Essays on the Labor Market for Internships

by

Eric J. Wilbrandt

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 5th, 2017

Keywords: Internships, Virtual Internships, Skill Demand

Approved by

R. Alan Seals Jr., Associate Professor of Economics
Michael Stern, Chair, Associate Professor of Economics
T. Randy Beard, Professor of Economics
Christopher Z. Vickers, Assistant Professor of Economics

Abstract

In chapter one we first describe the demand for interns. By processing the text of all available ads on a popular internship website, we are able to characterize the occupation that best matches the advertised internship. We also match each internship ad to its local labor market. We find that internships in loose labor markets are more likely to be unpaid. Paid internships are in occupations that have higher wages, require less on the job training, and are more quantitative. We then conduct an audit study with more than 11,500 resumes, where we randomly assign student characteristics and apply to internships. We find that employers are more likely to respond to an application when they are looking for an unpaid intern. We find little effect of major field of study, volunteer experience, and college work experience on employer callbacks and some evidence that a higher GPA and previous internship experience increases positive employer responses. Black applicants receive fewer positive responses than white applicants, but this is entirely driven by greater discrimination against black-named applicants living far away from employers.

Chapter two continues on this work by describing the demand for skills for virtual internsships. Using the available database from all available ads used in chapter one, I focus on a small subset containing only virtual internships. Following Deming and Kahn (2016), I process the text of the advertisements using key words aligned with specific skill sets. From this data set I assess the skills associated with paid, unpaid, part-time, and full time internships. In addition, I link each internship ad to its local labor market to assess the impact of unemployment. I find an increase in the unemployment rate decreases the probability that an internship is paid. Finally, I assess the association between different skill sets and downstream wages. Using the Occupation Employment Statistics data set, I link the ad's to downstream occupations using a machine learning algorithm. I find that cognitive and financial skill sets are statistically significant and positively associated with the downstream wage. Most surprisingly, words associated with positive character traits are negatively associated with downstream wage.

Acknowledgments

I have received unwavering support over the last four years at Auburn University. First, I would like to thank the Department of Economics at Auburn University and the Department of Economics at the University of Wisconsin-LaCrosse for providing the financial support needed to complete this dissertation. Second, I would like to thank my committee: Associate Professor R. Alan Seals, Associate Professor Michael Stern, Professor T. Randy Beard, and Assistant Professor Christopher Vickers. They have provided many opportunities and unconditional technical support throughout the entire process. Third, I would like to thank Associate Professor John Nunley of The University of Wisconsin La-Crosse and Professor David A. Jaeger of The City University of New York Graduate School. Without either of you this project would not have reached its full potential. Finally, I would like to thank my family and fiance who provided the emotional support throughout my time at Auburn. I hold you all in the highest regard.

This dissertation is dedicated to my father, Paul Wilbrandt. Without whom I would have never have been taught the values of hard work, patience, and decency.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Auburn	Auburn University
SOC	Standard Occupation Classification
MLS	Metropolitan Statistical Area
OES	Occupation Employment Statistics
ACS	American Community Survey

Chapter 1

The Demand for Interns

Internships often provide students first experiences in the skilled labor market, and they are increasingly perceived as a requirement for regular employment after college. More than 60 percent of graduates have held an internship at some point during their college careers, more than double the rate in the 1980s [1]. In response to increasing mismatch between educational attainment and jobs for new college graduates since the early-2000s (Abel, Deitz, and Su 2014) and the prioritization of relevant experience in hiring decisions (Finch, et al. 2013; Cappelli 2014; Nunley, et al. 2016a; Nunley, et al. 2016b) students have come under increasing pressure to obtain relevant work experience, and internships may offer students a way to enhance their resumes. Empirical evidence on the labor-market returns to internship experience indicates improved early-career employment prospects (Nunley, Pugh, Romero and Seals 2016a) and higher earnings later in life (Saniter and Siedler 2014).

The U.S. Department of Labor characterizes internships as a means to provide short-term, practical experience for students, recent graduates, and people changing careers.[2] Although the durations of internships are usually less than one year and apprenticeships can last many years, internships have assumed many of the features of apprenticeships as the size of the internship market has grown. A 2015 survey of large firms in the United States, for example, indicated that 90 percent of interns who return for a subsequent internship were offered full-time employment[3]. Among the 90 percent of returning interns, almost 90 percent of them accepted the employment offers from the firms for which they had completed internships. The

[1] See $http://www.naceweb.org/s10072015/internship-co-op-student-survey.aspx$ (last seen 19 March 2017) and the Lindquist-Endicott Report (1992).

[2] See https://www.bls.gov/careeroutlook/2006/summer/art02.pdf (last seen 19 March 2017).

[3] The National Association of Colleges and Employers 2015 Internship & Co-op Survey.

high rate of transition from intern to full-time employee seems to indicate that an internship, like an apprenticeship, serves an important role in determining eventual employment[4].

Another institutional feature of apprenticeships not paying a wage for the apprentices work has also become a common, if controversial, feature of internships[5]. The Department of Labor (DOL) has stated that it is possible for an intern not to be considered an employee under the Fair Labor Standards Act (FLSA) if the internship is similar to training obtained in an educational environment and it is understood the intern is not entitled to pay or regular employment with the firm in the future[6]. Whether interns are employees has been the subject of a recent lawsuit, however, that may be eventually argued before the U.S. Supreme Court[7].

We first describe the internship market by using web-scraped advertisements from a prominent website. The text from these internship ads is used as input in a machine-learning algorithm that assigns to each an O*NET Standard Occupational Classification (SOC) code, allowing us to link the advertisement data to the American Community Survey (ACS) and O*NET. We also use the geographic location of the firms seeking interns to examine how local labor market conditions may affect the characteristics of the demand for interns. After combining the various data sources, we summarize the characteristics of the internships in our sample overall and by major occupation categories, which provides a snapshot of the skills demanded and expected earnings associated with internships classified into different occupation categories. We emphasize the factors that are correlated with the paid/unpaid status of internships: the roles of (a) tight versus loose conditions in local labor markets, (b) general versus job-specific training (Acemoglu and Pischke 1998, 1999a, 1999b; Becker 1964), and (c) different skill demands

[4]Fersterer, Pischke, and Winter-Ebmer (2008) find annual earnings returns to apprenticeships in Austria of approximately four percent, and also note that there is substantial heterogeneity in quality across industries and occupations in apprenticeships. Saniter and Siedler (2015) find positive wage returns of roughly six percent associated with internships. Heterogeneity in the return on internships depends, however, on the strength of labor market orientation for specific majors.

[5]Apprenticing under the supervision of a more experienced artisan for little or no pay is an ancient practice (Wallis 2008). Apprenticeship contracts often included bonding schemes, by which an apprentice might post a cash amount to compensate the master in the event the apprentice quits before the term is over (Elbaum 1989).

[6]See https://www.dol.gov/whd/regs/compliance/whdfs71.pdf (last seen 19 March 2017).

