we use cities in all 50 states to maximize external validity and to include

---

<sup>14</sup> As additional evidence, Rutledge, Sass, and Ramos-Mercado (2016) compute the ratio of older (50-64) to prime age (30-49) hires in detailed occupations. Retail sales is in the top 10, based on 1996-2012 CPS data. They also report that the jobs into which older workers tend to be hired are much narrower for less-educated workers. Thus, although the study was never meant to provide representative evidence on all older job seekers, it seems to point to a significant part of the labor market, especially for less-skilled older workers.

variation in antidiscrimination laws across all states. This contrasts quite sharply with two of the past studies, which used only one or two cities (Lahey, 2008b; Bendick et al., 1999). Because low-skill workers have low geographic mobility (Molloy, Smith, and Wozniak, 2011), we also target the resumes to retail jobs in specific cities (one per state, see Table 2), with the job and education history on each resume matching the city from which the job ad to which we apply originates. This was a factor underlying our decision to limit the analysis in this paper to retail sales jobs.

### Job Histories

We relied on the actual job histories from the resume database, as well as other data sources, to create realistic job histories on our resumes. Examination of our scraped resumes indicated that even in the low-skilled retail sales jobs we study, resumes are tailored to the jobs. To construct the job histories, we first pool job titles and descriptions from the actual resumes to create a set of entries in the retail field, with only minor changes such as phrasing or grammar for consistency. We combined these job descriptions using the resume characteristic randomizer program created by Lahey and Beasley (2009). The program randomized the combination of job titles and descriptions, and job tenures. The program runs backward from the most current job to the beginning of the potential job history. We had to build in a probability of a job ending, and experimented with the randomizer to choose a probability that appeared to create job histories similar to the resumes we downloaded, in terms of number of jobs held and average tenure on a job; this iterative process led us to choose a 15 percent annual probability that the program will end the current job and move on to the next randomly assigned job.

We used the resume randomizer to produce a large number of job histories, and then selected a smaller set that looked the most realistic based on the resumes found on the job-hunting website. In particular, we dropped those that had very high levels of turnover. From this

sample of acceptable histories, we created four job histories for each city (and for each resume style we create). We added employer names and addresses randomly to each job in our final job histories. We identified 15 possible employers for each city and assigned each employer to a job description such that no employer is used more than once on the same resume, or more than once across resumes in the quadruplet of resumes that are sent to each employer. We ensured that the job title and description was realistic for the employer. In addition, we used employers that were active at the time and in the region listed, relying mainly on national chains that had stores in many cities.

To mimic the seasonal pattern of job changes, we randomly drew the separation month for each job, except the most recently held job, from the distribution of job separation dates from the Job Openings and Labor Turnover Survey (JOLTS). We use the distribution specific to “Retail Trade.” During the course of the field experiment, every month we moved the ending date of the most recent job forward one month, so that durations did not lengthen during the time the experiment was in the field. We distinguish resumes based on whether applicants are currently unemployed. We assign all applicants within each quadruplet as either employed (the most recent job end date listed as “Present”), or unemployed, with 50 percent probability for each.<sup>15</sup> When applicants are unemployed, the resumes indicate that their last job ended in the month prior to the job application.

### Skills

To address the Heckman critique, we designate half the resume quadruplets to be high-skilled and half to be low-skilled.<sup>16</sup> For each type of high-skill resume, there are seven possible skills, five of which are chosen randomly (so that they are not perfectly collinear within a job).

---

<sup>15</sup> We did not want random assignment of unemployed or employed resumes within a quadruplet to dominate the effect of age.

<sup>16</sup> Like for unemployment, we make the set of resumes sent to each employer uniformly high-skill or low-skill because skill and age define different treatment groups.

Included in the skill vector are five general skills: a Bachelor of Arts degree; fluency in Spanish as a second language; an “employee of the month” award on the most recent job; one of three volunteer activities (food bank, homeless shelter, or animal shelter); and an absence of typographical errors.<sup>17</sup> Two skills of the seven are specific to retail sales, including Microsoft Office and programs used to monitor inventory (VendPOS, AmberPOS, and Lightspeed).

### Additional Resume Elements

There are a number of additional resume elements that we added. Residential addresses with regard to socioeconomic status, demographic characteristics, and distance to jobs, were selected to be realistic for both older and younger applicants and the jobs to which we were applying, did not signal a race other than white, and were not likely to send an unusual signal (positive or negative) about the applicant. The addresses were randomly assigned with respect to age, so there is no association between socioeconomic status of the neighborhood and age of applicant.

We randomly assign high schools, and colleges and universities for the high-skilled resumes, for each city, to each applicant in our quadruplet. We use local schools, colleges, and universities that were in operation since 1960 so that there is no possibility that an applicant attended a school that was not operational at the time. We also restrict our schools to those with a significant share of white students. For smaller cities, this often limits the number of high schools or colleges that were available. In six cities, we were only able to find two high schools that fit our criteria. For the rest, we selected three different high schools.<sup>18</sup> We avoided top-tier/flagship universities whenever possible. We also restricted our schools to not include

---

<sup>17</sup> Thus, all low-skill resumes and the high-skill resumes not assigned this skill include two typos. We use a missing space and a missing period, with one of these appearing for the most recent job, which employers are most likely to read. These kinds of errors were more common on actual resumes than spelling errors.

<sup>18</sup> There were many cities for which we could not identify four high schools. In such cases, employers should not be surprised to get two resumes listing the same high school.

Historically Black Colleges. In two states (Wyoming and Delaware), there was only one university that fit our criteria.

### *Resume Quadruplets*

Each of the four resumes in the quadruplet was randomly assigned a different resume template, which ensured that all four resumes looked different. Most other characteristics were randomly and uniquely assigned to each resume in each quadruplet to further ensure that the applicants were distinguished from each other, and that any resume characteristics that inadvertently were more or less appealing to employers were distributed randomly with respect to the four applicants in each quadruplet. These characteristics included first and last names, school names, addresses, phone numbers, email address formats and domains, cover letter style, and the language describing jobs and skills.<sup>19</sup>

### *Applying for Jobs*

We identify jobs to apply for using a common job-posting website. Research assistants read the posts regularly to select jobs for the study, using a well-specified set of criteria. Jobs had to be entry level (e.g., not managers or supervisors), and the ads could not require in-person applications, inquiries by phone, or require applicants to use an external website. The ads could not require additional documents we had not prepared (e.g., a salary history, etc.), or skills that our resumes did not have.<sup>20</sup>

Research assistants saved the list of jobs to apply for in a shared folder. We wrote Python code to automate the application process from the jobs put in this shared folder. This substantially reduced labor costs, removed human error such as attaching the wrong resume, and ensured that jobs applications used a uniform procedure. Using SQL, the code matched the job

---

<sup>19</sup> The on-line appendix from Neumark et al. (2015) provides examples of resume types exhibiting these and other variations.

<sup>20</sup> A number of other exclusion criteria are outlined in the on-line appendix, as are other quality-control procedures we implemented, and checks on them, regarding the job application process.

ad data to the applicant based on city and date. Each day was randomly assigned a different quadruplet of resumes in terms of skill levels, and employed or unemployed. Within each quadruplet the order of resumes was randomized. The code ran every other day and added 7 to 8 hour delays between applications to the same jobs.

### *Sample Size*

In an experiment, it is important not to continue to collect data until the estimated differences become statistically significant. We had an explicit data collection plan that covered two academic quarters, in which we collected as much data as the available job ads would allow. No data were analyzed until the data collection was complete. We ended up sending out 14,428 applications to 3,607 jobs.

