HOW DO ONLINE DEGREES AFFECT LABOR MARKET PROSPECTS? EVIDENCE FROM A CORRESPONDENCE AUDIT STUDY

CONOR LENNON*

This article 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 fictitious applicant profiles. The applicant profiles are designed to be representative of recent college graduates from established universities. Using random assignment to degree type, applicant profiles that indicate a traditional (in-person) degree receive nearly twice as many callbacks as those that indicate an online degree. Findings suggest that, at least currently, completing an online degree program would significantly limit the labor market prospects of typical college students.

In the fall of 2016, 6.4 million college students in the United States took at least one online class and 3 million took only online classes (Allen, Seaman, and Seaman 2018).[1] Indeed, many well-known schools—including Arizona State University, the Ohio State University, Pennsylvania State University, and Northeastern University—offer dozens of degrees completely online.[2] Although only a fraction of students currently in such programs fit the typical undergraduate student demographic, the growth of online classes and degree programs naturally raises several questions: What will higher

[1]Allen et al. used data from the Integrated Postsecondary Education Data System (IPEDS), which distinguishes between in-person instruction and "distance education." These distance education classes are typically delivered online.

[2]See https://www.worldcampus.psu.edu/, for example. Schools typically emphasize the similarity of the training, quality, and admissions requirements between their online and in-person degree programs.

*Conor Lennon (https://orcid.org/0000-0002-4131-5434) is an Assistant Professor at the University of Louisville.

Thanks to Nora Fergany and two upwork.com freelancers (who wish to remain nameless) for their help on this project. Thanks to Keith Teltser and seminar participants at the University of Louisville, Purdue University, and Auburn University for their helpful comments. For information regarding the data and/or computer programs used for this study, please address correspondence to the author at conor.lennon@louisville.edu.

Keywords: online education, correspondence study, human capital, recruitment, job search, signaling, online degree, hiring, audit study

ILR Review, 74(4), August 2021, pp. 920–947
DOI: 10.1177/0019793919899943.
Journal website: journals.sagepub.com/home/ilr
Article reuse guidelines: sagepub.com/journals-permissions

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?

Economic theory suggests that the answers to these questions depend, at least in part, on how employers view graduates from online degree programs. Little is known, however, about how employers perceive such graduates, particularly relative to those who pursue a traditional in-person degree program. For that reason, this article reports the findings from a correspondence audit study designed to answer the following research question: "How do online degrees affect employment prospects?"

Specifically, I examine employer responses to 1,891 job applications using 100 unique fictitious applicant profiles. The fictitious profiles are based on real résumés, gathered from a major online jobs website, and represent recent college graduates in four broad areas: business, engineering, nursing, and accounting. For each real résumé, names, dates, contact information, addresses, and previous employer and education details were anonymized. At random, for 50 of these résumés, the researcher added the word "online" in parentheses next to the name of the listed college or university. The researcher then used these résumés to apply for suitable job openings. For each of the study's fictitious job candidates, I provided an e-mail address and a phone number (connected to an online voicemail service) that employers could call if they wished to pursue the candidate's application. Because employers typically left voicemail messages without specifically offering an interview time, any positive personalized contact is considered a "callback." Of course, callbacks for interviews are not a perfect measure of labor market success, but applicants with more job interviews should face shorter spells of unemployment, more job offers, and higher wages, all else being equal.

Note that for each of the study's fictitious applicants, I indicated a bachelor's degree from an established four-year nonprofit school. In all cases, the school offered both an online and in-person version of the indicated degree. For-profit schools do not primarily serve the population on which the article is focused: young adults who attend college shortly after graduating from high school.[3] For that reason, the study can say nothing about differences in outcomes for graduates from traditional universities versus for-profit schools, such as the University of Phoenix or DeVry. Instead, the goal of this article is to examine how employment outcomes would vary for high school graduates if they were to pursue an online degree compared to an in-person degree, all else being equal. This comparison matters because it is the choices of these kinds of students that will shape higher education over the coming decades.

[3]For many students who complete an online degree at a for-profit institution, the choice is between an online degree and no degree rather than between a traditional degree and an online degree.

Online Instruction and the Value of Correspondence Studies

Lack (2013) reviewed how learning outcomes are affected by online instruction in sociology (Driscoll et al. 2012), accounting (Rich and Dereshiwsky 2011), and management (Daymount and Blau 2008) settings. Across these studies, Lack's review found no evidence that, controlling for observable characteristics, students learn less effectively when the medium of instruction is online rather than in-person. The conclusions that can be drawn from these early studies of online learning are clouded, however, by differences in research methods, subject attrition, treatment and control group cross-contamination, small sample sizes, and different populations of interest, along with each study having a unique institutional setting and time frame.