[7]A U.S. district court ruled in 2015 that unpaid interns were employees under the FSLA and, therefore, were entitled to (at least minimum wage) compensation (Glatt v. Fox Searchlight Pictures, Inc. 2015). This ruling was in turn recently vacated by a Federal Appeals Court decision that created a new set of criteria that determines whether or not the unpaid intern is the primary beneficiary of the internship (Glatt v. Fox Searchlight Pictures, Inc. 2016). Hacker (2016) argues that the latter ruling generates an ambiguous test for employee status.

(Autor, Levy, and Murname 2003; Autor and Price 2013; Deming 2016). To examine how student characteristics affect success in acquiring internships, we conducted a large-scale resume audit of the market for interns.

We find that the majority of internships advertised are unpaid and that there is a strong positive relationship between paid status and and full time status. The availability of internships in general and paid/unpaid internships in particular varies widely across states[8].

In terms of the econometric evidence based on the advertisement data, we find that a one-unit increase in the unemployment rate triggers a nine percent reduction in the probability that an internship is paid. The relationship between paid status and months of on-the-job training (OJT) required is convex in its shape; that is, the probability that an internship is paid falls at low levels of OJT but begins to increase around 17 months of required OJT. We find no statistical evidence linking particular skills demands, such as social skills, routine tasks or service skills, to the paid status of internships. The exception is nonroutine analytical skills, which is associated with a 35 percent higher probability of paid status. The evidence for expected earnings indicates a positive relationship: the probability of paid status is around one percentage point higher when expected earnings rises by five percent.

The econometric evidence from the resume audit reveals a number of noteworthy findings. Unpaid internships have higher positive employer response rates than paid internships, suggesting that firms may need to cast a wider net to fill unpaid internships. However, the employer response gap between paid and unpaid internships is eliminated for internships that correspond to high-paying occupations. Part-time internships carry similar response rates to full-time internships for unpaid internships, but they have lower response rates for paid internships. We tend to find no effects for major field of study, volunteer experience, and college work experience. We find limited statistical evidence that having a high grade point average or working as an intern in the past improves internship opportunities. The only resume attributes that have robust effects on employer responses are race and distance from employers. Black applicants

[8] More populated states, such as California and New York, tend to have large numbers of internships available, but they tend to be disproportionately unpaid. By contrast, less populated states, such as Utah and Vermont, tend to have more paid than unpaid internships.

receive approximately 27 percent fewer positive responses than their white counterparts. Applicants living 500 miles or more away from employers are about 50 percent less likely to receive positive responses. The overall racial difference in positive responses rates is, however, driven entirely by greater discrimination against black-named applicants living 500 or more miles away from employers. This finding is supported by an analysis of the text from employer responses in which black-named candidates are more likely to receive a request for information concerning their current location. These findings point toward statistical discrimination as the underlying mechanism for racial differences and contest the idea that the discrimination identified is based on implicit biases (See Bertand, Chugh and Mullainathan 2005).

1.1 Characterizing the Demand for Interns

To characterize the demand side of the internship market, we use data web-scraped internship ads posted on a prominent website in fall 2016 and then again in spring 2017[9]. The initial data set included 43,319 internship ads. The ads provide pertinent information regarding the internship opening, including whether the internship is paid or unpaid, whether the internship is part time or full time, the internship title, descriptions of duties, responsibilities and requirements, the advertising firms geographic location, the month, day and year the ad was posted online, the application deadline, and the advertising firms name. The detailed information provided allows us to examine various internship characteristics as well as link the ad data to external data sources at the occupation and metropolitan-area levels. After linking the ad data to various external data sources, the data set, which we use to describe the demand side of the internship market, includes 36,669 internship ads with complete records[10].

[9] We web-scraped all internship advertisements posted on the website on 14 November 2016 and 18 March 2017. Per our Institutional Review Board (IRB) agreement, we are unable to disclose the name of the website or the names of firms seeking interns.

[10] In the process of linking the ad data to external data sources at the occupation and metropolitan-area levels, we have incomplete records for 6,650 ads. Over 80 percent of the observations with incomplete records are those for which an occupation-classification code could not be assigned. The remaining observations result from one of four problems: (i) the ad could not be linked to the ACS data, (ii) the ad could not be linked to the O*NET data, (iii) the ad could not be linked to a metropolitan area, or (iv) the unemployment rate was unavailable for the metropolitan area in which the firm is located. See Appendix A for more details on how the analysis data set was constructed.

Table 1.1: Cross Tabulation Between Paid/Unpaid and Part-Time/Full-Time Statuses

	Part-Time Internships	Full-Time Internships	All Internships
	(1)	(2)	(3)
Paid Internships	19.21%	21.29%	40.49%
Unpaid Internships	51.06%	8.45%	59.51%
All Internships	70.27%	29.73%	100.00%

With the ad data, we are able to differentiate between paid/unpaid and part-time/full-time internships. A two-way cross-tabulation between paid/unpaid and part-time/full-time statuses reveals a clear pattern: Unpaid internships are more common than paid ones, and paid internships are more likely to be full time (Table 1)[11]. Using the ad text, we create variables that measure firm-level demands for particular skills and intern attributes. Deming and Kahn (2017) search for particular words or phrases that indicate particular skills or attributes (See Appendix Table A1; recreation of Deming and Kahn). Indicator variables for 10 different types of skills and intern attributes are created from the text search: cognitive, social, character, writing, customer service, project management, people management, financial, computer (general), and software (specific). Over half of the internship ads include words and phrases that are indicative of a particular demand for social skills as well as customer service (Table 2).

[11] The Pearson χ^2 test for independence rejects the null hypothesis of independence between paid/unpaid and part-time/full-time statuses at the one-percent level, an indication that paid/unpaid and part-time/full-time statuses are, to some extent, jointly determined.

Table 1.2: Summary Statistics for Firm-Level Demands

	All Internships	Unpaid Internships	Paid Internships	Part-Time Internships	Full-Time Internships
	(1)	(2)	(3)	(4)	(5)
Cognitive	0.369	0.381	0.352	0.371	0.365
	(0.483)	(0.486)	(0.478)	(0.483)	(0.481
Social	0.558	0.523	0.609	0.527	0.631
	(0.497)	(0.499)	(0.488)	(0.499)	(0.483)
Character	0.296	0.306	0.281	0.306	0.272
	(0.457)	(0.461)	(0.449)	(0.461)	(0.445)
Writing	0.423	0.474	0.346	0.461	0.333
	(0.494)	(0.499)	(0.476)	(0.498)	(0.471)
Customer Service	0.575	0.469	0.731	0.507	0.733
	(0.494)	(0.499)	(0.443)	(0.500)	(0.442)
Project Management	0.033	0.030	0.038	0.028	0.044
	(0.179)	(0.171)	(0.190)	(0.166)	(0.205)
People Management	0.476	0.448	0.518	0.444	0.553
	(0.499)	(0.497)	(0.500)	(0.497)	(0.497)
Financial	0.137	0.124	0.157	0.122	0.173
	(0.344)	(0.329)	(0.364)	(0.328)	(0.378)
Computers (general)	0.545	0.550	0.537	0.544	0.546
	(0.498)	(0.498)	(0.499)	(0.498)	(0.498)
Software (specific)	0.029	0.025	0.034	0.022	0.046
	(0.168)	(0.157)	(0.181)	(0.145)	(0.210)
Observations	43,319	25,769	17,550	30,354	12,965