### *Collecting Responses*

Responses to job applications could be received by email or phone. All responses were forwarded to a central email account, with voicemails arriving as attachments. We then read each email and listened to each voicemail to record the response. We then used additional information to match a response to a specific job ad, using information on the job ads recorded during the job application process.

If the email was sent as a reply to the job-listing website submission, then the email also contained a unique id number for the job ad. Sometimes firms responded directly to the individual, in which case we had to use other information to match to the specific job. Phone call responses conveyed less information. Every voicemail contained the phone number of the firm calling and the phone number on the resume they were trying to contact. The automated voicemail message instructed firms to include their name and their number in their message. Identifying information that was extracted from a voicemail included, when possible, the firm name, applicant name, the job title, and any other information that could be used to narrow down

the list of possible job ads (e.g., how long ago they received the resume). The information extracted from the voicemail was used to match each voicemail to a job ad. Table 1 reports the distribution of responses by phone or email (or both).

Each response was coded as an unambiguous positive response (e.g. “Please call to set up an interview”), an ambiguous response (e.g. “Please return our call, we have a few additional questions”), or an unambiguous negative response (e.g. “Thank you for your interest, but the job has been filled”). To avoid having to classify subjectively the ambiguous responses, they were treated as callbacks;<sup>21</sup> the negative responses were treated the same as no callbacks.

### **Nonexperimental Evidence**

Although our paper emphasizes experimental evidence from the correspondence study, we also present some parallel evidence on hiring behavior using data from the Quarterly Workforce Indicators (QWI)<sup>22</sup> (U.S. Bureau of the Census, 2016). The QWI are a set of economic indicators including employment, job creation, and employment flows. The version of the QWI that we use reports by age group and sex, tabulated to the state level. We use the hiring rate variable that measures hires as a percent of average employment.<sup>23</sup> The data only permit broader age groups than the ages we use in our experiment (29-31 and 64-66). We use 25-34 year-olds as our young group, and two alternative older groups: 55-64, and 65-99. We restrict attention to data on the retail sales industry (NAICS codes 44-45), to match the experimental data. We use data for the first three quarters of 2014 — the most current data recently available; the experimental data were collected from early February to early July of 2016.

The QWI data on hiring do not provide the same information as the experimental

---

<sup>21</sup> The ambiguous responses are 7.8% of all cases coded as positive callbacks.

<sup>22</sup> These were downloaded from Cornell University’s Virtual Research Data Center (R2015Q2 release). By downloading data from the Cornell Virtual RDC Web site, we acknowledge support from NSF grant #SES-0922005 that made these data available.

<sup>23</sup> This and other QWI variables are discussed in-depth in [http://lehd.ces.census.gov/doc/QWI\\_101.pdf](http://lehd.ces.census.gov/doc/QWI_101.pdf) (viewed September 25, 2016).

evidence on callback rates. In particular, the QWI hiring rates do not hold characteristics of young and old applicants fixed (which we do on the resumes, in the experimental data). In addition, the QWI data can be influenced by whether people in different age groups look for new jobs, whereas in the experimental data this does not affect the results since we send out applicants of both ages.<sup>24</sup>

Nonetheless, it is of interest to compare the results using the two different data sources. If the evidence is similar, it might suggest that we do not necessarily need experimental data to study the effects of antidiscrimination laws on discriminatory behavior. On the other hand, the analysis using the two data sources might be viewed as answering different questions. Policymakers focused just on boosting hiring of older workers might in fact be more interested in the effects of antidiscrimination laws on hiring of older workers relative to younger workers — without regarding to changes in the composition of who looks for work, etc. — than in how otherwise identical applicants are treated.

This does raise one important caution, however, about our analysis of both the experimental and the nonexperimental data. In particular, the measured variation is cross-sectional, not longitudinal. In the experimental data, this is dictated by the collection of data over a short period.<sup>25</sup> However, the ability to study the effects of current variations in state antidiscrimination laws is severely limited, because there are very few changes in these laws in recent decades (Neumark and Song, 2013; Neumark et al., forthcoming). Thus, both types of evidence can potentially reflect other factors correlated with both outcomes for older workers

---

<sup>24</sup> As a concrete example, part of the motivation for the restaurant audit study in Neumark (1996) was that expensive restaurants claimed they do not hire women as wait staff because women do not apply. By sending both male and female applicants in the experiment, the paper showed that to some extent, at least, there was discrimination against female applicants when they did apply.

<sup>25</sup> For an interesting example of correspondence study evidence collected before and after a policy change (in the context of hiring differences of those with and without criminal backgrounds), see Agan and Starr (2016).

and antidiscrimination laws.

### **Coding of Antidiscrimination Laws**

Our coding of age discrimination laws and disability discrimination laws was developed, and is fully described in Neumark and Song (2013) and Neumark et al. (forthcoming); these papers also report some analyses of the effects of these laws, albeit using only nonexperimental data. The compilation of information on these laws entailed extensive background research on state statutes and their histories, culled from legal databases including Lexis-Nexis, Westlaw, and Hein Online, as well as many other sources.

The current laws are reported in Table 2.<sup>26</sup> We focus on the two aspects of age discrimination laws that the past research suggested were important. The first is the minimum firm-size cutoff for the law to apply.<sup>27</sup> We use a firm-size cutoff of fewer than 10 workers to capture state laws that extend to substantially smaller firms (the minimum for the ADEA to apply is 20). The second is whether compensatory or punitive damages are allowed, which they are not under federal law.<sup>28</sup>

State disability discrimination laws are sometimes stronger than the federal ADA in three principal ways, all captured as well in Table 2. As with age laws, there is a minimum firm size

---

<sup>26</sup> Table 2 reveals that the distribution of stronger protections across states does not reflect the usual pattern related to generosity of social programs, minimum wages, etc. For example, some southern states have among the strongest anti-discrimination protections.

<sup>27</sup> For example, in Florida a worker who works at a firm that employs fewer than 15 employees is not covered under the Florida state law. On the contrary, all employees in Colorado are covered by state law because it is applicable to all firms with at least 1 employee.

<sup>28</sup> See United States Equal Employment Opportunity Commission (2002). Some states require proof of intent to discriminate in order for compensatory or punitive damages to be awarded, whereas others require “willful” violation. Because the federal law allows additional liquidated, nonpunitive damages (double back pay and benefits) when there is “willful” violation, the question of whether the state requires intent or willful violation may seem to be potentially relevant in deciding whether a state law offers greater protection. However, willful violation is a much stricter standard than intent (Moberly, 1994). Moreover, compensatory or punitive damages are almost certainly greater than liquidated damages, and they can be much greater. As a consequence, a state law that provides compensatory or punitive damages, whether or not this requires proof of intent or willful violation, clearly entails stronger remedies than the federal law.

to which disability discrimination laws apply. The minimum firm size for the ADA to apply is 15; in our analysis we distinguish states with a firm size minimum lower than 10, the same as for age-discrimination laws. There is also variation in damages, through higher or uncapped compensatory and punitive damages, relative to the capped damages available under the ADA. We distinguish states with larger damages than the ADA; we base this classification on punitive rather than compensatory damages, since punitive damages are likely to drive large judgments.

Finally, state laws vary in terms of the definition of disability. Most states adopt the ADA definition, either explicitly or via case law. Some states use a laxer definition, changing a key part of the definition of disability from “substantially limits one or more major life activities” to either “materially limits” (Minnesota) or just “limits” (California). Other states vary the definition of disability by requiring that the disability be “medically diagnosed” without regard to whether the impairment limits major life activities (Long, 2004). Table 2 includes information on both dimensions of the definition of disability, and we use both in our analysis.