In response, researchers have used larger sample sizes and controlled variation to examine how online instruction affects learning. For example, Figlio, Rush, and Yin (2013) examined the effects of watching online rather than attending introductory economics lectures at a large selective research institution in the United States. The authors observed that regardless of sex or race, average test scores were higher for those who were randomly assigned to "live lectures." The effects were modest (between 1.9 and 3 points out of 100), however, and not always statistically different from zero. Bowen, Chingos, Lack, and Nygren (2014) focused on learning outcomes in statistics classes but allowed for the online instruction to be augmented by an interactive learning platform. They found that exam pass rates and final exam scores were not affected by the mode of instruction. A notable exception in the literature is Alpert, Couch, and Harmon (2016), who found that performance on a cumulative microeconomics final exam was 5 to 10 points (out of 100) lower for students randomly assigned to an online section.

Although 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. 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. suggested that "Internet-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" (2013: 764). Although there may be a pedagogical benefit, it is unclear whether students would want that benefit if it comes at the cost of diminished labor market prospects. Of course, labor market outcomes are unlikely to be affected by a change in the mode of instruction of one or even a handful of college courses. Nonetheless, that does not mean labor market outcomes can be safely ignored.

The lack of discussion and research on the impact of online education on labor market outcomes motivates this article. 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 toward potential employees with degrees that are earned online.[4] 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) provided the ideal example of the value and purpose of such studies. These authors examined whether employers screen résumés using indicators for race (such as names). Their article's title, "Are Emily and Greg More Employable than Lakisha and Jamal?," is self-explanatory. Focusing on low-skilled positions, Bertrand and Mullainathan found that white-sounding names received 50% more callbacks for interviews, all else being equal. Note that because 50 of the fictitious résumés in the present study feature ethnic applicant names (see Appendix A), the study can extend Bertrand and Mullainathan's findings to positions for which a bachelor's degree is required. As a preview, estimates suggest that 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 the Findings section for more detail).[5]

Last, this article makes a unique contribution to the literature on the effects of college "quality" on students' labor market outcomes. Dale and Krueger (2002, 2014) provided a detailed analysis of that literature and found that selectivity, as proxied by undergraduate acceptance rates, positively affects labor market outcomes. Dale and Krueger also found that the effect of a highly selective or elite college on labor market outcomes dissipates after controlling for student quality and subsequent selection into higher-quality schools. Dale and Krueger focused on differences in earnings using administrative and survey data. More recently, several correspondence studies, including Gaddis (2015), Darolia et al. (2015), Deming et al. (2016), and Deterding and Pedulla (2016), examined how college or university quality affects labor market outcomes. Gaddis (2015) found that fictitious applicants with a degree from an elite university receive more callbacks but the effect does not compensate (in terms of callbacks for interviews) for the effect of racial discrimination. Specifically, "black

[4]Gaddis (2018) examined the history of audit studies and explained the terminology, the breadth of personal characteristics, and the range of outcomes that can be examined using in-person and correspondence audits. Also, see Pager (2007) and Neumark (2018).

[5]A caution: Such findings should not be viewed as a causal estimate because I selected names nonscientifically. In particular, the study's fictitious applicant names come 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). Gaddis (2017a) noted that many correspondence studies focused on discrimination do not select names scientifically. Gaddis then used a large mTurk-based survey to examine this issue and found that the perception of black names used in previous correspondence audits (including Bertrand and Mullainathan 2004) varied significantly across individuals.

candidates from an elite university only do as well as white candidates from a less selective university" (2015: 1453). Darolia et al. (2015) and Deming et al. (2016) each examined how for-profit degrees are viewed by employers. Darolia et al. found "no evidence that employers prefer applicants with résumés listing a for-profit college relative to those whose résumés list either a community college or no college at all" (2015: 881). Deming and his coauthors implemented an audit study to examine the effect of having a degree from a for-profit school relative to a nonselective public school. They found 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 nonselective public institution" (2016: 778). Similarly, Deterding and Pedulla (2016) focused on callbacks for applicants with degrees earned at institutions that have open-door admission policies, including for-profit, nonprofit, and entirely fictitious institutions. They found no significant difference in callbacks across institution types.[6]

Note that existing work in this area focuses on differences between schools. By contrast, one can view this article as examining unique within-school variation in reputation created by offering online degree programs. Moreover, the existing studies examine low-prestige institutions (except for Gaddis 2015), generally comparing nonselective nonprofit to private for-profit schools. This article is different in that it focuses on relatively selective, established schools that offer both in-person and online versions of the same degree. Also, the résumés used to apply for positions are based on the résumés of real recent graduates, ensuring that the article's findings are relevant for typical college-age students. Most important, no study has attempted to isolate the effect of online degrees from any other type.