The internship title and its description are used as inputs in the O*NET-SOC AutoCoder[12], which assigns an 8-digit O*NET-SOC (henceforth, SOC) code to each ad. The majority of internships are assigned disproportionately to particular occupation categories: (i) business and finance, (ii) arts, design, entertainment, sports and media, (iii) sales, and (iv) office and administration (Table 3). The same internship-occupation categories also represent the largest shares of internships in the subsamples based on different internship statuses: paid (column 2),

[12]The O*NET-SOC AutoCoder, a proprietary machine-learning algorithm, that assigns O*NET-SOC codes to job or internship openings is accessible via the following website: http://www.onetsocautocoder.com/plus/onetmatch (last seen 19 March 2017). For details on the O*NET-SOC AutoCoder, click on the FAQ tab, which provides information on the uses, accuracy, and methodology. R.M. Wilson Consulting, Inc. owns the rights to the O*NET-SOC AutoCoder and provides access to it for fees that vary based on the number of records. Javed et al. (2015) notes that CareerBuilder, a large online job board, used the same algorithm to label job advertisements to be used in the classification of job titles for the creation of a large training data set for their proprietary algorithm Carotene.

unpaid (column 3), part time (column 4), and full time (column 5). The majority of internship-occupation categories are disproportionately unpaid internships (column 6). But there are exceptions, as sales, architecture and engineering, and construction and extraction, are more likely to be paid. All internship-occupation categories are comprised of a disproportionate number of part-time internships, except those belonging the sales category.

Table 1.3: Percentages of Different Types of Internships by Occupation Category

	Percentage of Ads within Each Internship Category						
	All	Paid	Unpaid	Part Time	Full Time	Share Unpaid	Share Paid
Major SOC Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Management (11)	6.20	7.51	5.30	5.24	8.44	49%	41%
Business and Finance (13)	22.10	21.27	22.66	22.78	20.49	39%	28%
Computer and Math (15)	5.87	6.67	5.32	5.28	7.26	46%	37%
Architecture and Engineering (17)	0.77	1.10	0.55	0.65	1.07	58%	41%
Life, Physical and Social Sciences (19)	0.65	0.41	0.81	0.64	0.67	26%	31%
Community and Social (21)	3.88	1.97	5.18	4.47	2.49	21%	19%
Legal (23)	0.55	0.28	0.72	0.57	0.49	21%	27%
Education, Training and Library (25)	2.24	1.52	2.72	2.72	1.10	28%	15%
Arts, Design, Entertainment, Sports and Media (27)	30.04	18.24	38.08	35.50	17.14	25%	17%
Healthcare Practitioner (29)	0.34	0.36	0.33	0.35	0.33	42%	29%
Healthcare Support (31)	0.31	0.16	0.41	0.35	0.20	21%	19%
Protective (33)	0.07	0.05	0.08	0.08	0.04	32%	16%
Food Preparation and Serving-Related (35)	0.49	0.46	0.51	0.45	0.57	38%	35%
Building/Grounds Cleaning and Maintenance (37)	0.08	0.07	0.10	0.08	0.09	32%	32%
Personal Care (39)	0.72	0.57	0.83	0.69	0.80	32%	33%
Sales (41)	15.23	27.78	6.68	9.63	28.45	74%	56%
Office and Administration (43)	9.54	10.82	8.67	9.52	9.58	46%	30%
Farm, Fish and Forestry (45)	0.11	0.07	0.14	0.10	0.13	27%	34%
Construction and Extraction (47)	0.11	0.16	0.08	0.11	0.11	58%	29%
Install and Repair (49)	0.05	0.06	0.04	0.05	0.06	50%	33%
Production (51)	0.58	0.37	0.72	0.65	0.42	26%	22%
Transportation and Moving (53)	0.08	0.08	0.08	0.08	0.08	41%	31%
Total Number of Advertisements	36,657	14,844	21,813	25,758	10,899		–

Another statistic produced by the O*NET-SOC AutoCoder is an occupation-match score, which is effectively a measure of how reliably the O*NET-SOC AutoCoder assigns the proper SOC code. When the occupation-match score is high, the classification of an internship into a particular occupation category is more accurate than when the occupation-match score is low. For example, scores of 70 or above imply that the algorithm correctly predicts the O*NET-SOC code to which the advertisement belongs at least 70 percent of the time. Figure A.4 in the appendix plots the kernel density estimate for the occupation-match score. From the plot, the scores range from low (50s and 60s) to high (80s and 90s) values. The average occupation-match score is slightly above 74 with a standard deviation of almost 11. Slightly more than 65 percent of the internships in our sample were assigned to occupation categories with scores in excess of 70[13].

After assigning SOC codes to the internship ads, we then link the ads to occupation-level data from the American Community Survey (ACS) and O*NET via the SOC codes. To construct a measure of expected earnings, we compute natural logarithm of average earnings for each detailed occupation category using four years of pooled cross-section data (2012-2015) from the ACS. The sample used to compute the earnings measure is based on employed workers between 24- and 28-years-old with Bachelors degrees.

Because we are interested in examining the general- versus specific-training aspects of internships, we use the on-the-job training (OJT) variables available from O*NET, which consists of nine OJT variables that are broken down into different amounts of required training (e.g., none, less than one month, one to three months, three to six months, six months to one year, one year to two years, two years to four years, four to 10 years, and over 10 years). For each category, O*NET reports the percentage of respondents that report a particular amount of required training. Based on this information, we construct a continuous measure of OJT. In particular, we use the midpoint in the range (e.g., two months for the one to three months category, 36 months for the two to four years category), and then multiply the midpoint in each OJT category by the percentage of respondents claiming that a particular amount (or range) of training is required for the occupation.

[13] According to the O*NET-SOC AutoCoder website, scores below 60 should be considered relatively unreliable. Due to the lower degree of reliability for scores less than 60, we check the sensitivity of our estimates by re-estimating all models using a subset of the data set that excludes observations with occupation-match scores below particular thresholds (e.g., 60, 70, 75, and 80). The estimated effects of the explanatory variables are not particularly sensitive to the threshold cutoffs, as the coefficients are the same in terms of sign and similar in terms of magnitude.

The remaining variables from O*NET proxy for the following skill demands: routine tasks, non-routine analytical tasks, social skills, service tasks. We follow Deming (2016) and use the same survey questions from O*NET to create these variables. With the exception of the variables used to construct the routine-task measure, the other variables are provided on a one-to-five scale, in which, for example, one is not important and five is extremely important. Each internship ad is assigned the average value for each skill demand. We then create an indicator variable for each skill demand that equals one when the value corresponding to the internship ad is above the median in the O*NET data. We follow a similar approach for the routine-task variable, but the data are based on the percentage of respondents claiming a particular degree of automation/routineness in lieu of a one-to-five scale. We create an indicator variable that equals one if the percentage of respondents in a particular occupation assigned to each ad is greater than the median value in the O*NET data.