## **Results**

### *Basic Callback Rates*

Table 3 reports raw differences in callback rates by age, and statistical tests of whether callback rates are independent of age.<sup>29</sup> In Panel A, for males, we find strong overall evidence of age discrimination, with callback rates statistically significantly lower by 7.6 percentage points for older workers compared to younger workers, or 30.4 percent lower. The evidence in Panel B, for females, similarly points to age discrimination. The absolute difference is a bit larger (8.5 percent), although it is more similar in relative terms because the callback rate is about 3.5 percentage points higher for women than for men. These results are similar to those in Neumark

---

<sup>29</sup> This test treats the observations as independent. In the regression (probit) analyses that follow, the standard errors are clustered appropriately.

et al. (2015), although there the callback differential was larger for women (about 10 percent versus 6 percent for men).<sup>30</sup>

In correspondence studies, there is a question of what evidence on callbacks tells us about hiring. For example, if employers believe there is age discrimination, then they may expect older applicants to be more likely to respond positively to a callback. Nondiscriminatory employers might then direct more callbacks to older workers, which would generate a bias against finding evidence of age discrimination (although employers might do the opposite if they have a target share of older worker hires). However, there is evidence that differences in callback rates accurately reflect hiring discrimination. The Bendick et al. (1999) audit study that captured differences in outcomes at different stages of the application process found that three-quarters of the overall discriminatory difference in treatment occurred at the preinterview stage.<sup>31</sup> Thus, there is good justification for assuming that our results for callbacks would carry over to job offers, although of course the magnitudes could differ.

### *Multivariate Estimates*

Table 4 reports results of probit estimates for callbacks (equation (1)), showing marginal effects. In each case, we first report results with controls for the state, the order in which applications were submitted, current employment/unemployment, and skills. We then add controls for an extensive set of resume features listed in the table notes. The random assignment of age to resumes in AC implies that the controls should not affect the estimated differences associated with age, and that is reflected here, as the estimates in Table 4 are very similar to those in Table 3, with an estimated percentage point shortfall in callbacks of 7.5-7.7 percentage

---

<sup>30</sup> Note that the callback rates at all ages are higher for women than for men. Similarly, Neumark et al. (2015) and Bertrand and Mullainathan (2004) did not find discrimination against women in retail.

<sup>31</sup> An employer is more likely to discriminate at the preinterview (callback) stage than at the interview stage. Because company personnel systems often create data records for those interviewed, discrimination in offering jobs to applicants may be much easier to detect than discrimination in deciding who to call back for an interview.

points for men, and 8.6-8.9 percentage points for women.

In this and subsequent tables analyzing the experimental data, the estimates are clustered at the age-by-state level. We do this because the policy variation we study when we estimate the effects of state antidiscrimination laws on callbacks varies by state and by age (since we include age-by-state interactions), and this clustering will then exactly parallel what we do with the nonexperimental data.<sup>32</sup>

Given that the additional resume feature controls make essentially no difference to the estimates, nor should they, going forward we use the more parsimonious specifications in columns (1) and (3). These specifications retain the skill variables we added to address the Heckman critique (as well as the unemployment and order of application variables, which may also function like the skill variables).

#### *Adding State Antidiscrimination Laws*

We next turn to the main contribution of this paper – the estimation of the effects of state antidiscrimination laws protecting older workers on callback rates for older relative to younger workers. We do this by modifying equation (1) to include interactions of dummy variables for these state laws (in some cases a vector) with the dummy variable for older applicants. Because we include state dummy variables, we do not include the main effects of the state antidiscrimination laws. Excluding the state dummy variables and including the main effects of the laws would result in a less saturated model, whereas the models we estimate allow more flexibility for differences in callback rates for younger workers across states than only variation correlated with the state antidiscrimination laws. Of course, we have to assume that state-by-age interactions are excluded from the model to estimate the interactive effects of interest. Adding to

---

<sup>32</sup> Absent this consideration, one might want to cluster at the level of the resume or the job ad. In Neumark et al. (2015) we verified that the two alternatives have virtually no effect on the standard errors.

equation (1) an ‘s’ subscript to denote states, and defining  $A_s$  as the dummy variable (or vector of dummy variables) capturing state antidiscrimination laws, we augment the model to be

$$(2) \quad C_{is}^* = \alpha + \gamma S_{is} + S_{is} \cdot A_s \gamma' + X_{is} \delta + \varepsilon_{is},$$

where, recall,  $X$  includes the state dummy variables. Our interest centers, of course, on whether stronger state antidiscrimination laws are associated with differences in the relative callback rate of older workers, captured in  $\gamma'$ .

The first estimates we report, in Table 5, add, separately, the two features of age discrimination laws on which we focus — a smaller firm-size cutoff and larger damages. In Table 5, the main effects of “Old” refer to states where the federal law binds, and the interaction with the feature of the law considered captures the differential in the relative callback rate where there is a stronger state law. For a lower firm-size cutoff, the estimated interaction for men is negative but insignificant, while the estimate for women is positive and statistically significant. The estimates for women imply that in the states where the federal law binds, the callback differential by age is quite a bit larger than for men (10.6 versus 6.3 percentage points lower). But it is not clear, *a priori*, why the estimated effect of the lower firm-size cutoff would be different for women than for men. In the actual labor market, it is possible that older women on average apply to work at smaller firms. But in the correspondence study that should not play a role, since all job ads receive two male and two female applicants. The estimated interactions with the dummy variable for larger damages are small and insignificant (and negative) for both men and women.

Table 6 turns to state disability discrimination protections. This table is more complicated because there is a third feature of the laws that we study — the definition of disability — and because there are two different classifications of this definition. Looking first at the results for the lower firm-size cutoff and larger damages, which are comparable to the

features of age discrimination laws considered in Table 5, we find small and insignificant effects in three cases, but a positive and statistically significant effect of larger damages for older women. This is different from the result for age discrimination laws where we found a positive and significant effect of a lower firm-size cutoff for women. But it is similar in the sense that for women, but not for men, we find a positive and significant effect of some feature of state antidiscrimination laws that strengthens the law relative to the federal law.

Columns (3), (4), (7), and (8) report estimates of the effects of a broader definition of disability on the relative callback rate for older workers. Here, all the estimates are negative, but only one (for men, and for the medical only definition) is statistically significant. The effect of a broader definition can, of course, cut two ways. On the one hand it can extend protections and increase hiring. But on the other hand, it could make employers warier of hiring an older worker who might suffer a health decline and become subject to state disability discrimination protections – in this case, more easily because of the broader disability definition.

The previous two tables may not estimate the independent effects of each variation in state antidiscrimination protections, because the presence or absence of different features of state laws are correlated across states, as Table 2 indeed suggests.<sup>33</sup> Thus, in Table 7 we add all the law interactions simultaneously. Perhaps unexpectedly, the estimates are still relatively precise. For the age discrimination laws, we now get a clearer message, albeit still one that points to different effects for women versus men. In particular, for women there is strong evidence that a lower firm-size cutoff increases the relative callback rate of older workers, and indeed the magnitude offsets a large share of the callback difference in states where the federal laws bind. For men, in contrast, the estimates are negative and statistically significant. For larger damages,

---

<sup>33</sup> For example, New Jersey has a lower firm-size cutoff and larger damages for both age discrimination and disability discrimination, and a broader medical definition, Rhode Island has the first four, but not a broader disability definition, and Nebraska has no stronger protections for either type of discrimination.

the estimates are small and statistically insignificant. For the disability discrimination laws, the effects of the firm-size cutoffs are reversed, with positive effects for men (significant only in column (1)) and negative effects for women (significant in both columns). For women, however, the effect of state disability discrimination laws providing for larger damages are positive, and significant at the 10 percent level. For the broader definition of disability, most of the estimates are small or insignificant, with the exception of the negative and significant effect for the medical-only broader definition of disability for men.