The Limits of Correspondence Studies

Correspondence studies are an excellent way to uncover the attitudes of employers toward specific employee characteristics. However, several caveats apply. First, it is not clear from these studies that fewer callbacks translate to lower wages and higher unemployment. Instead, information transmitted to employers via the résumé may improve matching and therefore could reduce wasteful and unnecessary interviews that would not result in a job offer anyway. In addition, although certain characteristics might reduce the total number of callbacks for a given applicant, they could increase the probability of getting the "right" callback.

Second, employers may fail to notice the experimental variation. Thus, the estimated effect of a characteristic could represent a lower bound on the true effect. On a related note, it is possible that résumés are positively

[6]Carbonaro and Schwarz (2018) reported on a challenging correspondence audit that examines how different high school characteristics (such as racial composition and selectivity) affect callbacks. Carbonaro and Schwarz's article also included a discussion of the challenges associated with audit study design.

or negatively filtered by software using trigger words, which could bias findings in either direction. This is less of a concern in this article 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 appears to fare poorly in a correspondence study setting may be able to improve their job prospects via alternative job search approaches.

Last, correspondence studies cannot be sure that their experimental variation does not interact with employers' experience. Take Bertrand and Mullainathan's (2004) study as an example. The article claimed to study the difference in callbacks between a résumé with a black-sounding name and an identical résumé with a white-sounding name. To be strict, however, the article studied the effect of having a black-sounding name, reporting it without alteration, and having a résumé that does not reflect changes that an employer may expect to see given that variation. That is, non-fictitious white males and black males might present very different résumés even if they had similar work histories and education.[7] If résumés 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 treatment effect is the combined effect of having a black-sounding name and having a résumé that does not seem like other résumés from black applicants. This example is not chosen at random. Bertrand and Mullainathan found that white applicants experience a much higher return to increased résumé quality (e.g., a larger increase in callbacks), which suggests that employers may be skeptical of high-quality black résumés.

Kroft, Lange, and Notowidigdo's (2013) audit study on the effects of unemployment duration on callbacks is subject to a similar critique. The authors identified not just the effect of unemployment duration but the combined effect of being unemployed and not providing 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 article is subject to similar unavoidable critiques. Specifically, the article's estimates should be interpreted as the impact on callbacks for interviews of having an online degree and telling the employer about it. Note that although telling an employer that an applicant has a degree

[7]Kang, DeCelles, Tilcsik, and Jun (2016) found, in interviews with racial minority university students, that minority students view resume whitening (removing indicators of minority status) as essential to success in the labor market except when applying to pro-diversity employers. Additionally, Kang et al. reported the findings from an audit study that examined how ''whitened'' resumes do, in terms of callbacks, when sent to employers that present themselves as valuing diversity. Paradoxically, the authors found that minorities may be particularly likely to experience disadvantage when they apply to ostensibly pro-diversity employers.

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.[8] This does not mean that all job applicants who have an online degree always choose to reveal that information to prospective employers. It eases concerns though that the résumés used in this study will stand out as extremely unique and/or unusual.

In addition, even if an applicant does not mention the online nature of their education in the résumé, it could become an issue later in the hiring process. For example, an applicant might be asked about work experience coincident with their college degree that was in another state. This means that although the effects of unemployment duration on callbacks and eventual employment might be eliminated with a one-sentence explanation, the effect of online degrees on labor market success is perhaps less avoidable. Alternatively, mimicking the concerns with Bertrand and Mullainathan's (2004) approach, employers may expect applicants with online degrees to take steps to compensate for their perceived deficiency via accomplishments in other areas. In the Findings section, I consider these issues 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 effect of these characteristics on callbacks 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.

Experimental Design

The procedure to generate résumés is similar across correspondence studies. For authenticity, the researcher creates a pool of résumés from real job seekers posted publicly on job-hunting websites. The real résumés are then deconstructed, anonymized, and reconstructed manually or via a computer program (see Lahey and Beasley 2009, 2018). Then, they are randomly assigned one of N possible variations in a characteristic of interest. The fictitious résumés 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 experimental variation.

[8]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.pdf and https://nces.ed.gov/fastfacts/display.asp?id=80 for more details.

This article uses a similar approach in which the medium of instruction for the applicant's education is the randomly assigned characteristic. The study differs from others in an important dimension, however. Many correspondence studies tend to focus on clerical, retail, and administrative openings to ensure that they can apply to many job openings with multiple résumés. 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 article focuses on several early career positions suitable for recent graduates in the business, engineering, IT, and medical professions. These positions represent 1) the types of jobs associated with degree programs offered online and in-person at many institutions, 2) 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 3) many jobs are typically advertised in these fields. A disclaimer that the findings may not generalize to other situations and professions applies.