Table 1.4: Summary Statistics for Internship-Occupation Characteristics

	Months of On-the-Job Training		Social Skills		Service Tasks		Routine Tasks		Non-Routine Analytical Tasks		Average Earnings	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Major SOC Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Management (11)	9.07	3.37	1.00	0.00	0.38	0.49	0.25	0.43	0.99	0.04	$59,465	$8,542
Business and Finance (13)	10.05	3.88	0.16	0.36	0.19	0.40	0.16	0.36	0.93	0.25	$55,484	$7,192
Computer and Math (15)	8.03	4.20	0.00	0.00	0.13	0.33	0.98	0.13	0.89	0.32	$56,334	$8,274
Architecture and Engineering (17)	17.38	6.84	0.16	0.36	0.01	0.07	0.50	0.50	1.00	0.00	$54,418	$9,731
Life, Physical and Social Sciences (19)	9.89	4.72	0.41	0.49	0.17	0.38	0.12	0.32	0.89	0.31	$43,145	$6,934
Community and Social (21)	5.32	3.43	1.00	0.00	1.00	0.00	0.86	0.34	0.00	0.00	$36,016	$735
Legal (23)	22.20	8.28	0.51	0.50	0.51	0.50	0.49	0.50	0.00	0.00	$67,748	$24,485
Education, Training and Library (25)	7.28	5.87	0.36	0.48	0.85	0.36	0.24	0.43	0.46	0.50	$32,348	$4,970
Arts, Design, Entertainment, Sports and Media (27)	5.29	3.19	0.64	0.48	0.58	0.49	0.32	0.47	0.02	0.13	$50,246	$4,297
Healthcare Practitioner (29)	7.61	4.17	0.62	0.49	0.94	0.24	0.51	0.50	0.49	0.50	$52,102	$19,519
Healthcare Support (31)	3.14	1.95	0.43	0.50	0.99	0.11	0.68	0.47	0.44	0.50	$31,885	$7,284
Protective (33)	9.80	8.22	0.53	0.52	0.87	0.35	0.80	0.41	0.13	0.35	$45,710	$10,892
Food Preparation and Serving-Related (35)	4.95	4.43	0.40	0.49	0.76	0.43	0.52	0.50	0.53	0.50	$27,367	$4,807
Building/Grounds Cleaning and Maintenance (37)	7.35	2.06	0.35	0.49	1.00	0.00	0.45	0.51	0.30	0.47	$30,081	$5,810
Personal Care (39)	3.97	3.35	0.22	0.41	0.98	0.13	0.24	0.43	0.03	0.18	$32,019	$4,472
Sales (41)	6.39	4.13	0.86	0.35	0.57	0.50	0.41	0.49	0.77	0.42	$59,931	$12,632
Office and Administration (43)	5.05	2.95	0.13	0.33	0.82	0.38	1.00	0.00	0.44	0.50	$36,442	$4,217
Farm, Fish and Forestry (45)	8.94	6.49	0.31	0.48	0.75	0.45	0.56	0.51	0.31	0.48	$38,209	$7,260
Construction and Extraction (47)	26.07	9.30	0.56	0.50	0.63	0.49	0.63	0.49	0.97	0.18	$52,307	$10,875
Install and Repair (49)	17.27	6.20	0.00	0.00	0.40	0.51	0.13	0.35	0.53	0.52	$39,165	$4,325
Production (51)	11.97	9.96	0.09	0.29	0.10	0.30	0.92	0.28	0.46	0.50	$31,111	$10,408
Transportation and Moving (53)	4.97	4.89	0.20	0.41	0.67	0.49	0.60	0.51	0.47	0.52	$39,916	$10,102
All Occupation Categories	7.17	4.63	0.49	0.50	0.50	0.50	0.43	0.49	0.50	0.50	$50,996	$11,277

Table 4 presents summary statistics (sample means and standard deviations) for each of the internship-occupation characteristics by major (2-Digit) SOC code. In some cases, there is substantial variation in the internship-occupation characteristics within the 2-digit SOC code. But, in other cases, there is little to no variation. As an example, consider Management Occupations. Within this occupation category, there is variation in terms of service tasks, but there is no variation in social skills and non-routine analytical tasks. Similar patterns in the data exist for other major SOC codes (e.g., computer and mathematics occupations, community and social occupations)[14].

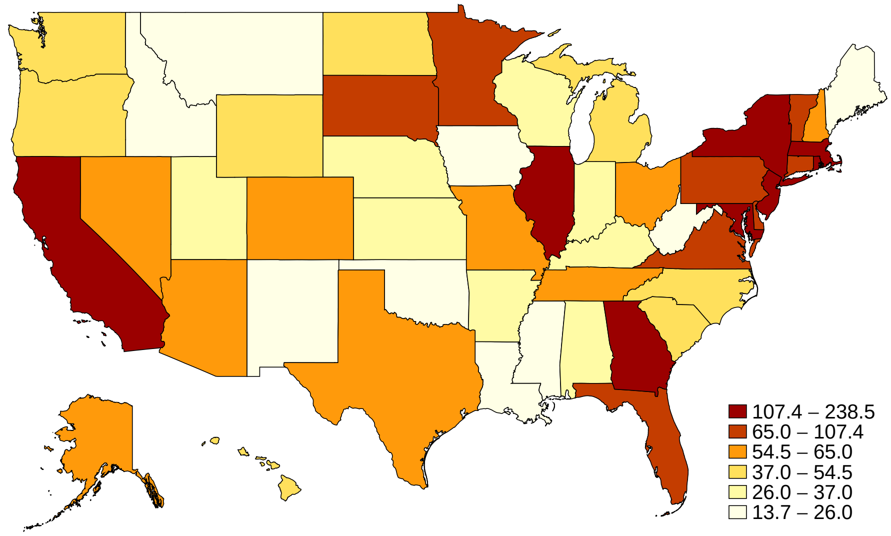
To construct a measure of expected earnings, we compute average earnings for each detailed occupation category using four years of pooled cross-section data (2012-2015). The sample used to compute the earnings measure is composed of workers between the ages of 24 and 30 with a Bachelors degree.

Table 4 also presents summary statistics (sample means and standard deviations) for each of the internship-occupation characteristics by major (2-Digit) SOC category. In some cases, there is substantial variation in internship characteristics within the major SOC category. But, in other cases, there is little to no variation. As an example, consider Management Occupations. Within this occupation category, there is variation in terms of service tasks, but there is no variation in social skills and nonroutine analytical tasks. Similar patterns in the data exist for other major SOC categories (e.g., computer and mathematics occupations, community and social occupations)[15].

[14] An implication of the lack of variation within major SOC categories is that including major SOC fixed effects in our regression models results in a high degree of collinearity among certain variables. Thus, standard errors increase substantially. As a result, our preferred set of estimates for the internship-occupation characteristics are based on between-SOC variation.

