How Does Online Education Affect Labor Market Prospects? Evidence From A Correspondence Audit Study

Conor Lennon*

March 19, 2019

Abstract

This paper reports the findings of a correspondence audit study that examines how online bachelor's degrees affect labor market outcomes. The study involves sending 1,891 applications for real job openings using 100 fictional applicant profiles. The applicant profiles are representative of recent college graduates, rather than those returning to college or switching careers. Using random assignment to degree type, fictional job applicants who list a traditional (in-person) degree receive nearly twice as many callbacks as those who list an online degree. The paper's findings suggest that, at least currently, completing an online degree program would significantly limit the labor market prospects of typical college students.

1 Introduction

In the Fall of 2016, 6.4 million college students took at least one online class and 3 million took only online classes (Allen et al., 2018).[1] While only a fraction of students in online degree programs fit the typical undergraduate student demographic, the growth of online classes as a medium of instruction naturally raises several questions: what will higher education look like 20 years from now? Will all students take at least some classes online? Will it eventually be the case that students do not take any in-person classes?

*University of Louisville, conor.lennon@louisville.edu. Thanks to Nora Fergany and two upwork.com freelancers (who wish to remain nameless) for their help on this project. Thanks also to the editor and two anonymous reviewers for their comments.

[1] Allen et al. use data from IPEDS, the Integrated Postsecondary Education Data System. IPEDS distinguishes between in-person instruction and "distance education." These distance education classes are typically delivered online as opposed to book- or DVD-based instruction.

Economic theory suggests that the answers to these questions depends, at least in part, on how employers value graduates from online degree programs relative to those from traditional in-person degree programs. Because so little is known about how employers value online degree program graduates, this paper reports the findings from a correspondence audit study that examines how online degrees affect employment prospects.[2] In particular, the findings presented in this paper are based on "callbacks" (interview requests) to 1,891 job applications using 100 unique fictional applicant profiles. The fictional applicant profiles are based on real resumes, gathered from a major online jobs website, for recent college graduates in four broad areas: business, engineering, nursing, and accounting. In each case, the original resume was anonymized by altering names, dates, contact information, addresses, and previous employer and education details. At random, for 50 of the study's fictional resumes, the researcher added the word "online" in parentheses next to the name of the applicant's college or university.[3]

Once the fictional resumes were ready, the researcher identified suitable job openings in each of the four areas and randomly selected one applicant (from the relevant area) to apply for each opening.[4] Then, the researcher monitored interview requests using an online voicemail and email service.[5] The identifying assumption is that any difference in callbacks is caused by differences in the randomly-assigned degree type. The study's design is similar to other correspondence audits including Bertrand and Mullainathan (2004), Kroft et al. (2013), Darolia et al. (2015), and Deming et al. (2016).[6] Section 3 (and Appendix A) provide details about the study's implementation.

The paper's findings suggest that traditional degree holders are almost twice as likely to be contacted for an interview as applicants who report an online degree.[7] Of course, callbacks are not a perfect measure of labor market success, but logic suggests that applicants with more job interviews should face shorter spells of unemployment, more job offers, and higher wages, all else equal. The observed effect on callbacks is significantly larger than the effect Deming et al. (2016)

[2]Gaddis (2018) examines the history of audit studies and explains the terminology, the breath of personal characteristics, and the range of outcomes that can be examined using in-person and correspondence audits.

[3]In all cases, the college/university listed on the resume offered both an online and in-person version of the stated degree program.

[4]Note that the unit of study in this paper is a business/firm. As a result, IRBs at the University of Pittsburgh and the University of Louisville (where the data presented in this study was "gathered") formally declined to review the project as it did not constitute human subjects research. This documentation is available upon request from the author.

[5]Note that requests for interview were politely declined as soon as was feasible.

[6]Baert (2018) surveys 90 correspondence studies from 2005 to 2018 that examine how personal characteristics affect interview requests and hiring decisions.

[7]Similarly to Deming et al. (2016), in this paper, a callback is considered any positive personalized contact. See Appendix A for more details.

find when they implement an audit study to examine the effect of having a degree from a for-profit school. Specifically, they find that “a business bachelor’s degree from a for-profit online institution is 22 percent less likely to receive a callback than one from a non-selective public institution.” Darolia et al. (2015) examine the same issue and find “no evidence that employers prefer applicants with resumes listing a for-profit college relative to those whose resumes list either a community college or no college at all.” For further context, the magnitude of the effect in this study is at least as large as the effect on callbacks in studies examining race (Bertrand and Mullainathan, 2004), age (Lahey, 2008), and extended unemployment duration (Kroft et al., 2013).

In audit study terminology, this study’s design was “unmatched” (Vuolo et al., 2018). That is, for each identified job opening, only one application was submitted. While Vuolo et al. demonstrate that using a matched design (sending pairs of applications to a single opening) is neither necessary nor always the most efficient option, the primary reason for using an unmatched design was to avoid detection. In many correspondence studies, the investigators apply to administrative and clerical positions that do not require a college education. In this study, the job openings are in relatively skilled occupations. Given the limited number of suitable candidates for these positions, the likelihood of detection was a significant concern.

It is important to note that this paper examines outcomes for (fictional) graduates who earn degrees at established four-year universities and colleges. Many well-known schools, such as Arizona State University, the Ohio State University, Penn State University, and Northeastern University, offer dozens of completely online degrees. Only a fraction of the students currently enrolled in such programs are traditional college-age students who choose to do a degree online rather than in-person. However, this paper examines what the consequences might be for traditional college-age students who choose that route. As a consequence, the comparison of interest in this study is not between graduates from traditional universities and for-profit schools like the University of Phoenix or DeVry. For-profit schools do not primarily serve the population on which the paper is focused: young adults who attend college shortly after graduating from high school (see Darolia et al., 2015 and Deming et al., 2016). In addition, the paper does not examine the benefits of online degree programs for non-traditional students, such as those who are retraining after several years in the workforce. For many of those students the choice is between an online degree and no degree rather than between a traditional degree and an online degree.

This paper complements the literature on learning outcomes for typical undergraduate students in online settings. Bennett et al. (2007), Ary and Brune (2011), Hernandez-Julian and Peters (2012), Figlio et al. (2013), and Bowen et al. (2014) are just a few examples. Generally, researchers have found that student learning (measured in a variety of ways) is affected mildly or not at all by the medium of instruction. A notable exception to this pattern is Alpert et al. (2016) who find that performance on a cumulative microeconomics final exam was 5 to 10 points (out of 100) lower for students randomly assigned to the online section. While it is important to understand how online instruction affects learning, the literature has either ignored the effect online education might have on labor market outcomes or implicitly assumed that any effect is captured by differences in human capital accumulation. However, such an approach ignores education's role as a signal of productivity (Spence, 1973). Indeed, if there are no human capital accumulation differences between online and in-person degree programs, then fewer callbacks for those with online degrees would strongly support the idea that employers view a traditional degree as a better signal of productivity.

Section 2 examines the relative merits of online instruction along with the benefits and drawbacks of using correspondence studies. Section 3 describes the correspondence study procedure in detail. Section 4 summarizes the data and checks on the experimental randomization. Section 5 presents the main estimates and examines heterogeneity by perceived race, gender, resume quality, and career area. Section 6 concludes.

2 Online Instruction and the Value of Correspondence Studies

Lack (2013) provides a review of the available research on the learning outcomes associated with online instruction. The review details large and small studies in several areas including sociology (Driscoll et al., 2012), accounting (Rich and Dereshiwsky, 2011), management (Daymount and Blau, 2008), and engineering (Enriquez, 2010). Lack's review finds no evidence that students, controlling for observable characteristics, learn less effectively when the medium of instruction is online rather than in-person. Unfortunately, the conclusions that can be drawn from these studies are complicated by differences in research methods, subject attrition, treatment and control group cross-contamination, small sample sizes, different populations of interest, along with each study having a unique institutional setting and time-frame.