This study's design is comparable to other correspondence audits including Bertrand and Mullainathan (2004), Kroft et al. (2013), Darolia et al. (2015), and Deming et al. (2016).[9] It is worth noting here, however, that—in audit study terminology—the study's design was "unmatched" (Vuolo, Uggen, and Lageson 2018); that is, only one application was sent to each job opening.[10] Although Vuolo, Uggen, and Lageson (2016) demonstrated that using a matched design (sending pairs, or more, 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 audit studies, the investigators apply to administrative and clerical positions that do not require a college education. In this study, the job openings are in skilled occupations. Given the limited number of suitable candidates for these kinds of positions, the likelihood of detection was a significant concern. Moreover, Phillips (2019) showed that matched audit designs can distort audit study findings because of strong spillover effects. Phillips showed that matched approaches "confound discrimination against an individual's characteristics with employers' responses to the composition of the applicant pool" and that "adjusting for applicant pool composition increases measured discrimination by 30% on average" (2019: 2240). Phillips's findings also added further nuance to the challenge of power calculations in audit studies.

Because the procedural details are similar across correspondence studies, Appendix A explains how I created the pool of 100 résumés and cover

---

[9]Baert (2018) surveyed 90 audit studies (dating from 2005 to 2018) that examine how personal characteristics affect interview requests and hiring decisions.

[10]Note that the unit of study in this article is a business/firm. As a result, Institutional Review Boards at the University of Pittsburgh and the University of Louisville (where the researcher worked during the study) declined to review the project as it did not constitute human subjects research. This documentation is available upon request from the author.

letters used in the study. It also explains how these fictitious applicants were randomly assigned to "online degree" status, how and when applications were completed, and how callbacks for interviews were handled.

Data and Estimation

Data

Table 1 presents summary statistics on the demographic characteristics of the study's fictitious applicants. The table also presents sample characteristics stratified by degree assignment (online or traditional) to examine how the randomization 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 have attended a less-selective college (as measured using U.S. News & World Report undergraduate admission rates—higher numbers indicate less selectivity).[11] The table then presents the same summary statistics stratified by both race and gender. Several clear differences emerge in the demographic characteristics and callback rates of the various groups. For example, among the résumés 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 résumé characteristics in regression estimates.

In addition, Table 1 shows an overall callback rate of 12.2% (231 callbacks from 1,891 applications). By contrast, Kroft et al. (2013) had a callback rate of just 4.7% whereas Bertrand and Mullainathan's (2004) callback rate was 8.05%. Kroft et al.'s extremely low callback rate is likely because their résumés portrayed unemployed applicants. The higher callback rate in this study is likely also due to design choices. For example, fictitious applicants in this study are particularly well matched to available positions, have résumés that reflect relevant experience, generally possess high GPAs, and have sharp and succinct cover letters. In addition, job openings were less than 48 hours old at the time of application. Résumés reflecting quality candidates should generate more callbacks and ensure adequate statistical power. Conversely, other correspondence audits achieve sufficient power by applying for many low-skilled positions, often sending several résumés to a single employer (see Vuolo et al. 2016 for more on the difficulty of calculating statistical power in correspondence study settings).

Estimation

Given that 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:

---

[11]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)	SD	(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)	SD	(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)	SD	(0.28)		(0.33)	(0.20)	(0.79)
Difference in means	(p value, t-test)	0.002	0.322	0.5357	1.000	0.5199
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)	SD	(0.34)		(0.22)	(0.20)	(0.66)
Black	Mean	0.09	0.64	3.20	0.537	1.72
(N = 25, n = 410)	SD	(0.29)		(0.40)	(0.20)	(0.87)
Hispanic	Mean	0.115	0.44	3.27	0.601	2.07
(N = 25, n = 427)	SD	(0.32)		(0.42)	(0.20)	(0.80)
Difference in means	(p value, F-test)	0.3023	0.3139	0.0065	0.5782	0.0130
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)	SD	(0.26)		(0.33)	(0.20)	(0.78)
Female	Mean	0.122	0	3.33	0.573	1.65
(N = 49, n = 964)	SD	(0.37)		(0.327)	(0.20)	(0.76)
Difference in means	(p value, t-test)	0.4392	—	0.5104	0.8316	0.2790

Notes: In the table "N" refers to the number of résumés/profiles, whereas "n" refers to the number of applications completed using those N résumés. Selectivity is the only piece of information not provided on the résumés and is the undergraduate admission rate reported by U.S. News & World Report. Higher values indicate less selective institutions. For the summary statistics that are stratified by degree type, sex, and race, the final row in the subsection of the table presents the p value from a t-test or F-test for equality of means as appropriate. Values close to zero indicate a statistically significant difference in means among the relevant groups. SD = standard deviation.