Overall, the examination of the effects of stronger state age and disability discrimination protections on hiring of older workers provides a somewhat mixed message. For men, most of the evidence points to negative effects; this is true for a lower firm-size cutoff for age discrimination, and the definition of disability that extends to medical issues. There is one estimate that points to a positive impact for the firm-size cutoff for disability discrimination, but this finding is not robust. Thus, for men, the preponderance of the evidence suggests that stronger antidiscrimination protections for older workers may deter hiring. For women, the evidence is more mixed and harder to reconcile. Indeed, the strongest evidence is for firm-size cutoffs, but the estimated effects are of different signs for age discrimination (positive) versus disability discrimination (negative). And there is weaker evidence that larger damages for disability discrimination may increase hiring. From this, we draw two key conclusions: (1) there clearly is *not* unambiguous evidence that stronger age and disability discrimination protections boost hiring of older workers; and (2) there is some evidence – most clear for men – that these laws may be more likely to reduce hiring of older workers.

#### *Correcting for Bias from Differences in the Variances of Unobservables*

We next turn to the estimates that are intended to eliminate the bias identified by the Heckman critique. To briefly explain the procedure, we first estimate a probit model with the

controls and their interactions with “Old” included.<sup>34</sup> We then test the overidentifying restriction for the controls, to see whether the data are consistent with the effects for young and old differing in a way that is driven only by the difference in variance of the unobservables (that is, the ratios of effects for young and old workers are equal).<sup>35</sup> It turns out that the overidentifying restrictions using all of the controls are not rejected by the data, so we do not have to narrow down the set of variables used to identify the relative variance. We then estimate a heteroskedastic probit model that imposes equal coefficients of the controls in the latent variable model, with the variance of the residual differing between young and old workers. The estimates of this model are used to estimate marginal effects, and to decompose the marginal effects to isolate the effects of the variables on the level of the latent variable, which are the unbiased estimates of discrimination. (The decomposition also identifies the effect of “Old” via the variance, which, as explained in Neumark (2012), is an artifact of the study design using a very narrow range of resume quality.)<sup>36</sup>

The results are reported in Table 8. The upper rows of the table report the marginal effects corrected for bias. The specifications are otherwise the same as those in Table 7, using all the laws simultaneously, and hence can be compared directly. One result is that the estimates for the main effects of “Old,” which measure age discrimination in the states where the federal laws bind, becomes a little bit larger in absolute value for men (at least in column (1)), but quite

---

<sup>34</sup> We do not report these results here. It turns out that the skill variables have stronger effects on callback probabilities than we obtained in Neumark et al. (2015) using the same variables in our job applications for sales jobs. That could be because of the smaller number of cities (12, in 11 states) to which we sent applications, especially given that New York and Los Angeles provided very large numbers of observations.

<sup>35</sup> To identify the effect of the old-state law interactions, we have to assume equal coefficients for the state dummy variables, so this restriction is simply imposed. The overidentification test we use pertains to all of the other controls.

<sup>36</sup> This decomposition is unique when using the calculation of marginal effects that treats the variables as continuous, which is not standard, but has virtually no effect on the estimated marginal effects. The standard calculation of the marginal effect for discrete variables does not yield a unique decomposition.

a bit larger for women – by around 25 percent. Thus, for these states the evidence of age discrimination always strengthens, although appreciably only for women.

The third panel of the table reports the ratio of the standard deviation of the unobservable for old relative to young workers. For men, this ratio exceeds 1.1 in column (1), and for women, it is larger still in both columns. The larger standard deviations for older workers, coupled with the larger estimates of discrimination, are consistent with the resumes on average being of lower quality. However, the estimated interactions between “Old” and the features of state antidiscrimination laws do not change much from correcting for the bias from different variances of the unobservables, consistent with the relative standard deviations of the unobservables being relatively close to 1.<sup>37</sup>

#### *Results for Nonexperimental Data*

Finally, we report results for the nonexperimental QWI hiring data. Table 9 reports analyses that parallel Table 5 – considering each type of age discrimination law in isolation. Tables 10A and 10B report analyses that parallel Table 6 – considering each type of disability discrimination law in isolation. And Table 11 parallels Table 7 in including all the laws simultaneously, alternating between the two ways that state disability discrimination laws broaden the definition of disability.

Tables 9, 10A, and 10B paint a clear and unambiguous picture. In every column, the provision of state age or disability discrimination laws that we consider has an estimated positive interaction with “Old,” implying that these provisions boost the relative hiring of older workers.

---

<sup>37</sup> We found more evidence of bias in Neumark et al. (2015) – which only estimated the effects of age – with the bias correction strengthening the evidence of discrimination for women, and weakening it substantially for men. Those estimates may have been less robust because of using many fewer states (cities), as well as because the skill variables had weak predictive power for callbacks in sales. In addition, it is possible there was more bias because with fewer states, the resumes we sent out may have more uniformly been on one side of the distribution of resume quality that employers observe. In both studies, our resumes used for different states were of uniform quality. But if applicant quality differs across the states/cities, then by using more of them we may have reduced the bias.

Moreover, nearly every estimated interaction is significant at the 5 percent level or less.

Table 11, however, provides more reliable evidence by controlling of the different provisions of these laws simultaneously. Interestingly, once we do this, the estimated interactions for the age discrimination law provisions are near zero and never statistically significant. In contrast, a number of the estimated interaction effects for state disability laws are positive and significant, and many are quite a bit larger in magnitude than the estimates for age discrimination laws. In particular, a lower firm-size cutoff for disability laws is associated with higher relative hiring of men, although this result is not significant or is only marginally significant in columns (3) and (4). More robust is the evidence that larger damages for disability discrimination boost the relative hiring of older workers. This finding is statistically significant in every case (for men and women) for the younger of the two older groups (ages 55-64), and for women, using both definitions of the older group the finding, is always significant at the 5 percent level or better. Finally, either of the two broader definitions are associated with higher relative hiring of older workers. However, the effect is of a much larger magnitude and is much more statistically significant (always at the 1 percent level) using the broader definition that uses the laxer definition of limits. For the broader definition based on the medical definition of disability only, only one estimate is statistically significant (for women 55-64, at the 5 percent level). Thus, the nonexperimental data clearly provide stronger and more consistent evidence that state laws protecting older workers from discrimination boost hiring, compared to the results for the callback rates estimated from the experimental data. Moreover, this evidence arises only for disability discrimination protections.

## **Conclusions and Discussion**

In this study, we provide evidence from a field experiment — a correspondence study — on age discrimination in hiring for retail sales jobs. The unique contribution of this paper is to

collect experimental data in all 50 states, and then to relate the measure of age discrimination — the difference in callback rates between old and young applicants — to variation across states in antidiscrimination laws offering protections to older workers that are stronger than the federal laws. We study both age discrimination and disability discrimination laws. While age discrimination laws explicitly target discrimination against older workers, we argue that it is natural to expect disability discrimination laws to do far more to protect older workers than younger workers. Finally, in addition to analyzing the experimental data, we also study nonexperimental data on differences across states in hiring rates of older versus younger workers.

The experimental evidence points consistently to evidence of hiring discrimination against older men and more so against older women. However, the evidence on the relationship between hiring discrimination against older workers and state variation in age and disability discrimination laws is not so clear. Some protections appear to exacerbate the callback difference between older and younger workers — in particular, a lower firm-size cutoff for age discrimination laws and a broader definition of disability for men, and a lower firm-size cutoff for disability discrimination laws for women. In contrast, there is some evidence that stronger protections are associated with less discrimination against older workers – for a lower firm-size cutoff for age discrimination laws and larger damages for disability discrimination laws for women.