In more recent years, researchers have used larger sample sizes and controlled variation to examine how online instruction affects learning. Figlio et al. (2013) examine the effects of watching online rather than attending introductory economics lectures at a large selective research institution in the United States. The authors observe that regardless of sex or race, average test scores were higher for those who were randomly assigned to "live lectures." However, the effects were modest (between 1.9 and 3 points out of 100) and not always statistically different from zero. Bowen et al. (2014) focus on learning outcomes in statistics classes but allow for the online instruction to be augmented by an interactive learning platform. They find that students "are not harmed by this mode of instruction in terms of pass rates, final exam scores, and performance on a standardized assessment of statistical literacy." Alpert et al. (2016) also find that students in this kind of "hybrid" class do as well as those in classes that follow a traditional twice-weekly meeting schedule.

While the research on learning outcomes suggests that online coursework might be a valid alternative to traditional in-person instruction, no study has been able to examine how a purely online degree program affects learning outcomes relative to a traditional degree.[8] More problematically, the existing work on this topic implicitly assumes that the right metric for judging the success or failure of online coursework is human capital accumulation as measured by performance on assignments and examinations. For example, Figlio et al. suggest "[i]nternet-based classes may even dominate live-lecture classes, as they offer students more flexibility in the timing of attendance as well as the opportunity to review lectures to clear up confusing points." Of course, labor market outcomes would be impossible to convincingly relate to a change in the mode of instruction of a single college course. However, that doesn't mean labor market outcomes should be ignored.[9] The lack of discussion and research on the impact of online education on labor market outcomes motivates this paper. The fact that information on the medium of instruction is not recorded by labor market surveys means that a correspondence study is likely the only effective way to causally relate the attitudes of real employers towards potential employees with degrees that are earned online.

[8]To be fair, to validly examine the outcomes from entirely online programs versus traditional degrees, experimental variation at a higher level would be required. However, volunteer subjects would likely form a small and unique group and students who were randomly assigned into online degrees may take other steps to mitigate its effects over a period of four or more years. In any case, such a study, given the significance of the intervention it entails, would be unlikely to obtain approval from an Institutional Review Board.

[9]Rechlin and Kraiger (2012) appear to be the only authors who have considered this issue. However, their study examines the attitudes of just 23 employed Industrial-Organization (I-O) Psychologists towards job applicants who have completed a psychology degree online and demand effects bias their survey instrument.

More generally, correspondence studies are a reliable solution when crucial information available to employers is not available to or cannot be controlled by researchers. Bertrand and Mullainathan (2004) provide the ideal example of the value and purpose of such studies. These authors examine if employers screen resumes using indicators for race (such as names). Their paper's title "Are Emily and Greg More Employable than Lakisha and Jamal?" provides the clearest illustration of their approach. Focusing on low-skilled positions, Bertrand and Mullainathan find that white-sounding names received 50 percent more callbacks for interview, all else equal. While not the main focus, this paper's audit study can extend Bertrand and Mullainathan's findings to positions for which a bachelor's degree is required because 50 of the fictional resumes in this study featured ethnic applicant names (see Appendix A).[10]

This paper's approach and research question are complementary to Darolia et al. (2015) and Deming et al. (2016). Each examines how for-profit degrees (employers are not informed that many of these programs are delivered online) are viewed by employers using a correspondence study. Darolia et al.'s findings suggest that a for-profit college education is no better (in terms of receiving callbacks for interview) than community college or no college at all. Deming and his co-authors find that for-profit degrees receive slightly fewer callbacks than degrees from non-selective public institutions. Both studies examine low-prestige institutions, comparing public non-profit to private for-profit degree programs. Moreover, neither isolate the effect of online instruction itself. This paper is different in that it focuses on relatively selective, established schools that offer both in-person and online versions of the same degree. In addition, the resumes used to apply for positions are based on the resumes of real recent graduates ensuring that the paper's findings are relevant for typical college-age students.

Lastly, this paper makes a contribution to the literature on the effects of college "quality" on students' labor market outcomes. Dale and Krueger (2014) provide a detailed analysis of that literature and find that selectivity, as proxied by undergraduate acceptance rates, positively affects labor market outcomes. However, the effect on labor market outcomes dissipates after controlling for student selection into better schools. While research on this topic is usually based on differences

[10]As a preview, estimates suggest a difference in callback rates between races persists for higher-skilled positions but the difference is not as large as Bertrand and Mullainathan estimated for low-skilled positions (see Section 5). However, these findings should not be viewed as causal because the selection of black, white, and hispanic names was based on convenience rather than any attempt to have a representative set of names. In particular, the names chosen came from a mix of babycenter.com lists of popular names by race, the list of "whitest" and "blackest" names in Freakonomics (Levitt and Dubner, 2005) and a list of the Top 400 1990s names provided by the Social Security Administration (see https://www.ssa.gov/oact/babynames/decades/names1990s.html).

between schools, one can view this paper as examining within-school variation in reputation created by offering online degree programs.

2.1 The Limits of Correspondence Studies

Correspondence studies are an excellent way to uncover the attitudes of employers towards specific employee characteristics. However, a number of caveats apply. First, callbacks for interviews do not pay bills, and it is not clear from these studies that fewer callbacks actually translates to lower wages and higher unemployment. Instead, information transmitted to employers via the resume may improve matching and therefore could reduce wasteful and unnecessary interviews that would not result in a job offer anyways. In addition, the revelation of some characteristics perhaps allows an applicant to gain access to an interview they would not have otherwise earned that ultimately leads to a preferred job offer. Some employers may seek black, female, older, or homosexual workers, or might empathize with those who faced a spell of unemployment. Essentially, certain characteristics might reduce the total number of callbacks for a given applicant but increase the probability of getting the "right" callback.

Second, employers may fail to notice the experimentally-varied characteristic. Thus, the estimated effect of a particular characteristic on callbacks would represent a lower bound on the true effect. On a related note, it is possible that resumes are positively or negatively filtered by software using trigger words, which could bias findings in either direction. This is less of a concern in this paper because the word "online" is the only variation introduced and is likely too common to base any filter upon.

Third, applying for jobs posted in newspapers and online is only one way to secure employment. Social networks and connections, internships, and personal recommendations may compensate for or exacerbate the effects seen in correspondence studies. An individual who fares poorly in a correspondence study may be able to improve their job prospects via alternative approaches to job search.

Last, correspondence studies cannot be sure that their experimental variation does not interact with employers' experience. Take Bertrand and Mullainathan's study as an example. The paper claims to study the effect of having a black-sounding name compared to an identical resume with a white-sounding name. However, to be strict, their paper studies the effect of having a black-sounding name, reporting it without alteration (Jamal Jones could easily present himself on his resume as Jay or J. Jones) *and* having a resume that does not reflect changes that an employer may *expect* to see given that variation. That is, non-fictional white males and black males might present very different resumes even if they had similar work histories and education. If resumes from otherwise similar whites and blacks are systematically different in the population, those differences are part of the experimental variation. In such a case, the effect reported in the paper is the combined effect of having a black-sounding name but having a resume that doesn't seem like other resumes from black applicants. This example is not chosen at random. Bertrand and Mullainathan find that white applicants experience a much higher return to increased resume quality, which suggests that employers may be skeptical of high-quality black resumes.