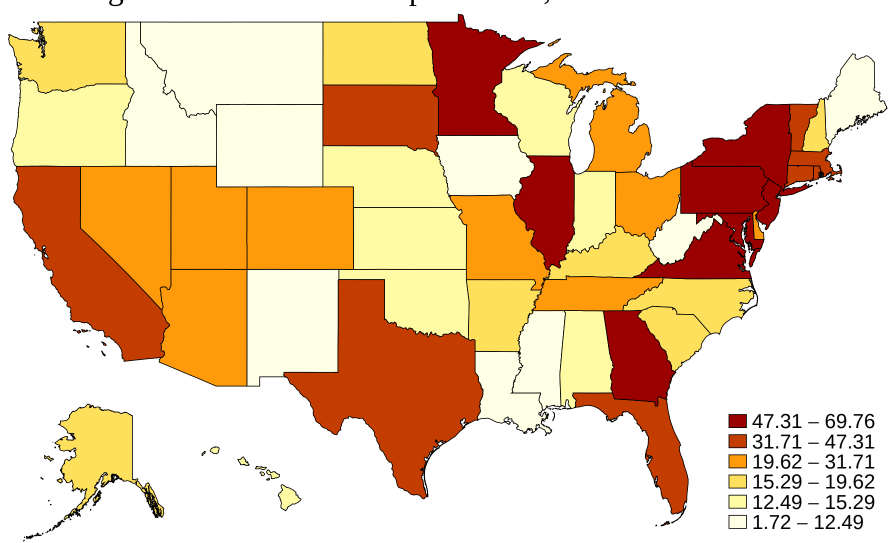
[15] An implication of the lack of variation within major SOC categories is that including major SOC fixed effects in our regression models results in a high degree of collinearity among certain variables. Thus, standard errors increase substantially.

Figure 1.1: Internships Per 100,000 19-24 Year Olds



Using the geographic locations of firms, we link the advertising firm to a metropolitan area (i.e. MSA), which enables us to examine the availability of internships across the U.S. Internships are geographically disperse, but are concentrated on the coasts and upper Midwest. In Figure 1.1, we show the number of internship per 100,000 of the population of 19-24 year olds by metropolitan area[16].

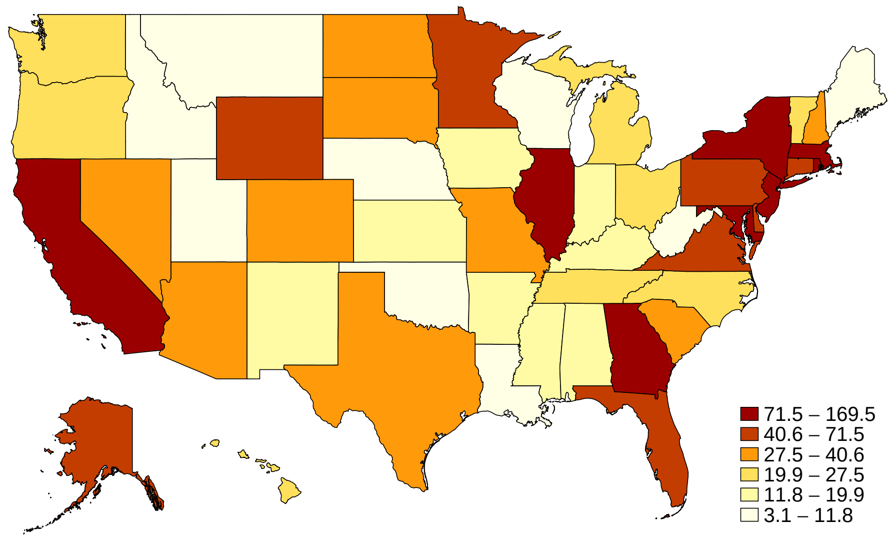
Figure 1.2: Paid Internships Per 100,000 19-24 Year Olds



[16]We use the Open Cage geocoder (http://geocoder.opencage.com, last seen 19 March 2017) through the opencagegeo Stata module to geocoded the addresses of the firms and obtained the state and county in which the firm is located. For the analysis below using the unemployment rate, we were able to match 100 percent of the county/state combinations to county level FIPS codes. Using a crosswalk from the National Bureau of Economic Research (NBER), we then match the county level FIPS codes to the area level FIPS codes (Core Based Statistical Areas, or CBSAs, and New England City and Town Areas, NECTAs). We are able to match 98.2 percent of county level FIPS codes to the corresponding CBSA or NECTA code

In figures 1.2 and 1.3 we show the distribution of paid and unpaid internships, respectively. We observe internship openings in all 50 states, but they are more prevalent in the mid-Atlantic states, California, Georgia. Paid internships are more common in Utah, Vermont, and Montana, while unpaid internships are more common in New York Geographic differences also exist in the number of paid and unpaid internship openings across states. For example, paid internships are more common in Utah, Vermont and Montana, while unpaid internships are more common New Jersey, New York and California.

Figure 1.3: Unpaid Internships Per 100,000 19-24 Year Olds



To examine the relationship between the regular labor market and the demand for internships, we use the precise date the advertisement was posted online and metropolitan area in which the advertising firm is located to link the internship ads to the corresponding metropolitan-area unemployment rate for the month/year in which the ads were published on the website[17]. For our sample period, which runs from November 2014 through November 2016, the average unemployment rate is 4.9 percent with a standard deviation of 0.9 percent. Assigning the area unemployment rates by the advertisement posting date generates variation within and between local labor markets.

[17]We match the FIPS code for each ad to the database of seasonally adjusted civilian labor force and unemployment variables from the Bureau of Labor Statistics. We were unable to match 420 observations to the corresponding unemployment rate, less than two percent of the original sample.

1.1.1 Determinants of Paid Status

The controversial issues surrounding unpaid internships makes an analysis of the determinants of paid-unpaid status particularly important for policy. Using an indicator variable for paid status as the dependent variable, we estimate the relationship between other ad characteristics, the characteristics of occupations that correspond to the internship openings, and metropolitan-area labor market conditions.

In particular, we study the relationship between particular skill demands at the firm level and paid-unpaid status, required skills at the occupation level Our analysis focuses on the impact of particular internship/occupation and environmental characteristics, such as internship attributes (skill demands, part-time/full-time status), a measure of expected earnings, the amount of on-the-job training required, the state of the labor market (loose or tight conditions), and particular skill/task demands (routine, nonroutine-analytical, social, and service tasks) at the occupation level. The regression model used to investigate the determinants of paid status is:

$$paid_{j,o,l,d} = \alpha + X_j'\beta + Z_o'\gamma + \delta unemp_{l,d} + \theta\overline{wage}_{o,l} + \phi_o + \phi_l + \varepsilon_{j,o,l,d}$$

The terms j, o, l, and d index ads, occupation categories (i.e. SOC codes), local labor markets, the month-year the ad was posted online, respectively. The variable paid equals one when an internship is paid and zero otherwise; X is a vector of ad-specific explanatory variables, including an indicator variable for part-time status and a set of dichotomous variables that proxy for other attributes and skills demanded by advertising firms (summarized in Table 3); Z is a vector of occupation-specific explanatory variables, which capture expected earnings and skill/task demands (routine, nonroutine-analytical, social, and service); unemp is the unemployment rate at the local-labor-market level at the posting date of the internship opening; $\overline{wage}$is the average wage in each occupation category measured at the metropolitan-area level; ϕ_o represents 2-digit SOC fixed effects; ϕ_l represents local-labor-market fixed effects; and ε represents factors not accounted for in equation (1) that predict the paid/unpaid status of the internships. The α, β, γ, δ and θ are parameters to be estimated.