To summarize the experimental evidence, clearly this evidence does not support a general conclusion that stronger antidiscrimination protections reduce measured hiring discrimination against older workers. Indeed, somewhat more evidence suggests that these stronger protections under state laws increase measured discrimination. This latter effect is possible, because protections that might make it more difficult to terminate an older worker, or in the case of

disability, raise future accommodation costs for employers, can deter hiring of the protected group, especially if the antidiscrimination laws are relatively ineffective at reducing discrimination in hiring while being more effective with regard to terminations. However, the absence of consistent evidence in this direction ultimately makes us reluctant to draw strong conclusions from the relationship between measured discrimination against older workers and state antidiscrimination laws – except the “negative” conclusion that there is not a compelling case that these laws reduce hiring discrimination against older workers.

The evidence from the nonexperimental data on hiring is quite different. In particular, while the data on hiring rates yields little indication that stronger state age discrimination laws are associated with higher relative hiring of older workers, it generates quite unambiguous evidence that the relative hiring of older workers is higher in states with stronger protections against disability discrimination.

The obvious question is why the answers from the experimental and nonexperimental data are different. We do not necessarily anticipate the same answer. If we did, there would be little need to carry out a correspondence study. The experimental evidence provides a direct measure of hiring discrimination, whereas the hiring differences captured by the nonexperimental evidence can reflect compositional variation across states in which older workers look for work, as well as differences across states in the likelihood that older individuals look for work; the experimental variation eliminates both of these sources of variation.

Perhaps the most natural explanation for the different results is that stronger state laws protecting older workers from discrimination do not have a clear causal effect on measured discrimination, but are more likely to be adopted where more older workers are looking for work, generating spurious evidence of positive effects on hiring rates. It may make sense that these laws are less endogenous with respect to the age discrimination we measure with the

correspondence study, since this discrimination is unlikely to be easily observed by policymakers, and may not be strongly correlated (and perhaps could be negatively correlated) with employment or hiring rates of older workers. Alternatively, stronger state discrimination laws may have a positive causal effect on hiring through changing the composition of which older workers seek employment or more generally encouraging older workers to work, even though the laws do not reduce the discrimination between otherwise identical older and younger job applicants that the correspondence study measures.

Under either interpretation, there is little basis from our evidence for concluding that stronger state laws protecting older workers from discrimination reduce age discrimination in hiring; under the second interpretation, however, they may still increase hiring of older workers via other channels. It is possible that the latter effect is of more importance to policymakers trying to increase the employment of older workers, although the case for interpreting it as causal is weakened by the fact that we have only cross-sectional variation in state age discrimination laws. There is past work on longitudinal variation in age and disability discrimination laws (from the advent of state laws and then federal laws) that likely provides better evidence on causal effects — albeit not with respect to the features of state laws we study in this paper. This past work indicates that adoption of age discrimination laws boosted employment of older workers (Adams, 2004, Neumark and Stock, 1999). However, the evidence is less clear for disability discrimination laws (e.g., Beegle and Stock, 2003; Kruse and Schur, 2003; Button, forthcoming), suggesting that the relationship we find between state disability protections and relative hiring of older workers may not be causal. In that case, the evidence from our experimental data may, in fact, be more definitive, and there may not be much case for concluding that stronger state age and disability discrimination laws either reduce age discrimination in hiring against older workers, or more generally, increase hiring of older

workers.

## References

- Adams, Scott J. 2004. "Age Discrimination Legislation and the Employment of Older Workers." *Labour Economics*, Vol. 11, pp. 219-41.
- Agan, Amanda Y., and Sonja B. Starr. 2016. "Ban the Box, Criminal Records, and Statistical Discrimination: A Field Experiment." University of Michigan Law & Economics Research Paper No. 16-012.
- Ameri, Mason, et al. 2015. "The Disability Employment Puzzle: A Field Experiment on Employer Hiring Behavior." NBER Working Paper No. 21560.
- Beegle, Kathleen, and Wendy A. Stock. 2003. "The Labor Market Effects of Disability Discrimination Laws." *Journal of Human Resources*, Vol. 38, pp. 806-59.
- Bendick, Marc, Jr., Lauren E. Brown, and Kennington Wall. 1999. "No Foot in the Door: An Experimental Study of Employment Discrimination Against Older Workers." *Journal of Aging & Social Policy*, Vol. 10, pp. 5-23.
- Bendick, Marc, Jr., Charles W. Jackson, and J. Horacio Romero. 1997. "Employment Discrimination Against Older Workers: An Experimental Study of Hiring Practices." *Journal of Aging & Social Policy*, Vol. 8, pp. 25-46.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, Vol. 94, pp. 991-1013.
- Bloch, Farrell. 1994. Antidiscrimination Law and Minority Employment. Chicago: University of Chicago Press.
- Button, Patrick. "Expanding Employment Discrimination Protections for Individuals with Disabilities: Evidence from California." Forthcoming in *Industrial and Labor Relations Review*
- Cahill, Kevin E, Michael D. Giandrea, and Joseph F. Quinn. 2006. "Retirement Patterns from Career Employment." *The Gerontologist*, Vol. 46, pp. 514-23.
- Colker, Ruth. 1999. "The Americans with Disabilities Act: A Windfall for Defendants." *Harvard Civil Rights Civil Liberties Law Review*, Vol. 34, pp. 99-162.
- Farber, Henry S., Dan Silverman, and Till von Wachter. 2015. "Factors Determining Callbacks to Job Applications by the Unemployed." NBER Working Paper No. 21689.
- Fix, Michael, and Raymond Struyk. 1993. Clear and Convincing Evidence: Measurement of Discrimination in America. Washington, DC: The Urban Institute Press.
- Gardner, Russell H., and Carolyn J. Campanella. 1991. "The Undue Hardship Defense to the Reasonable Accommodation Requirement of the Americans with Disabilities Act of 1990." *Labor Lawyer*, Vol. 7, pp. 37-51.
- Heckman, James J. 1998. "Detecting Discrimination." *Journal of Economic Perspectives*, Vol. 12, pp. 101-16.
- Heckman, James J., Lance L. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In Hanushek and Welch, eds., Handbook of the Economics of Education, Volume 1, Chapter 7, pp. 307-458.
- Heckman, James, and Peter Siegelman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." In Fix and Struyk, eds., Clear and Convincing Evidence: Measurement of Discrimination in America. Washington, D.C.: The Urban Institute Press, pp. 187-258.
- Johnson, Richard W. 2014. "Later Life Job Changes before and after the Great Recession." Draft final report to AARP.
- Johnson, Richard W., Janette Kawachi, and Eric K. Lewis. 2009. "Older Workers on the Move: Recareering in Later Life." Washington, DC: AARP Public Policy Institute.
- Jolls, Christine, and J. J. Prescott. 2005. "Disaggregating Employment Protection: The Case of Disability Discrimination." Harvard Public Law Working Paper No. 106.
- Kruse, Douglas, and Lisa Schur. 2003. "Employment of People with Disabilities Following the ADA."