Kroft et al. (2013)'s audit study on the effects of unemployment duration on callbacks is subject to a similar critique. The authors are identifying not just the effect of unemployment duration but the combined effect of being unemployed and being foolhardy enough to not have a good explanation (even if contrived or completely fabricated and even if the employer knows that it is fabricated) for the spell of unemployment.

This paper is subject to similar unavoidable critiques. Specifically, the paper's estimates should be interpreted as the impact on callbacks for interview from having an online degree *and* telling the employer about it. Note that while telling an employer that an applicant has a degree earned online may seem contrived, a 2010 survey by the Society of Human Resource Managers found that only 17% of human resource professionals had never seen an applicant clearly indicate an online degree.[11] This does not mean that all job applicants who have an online degree always choose to reveal that information to prospective employers. However, it eases concerns that the resumes used in this study will stand out as extremely unique and/or unusual. The main concern would be that unique resumes might bias employers in ways that are unrelated to the their opinion of the capability of the applicant.

In addition, even if an applicant does not mention the online nature of their education in the resume, the issue could come up during an interview. Consider an applicant who lists work experience coincident with their college degree in another state. Alternatively, the fact a degree was completed online likely will arise when the candidate is asked to answer location-specific questions at interview. This means that while the effects of unemployment duration on callbacks

[11] See http://www.shrm.org/research/surveyfindings/articles/pages/hiringpracticesandattitudes.aspx. In the years since that survey, the number of people completing online classes and degree programs has risen considerably. See http://nces.ed.gov/pubs2014/2014023.pdfandhttps://nces.ed.gov/fastfacts/display.asp?id=80 for more details.

and eventual employment may perhaps be eliminated with a one-sentence explanation, the effect of online degrees on labor market success is less avoidable. Alternatively, mimicking the concerns with Bertrand and Mullainathan's approach, employers may expect applicants with online degrees to take steps to compensate for their perceived deficiency via improvements in other areas. These issues are considered in Section 5 by examining how the effect of online education varies with respect to compensating factors such as GPA and work experience. Intuitively, the idea is that if employers are expecting factors that compensate for having an online degree - more work experience or a higher GPA - that they do not see, the returns to these factors will differ by degree type.

These methodological caveats do not invalidate the correspondence study method nor the causal relationships uncovered. Instead, they should be viewed as qualifications that delineate what is being explained and how it might be interpreted differently under alternative circumstances.

3 Experimental Design

The procedure to generate resumes is similar across correspondence studies. For authenticity, the researcher creates a pool of resumes using the resumes of real job-seekers posted publicly on job hunting websites. The real resumes are deconstructed, anonymized, and reconstructed manually or via a computer program (see Lahey and Beasley, 2009 and Lahey and Beasley, 2018). Then, they are randomly assigned one of N possible variations in a characteristic of interest. The fictional resumes are used to apply for real jobs and the researcher tracks "callbacks" (interview requests). Because the randomization is orthogonal, by construction, to other characteristics, differences in callbacks can be considered causally-related to the associated experimental variation.

This paper uses a similar approach where the medium of instruction for the applicant's education is the randomly-assigned characteristic. However, the study differs from others in an important dimension. Many correspondence studies tend to focus on clerical, retail, and administrative openings to ensure they can apply to many job openings with multiple resumes. This study is focused on differences in callbacks for those who have bachelor's degrees. As a result, entry-level clerical and retail jobs are not realistic options. Instead, the paper focuses on several early-career positions suitable for recent graduates in the business, engineering, IT, and medical professions. These positions represent the types of jobs associated with degree programs offered online and in-person at many institutions, a bachelor's degree in these fields is linked to employment in a particular well-defined job (such as software engineer, nurse, accountant, or business analyst as opposed to english, history, and sociology), and there are typically lots of jobs advertised in these fields.[12] A disclaimer that the findings may not generalize to other situations and professions applies.

Because the details are similar across correspondence studies, Appendix A explains how the researcher created the pool of 100 resumes and cover letters used in the study. It also explains how these fictional applicants were randomly assigned to "online degree" status, how and when applications were completed, and how callbacks for interview were handled.

4 Data and Estimation

4.1 Data

Table 1 presents overall summary statistics on the demographic characteristics of the (fictional) applicants. Then, the table presents sample characteristics stratified by degree assignment (online or traditional) to examine how the randomization actually fared. Despite the randomization, those who are assigned to have a traditional degree are slightly more likely to be male, have less work experience, and attended a less selective college (as measured using U.S. News undergraduate admission rates - higher numbers indicate less selectivity).[13] The table then shows the same summary statistics stratified by both race and gender. There are several clear differences in the demographic characteristics and callback rates of the various groups. For example, among the resumes that reflect a black name, the proportion of men is larger than in the other ethnic groups. These imbalances illustrate the importance of controlling for observable resume characteristics in regression estimates.

In addition, the table shows that there was an overall callback rate of 12.2% (231 callbacks from 1,891 applications). In contrast, Kroft et al. have a callback rate of just 4.7% while Bertrand and

[12] As of mid-2018, U.S. News ranks online programs in Accounting, Business Administration and Management, Business Technology Management, Communication, Computer Science, Criminal Justice, Cybersecurity, Dietetics, Early Childhood Education, Electrical Engineering, Elementary Education, Engineering, Environmental Science, Family and Human Development, Finance, Graphic Design, Health Care Administration and Management, Health Science, Homeland Security, Information Technology, Interior Design, Liberal Studies, Marketing, Network Administration, Nursing, Paralegal Studies, Political Science, Psychology, Public Safety Administration, and Special Education. See http://www.usnews.com/education/online-education/bachelors.

[13] See http://colleges.usnews.rankingsandreviews.com/best-colleges.

Table 1: Summary Statistics

All	Statistic	Callback Proportion	Prop. Male	Grade Point Average (GPA)	Selectivity	Years of Experience
Entire Sample	Mean	0.122	0.51	3.36	0.575	1.71
(N=100, n=1,891)	Std. Dev.	(0.33)		(0.50)	(0.20)	(0.77)
By Type of Education	Statistic	Callback Proportion	Prop. Male	GPA	Selectivity	Years of Experience
Traditional Degree	Mean	0.159	0.53	3.38	0.583	1.69
(N=50, n=975)	St. Dev.	(0.37)		(0.33)	(0.20)	(0.76)
Online Degree	Mean	0.083	0.47	3.34	0.567	1.73
(N=50, n=916)	St. Dev.	(0.28)		(0.33)	(0.20)	(0.79)
By Race	Statistic	Callback Proportion	Prop. Male	GPA	Selectivity	Years of Experience
Caucasian	Mean	0.138	0.48	3.45	0.58	1.56
(N=50, n=1,054)	St. Dev.	(0.34)		(0.22)	(0.20)	(0.66)
Black	Mean	0.09	0.64	3.20	0.537	1.72
(N=25, n=410)	St. Dev.	(0.29)		(0.40)	(0.20)	(0.87)
Hispanic	Mean	0.115	0.44	3.27	0.601	2.07
(N=25, n=427)	St. Dev.	(0.32)		(0.42)	(0.20)	(0.80)
By Gender	Statistic	Callback Proportion	Prop. Male	GPA	Selectivity	Years of Experience
Male	Mean	0.121	1	3.39	0.577	1.77
(N=51, n=927)	St. Dev.	(0.26)		(0.33)	(0.20)	(0.78)
Female	Mean	0.122	0	3.33	0.573	1.65
(N=49, n=964)	St. Dev.	(0.37)		(0.327)	(0.20)	(0.76)

In the table“N” refers to the number of resumes/profiles. The small “n” refers to the number of applications completed using those N resumes. Selectivity is the only piece of information not provided by resumes. It is measured using U.S. News Undergraduate Admission Rates. Higher values indicate less selective institutions.