We estimate a variety of alternatives models that necessitate that equation (1) be written differently. In the interest of space, equation (1) depicts the final specification, which includes right-hand-side explanatory variables at the firm, occupation, metropolitan area and occupation-metropolitan area, respectively, that may predict paid status. The estimates for equation (1) and its variations are presented in Table 4.

The estimates presented in columns 1-6 differ in a several ways. First, we focus on the impact of skills and attributes demanded by firms and local-labor-market conditions in the metropolitan area (columns 1 and 2). Second, we incorporate skill demands measured at the occupation level, the amount of on-the-job training required for the corresponding occupation, and inflation-adjusted earnings at the occupation*metropolitan-area level. Third, some specifications (a) omit metropolitan-area fixed effects (columns 1, 3, and 6), (b) omit 2-digit SOC fixed effects (columns 1, 2, 3 and 4), and include both metropolitan-area and 2-digit SOC fixed effects (column 5).

In columns 1 and 2, the ad is nested within a labor market. As such, we compute cluster-robust standard errors with one-way clustering on the metropolitan area. However, in columns 3-6, there are non-nested clusters across labor markets and occupations. In these instances, we compute cluster-robust standard errors with three-way clustering on both the metropolitan area, occupation and occupation-metropolitan area (Cameron, Gelbach and Miller 2011). The ways in which standard errors are adjusted is listed at the bottom of Table 1.5.

Table 1.5: Determinants of Paid Status

	All Internships				Part-Time	Full-Time
Explanatory Variable	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment	-0.0393***	-0.0328***	-0.0395***	-0.0329***	-0.0243**	-0.0648*
	(0.0127)	(0.0119)	(0.0126)	(0.0117)	(0.0117)	(0.0351)
OJT	-0.0067**	-0.0056**	-0.0257**	-0.0224**	-0.0201**	-0.0257**
	(0.0031)	(0.0027)	(0.0118)	(0.0099)	(0.0099)	(0.0122)
OJT-Squared	–	–	0.0008***	0.0007**	0.0006**	0.0006
	–	–	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Social	0.0082	-0.0008	0.0127	0.0033	-0.0430	0.0788
	(0.0583)	(0.0521)	(0.0558)	(0.0501)	(0.0502)	(0.0511)
Analytical	0.1264***	0.1128***	0.1518***	0.1354***	0.1187***	0.1493***
	(0.0351)	(0.0307)	(0.0349)	(0.0306)	(0.0307)	(0.0365)
Service	0.0410	0.0402	0.0286	0.0290	0.0378	0.0424
	(0.0563)	(0.0478)	(0.0546)	(0.0463)	(0.0463)	(0.0473)
Routine	0.0532	0.0608*	0.0512	0.0588	0.0449	0.0852*
	(0.0378)	(0.0348)	(0.0392)	(0.0361)	(0.0362)	(0.0457)
Natural Log of Real Earnings	0.1484	0.1670*	0.1486	0.1663*	0.1172	0.2840***
	(0.1114)	(0.0895)	(0.1080)	(0.0872)	(0.0872)	(0.0982)
Part-Time Status	-0.4110***	-0.3872***	-0.4090***	-0.3856***	–	–
	(0.0355)	(0.0294)	(0.0352)	(0.0035)	–	–
Area FIPS Code Fixed Effects Included?	No	Yes	No	Yes	Yes	Yes

The impact of a one-unit percentage point increase in the area unemployment rate triggers a roughly three percentage point decrease in the probability that an internship is paid, depending on the model specification. The models that include area FIPS fixed effects produce our preferred estimates, as these regressions are likely closer to the ceteris paribus effect of unemployment on paid status than the specifications in which area FIPS fixed effects are omitted. The estimated effects of unemployment in columns (2) and (4) indicate an approximate nine percent decline in the probability that an internship is paid when the area unemployment rate rises by one percentage point. The estimated effect of unemployment on paid status is statistically at the one-percent level in each specification.

The relationship between OJT and paid status is nuanced. In columns (1) and (2), a one-month increase in OJT is associated with a 0.5-0.7 percentage point lower probability of paid status. However, when a squared term of OJT is added to the model, we find a convex relationship between paid status and OJT: the probability than an internship is paid tends to fall at low levels of OJT, but the relationship turns positive around 17 months of OJT[18].

In terms of the skill demands, the only variable that is statistically significant is nonroutine analytical tasks. We find a 35 percent higher probability of paid status when an internship-occupation attaches a relatively high level of importance to nonroutine analytical skills (e.g., problem solving using mathematics). The remaining variables measuring different skills demands, i.e. social skills, routine tasks and service tasks, are positively related to paid status, but none of these variables are statistically significant at conventional levels. However, the effects of skill demands are estimated imprecisely, perhaps due to a high degree of collinearity between variables.

In terms of expected earnings, we find a positive relationship between an increase in expected earnings and paid status. Using the estimates from column (4), a 10 percent increase in expected earnings is associated with a 1.7 percentage point higher probability that an internship is paid. The statistical significance of the estimated effect of expected earnings on paid status is sensitive the inclusion of area FIPS fixed effects. When these fixed effects are included, the coefficient increases in size enough to make it statistically significant at the 10-percent level.

Lastly, a strong negative correlation exists between part-time and paid status. Part-time status is large (in absolute value), negative and highly statistically significant in all model specifications. Part-time internships are between 38 and 41 percentage points less likely to be paid.

1.2 Resume Audit of the Internship Market

1.2.1 Experimental Design

Our experiment follows the traditional correspondence audit framework, which has been used to examine employment discrimination on the basis of race (Bertrand and Mullainathan,

[18]However, 17 months of required OJT is rare in our sample. In fact, such an amount of OJT is above the 95 percentile. Thus, for values common in our sample, the marginal effect of OJT on paid status is negative.

2004; Nunley et al. 2015), age (Lahey 2008; Neumark, Burn and Button, 2015), disability status (Ameri et al. 2015), and country of origin (Oreopoulos 2011)[19]. Recently, correspondence audits have been used to study the impact of traditional labor-market variables, such as unemployment duration (Farber, Silverman and von Wachter 2017; Kroft, Lange and Notowidigdo 2013; Nunley et al. 2016b), quality of post-graduation employment (Farber, Silverman and von Wachter 2017; Nunley et al. 2016b), and college type (i.e. for-profit versus nonprofit) (Darolia et al. 2015), college majors and specific college programs (Deming et al. 2014; Nunley et al. 2016a). Our experiment extends the effort to use correspondence studies as a means to study labor-market outcomes by auditing the market for interns.