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

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 fictitious applicant-specific characteristics X_i and an indicator D_i that equals 1 if the résumé indicates that individual i's degree was earned online. In this article X_i includes GPA, years of work experience, a measure of college selectivity (undergraduate selectivity—as used by Dale and Krueger 2002), gender, race (as indicated by name similarly to Bertrand and Mullainathan 2004), 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. A negative $\hat{\delta}$ would suggest that 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 because of concerns about endogeneity and omitted variable bias. These concerns cannot be driving the article's findings as the randomization of D_i avoids that problem by construction.

The study'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 résumé as controls in regression estimates.

Findings

Table 2 reports the article's main findings. The dependent variable is whether an application generated a callback (callback = 1 when a request for interview was received and 0 otherwise). The estimates in the table are post-estimation marginal effects from a pooled sample probit regression with standard errors clustered at the applicant level.[12] The reported coefficients should be considered percentage-point differences in the probability of a callback.

The table presents 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. Specifically, the estimates in the final column suggest a 7.3 percentage-point difference in callback rates between

[12]The audit procedure creates an unconventional panel data set: There are repeated observations for each job "applicant" but no time component. Appendix B presents random effects estimates. They differ only marginally from the pooled-sample estimates. A fixed-effects specification is not feasible given that, for each fictitious resume, the independent variables do not change.

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	No	Yes	Yes
Career and school characteristics	No	No	Yes

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

traditional and online degree holders, all else being equal.[13] Given that the mean callback rate for online degree holders is 8.3%, a 7.3 percentage-point difference suggests that a résumé reflecting a traditional degree will receive almost twice as many callbacks for interviews as a résumé reporting an online degree, all else being equal.

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 different from zero. 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 is positively associated with a higher callback rate. Selectivity is measured using undergraduate acceptance rates from U.S. News & World Report and the negative coefficient implies that attending a selective school matters: For each one-percentage-point decrease in selectivity, a .2-percentage-point reduction

[13]Of course, there are several issues with computing marginal effects when estimation involves several 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 (≈ 0.5) in the data. The raw probit estimates are available from the author upon request.

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	Yes	Yes	Yes	Yes
Career and school characteristics	Yes	Yes	Yes	Yes

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

occurs in the probability of a callback. 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.

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. The difference is important though 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, two fictitious female nurses in the sample, who were randomly assigned an online degree and had two full years of relevant experience, had a callback rate of more than 25%. That is a higher callback rate than any other fictitious 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 especially for females.

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 (years)	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	Yes	Yes	Yes
Career and school characteristics	Yes	Yes	Yes

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

These two fictitious candidates are also driving the large and statistically significant coefficient on experience in the first column of Table 3.[14] As Deming et al. (2016) noted, 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.

Table 4 presents estimates separately for business applicants (accountants and business analysts), engineers (software and mechanical), and nurses. Because of the high callback rate for the two fictitious nurses, nurses with online degrees appear to experience no difference in callbacks. Additionally, these two applicants happen to have traditionally Caucasian names, which is driving the large negative coefficient for minority nurses. A larger sample would be less sensitive to such issues.

[14]Estimates from a pooled specification that interacts each reported covariate with an indicator for male can be used to test for differences between the estimates in columns (1) and (2). The null hypothesis is that the coefficients on the interaction terms in such a specification are zero. As the coefficients are probit estimates, the test statistic has a chi-squared distribution. The p value for the test was 0.0694, indicating that the null of jointly zero cannot be rejected at the 5% level. For a similar exercise for columns (3) and (4), the p value for the test was 0.0967. Again, the null of zero difference between the estimates in columns (3) and (4) cannot be rejected at the 5% level.

Table 5. Pooled-Probit Estimates: Returns to Résumé 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	Yes	Yes	Yes	Yes	Yes
Career and school characteristics	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the applicant level. 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.
$*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Returns to Résumé Characteristics

As mentioned earlier, a potential concern with audit studies is that findings are driven not only by the experimental variation but also by how differences in the fictitious résumés compare to the differences employers expect to see in real résumés. 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 perhaps should not result in all else remaining equal. Empirically, these kinds of concerns should lead to different "returns" to aspects of résumé quality for online degree holders. 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:

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

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 between having an online degree ($D_i = 1$) and $Characteristic_i$. Table 5 presents only the marginal effects on callbacks for the terms of interest by degree type. In column (1) 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 are 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 column (2), 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 column (3), GPA matters significantly but only for in-person degree holders. Put another way, if you earn an online degree, even a 4.0 GPA will not help all that much. This estimate is a confirmation of the main takeaway of this article: Employers currently do not appear to trust online education.