Mullainathan’s callback rate is 8.05%. Kroft et al.’s extremely low callback rate is likely because their resumes portrayed unemployed applicants. The higher callback rate in this study is likely also due to design choices. For example, fictional applicants in this study are particularly well-matched to available positions, have a resume that reflects relevant experience, generally possess high GPAs, and have a sharp and succinct cover letter. In addition, job openings were less than 48 hours “old” at the time of application. Resumes reflecting quality candidates should generate more callbacks and ensure adequate statistical power. In contrast, other correspondence audits achieve sufficient power by applying for many low-skilled positions, often sending several resumes to a single employer (see Vuolo et al., 2016 for more on the difficulty of calculating statistical power in correspondence study settings).

4.2 Estimation

As the assignment to an online degree is random, the estimate of δ from a regression of the following form can be viewed as the difference in callback probability between applicants who earn their degree online rather than at a traditional/in-person degree program:

$$y_{i,k} = \beta X_i + \delta D_i + \epsilon_i \tag{1}$$

In equation (1), $y_{i,k}$ takes on the value of 1 if employer k responds positively to applicant i's application. The employer's decision is assumed to be related to the fictional applicant-specific characteristics X_i and an indicator D_i that equals 1 if the resume indicates that individual i's degree was earned online. In this paper X_i includes GPA, years of work experience, a measure of college selectivity (undergraduate selectivity - as used by Dale and Krueger), gender, race (as indicated by name similarly to Bertrand and Mullainathan), and a binary indicator for the applicant's industry (business, engineering, nursing, and so on). The coefficient associated with D_i, $\hat{\delta}$, is the main estimate of interest. In particular, a negative $\hat{\delta}$ would suggest the likelihood of getting a callback is lower for online degree holders, even after accounting for other factors. Such an empirical approach would not be feasible using labor market survey data due to concerns about endogeneity and omitted variable bias. These concerns cannot be driving the paper's findings as the randomization of D_i avoids that problem by construction.

The paper's "unmatched" design ensures that $\hat{\delta}$ is the difference in the mean callback rate between the group of people randomly assigned to have an online degree and those assigned to have a traditional degree. The unmatched design means that differences between groups could be driving any observed effects *if* the randomization were to "fail" in some way. For that reason, it is important to include the objective characteristics of each resume as controls in regression estimates.

5 Findings

Table 2 reports the paper's main findings. The estimates in the table are post-estimation marginal effects from a probit regression with standard errors clustered at the applicant level.[14] The dependent variable is whether or not an application generated a callback (Callback=1 when a request for interview was received). The table reports three specifications with additional controls added sequentially. The preferred specification in the final column includes controls for all available covariates: race, sex, experience, college selectivity, career/field, and GPA.

The effect of having an online degree is large and negative in all specifications. The coefficients in the table should be considered percentage point differences. Specifically, the estimates in the final column suggest that there is a 7.3 percentage point difference in callback rates between traditional and online degree holders, all else equal.[15] Given that the mean callback rate for online degree holders is 8.3 percent, a 7.3 percentage point difference suggests that a resume reflecting a traditional degree will receive almost twice as many callbacks for interview as a resume reporting an online degree, all else equal. For context, Bertrand and Mullainathan found whites were about 1.5 times more likely to receive a callback for interview as blacks, and Kroft et al. found that someone who is just one month unemployed is about 1.8 times more likely to receive a callback for interview than someone who has been unemployed for eight months. On the other hand, Darolia et al. (2015) and Deming et al. (2016) find little to no effect on callbacks when comparing for-profit degree holders to those with a community college or only a high school education.[16]

The coefficients on the covariates in Table 2 tend to follow the stylized facts of the labor market. African and Hispanic Americans fare slightly worse than Caucasians but the difference in callback rates is not statistically significant. In particular, the effect of an African-American name is mild in each specification. Ethnicity is indicated only via name (see Appendix A) and it is possible that Latin American names were more identifiable to employers. Unsurprisingly, years of experience

[14]The audit procedure creates an unconventional panel data-set: there are repeated observations for each job "applicant" but no time component. Random effects estimates are presented in Appendix B. They differ only marginally from the pooled-sample estimates presented here. A fixed-effects specification is not feasible as, for each fictional resume, the independent variables do not change.

[15]Of course, there are several issues with computing marginal effects when estimation involves a number of indicator variables. The main issue is that the procedure crudely considers the effect of the variable of interest at the average of variables that have no such interpretation. For example, the process sets the value of "gender" to its average value ($\approx$0.5) in the data. The raw probit estimates are available from the author upon request.

[16]The issue of how educational prestige affects labor market outcomes also interests sociologists. For example, Deterding and Pedulla (2016) focus on the shifting landscape in U.S. higher education and find that "employers responded similarly to applicants listing a degree from a fictional college and applicants listing a local for-profit or nonprofit institution."

Table 2: Callback Rate - Pooled-Probit Estimates

	(1) Callback	(2) Callback	(3) Callback
Online	-7.6005***	-6.7869***	-7.3114***
	(2.4670)	(2.1803)	(1.8886)
Male		0.1539	1.0081
		(2.3790)	(1.8955)
Black		-1.6354	0.3697
		(2.2319)	(2.1736)
Hispanic		-1.3356	-3.5182
		(2.8675)	(2.5662)
Grade Point Average		12.8895***	12.8974***
		(3.1807)	(3.2645)
Experience (Years)		2.1141	2.5627*
		(1.6925)	(1.4360)
Selectivity (U.S. News Acceptance Rate)			-0.1986***
			(0.0606)
Observations	1,891	1,891	1,891
No. of Applicant Profiles	100	100	100
Individual Characteristics	N	Y	Y
Career and School Characteristics	N	N	Y

Standard errors are clustered at the applicant-level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table reports marginal effects from pooled-probit estimations with controls added sequentially. The coefficient estimates on career/field fixed effects are not reported to economize on space. The coefficients can be interpreted as percentage point differences in callback rates for a one unit change in the variable of interest.

is positively associated with callback likelihood. Selectivity is measured using undergraduate acceptance rates from U.S. News and the negative coefficient implies attending a selective school matters. For each one percentage point decrease in selectivity, there is a .2 percentage point drop in callbacks. Higher GPAs are also associated with higher callback rates: the estimates suggest that going from a reported 3.0 GPA to a 3.1 GPA is associated with a 1.3 percentage point increase in callbacks. Lastly, there is no statistically significant difference between male and female callback rates.

Table 3: Pooled-Probit Estimates - By Subgroup

	(1) Callback Females Only	(2) Callback Males Only	(3) Callback Caucasians Only	(4) Callback Minorities Only
Online	-4.4671*	-9.3566***	-11.3387***	-3.0097
	(2.4440)	(2.5105)	(2.6966)	(2.3910)
Male			-0.2307	2.5549
			(2.5146)	(2.0827)
Black	-3.0739	1.6074		
	(2.9837)	(2.9569)		
Hispanic	-5.1300*	-2.3238		
	(2.7074)	(4.7602)		
Grade Point Average	11.9997**	11.6436***	13.4993**	13.1791***
	(4.7357)	(4.3624)	(5.5316)	(2.8524)
Experience (Years)	3.9105**	1.0277	0.6573	2.4714*
	(1.7001)	(2.0886)	(2.3394)	(1.3214)
Selectivity (Acceptance Rate)	-0.1290*	-0.2370**	-0.3219***	-0.0656
	(0.0708)	(0.0956)	(0.0801)	(0.0493)
Observations	927	964	1,054	837
No. of Applicant Profiles	49	51	50	50
Individual Characteristics	Y	Y	Y	Y
Career and School Characteristics	Y	Y	Y	Y

Standard errors are clustered at the applicant-level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table reports the marginal effects from pooled-probit estimations. The coefficient estimates on career/field fixed effects are not reported to economize on space. The coefficients can be interpreted as percentage point differences in callback rates for a one unit change in the variable of interest.