We submitted fictitious resumes to internship openings during the fall and spring semesters of the 2015-2016 academic year via a widely-used online internship search board[20]. Unlike traditional audit studies, we created online profiles for the fictive applicants, which were then reused during the semester in which resumes were submitted to internship postings[21]. Each profile was used to apply to 10 different internships and, as a result, the experiment has a clustered design. In particular, we created 576 profiles in each semester, with each profile forming a unique cluster along three dimensions: (i) applicant name, (ii) major field of study, and (iii) university. We used two distinctively white names (Wyatt Schmidt and Colin Johannson) and two distinctively black names (Darius Jackson and Xavier Washington)[22]. Each racially-distinct name was assigned six major fields of study (Biology, Economics, Business

[19]See Pager (2007), Bertrand and Duflo (2016) and Neumark and Rich (2016) for recent surveys of the audit literature.

[20]Our IRBs prohibit us from revealing the names of any firms or universities used in the experiment. Likewise, we are prohibited from providing information on the website used to submit fictive resumes

[21]Two audits of the real-estate rental market use online profiles of fictitious, prospective renters using ethnic-sounding names (Gaddis and Ghoshal 2015; Edelman, Luca and Svirsky 2015).

[22]In response to the critique provided in Charles and Guryuan (2011), it is important to use names that are common in the populations of both white and black babies. Because our fictive applicants would tend to be in their early-20s when apply for the internships, we used the 1984 Social Security Name Database, which provides rankings on the popularity of particular first names (overall, not by race). Babies born in 1984 would be 21-years-old in 2015 when our experiment began. We then searched for the distinctively black and white names listed in Levitt and Dubner (2005). Xavier and Darius are ranked 139 and 155, and Wyatt and Colin are ranked 124 and 197. These rankings indicate that the names chosen for our experiment were given to babies at similar rates. Thus, it is likely that the names chosen for our study would be perceived as common. The names, however, should signal race, particularly when combined with the racially distinct last names. The distinctively black last names are from the following website, which based its report on the 2000 Census: https://hailtoyou.wordpress.com/2010/11/24/the-blackest-surnames-in-the-usa/. The distinctively white last names are ranked in the top five of surnames in countries with large white populations (Germany and Sweden): http://worldnames.publicprofiler.org/SearchArea.aspx.

Administration, Marketing, Psychology, and English) at twenty-four, large, public universities that span the U.S., yielding 24 profiles at each of the 24 universities, for a total of 576. In effect, our data includes 1152 clusters because each name-major-university cluster is assigned a different randomly-generated resume in the fall and spring semesters[23]. Each resume included an email address to which prospective firms could respond.

We applied to 5,760 internship advertisements in both the spring and fall semesters of the 2015-2016 academic year, giving a total of 11,520 resumes submitted. We used the program created by Lahey and Beasley (2009) to randomize other attributes included on the resumes, including grade point average (GPA), work experience while in college, past internship experience, volunteer experience, and computer skills[24]. The combination of the name-major-university clusters and the random assignment of the resume attributes ensures that so-called template bias is avoided (Lahey and Beasely 2009).

To produce a random sample to which we could submit the fictive resumes, we followed a five-step process

1. Randomly select an internship category (marketing, research, or business) from which to choose an internship[25].

2. Randomly select a part-time or full-time internship.

3. Randomly select a paid or unpaid internship.

4. Randomly select the webpage the internship is to be chosen from[26].

[23]Assuming a cluster size of 10, our initial power calculations for a randomized control trial with a clustered design indicate that for a detectable difference of one-percentage point with 80 percent power, we would need 47 clusters per arm (94 clusters total).

[24]In Appendix Table A1, we present summary statistics for the resume attributes for all internships (column 1), unpaid internships (column 2), paid internships (column 3), part-time internships (column 4), and full-time internships. A comparison of the columns indicates balance in the random assignment of the resume attributes to the fictive applicants resumes across the different types of internships to which applications were submitted.

[25]We chose the marketing, research and business internship categories for two reasons. First, these categories included large numbers of both paid and unpaid internships as well as part- and full-time internships consistently across cities. Second, the experimental design allows us to study mismatch in qualifications, as it is common for business-related majors (business administration, economics and marketing) to intern in business-related fields.

[26]After completing steps 1-3, there are multiple numbered pages of internships available when the internship advertisements are displayed. Rather than select internships on the first page, we generate a random number that depends which page we choose the internship from.

5. Randomly select the internship on the webpage to which a resume would be submitted[27].

Random selection of internship postings also generates variation in the distance the applicant is from the employer. The internship advertisements to which we sent the fictive resumes were chosen at random in labor markets that were near the university (within roughly 0-500 miles), at an intermediate distance from the university (about 750-1250 miles), or far from the university (1250-2250 miles) of the resume in question.

After submitting a resume to a particular internship opening, we monitored the accounts for responses from employers. In general, we identified three types of employer responses: expressions of interest, interview requests, and location inquiries. In an effort to streamline the main findings from the experiment, we create an outcome variable that combines both expressions of interest and interview requests into one variable. We refer to this outcome as a positive response. We use the positive-response variable as primary outcome variable in the analysis of employer responses.

Table 1.6: Sample Means and Standard Deviations for Different Employer Responses

	Type of Internship				
	All Internships	Unpaid Internships	Paid Internships	Part-Time Internships	Full-Time Internships
	(1)	(2)	(3)	(4)	(5)
Positive Response	0.0595	0.0791	0.0403	0.0565	0.0624
	(0.2365)	(0.2699)	(0.1967)	(0.2309)	(0.2420)
Expresses Interest	0.0291	0.0360	0.0223	0.0352	0.0230
	(0.1680)	(0.1864)	(0.1476)	(0.1844)	(0.1499)
Interview Opportunity	0.0304	0.0431	0.0180	0.0272	0.0335
	(0.1716)	(0.2030)	(0.1330)	(0.1672)	(0.1800)
Location Inquiry	0.0093	0.0144	0.0043	0.0096	0.0090
	(0.0959)	(0.1192)	(0.0653)	(0.0975)	(0.0944)

We present summary statistics in Table 1.6 for the different types of employer responses, including the positive response rate (our primary dependent variable), expressions of interest, interview opportunities, and location inquiries. The employer response rates, regardless of the type of employer response, are higher for unpaid internships than they are for paid internships

[27]After step 4 is completed, the webpage displays 10 internship advertisements. Thus, the random number generated is between 1 and 10. For example, a 1 indicates that we will submit a resume to the first internship advertisement displayed.

(columns 2 and 3). No clear pattern in terms of employer responses is present between part- and full-time internships (columns 4 and 5).