In the final two columns, degree type does not matter much for black workers but does for Hispanic workers. An online degree does not seem to affect the probability of callback for Hispanic applicants (the omitted category is Caucasian applicants). The estimates suggest, however, that Hispanic applicants with traditional in-person degrees will receive fewer callbacks than do 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 article (U.S. News & World Report undergraduate acceptance rates). That is, although 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 finding 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 résumés in this correspondence study, and these kinds of schools are neither the most nor least prestigious. In addition, the U.S. News–reported undergraduate acceptance rate may not fully capture the differences among these schools.

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 that students care about. Little is known, however, about how online degree programs affect labor market prospects. For that reason, I report the findings from a correspondence audit study that tests how employers view degrees earned online.

The study's findings suggest that traditional degree holders are almost twice as likely to be called back compared to applicants who report an online degree. In particular, the callback rate was 8.3% for fictitious applicants with an online degree but 15.9% for those listing a traditional in-person degree. The identifying assumption is that the difference in callbacks between those listing an online degree and those listing an in-person degree is only because of differences in the randomly assigned degree type. If that assumption holds, the article's findings suggest that employers do not currently view online degree holders as comparable to those with an in-person degree. The obvious takeaway for job applicants who earn a degree online is not to inform employers about the medium of instruction. But, that approach will be helpful to online degree holders' job prospects only if employers do not react negatively to the information later in the hiring process.

The article's findings complement the work of Darolia et al. (2015), Deming et al. (2016), and Deterding and Pedulla (2016) who examined differences in outcomes for graduates from nonselective schools such as community colleges and for-profit schools. In future work, it would be useful to examine how outcomes for graduates from online degree programs from established universities compare to those who have degrees from nonselective schools (both online and in-person).

Note that this study uses fictitious résumés that represent young recent graduates from traditional and online degree programs at well-known schools. The article does not examine the potential benefits of online degree programs for nontraditional students, such as those who are retraining after several years in the workforce. If online education is pursued solely by those returning to education and/or those who would never earn a traditional degree, then the findings are moot. Given the growth of online education in recent years (Allen et al. 2018), the article's findings should interest students, professors, and administrators because the estimates confirm that, from an employer standpoint, traditional modes of education are still viewed as superior to online education.

That said, it is not clear from this study what aspect of a traditional college education employers are responding favorably toward. Because learning outcomes appear to vary little between in-person and online instruction (Bennett, Padgham, McCarty, and Carter 2007; Ary and Brune 2011; Hernandez-Julian and Peters 2012; Figlio et al. 2013; and Bowen et al. 2014), fewer callbacks for those with online degrees would support the idea that employers view having a traditional degree as a better signal of employability (Spence 1973). Alternatively, employers may be inferring some socioeconomic characteristics, or they may believe that human capital formation is diminished in online programs relative to traditional degrees (even if it is not), that the individual will be less socially adept, or that a traditional college education gives students something more than just grades written on a piece of paper. This article asks only whether 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.

Appendix 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 article, 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 fictitious applicant profiles (résumés and associated cover letters). Second, the researcher identified and then used the fictitious applicant profiles to apply for suitable real job openings. Third, the researcher monitored voicemail and e-mail 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 whether 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.

Résumé Generation

This study's findings are based on applications to real jobs using 100 different fictitious applicant profiles (a profile consists of a résumé and a cover letter). These profiles represented workers in business, software engineering, mechanical engineering, nursing, and accountancy. The résumés used in the study are based upon publicly posted résumés of recent college graduates on a major job-hunting website. Only résumés 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 résumés for which 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 résumé, 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 résumé were not altered. These changes were intended to protect identities while preserving the overall authenticity and quality of the résumé. The résumés were then further anonymized by mixing and matching résumé details within subgroups of similar applicants (nurses, engineers, and so on). This process ensured that the résumés used to apply for job openings did not resemble the actual résumé of any real-world job seeker.

The résumés 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 United States, 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 résumés chosen were limited to those currently employed (and therefore having at least *some* experience) to ensure that 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 fictitious résumé 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.[15] 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 from the transcript of those who completed their degree on campus.[16] Because the degree program on each résumé 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 résumés.

In particular, the researcher entered the details of each résumé into a spreadsheet. Then, the spreadsheet program generated a random number between 1 and 100 (with replacement) for each résumé. The résumés associated with the 50 smallest numbers were assigned to have an online degree. For these résumés, the researcher added the word "online" in parentheses next to the name of the college or university. On a résumé, this appeared as "[Name of University or College] (online)." That is the only difference potential employers would see on an applicant's résumé. Note that this process requires that for each fictitious résumé, the type of education received did not vary across applications. In addition, because these jobs are not low skill, each application was accompanied by a cover letter. This is unusual in correspondence studies.