Table 3 presents estimates separately for females, males, whites, and minorities. Each specification clusters standard errors at the applicant level and includes all available covariates. The negative effect on callbacks is smaller for females and minorities who earn an online degree relative to males and Caucasians. The estimated effect for minorities appears small and is statistically insignificant. However, the difference is important because the overall callback rate is lower for minorities regardless of type of degree conveyed to the employer (see Table 1). The disparity between male and female callback differences is unexpected. Upon further examination, there are two fictional female nurses in the sample, who were randomly-assigned an online degree, who have two full years of relevant experience. Both of these fictional applicants had a callback rate of

Table 4: Pooled-Probit Estimates - By Field/Profession

Marginal Effect on Callbacks for:	(1) Callback Business/Accounting	(2) Callback Engineering	(3) Callback Nursing
Online	-7.5127**	-6.3471***	-1.1830
	(3.3295)	(1.7095)	(4.5379)
Male	1.4764	1.2731	-5.7055*
	(3.3140)	(1.9083)	(3.2254)
Black	5.4229	-0.8200	-14.9417***
	(3.9352)	(1.9957)	(5.7588)
Hispanic	-2.6445	19.5812***	-13.1596***
	(3.3556)	(6.1445)	(4.8719)
Grade Point Average	12.0115**	6.3524**	22.4570***
	(5.0768)	(3.1961)	(7.4622)
Experience (Yrs)	1.2797	0.8296	7.4955*
	(2.1499)	(0.6795)	(4.1476)
Selectivity	-0.2441***	-0.0872**	0.0343
	(0.0786)	(0.0386)	(0.1745)
Observations	849	585	457
No. of Applicant Profiles	47	33	20
Individual Characteristics	Y	Y	Y
Career and School Characteristics	Y	Y	Y

Standard errors are clustered at the applicant-level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table reports the marginal effects from pooled-probit estimations. The coefficient estimates on career/field fixed effects are not reported to economize on space. The coefficients can be interpreted as percentage point differences in callback rates for a one unit change in the variable of interest.

more than 25%. That is a higher callback rate than any other fictional candidate male or female, in any profession, regardless of online or in-person degree type. Without those two well-qualified applicants, the disparity between callback rates would be larger both overall and for females, in particular. These two fictional candidates are also driving the large and statistically significant co-efficient on experience in the first column of Table 3. As Deming et al. (2016) note, nursing typically requires an occupational license that diminishes the role of academic qualifications in the screening process, which is true for online nursing degrees too

In Table 4, estimates are presented separately for business applicants (accountants and business analysts), engineers (software and mechanical), and nurses. Because of the high callback rate for those two nurses assigned an online degree, nurses with online degrees appear to experience no difference in callbacks. Additionally, these two applicants happen to convey a Caucasian name and that is driving the large negative coefficient for minority nurses. A larger sample would be less sensitive to such issues.

5.1 Returns to Resume Characteristics

As mentioned in Section 2, a potential concern with audit studies is that findings are driven not only by the experimental variation but also by how differences in the experimental resumes compare to the differences employers expect to see in real resumes. Take Bertrand and Mullainathan's finding that Black applicants receive fewer callbacks. If otherwise-identical Black and Caucasian resumes are interpreted differently, because employers have learned that one of the groups tends to present the same experience and skills differently, then the effects Bertrand and Mullainathan find could just be an artifact of a labor market norm they (and perhaps employers, at least subconsciously) are unaware of. That is, perhaps employers are not calling Black applicants with a given resume quality because they expect such applicants to "oversell" their experience and abilities in order to combat expected discrimination. In that case, the signal of ability the employer takes from each resume is different, not because the employer is discriminating, but because of the actions of other applicants. This is something that a researcher cannot control.

A similar concern arises in this paper. An employer may expect a person who has an online degree to report other compensating characteristics. When they do *not* see this, they infer something about the candidate's ability that the researcher is not controlling for. Essentially, the researcher is holding all else equal but the changes should not result in all else remaining equal. Empirically, these kinds of concerns should lead to different "returns" to aspects of resume quality for online degree holders, all else equal. For that reason, Table 5 reports post-estimation marginal effects from specifications where the main treatment variable is interacted with applicant characteristics. Specifically, Table 5 reports estimates from the following estimating equation;

$$y_{i,k} = \beta X_i + \lambda D_i + \gamma Characteristic_i + \delta D_i \times Characteristic_i + \epsilon_i \qquad (2)$$

In equation (2), $Characteristic_i$ is a placeholder for sex, GPA, experience, and race for person i and the other terms are as described in equation (1). The coefficient of interest is the interaction

Table 5: Pooled-Probit Estimates - Returns to Resume Characteristics

Marginal Effect on Callbacks for:	(1) Callback Male	(2) Callback Experience	(3) Callback GPA	(4) Callback Black	(5) Callback Hispanic
Online	-1.2429	2.4531	4.1134	-0.5348	1.1171
	(1.9736)	(2.5495)	(2.8534)	(3.5691)	(2.6064)
In-Person	3.403	2.6899**	21.8596***	1.7177	-7.7075*
	(3.3319)	(1.2993)	(4.9221)	(2.3708)	(4.2464)
Observations	1,891	1,891	1,891	1,891	1,891
No. of Applicant Profiles	100	100	100	100	100
Individual Characteristics	Y	Y	Y	Y	Y
Career and School Characteristics	Y	Y	Y	Y	Y

Standard errors are clustered at the applicant-level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table reports the marginal effects from pooled-probit estimations. The coefficients can be interpreted as percentage point differences in callback rates for a one unit change in the variable of interest.

between having an online degree ($D_i = 1$) and $Characteristic_i$. Table 5 presents only the marginal effects for the terms of interest. In particular, the table reports the marginal effect of a particular characteristic on callbacks by degree type. In the first column of the table the focus is on males versus females. The first estimate suggests that males with an online degree are slightly less likely to be called back than females. However, the estimate is not statistically different from zero. The second estimate suggests that males with an in-person degree are slightly more likely to be called back. Again, however, the estimate is not statistically different from zero. The estimates in the remaining columns should be interpreted similarly. In the second column, the estimates suggest that experience helps both types of degree holders (but the effect is statistically different from zero only for in-person degree holders). In the third column of the table, GPA matters significantly but only for in-person degree holders. Put another way, if you earn an online degree, even a 4.0 GPA won't help all that much. This estimate is essentially a confirmation of the main takeaway of this paper: employers currently do not appear to trust or value online education. In the final two columns, degree type does not matter much for Black workers but does for Hispanic workers. In particular, an online degree does not seem to affect the probability of callback for Hispanic applicants (the omitted category is Caucasian applicants). However, the estimates suggest that Hispanic applicants with traditional in-person degrees will receive fewer callbacks than Caucasian workers with traditional in-person degrees.