- Industrial Relations*, Vol. 42, pp. 31-66.
- Lahey, Joanna. 2008a. "State Age Protection Laws and the Age Discrimination in Employment Act." *Journal of Law and Economics*, Vol. 51, pp. 433-60.
- Lahey, Joanna. 2008b. "Age, Women, and Hiring: An Experimental Study." *Journal of Human Resources*, Vol. 43, pp. 30-56.
- Lahey, Joanna N., and Ryan A. Beasley. 2009. "Computerizing Audit Studies." *Journal of Economic Behavior & Organization*, Vol. 70, pp. 508-14.
- Long, Alex. 2004. "State Anti-Discrimination Law as a Model for Amending the Americans with Disabilities Act." *University of Pittsburgh Law Review*, Vol. 65, pp. 597-653.
- Maestas, Nicole. 2010. "Back to Work: Expectations and Realizations of Work after Retirement." *Journal of Human Resources*, Vol. 45, pp. 718-48.
- Mincer, Jacob. 1974. Schooling, Experience, and Earnings. New York: Columbia University Press.
- Moberly, Michael D. 1994. "Reconsidering the Discriminatory Motive Requirement in ADEA Disparate Treatment Cases." *New Mexico Law Review*, Vol. 24, pp. 89-124.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, Vol. 25, pp. 173-96.
- Neumark, David. "Experimental Evidence on Labor Market Discrimination." Forthcoming in *Journal of Economic Literature*.
- Neumark, David. 2012. "Detecting Evidence of Discrimination in Audit and Correspondence Studies." *Journal of Human Resources*, Vol. 47, pp. 1128-57.
- Neumark, David. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics*, Vol. 111, pp. 915-41.
- Neumark, David, Ian Burn, and Patrick Button. 2015. "Is It Harder for Older Workers to Find Jobs? New and Improved Evidence from a Field Experiment." NBER Working Paper No. 21669.
- Neumark, David, and Joanne Song. 2013. "Do Stronger Age Discrimination Laws Make Social Security Reforms More Effective?" *Journal of Public Economics*, Vol. 108, pp. 1-16.
- Neumark, David, Joanne Song, and Patrick Button. "Does Protecting Older Workers from Discrimination Make It Harder to Get Hired? Evidence from Disability Discrimination Laws." Forthcoming in *Research on Aging*.
- Neumark, David, and Wendy A. Stock. 1999. "Age Discrimination Laws and Labor Market Efficiency." *Journal of Political Economy*, Vol. 107, pp. 1081-125.
- Posner, Richard A. 1995. Aging and Old Age. Chicago: University of Chicago Press.
- Riach, Peter A., and Judith Rich. 2010. "An Experimental Investigation of Age Discrimination in the English Labor Market." *Annals of Economics and Statistics*, No. 99/100, pp. 169-85.
- Riach, Peter A., and Judith Rich. 2006. "An Experimental Investigation of Age Discrimination in the French Labour Market." IZA Discussion Paper No. 2522.
- Rowe, John W., and Robert L. Kahn. 1997. "Successful Aging." *The Gerontologist*, Vol. 37, pp. 433-40.
- Rutledge, Matthew S., Steven A. Sass, and Jorge D. Ramos-Mercado. 2016. "How Does Occupational Access for Older Workers Differ by Education?" Center for Retirement Research at Boston College.
- Starkman, Paul E. 1992. "Alleging a 'Pattern or Practice' under ADEA: An Analysis of the Impact and Problems of Proof." *Labor Lawyer*, Vol. 8, pp. 91-123.
- Sterns, Harvey L., and Suzanne M. Miklos. 1995. "The Aging Worker in a Changing Environment: Organization and Individual Issues." *Journal of Vocational Behavior*, Vol. 47, pp. 248-68.
- Stock, Wendy A., and Kathleen Beegle. 2004. "Employment Protections for Older Workers: Do Disability Discrimination Laws Matter?" *Contemporary Economic Policy*, Vol. 22, pp. 111-26.
- Tilcsik, András. 2011. "Pride and Prejudice: Employment Discrimination against Openly Gay Men in the United States." *American Journal of Sociology*, Vol. 117, 586-626.
- Tinkham, Thomas. 2010. "The Uses and Misuses of Statistical Proof in Age Discrimination Claims." William Mitchell College of Law, Legal Studies Research Paper Series 2010-20.
- U.S. Bureau of the Census. 2016. "Quarterly Workforce Indicators [Release R2015Q2] [computer file]." Ithaca: Cornell University, Labor Dynamics Institute.

U.S. Equal Employment Opportunity Commission. 2002. "Federal Laws Prohibiting Job Discrimination Questions and Answers," available at [www.eeoc.gov/facts/qanda.html](http://www.eeoc.gov/facts/qanda.html), viewed April 13, 2009.

**Table 1: Level of Matching of Callbacks**

	<b>Matched positive responses</b>	<b>No responses</b>	<b>Total</b>
<b>Voicemail</b>	1,614	N.A.	1,614
<b>Email</b>	1,218	N.A.	1,218
<b>Both</b>	438	N.A.	438
<b>All</b>	3,270	11,158	14,428

Notes: There are 3,270 matched responses to 14,428 resumes that were sent out. For responses received from employers, we tried to match each response to a unique job identifier. We received three voicemails that we were unable to match to either a unique job identifier or to the resume that was sent.

**Table 2: State Disability and Age Discrimination Laws, 2016**

State (City)	Age discrimination laws		Disability discrimination laws		
	Minimum firm size	Larger damages than ADEA	Minimum firm size	Larger damages than ADA	Broader (medical) definition of disability
Alabama (Birmingham)	20	No	No law	No law	No law
Alaska (Anchorage)	1	Yes	1	Yes	No
Arizona (Phoenix)	15	No	15	No (no punitive)	No
Arkansas (Little Rock)	No law	No law	9	No (same as ADA)	No
California (Los Angeles)	5	Yes	5	Yes (uncapped)	No (“limits” only)
Colorado (Denver)	1	No	1	No (same as ADA)	No
Connecticut (Hartford)	3	No	3	No (no punitive)	Yes
Delaware(Wilmington)	4	Yes	15	No (same as ADA)	No
Florida (Miami)	15	Yes	15	No (punitive capped at \$100k)	No
Georgia (Atlanta)	1	No	15	No (no punitive)	No
Hawaii (Honolulu)	1	Yes	1	Yes (uncapped)	No
Idaho (Boise)	5	Yes	5	No (punitive capped at \$10k)	No
Illinois (Chicago)	15	Yes	15	No (no punitive)	Yes
Indiana (Indianapolis)	1	No	15	No (no punitive)	No
Iowa (Des Moines)	4	Yes	4	No (no punitive)	No
Kansas (Wichita)	4	Yes	4	No (damages capped at \$2k)	No
Kentucky (Louisville)	8	Yes	15	No (no punitive)	No
Louisiana (New Orleans)	20	Yes	20	No (no punitive)	No
Maine (Portland)	1	Yes	1	Yes	No
Maryland (Baltimore)	15	Yes	15	No (same as ADA)	No
Massachusetts (Boston)	6	Yes	6	Yes (uncapped)	No
Michigan (Detroit)	1	Yes	1	No (no punitive)	No
Minnesota (Minneapolis)	1	Yes	1	No (punitive capped at \$25k)	No (“materially limits” only)
Mississippi (Jackson)	No law	No law	No law	No law	No law
Missouri (Kansas City)	6	Yes	6	Yes (uncapped)	No
Montana (Billings)	1	Yes	1	No (no punitive)	No
Nebraska (Lincoln)	20	No	15	No (no punitive)	No
Nevada (Las Vegas)	15	No	15	Yes	No
New Hampshire (Manchester)	6	Yes	6	No (no punitive)	No
New Jersey (Trenton)	1	Yes	1	Yes (uncapped)	Yes
New Mexico (Albuquerque)	4	Yes	4	No (no punitive)	No
New York (New York)	4	Yes	4	No (no punitive)	Yes
North Carolina (Charlotte)	15	No	15	Yes	No
North Dakota (Bismarck)	1	No	1	No (no damages)	No
Ohio (Columbus)	4	Yes	4	Yes (uncapped)	No
Oklahoma (Oklahoma City)	15	No	15	No (no punitive)	No
Oregon (Portland)	1	Yes	6	Yes (uncapped)	No
Pennsylvania (Pittsburgh)	4	No	4	No (no punitive)	No
Rhode Island (Providence)	4	Yes	4	Yes (uncapped)	No
South Carolina (Columbia)	15	No	15	No (same as ADA)	No