Cover Letter Generation

Job postings for skilled positions typically request and almost always allow a cover letter. Therefore, for each résumé, 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

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

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

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 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 an interview.

Sample Résumé and Cover Letter

This subsection provides a sample résumé (Figure A.1) and associated cover letter (Figure A.2). The sample résumé 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 résumés and in cover letters were changed to be closer to the "current date" (the date of the application) as the study progressed.[17]

Names and Signals of Ethnicity

Fifty of the résumés used in the study conveyed a Caucasian applicant (with names such as Matthew, David, Katie, and Jessica accompanied by last names of European origin such as Smith, Mueller, Allen, and Schwartz). Of the other 50 résumés, 25 were African American (with names such as DeShawn, Shanice, and Jasmine accompanied by last names such as Wilson, Jackson, and Jones) and 25 were identifiably Latin American (for example, Juan Pablo, Agustin, Gabriela, and Sofia accompanied by last names such as Lopez, Gomez, Fernandez, and Ximenez). As mentioned in the body of the article, the 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/

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

Figure A.1.

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

Figure A.1. Continued

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)

babynames/decades/names1990s.html). Gaddis (2017a) noted that many correspondence studies that focus on discrimination do not scientifically select the names they use to signal race. Gaddis (2017a) then used a large mTurk-based survey to examine the consequences of this issue and found that perceptions of black names used in previous correspondence audits (including Bertrand and Mullainathan 2004) varied across individuals. Gaddis (2017b) repeated this study with Hispanic names and found that Hispanic last names are key to generating a perception of Hispanic descent.

Job Openings and Applications

To apply for positions, the researcher first identified recently advertised (less than 48 hours old at the time of application) positions suitable for any member of each subgroup of fictitious 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 that all of the candidates in a subgroup 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, 2018 for more details on power calculations in audit studies). Then, the researcher randomly selected one fictitious applicant to apply to each selected position. In correspondence study terminology, and as mentioned in the body of the article, this means that the study's design was "unmatched" (Vuolo et al. 2018). The main 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.[18] As the study progressed,

[18]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 that callbacks could be monitored and responded to.

Figure A.2.

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

dates of employment and graduation were changed to ensure résumés always reflected recent graduates with some but not a lot of work experience at the time of an application. Because the fictitious applicant for each opening was randomly selected, there is variation in the total number of applications sent for each fictitious applicant. The researcher made no attempt to ensure that the characteristics that can vary across résumés were equally represented in job applications. Instead, the article's regression estimates control for each of these characteristics to ensure that a composition effect does not bias the article's findings. 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 e-mail 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 résumé. This restriction led to many abandoned applications as job application systems often require more than a résumé 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 résumé 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 résumés, the researcher created a unique e-mail address (generally, first name, middle initial, last name "at" some Internet domain, or a slight variation if that was not available). The e-mail addresses were then associated with "virtual" phone numbers and voicemail services. Creating a unique phone number for each résumé 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 voicemail message was left as the default computerized greeting. That is, any message an employer heard when calling was the same regardless of résumé 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 in which they did not, the researcher figured out which position they were calling about by completing an Internet search for the caller's phone number.

Monitoring Callbacks

After each batch of applications, the researcher monitored the relevant voice and e-mail inboxes. When a request for an interview was received, the researcher politely declined the request as soon as feasible. Similarly to Deming et al. (2016), a callback is considered any positive personalized contact. Employers who left a voicemail typically used words to the effect 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 résumé, these applications were abandoned. The article's estimates do not change appreciably by including or excluding these as callbacks.

Although each résumé reports a postal address, the address is entirely fictitious and therefore any contact via postal mail would be missed. Bertrand and Mullainathan (2004) were concerned about this and contacted several human resources managers who suggested that postal requests for interviews were extremely rare. Given that Bertrand and Mullainathan's study was more than 15 years ago, any bias introduced by requests for interviews via postal mail can be ignored.

Appendix 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

Table B.1. Random Effects Estimates: Marginal Effects

	(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	No	Yes	Yes
Career and school characteristics	No	No	Yes

Notes: Standard errors (in parentheses) are clustered at the applicant level. The table reports the marginal effects from random effects probit estimations with controls added sequentially as indicated. To economize on space, the coefficient estimates on career/field fixed effects are not reported. The coefficients can be interpreted as percentage-point differences in callback rates for a one-unit change in the variable of interest.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

but that is less of a concern here because 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 whether 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 article, the results of an LM test suggest that a panel approach may be preferable but the test statistic was only borderline significant. For completeness, Table B.1 presents the same set of estimates as seen in Table 2 (in the body of the article) using a random-effects approach. Unsurprisingly, given that the LM test was only borderline significant, the estimates convey very little new information compared to Table 2.