In estimates not presented here, the interaction of college selectivity with having an online degree is essentially zero both in an absolute and in a statistical sense. This finding suggests that the effect of an online degree does not vary as a function of the measure of selectivity chosen for this paper (U.S. News undergraduate acceptance rates). That is, while those who attend a selective school fare better than those who do not (see Table 2) the relative negative effect of an online degree from such a school is neither more nor less pronounced than from any other school. This is perhaps a consequence of the mild variation in acceptance rates among the schools examined. Only schools that offer the same degree online and in-person appear on resumes in this correspondence study, and these kind of schools are neither the most nor least prestigious.[17]

6 Conclusion

Economic theory suggests that students will switch to online degree programs only if the benefits of doing so are greater than the cost. If employers do not value such credentials, then these programs may provide few benefits. However, little is known about how online degree programs affect labor market prospects. For that reason, this paper reports the findings from a correspondence audit study that tests how employers view degrees earned online. The paper's findings show that traditional degree holders are almost twice as likely to be contacted for an interview as applicants who report an online degree. An important caveat is that some of the estimates suggest that callback rates will be lower for those who pursue an online education *and* do not take any steps to counteract that decision. Despite such a caveat, the paper's findings clearly show that employers currently do not find applicants with online degrees to be as attractive as those with traditional degrees. The obvious take-away for job applicants (who earn a degree online) is not to inform employers about the medium of instruction. However, that is only helpful to online degree holders' job prospects if employers would not find out and react to this information later in the hiring process. A similar argument applies to many other correspondence studies: for example, the takeaway from Kroft et al. (2013), who study how spells of unemployment affect job prospects, would be for applicants to disguise spells of unemployment on their resume.

It is important to note that this study compares fictional job applicants with traditional and online degrees earned at well-known schools. If online education is pursued solely by those

[17] In addition, the US News-reported undergraduate acceptance rate may not fully-capture the actual prestige differences among these schools.

who would never earn a traditional degree, then the findings are somewhat moot. However, given the growth of online education reported by Allen et al. (2018), the paper's findings should interest students, professors, and administrators because the estimates confirm traditional modes of education are still viewed as superior to online education from an employer viewpoint.

Lastly, it is not clear what aspect of a traditional college education employers are responding favorably towards. They may believe human capital formation is diminished in online programs relative to traditional degrees (even if it is not), they may believe the individual will be less socially adept, are inferring some socioeconomic characteristics, or they may feel a traditional college education gives students something more than just grades written on a piece of paper. This paper asks only *if* labor market outcomes are diminished for those who earn online degrees. Understanding exactly why students with online degrees fare poorly in the labor market is a topic for future work.

References

Allen, I. E., Seaman, J., and Seaman, J. E. (2018). Grade Increase: Tracking Distance Education in the United States. Technical report, Babson Survey Research Group.

Alpert, W. T., Couch, K. A., and Harmon, O. R. (2016). A randomized assessment of online learning. *American Economic Review Papers and Proceedings*, 106:378–82.

Ary, E. J. and Brune, C. W. (2011). A comparison of student learning outcomes in traditional and online personal finance courses. *MERLOT Journal of Online Learning and Teaching*, 7(4):465–474.

Baert, S. (2018). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In Gaddis, S. M., editor, *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Springer International Publishing AG.

Bennett, D. S., Padgham, G. L., McCarty, C. S., and Carter, M. S. (2007). Teaching principles of economics: Internet vs. traditional classroom instruction. *Journal of Economics and Economic Education Research*, 8(1):21–31.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4):991–1013.

Bowen, W. G., Chingos, M. M., Lack, K. A., and Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management*, 33(1):94–111.

Dale, S. B. and Krueger, A. B. (2014). Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of Human Resources*, 49(2):323–358.

Darolia, R., Koedel, C., Martorell, P., Wilson, K., and Perez-Arce, F. (2015). Do employers prefer workers who attend for-profit colleges? evidence from a field experiment. *Journal of Policy Analysis and Management*, 34(4):881–903.

Daymount, T. and Blau, G. (2008). Student performance in online and traditional sections of an undergraduate management course. *Journal of Behavioral and Applied Management*, 9(3):275–294.

Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., and Katz, L. F. (2016). The value of postsecondary credentials in the labor market. *American Economic Review*, 106(3):778–806.

Deterding, N. M. and Pedulla, D. S. (2016). Educational authority in the "open door" marketplace: Labor market consequences of for-profit, nonprofit, and fictional educational credentials. *Sociology of Education*, 89(3):155–170.

Driscoll, A., Jicha, K., Hunt, A. N., Tichavsky, L., and Thompson, G. (2012). Can online courses deliver in-class results? A comparison of student performance and satisfaction in an online versus a face-to-face introductory sociology course. *Teaching Sociology*, 40(4):312–331.

Enriquez, A. (2010). Assessing the effectiveness of synchronous content delivery in an online introductory circuits analysis course. *Proceedings of the annual conference of the American Society* *for Engineering Education*.

Figlio, D., Rush, M., and Yin, L. (2013). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, 31(4):763–784.

Gaddis, S. M. (2018). An introduction to audit studies in the social sciences. In Gaddis, S. M., editor, *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Springer International Publishing AG.

Hernandez-Julian, R. and Peters, C. (2012). Does the medium matter? Online versus paper coursework. *Southern Economic Journal*, 78(4):1333–1345.

Kroft, K., Lange, F., and Notowidigdo, M. J. (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*, 128(3):1123–1167.

Lack, K. A. (2013). Current status of research on online learning in postsecondary education. Technical report, ITHAKA S+R.

Lahey, J. and Beasley, R. (2018). Technical aspects of correspondence studies. In Gaddis, S. M., editor, *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Springer International Publishing AG.

Lahey, J. N. (2008). Age, women, and hiring: An experimental study. *The Journal of Human Resources*, 43(1):30–56.

Lahey, J. N. and Beasley, R. A. (2009). Computerizing audit studies. *Journal of Economic Behavior and Organization*, 70(3):508–514.

Levitt, S. D. and Dubner, S. J. (2005). *Freakonomics: a rogue economist explores the hidden side of everything*. New York: William Morrow.

Rechlin, A. M. and Kraiger, K. (2012). The effect of degree characteristics on hiring outcomes for I-O psychologists. *The Industrial-Organizational Psychologist*, 49(4):37–47.

Rich, A. J. and Dereshiwsky, M. I. (2011). Assessing the comparative effectiveness of teaching undergraduate intermediate accounting in the online classroom format. *Journal of College Teaching and Learning*.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87 (3):355–374.

Vuolo, M., Uggen, C., and Lageson, S. (2016). Statistical power in experimental audit studies: Cautions and calculations for matched tests with nominal outcomes. *Sociological Methods and Research*, 45(2):260–303.

Vuolo, M., Uggen, C., and Lageson, S. (2018). To match of not to match? statistical and substantive considerations in audit design and analysis. In Gaddis, S. M., editor, *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Springer International Publishing AG.

A Audit Procedure

Completing a correspondence audit study requires three major steps. First, the researcher must create a pool of materials that can be used to study the research question at hand. In this paper, the goal is to study how online degree programs would affect the labor market prospects of typical graduates. Therefore, the researcher first had to generate a set of fictional applicant profiles (resumes and associated cover letters). Second, the researcher identified and then used the fictional applicant profiles to apply for suitable real job openings. Third, the researcher monitored voicemail and email inquiries ("callbacks"). Many employers left voicemails such as "Hi, we received your application for position X, we would like to speak with you about the position." It is difficult to know if that means the employer would like to interview the applicant. For that reason, any positive response to an application (basically anything other than a "no, thank you") was recorded as a callback. The following subsections explain the details of each of these steps.