State (City)	Age discrimination laws		Disability discrimination laws		
	Minimum firm size	Larger damages than ADEA	Minimum firm size	Larger damages than ADA	Broader (medical) definition of disability
South Dakota (Sioux Falls)	No law	No law	1	No (no punitive)	No
Tennessee (Memphis)	8	Yes	8	No (no punitive)	No
Texas (Houston)	15	Yes	15	No (same as ADA)	No
Utah (Salt Lake City)	15	No	15	No (no punitive)	No
Vermont (Burlington)	1	Yes	1	Yes (uncapped)	No
Virginia (Virginia Beach)	6	No	1	No (no punitive)	No
Washington (Seattle)	8	Yes	8	No (no punitive)	Yes
West Virginia (Charleston)	12	No	12	Yes (uncapped)	No
Wisconsin (Milwaukee)	1	No	1	No (no damages)	No
Wyoming (Cheyenne)	2	No	2	No (no punitive)	No

Notes: State laws are as of 2016. Age discrimination laws are from Neumark and Song (2013) and disability discrimination laws are from Neumark et al. (forthcoming), but are updated. For the states listed as “Yes” under Larger Damages than ADA, but not uncapped, details are as follows: Alaska – uncapped compensatory damages, punitive damages capped above ADA levels; Maine – exceeds ADA cap for firms of 201+ employees; Nevada – uncapped compensatory damages except against government, punitive damages capped at maximum of \$300k and three times compensatory damages; North Carolina – uncapped compensatory damages except against government, punitive damages capped at maximum of \$250k and three times compensatory damages.

**Table 3: Callback Rates by Age**

		<b>Young (29-31)</b>	<b>Old (64-66)</b>	<b>Absolute (percentage point) difference in callback rate for old</b>	<b>Percent difference in callback rate for old</b>
<i>A. Males (N=7,212)</i>					
<i>Callback (%)</i>	No	75.01	82.61	-7.60	-30.42%
	Yes	24.99	17.39		
<i>Tests of independence (p-value), young vs. old</i>		0.00			
<i>B. Females (N=7,212)</i>					
<i>Callback (%)</i>	No	71.58	80.12	-8.54	-30.05%
	Yes	28.42	19.88		
<i>Tests of independence (p-value), young vs. old</i>		0.00			

Notes: The p-values reported for the tests of independence are from Fisher's exact test (two-sided). There were no positive responses for West Virginia, so it drops out of the probit analysis in subsequent tables. We therefore also drop West Virginia from this table to have results for the same sample; this has virtually no impact on the estimates in this table.

**Table 4: Probit Estimates for Callbacks by Age, Marginal Effects**

	Males		Female	
	(1)	(2)	(3)	(4)
<i>Callback estimates</i>				
Old (64-66)	-0.077*** (0.007)	-0.075*** (0.006)	-0.086*** (0.006)	-0.089*** (0.006)
<i>Controls</i>				
State, order, unemployed, skills	X	X	X	X
Resume features		X		X
<i>Callback rate for young (29-31)</i>	24.99%		28.42%	
<i>N</i>	7,212		7,212	
<i>Clusters</i>	3,607		3,607	

Notes: Marginal effects are reported, computed as the discrete change in the probability associated with the dummy variable, evaluating other variables at their means. Standard errors are clustered at the age-by-state level. Significantly different from zero at 1-percent level (\*\*\*), 5-percent level (\*\*) or 10-percent level (\*). Resume features include: template; email script; email format; script subject, opening, body, and signature; and file name format. See notes to Table 3.

**Table 5: Probit Estimates for Callbacks by Age, with Effects of State Age Antidiscrimination Laws Added, Marginal Effects**

	Males		Female	
	(1)	(2)	(3)	(4)
<i>Callback estimates</i>				
Old (64-66)	-0.063 <sup>***</sup> (0.011)	-0.067 <sup>***</sup> (0.010)	-0.106 <sup>***</sup> (0.011)	-0.085 <sup>***</sup> (0.007)
Old (64-66) x Firm-size cutoff < 10	-0.020 (0.014)		0.028 <sup>**</sup> (0.013)	
Old (64-66) x Larger damages		-0.014 (0.013)		-0.002 (0.010)
<i>Controls</i>				
State, order, unemployed, skills	X	X	X	X
<i>Callback rate for young (29-31)</i>	24.99%		28.42%	
<i>N</i>	7,212		7,212	

Notes: See notes to Tables 3 and 4.

**Table 6: Probit Estimates for Callbacks by Age, with Effects of State Disability Antidiscrimination Laws Added, Marginal Effects**

	Males				Female			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Callback estimates</i>								
Old (64-66)	-0.070 <sup>***</sup> (0.011)	-0.076 <sup>***</sup> (0.009)	-0.070 <sup>***</sup> (0.007)	-0.072 <sup>***</sup> (0.007)	-0.084 <sup>***</sup> (0.010)	-0.092 <sup>***</sup> (0.007)	-0.083 <sup>***</sup> (0.006)	-0.081 <sup>***</sup> (0.006)
Old (64-66) x Firm size < 10	-0.010 (0.014)				-0.003 (0.012)			
Old (64-66) x Larger damages		-0.004 (0.015)				0.022 <sup>**</sup> (0.013)		
Old (64-66) x Broader disability definition (medical only)			-0.034 <sup>**</sup> (0.015)				-0.012 (0.015)	
Old (64-66) x Broader disability definition (medical or limits)				-0.017 (0.016)				-0.017 (0.014)
<i>Controls</i>								
State, order, unemployed, skills	X	X	X	X	X	X	X	X
<i>Callback rate for young (29-31)</i>	24.99%				28.42%			
<i>N</i>	7,212				7,212			

Notes: See notes to Tables 3 and 4.

**Table 7: Probit Estimates for Callbacks by Age, with Effects of State Age and Disability Antidiscrimination Laws Added Together, Marginal Effects**

	Males		Female	
	(1)	(2)	(3)	(4)
<i>Callback estimates</i>				
Old (64-66)	-0.057*** (0.014)	-0.060*** (0.014)	-0.099*** (0.011)	-0.099*** (0.010)
Old (64-66) x Age firm size < 10	-0.047*** (0.011)	-0.035*** (0.012)	0.070*** (0.011)	0.065*** (0.011)
Old (64-66) x Age larger damages	-0.002 (0.014)	-0.007 (0.015)	-0.004 (0.012)	-0.000 (0.012)
Old (64-66) x Disability firm size < 10	0.044*** (0.014)	0.029 (0.016)	-0.059*** (0.012)	-0.051*** (0.012)
Old (64-66) x Disability larger damages	-0.014 (0.014)	-0.005 (0.016)	0.025* (0.014)	0.020* (0.012)
Old (64-66) x Broader disability definition (medical only)	-0.049*** (0.012)		0.014 (0.013)	
Old (64-66) x Broader disability definition (medical or limits)		-0.021 (0.019)		-0.002 (0.014)
<i>Controls</i>				
State, order, unemployed, skills	X	X	X	X
<i>Callback rate for young (29-31)</i>	24.99%		28.42%	
<i>N</i>	7,212		7,212	

Notes: See notes to Tables 3 and 4.