References

Allen, I. Elaine, Jeff Seaman, and Julia E. Seaman. 2018. Grade increase: Tracking distance education in the United States. Technical report. January. Babson Survey Research Group, Babson College.

Alpert, William T., Kenneth A. Couch, and Oskar R. Harmon. 2016. A randomized assessment of online learning. *American Economic Review Papers and Proceedings* 106: 378–82.

Ary, Eddie J., and Christopher W. Brune. 2011. A comparison of student learning outcomes in traditional and online personal finance courses. *MERLOT Journal of Online Learning and Teaching* 7(4): 465–74.

Baert, Stijn. 2018. Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In S. Michael Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 63–77. Cham, Switzerland: Springer International Publishing AG.

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

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* 94(4): 991–1013.

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

Carbonaro, William, and Jonathan Schwarz. 2018. Opportunities and challenges in designing and conducting a labor market résumé study. In S. Michael Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 143–58. Cham, Switzerland: Springer International Publishing AG.

Dale, Stacy B., and Alan B. Krueger. 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics* 117(4): 1491–527.

———. 2014. Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of Human Resources* 49(2): 323–58.

Darolia, Rajeev, Cory Koedel, Paco Martorell, Katie Wilson, and Francisco Perez-Arce. 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, Thomas, and Gary Blau. 2008. Student performance in online and traditional sections of an undergraduate management course. *Journal of Behavioral and Applied Management* 9(3): 275–94.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. 2016. The value of postsecondary credentials in the labor market. *American Economic Review* 106(3): 778–806.

Deterding, Nicole M., and David S. Pedulla. 2016. Educational authority in the "open door" marketplace: Labor market consequences of for-profit, nonprofit, and fictional educational credentials. *Sociology of Education* 89(3): 155–70.

Driscoll, Adam, Karl Jicha, Andrea N. Hunt, Lisa Tichavsky, and Gretchen Thompson. 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): 312–31.

Figlio, David, Mark Rush, and Lu Yin. 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): 763–84.

Gaddis, S. Michael. 2015. Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Social Forces* 93(4): 1451–79.

———. 2017a. How black are Lakisha and Jamal? Racial perceptions from names used in correspondence audit studies. *Sociological Science* 4: 469–89.

———. 2017b. Racial/ethnic perceptions from Hispanic names: Selecting names to test for discrimination. *Socius: Sociological Research for a Dynamic World* 3: 1–11.

———. 2018. An introduction to audit studies in the social sciences. In S. Michael Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 3–44. Cham, Switzerland: Springer International Publishing AG.

Hernandez-Julian, Rey, and Christina Peters. 2012. Does the medium matter? Online versus paper coursework. *Southern Economic Journal* 78(4): 1333–45.

Kang, Sonia K., Katherine A. DeCelles, Andras Tilcsik, and Sora Jun. 2016. Whitened résumés: Race and self-presentation in the labor market. *Administrative Science Quarterly* 61(3): 469–502.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo. 2013. Duration dependence and labor market conditions: Evidence from a field experiment. *Quarterly Journal of Economics* 128(3): 1123–67.

Lack, Kelly A. 2013. Current status of research on online learning in postsecondary education. Research report. March. New York: Ithaka S+R.

Lahey, Joanna N., and Ryan A. Beasley. 2009. Computerizing audit studies. *Journal of Economic Behavior and Organization* 70(3): 508–14.

———. 2018. Technical aspects of correspondence studies. In S. Michael Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 81–101. Cham, Switzerland: Springer International Publishing AG.

Levitt, Steven D., and Stephen J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: William Morrow.

Neumark, David. 2018. Experimental research on labor market discrimination. *Journal of Economic Literature* 56(3): 799–866.

Pager, Devah. 2007. The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science* 609(1): 104–33.

Phillips, David C. 2019. Do comparisons of fictional applicants measure discrimination when search externalities are present? Evidence from existing experiments. *Economic Journal* 129(621): 2240–64.

Rich, Anne J., and Mary I. Dereshiwsky. 2011. Assessing the comparative effectiveness of teaching undergraduate intermediate accounting in the online classroom format. *Journal of College Teaching and Learning* 8(9): 19–28.

Spence, Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87(3): 355–74.

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

———. 2018. To match or not to match? Statistical and substantive considerations in audit design and analysis. In S. Michael Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, pp. 119–40. Cham: Switzerland: Springer International Publishing AG.