A.1 Resume Generation

This study's findings are based on applications to real jobs using 100 different fictional applicant profiles (a profile consists of a resume and a cover letter). These profiles represented workers in business, software engineering, mechanical engineering, nursing, and accountancy. The resumes used in the study are based upon publicly-posted resumes of recent college graduates on a major job-hunting website. Only resumes representing those who are recent graduates (obtained their BA/BS three or fewer years prior to application) were selected to be part of the study because the growth of online education is a relatively recent phenomenon. In addition, only resumes where the listed work experience matched the educational background were included. That is, a Registered Nurse with a nursing degree was working as a nurse, and software engineers selected were currently working in software development or some other information technology-related position. For each selected resume, the researcher then altered dates, names, contact information, address, and previous employer names and locations. The perceived gender, degree held (but not institution), current job title (but not employer name), and years of work experience reported on the resume were not altered. These changes were intended to protect identities while preserving the overall authenticity and quality of the resume. The resumes were then further anonymized by mixing and matching resume details within sub-groups of similar applicants (nurses, engineers, and so on). This process ensures that the resumes used to apply for job openings did not resemble the actual resume of any real-world job seeker.

The resumes used as a basis for the study's pool varied in almost every way one can imagine. The individuals lived in a variety of locations, had different work experience, attended different colleges throughout the US, had various degree titles (even within the same area of expertise), many listed internships or part-time employment in college, and some used personal statements or listed "headline" keywords while others did not. For practical reasons, the resumes chosen were limited to those currently employed (and therefore having at least *some* experience) to ensure it was possible to find enough suitable openings: openings for recent graduates with *no* experience are rare whereas those requiring one year of experience are relatively plentiful. It is open to debate, but it is likely that examining outcomes for those with no experience would see a similar or larger gap in callback rates if enough suitable openings could be identified.

Each fictional resume listed that the individual attended a university where the degree they earned was offered in both a traditional and an online-only format. As just one example, Penn State University offers 24 degree programs that can be completed online from anywhere in the world.[A1] Like many others, the Penn State World Campus FAQ page emphasizes that admissions standards are the same as for the rest of Penn State and that a transcript for an online degree will not be any different to the transcript of those who completed their degree on-campus.[A2] Because the degree program on each resume was always offered both online and in a traditional format at the listed college/university, the researcher could then use a random number generator to assign "online degrees" to a subset of the pool of resumes.

In particular, the researcher entered the details of each resume into a spreadsheet. Then, the spreadsheet program generated a random number between 1 and 100 (with replacement) for each resume. The resumes associated with the fifty smallest numbers were assigned to have an online degree. For these resumes, the researcher added the word "online" in parentheses next to the name of the college or university. On a resume this appeared as "[Name of University or College] (online)." That is the only difference potential employers would see on an applicant's resume.[A3] Note that this process requires that for each fictional resume, the type of education received did not vary across applications. In addition, because these jobs are not low-skill, each application was

[A1] The programs offered at Penn State's World Campus can be accessed via http://www.worldcampus.psu.edu.

[A2] For example, see the University of Florida's UFOnline FAQ page at http://ufonline.ufl.edu/resources/faqs/.

[A3] A sample is provided in Appendix A.

accompanied by a cover letter. This is unusual in correspondence studies. Indeed, this paper could be the first to focus exclusively on highly-skilled positions.

A.2 Cover Letter Generation

Job postings for skilled positions typically request and almost always allow a cover letter. Therefore, for each resume, the researcher created a cover letter that varied in content across workers but not in organization or intent. All cover letters contained four paragraphs. The first paragraph expressed interest in the available position. The second explained the candidate's current role, responsibilities, length of current job tenure, and expressed a desire to further their career in a new position. The third paragraph explained why the candidate would be a good fit for the available position but was not tailored to each position. Instead, the paragraph reminded the reader of the candidate's education along with their technical, analytical, communication, or other skills as relevant to the field. For example, all nurse cover letters suggested the candidate was kind, caring, and considerate. Software engineers were technically and analytically adept, and so on. These paragraphs did not vary meaningfully across individuals within the same field and so any bias introduced should be mitigated by including "career" fixed-effects in regression estimates (that is, indicator variables for nursing, engineering, and so on). The final paragraph of each cover letter reiterated the candidate's interest in the position and expressed a desire to discuss the position at interview.

A.3 Sample Resume and Cover Letter

This subsection provides a sample resume (Figure A1) and associated cover letter (Figure A2). The sample resume is presented as auto-formatted by the job application website used throughout the study. For obvious reasons, contact information is redacted. When applying for jobs, sometimes the cover letter could be attached as a document file, sometimes it had to be pasted into a text entry box. Dates on resumes and in cover letters were changed to be closer to the "current date" (the date of the application) as the experiment progressed.[A4]

[A4] There is no value in providing sample email and voicemail callbacks because most of the information - such as the job application website, voicemail and email provider, fictional applicant, and employer name - would have to be redacted.

Sample Applicant

Mechanical Engineer
Lexington, MA 02420
[redacted]@[redacted].com
(339) 970-[redacted]

BS degree in Mech. Engineering plus two years of hands-on experience.

Applications: SolidWorks, Visual Professional, AutoCAD, LabVIEW, and Microsoft Office.

Programming: C++, Python

Machining: ShopBot, Milling Machine, Lathe, Drill Press, shop tools and saws.

Work Experience

Mechanical Engineer
Sample Company A - Lexington, MA
July 2015 to Present

• Manage small and medium sized projects, from creating a work scope to completion and close out
• Using CAD create mechanical drawings, RFQs (request for quotes), inspect machined parts, and follow-ups
• Participation in completing reliability reports such as Risk Matrices, FTA (fault tree analysis), FMEAs (failure
modes and effect analysis), RBDs (reliability block diagrams), RCAs (root cause analysis)
• Resolved Root Cause Analyses on issues and failures, managed records for shared use by maintenance team
• Analyze machine performance by calculating MTBF (mean time between failure), MTTR (mean time to repair), and identify critical spare parts through frequency of failure
• Knowledge of reliability program to reduce risk, optimize cost, improve safety and quality in systems
• Implementing the Kaizen process, we have been able to increase OEE (overall equipment effectiveness) by various projects and focusing on large sources of downtime

Mechanical Engineer Internship
Sample Company B - Huntsville, AL
May 2014 to August 2014

• Shadowed aerospace engineering and computer science teams and helped prepare flight and advanced fabrication reports, debug MATLAB code, procure documents and spreadsheets of lab projects and results.

Education

Bachelor of Science in Mechanical Engineering (GPA 3.64)
Podunk State University (online) - Podunk, MA
September 2011 to May 2015

(a) Page 1

Skills

CNC (1 year), FABRICATION (1 year), LABVIEW (1 year), SOLIDWORKS (2 years), MATLAB (2 years), Microsoft Office (4 years)

Additional Information

Plays ultimate frisbee and enjoys soccer and basketball recreationally

Avid reader and founding member of local historical fiction book club

Volunteer for the Red Cross (spent 6 weeks volunteering as a construction worker for rural hospital in Kenya in 2013)

(b) Page 2

Figure A1: Resume for Sample Applicant

Sample Applicant
123 ABC St, Lexington, MA
(123) 456-7890, applicant@domain.com

March 25, 2017

Dear Sir/Madam:

Please accept this letter and accompanying resume as application for the advertised mechanical engineer position.

Since July of 2015 I have worked as a mechanical engineer at Sample Company A. I have developed a strong background in the methods and practices in mechanical system design, reliability testing, and troubleshooting failed mechanical equipment. In addition, my current role requires frequent use of Solid Works, AutoCAD, MATLAB, Microsoft Office, and more. I would enjoy the opportunity to improve my skills in a new position.