**Table 8: Probit Estimates for Callbacks by Age, with Effects of State Age and Disability Antidiscrimination Laws Added Together, Marginal Effects, with Correction for Bias from Different Variances of Unobservables for Young and Old Applicants**

	Males		Female	
	(1)	(2)	(3)	(4)
<i>Callback estimates (heteroskedastic probit, marginal effect via level)</i>				
Old (64-66)	-0.079** (0.034)	-0.062* (0.035)	-0.124*** (0.024)	-0.127*** (0.024)
Old (64-66) x Age firm size < 10	-0.044*** (0.013)	-0.035** (0.015)	0.079*** (0.014)	0.075*** (0.014)
Old (64-66) x Age larger damages	-0.006 (0.014)	-0.008 (0.015)	-0.008 (0.013)	-0.004 (0.012)
Old (64-66) x Disability firm size < 10	0.044*** (0.015)	0.028 (0.016)	-0.063*** (0.014)	-0.055*** (0.014)
Old (64-66) x Disability larger damages	-0.016 (0.015)	-0.005 (0.016)	0.022* (0.013)	0.017 (0.012)
Old (64-66) x Broader disability definition (medical only)	-0.060*** (0.018)		0.010 (0.013)	
Old (64-66) x Broader disability definition (medical or limits)		-0.022 (0.020)		-0.007 (0.014)
<i>Callback estimates (heteroskedastic probit, marginal effect via variance)</i>				
Old (64-66)	0.024 (0.036)	0.003 (0.036)	0.026 (0.024)	0.030 (0.024)
Overidentification test: ratios of coefficients on skills for old relative to young are equal (p-value, Wald test)	0.797	0.803	0.996	0.997
Standard deviation of unobservables, old/young	1.108	1.013	1.123	1.141
Test: standard vs. heteroscedastic probit (p-value, log-likelihood test)	0.545	0.935	0.405	0.349
<i>Controls</i>				
State, order, unemployed, skills	X	X	X	X
<i>Callback rate for young (29-31)</i>	24.99%		28.42%	
<i>N</i>	7,212		7,212	

Notes: In this table marginal effects are computed as the change in the probability associated with the dummy variable, using the continuous approximation, evaluating other variables at their means; we use the continuous version of the partial derivative, because this version gives an unambiguous decomposition of the estimates from the heteroscedastic probit model (Neumark, 2012). The overidentification test is based on interactions of the skill variables, order of application, and unemployment, with the dummy variable for old. See notes to Tables 3 and 4.

**Table 9: Regression Estimates for Hiring by Age, with Effects of State Age Antidiscrimination Laws Added, QWI Data**

	Males				Female			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Old age group</i>	55-64	65-99	55-64	65-99	55-64	65-99	55-64	65-99
<i>Hiring rate estimates</i>								
Old	-0.065*** (0.002)	-0.060*** (0.003)	-0.065*** (0.002)	-0.060*** (0.002)	-0.082*** (0.003)	-0.081*** (0.003)	-0.080*** (0.003)	-0.080*** (0.003)
Old x Firm-size cutoff < 10	0.006* (0.003)	0.010** (0.005)			0.012*** (0.004)	0.015** (0.007)		
Old x Larger damages			0.006** (0.003)	0.010** (0.004)			0.009** (0.004)	0.014** (0.006)
<i>Controls</i>								
State dummy variables	X	X	X	X	X	X	X	X
Female								
<i>Hiring rate for young (25-34)</i>	13.16%				13.75%			
<i>N</i>	294				294			

Notes: The sample includes hiring rate estimates from the Quarterly Workforce Indicators for 2014Q1, Q2, and Q3 for the retail sales industry (NAICS 44-45). Estimates are specific to four cells: younger men (age 25-34), younger women, older men (age 55-64 or 65-99), and older women. All regressions include state fixed effects and are weighted by the employment level in each age group, gender, and state cell. Standard errors are clustered at the age-by-state level. We drop West Virginia to match our experimental results; the results are almost identical with West Virginia included.

**Table 10A: Regression Estimates for Hiring by Age, with Effects of State Disability Age Antidiscrimination Laws Added, QWI Data, Males**

	Males							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Old age group</i>	55-64	65-99	55-64	65-99	55-64	65-99	55-64	65-99
<i>Hiring rate estimates</i>								
Old	-0.067*** (0.002)	-0.062*** (0.002)	-0.064*** (0.001)	-0.057*** (0.002)	-0.062*** (0.002)	-0.054*** (0.003)	-0.065*** (0.001)	-0.059*** (0.001)
Old x Firm size < 10	0.009*** (0.003)	0.014*** (0.004)						
Old x Larger damages			0.010** (0.004)	0.012* (0.007)				
Old x Broader disability definition (medical only)					0.006** (0.003)	0.007 (0.004)		
Old x Broader disability definition (medical or limits)							0.013*** (0.003)	0.019*** (0.006)
<i>Controls</i>								
State dummy variables	X	X	X	X	X	X	X	X
Female								
<i>Hiring rate for young (25-34)</i>	13.16%							
<i>N</i>	294							

Notes: See notes to Table 9.

**Table 10B: Regression Estimates for Hiring by Age, with Effects of State Disability Age Antidiscrimination Laws Added, QWI Data, Females**

	Female							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Old age group</i>	55-64	65-99	55-64	65-99	55-64	65-99	55-64	65-99
<i>Hiring rate estimates</i>								
Old	-0.084*** (0.003)	-0.083*** (0.003)	-0.079*** (0.002)	-0.077*** (0.003)	-0.077*** (0.003)	-0.073*** (0.005)	-0.081*** (0.002)	-0.079*** (0.002)
Old x Firm size < 10	0.015*** (0.004)	0.020*** (0.006)						
Old x Larger damages			0.015*** (0.005)	0.020** (0.010)				
Old x Broader disability definition (medical only)					0.012*** (0.003)	0.014** (0.006)		
Old x Broader disability definition (medical or limits)							0.021*** (0.004)	0.030*** (0.008)
<i>Controls</i>								
State dummy variables	X	X	X	X	X	X	X	X
<i>Hiring rate for young (25-34)</i>	13.75%							
<i>N</i>	294							

Notes: See notes to Table 9.

**Table 11: Regression Estimates for Hiring by Age, with Effects of State Age and Disability Antidiscrimination Laws Added Together, QWI Data**

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Old age group</i>	55-64	65-99	55-64	65-99	55-64	65-99	55-64	65-99
<i>Hiring rate estimates</i>								
Old	-0.069*** (0.002)	-0.065*** (0.003)	-0.067*** (0.002)	-0.064*** (0.003)	-0.087*** (0.003)	-0.088*** (0.005)	-0.085*** (0.003)	-0.086*** (0.004)
Old x Age firm size < 10	-0.003 (0.003)	-0.002 (0.003)	-0.002 (0.002)	-0.000 (0.003)	0.002 (0.004)	0.003 (0.005)	0.003 (0.003)	0.005 (0.005)
Old x Age larger damages	0.002 (0.002)	0.005 (0.003)	-0.000 (0.002)	0.002 (0.003)	0.002 (0.003)	0.006 (0.005)	-0.001 (0.003)	0.002 (0.004)
Old x Disability firm size < 10	0.008** (0.003)	0.012*** (0.004)	0.004 (0.003)	0.006* (0.003)	0.007* (0.004)	0.009 (0.006)	0.003 (0.004)	0.001 (0.006)
Old x Disability larger damages	0.008** (0.003)	0.008 (0.005)	0.006** (0.002)	0.007* (0.004)	0.012*** (0.004)	0.016** (0.008)	0.009*** (0.002)	0.013** (0.005)
Old x Broader disability definition (medical only)	0.003 (0.003)	0.001 (0.005)			0.009** (0.003)	0.010 (0.006)		
Old x Broader disability definition (medical or limits)			0.010*** (0.003)	0.014*** (0.005)			0.017*** (0.003)	0.025*** (0.006)
<i>Controls</i>								
State dummy variables	X	X	X	X	X	X	X	X
<i>Hiring rate for young (25-34)</i>	13.16%				13.75%			
<i>N</i>	294				294			

Notes: See notes to Table 9.