My Bachelor's degree in Mechanical Engineering, completed at Podunk State University, has given me a strong technical foundation. As part of my bachelor's degree I completed an internship with Sample Company B in Alabama. I have also volunteered as a construction worker in Kenya. This exposure to engineering work environments has been a tremendous benefit. These skills and experience along with strong communication, troubleshooting, and problem-solving skills will help me excel in a new role. In addition, my passion to learn new skills will make me an asset into the future.

A copy of my resume is enclosed for your review. I will be available for interview at a mutually convenient time to further discuss my capabilities and how I may be beneficial to your organization. I will be available to begin work at the end of April if you were to find me a good fit for this position. Please do not hesitate to contact me if you have any questions or need more information.

Sincerely,

____

Sample Applicant

Figure A2: Cover Letter for Sample Applicant

A.4 Names and Signals of Ethnicity

Fifty of the resumes used in the study conveyed a Caucasian applicant (with names like Matthew, David, Katie, and Jessica accompanied by last names of European origin such as Smith, Mueller, Allen, and Schwartz). Of the other 50, 25 were African American (with names like DeShawn, Shanice, and Jasmine accompanied by last names such as Wilson, Jackson, and Jones) and 25 were identifiably Latin American (with names lime Juan Pablo, Agustin, Gabriela, and Sofia accompanied by last names like Lopez, Gomez, Fernandez, and Ximenez). The paper's findings by race should not be viewed as a causal estimate of the effect of race because the selection of names was based on convenience rather than any attempt to have a representative set of names. In particular, the names chosen came from a mix of babycenter.com lists of popular names by race, the list of "whitest" and "blackest" names in Freakonomics (Levitt and Dubner, 2005) and a list of the Top 400 1990s names provided by the Social Security Administration (see https://www.ssa.gov/oact/babynames/decades/names1990s.html).

A.5 Job Openings and Applications

To apply for positions, the researcher first identified recently-advertised (less than 48 hours "old" in all cases) positions suitable for any member of each sub-group of fictional candidates (nurses, engineers, and so on) posted on a major job hunting website. For each opening, the researcher examined the text of the advertisement carefully to ensure all of the candidates in a sub-group were minimally-qualified. Focusing on recent job postings should maximize the chance of receiving a callback, providing more statistical power (see Vuolo et al., 2016 and Vuolo et al., 2018 for more details on power calculations in audit studies). Then, the researcher selected one fictional applicant to apply to each selected position. In correspondence study terminology, and as mentioned in the body of the paper, this means that the study's design was "unmatched" (Vuolo et al., 2018). The reason for using an unmatched design was to avoid detection.

Data collection (applying for jobs) began in the spring of 2015 and continued at various points until November 2017.[A5] As the study progressed, dates of employment and graduation were changed to ensure resume always reflected a recent graduate with some but not a lot of work

[A5]This time-frame was not intentional: from 2015 to 2017, the author changed institutional affiliation three times. For that reason, applications were completed in small batches to ensure callbacks could be easily monitored and responded to.

experience at the time of an application. There is variation in the total number of applications sent for each fictional applicant due to the random process used to select resumes for a given position. In addition, on a handful of occasions, an application became "unusable" after the application was submitted due to a canceled search (signaled by a brief email saying that the position was not going to be filled at this time).

To avoid bias, the study includes only openings that asked for information explicitly available in the existing cover letter and resume. This restriction lead to many abandoned applications as job application systems often require more than a resume and cover letter to be submitted. Unfortunately, it is rarely clear what will be asked when beginning a job application. Applications often appear to request just a resume and cover letter to be uploaded (or the information to be pasted into a firm-specific format) but upon clicking "submit" the system brings the applicant to another page of questions that can include basic personality tests or short essays specific to the firm, location, industry, or background of the potential applicant. To avoid the potential for bias from such essays and tests, the applications were abandoned at that point.

For each of the 100 resumes, the researcher created a unique email address (generally: first name, middle initial, last name "at" some internet domain, or a slight variation if that was not available). The email addresses were then associated with "virtual" phone numbers and voice-mail services. Creating a unique phone number for each resume was not feasible but also not necessary. Instead, the researcher created enough online voice mailboxes to ensure that no two profiles in the same career area or field had to share a phone number. The outgoing voice-mail message was left as the default computerized greeting. That is, any message an employer heard when calling was the same regardless of resume received and only differed by phone number. In almost all cases, the employer mentioned some combination of their name, the applicant's name, or the position title. For the handful of cases when they did not, the researcher figured out which position they were calling about by completing an internet search for the caller's phone number.

A.6 Monitoring Callbacks

After each batch of applications, the researcher monitored the relevant voice and email inboxes. When a request for interview was received the researcher politely declined the request as soon as feasible.[A6] Similarly to Deming et al. (2016), a callback is considered any positive personalized contact. Paraphrasing slightly, employers who left a voicemail typically stated that they wanted to “discuss” an application. Sometimes employers who contacted an applicant were clearly only calling to obtain more information (such as asking for additional documentation that they forgot to request via the job website) rather than expressing interest in interviewing. Because this information generally was not included in the cover letter or resume, these applications were abandoned. The paper’s estimates do not change appreciably by including or excluding these as callbacks.

While each resume reports a postal address the address is entirely fictitious (although it appears realistic) and any contact via postal mail would be missed. Bertrand and Mullainathan were concerned about this and contacted several human resources managers who suggested postal requests for interview were extremely rare. Given that Bertrand and Mullainathan’s study was completed 15 years ago the potential for bias introduced by requests for interview via postal mail can be ignored.

[A6] Thanks to Nora Fergany and an upwork.com freelancer for their help with responding to phone calls to “female” applicants.

B Panel Effects

Typically, if there is reason to believe that differences *across* entities have an influence on the dependent variable then a random effects approach to estimation is preferable. Random effects estimations assume that the error term is not correlated with independent variables to allow for values that are fixed for each individual to play a role as explanatory variables. Random effects specifications typically cause concerns about omitted variable bias but that is less of a concern here as there are no "missing" variables (by construction). A fixed effects approach is not feasible due to a lack of variation in the independent variables.

After estimating a random effects specification, a Breusch-Pagan Lagrange Multiplier (LM) test can examine if treating the data as a panel is appropriate. In particular, the null hypothesis in the LM test is that variances across entities is zero. That is, there is no significant difference across units of observation: no "panel effect." For the data used in this paper, the results of an LM test suggest a panel approach may be preferable but the test statistic was only borderline significant. For completeness, Table B1 presents the same set of estimates as seen in Table 2 (in the body of the paper) using a random-effects approach. Unsurprisingly, given the LM test was only borderline significant, the estimates convey very little new information compared to Table 2.

Table B1: Random Effects Probit Estimates

	(1) Callback	(2) Callback	(3) Callback
Online	-7.2392***	-6.6261***	-7.2319***
	(2.3202)	(2.0965)	(1.9315)
Male		1.1954	1.5268
		(2.2078)	(1.9670)
Black		-1.3512	0.3675
		(2.1997)	(2.1340)
Hispanic		-1.8710	-3.5225
		(2.8200)	(2.5930)
Grade Point Average		12.5944***	12.5765***
		(3.0886)	(3.1684)
Experience (Years)		2.6210*	2.7569*
		(1.5233)	(1.4115)
Selectivity (U.S. News Acceptance Rate)			-0.1980***
			(0.0587)
Observations	1,891	1,891	1,891
No. of Applicant Profiles	100	100	100
Individual Characteristics	N	Y	Y
Career and School Characteristics	N	N	Y

Standard errors are clustered at the applicant-level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table reports the marginal effects from random effects probit estimations with controls added sequentially as indicated. The coefficient estimates on career/field fixed effects are not reported to economize on space. The coefficients can be interpreted as percentage point differences in callback rates for a one unit change in the variable of interest.