1.2.2 Classifying Internships into Occupation Categories and Combining with Other Data Sources

Ninety-six percent (or 11,054) of the internships in our sample were assigned an 8-digit O*NET-SOC code via the O*NET-SOC AutoCoder. In Table 1.7, we present the percentage of resumes submitted to internship openings classified into different occupation categories for all internships (column 1). Similar to the ad data, we submit fictive resumes disproportionately to internships classified in the following occupation categories: (i) business and finance, (ii) arts, design, entertainment, sports and media, (iii) sales, and (iv) office and administration. In fact, resumes submitted to internships in these occupations represent over 80 percent of our sample.

Table 1.7: Percentage of Resumes Submitted to Internship Advertisements by Occupational Classification

		Internship Category			Type of Internship			
	All Types	Research	Marketing	Business	Paid	Unpaid	Part Time	Full Time
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Management (11)	5.3%	3.8%	5.6%	6.6%	6.2%	4.4%	5.0%	5.7%
Business and Finance (13)	34.5%	42.8%	30.8%	29.9%	34.1%	34.9%	36.5%	32.5%
Computer and Mathematical (15)	2.5%	2.3%	2.3%	2.9%	2.7%	2.3%	2.4%	2.6%
Community and Social Services (21)	1.6%	1.9%	1.3%	1.7%	0.8%	2.5%	1.5%	1.8%
Arts, Design, Entertainment,Sports and Media (27)	23.2%	21.4%	29.3%	18.8%	16.1%	30.5%	25.9%	20.5%
Sales (41)	17.3%	12.6%	18.0%	21.2%	25.2%	9.1%	13.6%	21.0%
Office and Administration (43)	8.5%	7.4%	6.6%	11.6%	8.3%	8.8%	7.9%	9.2%
Other	3.0%	4.8%	1.3%	2.8%	2.3%	3.6%	3.1%	2.8%
Unclassified	4.1%	2.8%	4.8%	4.5%	4.3%	3.8%	4.3%	3.8%
Number of Resumes Submitted	11,520	3,821	3,878	3,821	5,832	5,688	5,786	5,734

Columns (2), (3), and (4) illustrate that grouping internships by the three general categories (research, marketing and business) chosen as a part of the experimental design results in misclassification, as the internships in each general category span several of the occupation categories. For example, the Research category in column (2) comprises 42.8 percent in business and finance occupations, 21.4 percent in arts, design, entertainment, sports and media occupations, and 12.6 percent in sales occupations. Comparing column (5), paid internships, to column (6), unpaid internships, indicates that, for the most part, the share of internships is evenly split between paid and unpaid status within most of the major SOCs, except for (a) arts, design, entertainment, sports and media occupations and (b) sales occupations (Columns 5 and 6). The former category is disproportionately unpaid, while the latter category is disproportionately paid. With regard to part-time (column (7)) and full-time (column (8)) status, internships that map to business and finance occupations and arts, design, entertainment, sports and media occupations tend to be part time, while those that map to sales occupations tend to be full time. As a result, it could be that full-time internships are more likely to be paid than those that are part time[28]. The internships in the remaining categories are about evenly split between part-time and full-time statuses (Columns 7 and 8).

As noted in Section I, the O*NET-SOC AutoCoder assigns, in addition to the detailed SOC code, an occupation-match score, which is a measure of classification accuracy. The occupation-match score for the resume audit part of the study is shown in Figure A.5. From the plot, the scores range from low (50s) to high (90s) values, but many of the internship-occupation matches received scores in the 90s. The average score is slightly above 76, with a standard deviation of almost 12. The 25th percentile is 68; median is 78; and 75th percentile is 87. The O*NET-SOC Autocoder classified around 70 percent of the internships in our sample to occupations with scores in excess of 70.

Identical to our analysis of the web-scraped ad data, we combine data from the experiment, ACS and O*NET by merging on the detailed 6-Digit SOC code[29]. The combination of these data sources allows us to explore the impact of internship characteristics on employers

[28]The randomization process used to choose internship openings to submit resumes precludes us from examining the relationship between unpaid/paid status and part-time/full-time status, as we randomly determined whether to select a part-time or full-time internship and then randomly determined whether to select an unpaid or paid internship.

[29]When merging at the 6-digit level, we are able to match 100 percent of the internship openings to a detailed occupation code (excluding the 466 internships that could not be assigned a O*NET-SOC code). When combining data from the experiment with the O*NET data, we are able to match around 85 percent of the internships in our sample and the data from O*NET at the 8-digit level. When combining these data sources at the 6-digit level, we are able to match approximately 99 percent of the internships to the 6-digit SOC codes. Thus, the data set used for the estimates presented in Section 6 combine the experimental, ACS and O*NET data sets at the 6-digit level.

responses in ways that would not otherwise be possible. We use the classifications of the internships into detailed SOC codes to compute measures of expected earnings associated with the internships to which we send applications. We are also able to determine the characteristics (e.g., on-the-job training, skills demands) associated with the occupations that correspond to the internships in our sample. Lastly, the detailed SOC codes obtained from the machine-learning algorithm combined with data from the ACS provide a unique opportunity to assess external validity, as we are able to determine whether young men tend to work in occupations that correspond to the internships in our sample.

1.2.3 Empirical Framework

The baseline regression model used to analyze positive employer responses (i.e. callback) is:

$$positive_{i,j,o,l} = \alpha + X_j^{I'}\beta + X_j^{II'}\beta + X_o^{III'}\delta + \gamma unemp_l + \theta e\bar{a}rn_{o,l} + \varepsilon_{i,j,o,l}$$

The subscripts i, j, o and l index applicants, internship ads, occupation category to which the ad is assigned (i.e. 6-Digit SOC), and labor market (i.e. MSA) in which the advertising firm is located, respectively.

The variable callback equals one when an applicant receives a positive response (i.e. employer expresses interest or requests an interview) and zero otherwise[30]. X is a set of resume attributes assigned to the applicant, which includes dummy variables for (a) racially-distinct applicant names, (b) the university where the applicant is enrolled, (c) different major fields of study, (d) different grade point averages, (e) different types of work experience accumulated while the applicant is completing their degree, and (f) past internship experience; Z_j is a set of internship variables, which includes dummy variables for whether the internship is paid or unpaid and full time or part time; and $\epsilon_{i,j}$ represents variables that affect the dependent variable

[30]We were able to code alternative dependent variables, including whether the intern employer inquired about location. In the Appendix, we will present estimates for all dependent variables, but we chose to present positive responses in the manuscript due to the literatures reliance on positive responses as the primary dependent variable of interest.

not held constant; ϕ_s is a set of occupation-category fixed effects[31]; Z^I is a set of variables capturing internship characteristics, including dummy variables for whether the internship is paid or unpaid and full time or part time as well as a set of indicator variables that equal one when a particular word or phrase is used in the ads (See previous section for details; our approach, again, follows Deming and Kahn (2017)); and $\epsilon_{i,j}$ represents variables that affect the dependent variable not held constant; and α is the intercept term, and β and γ are the parameters of interest. In the process of analyzing the data, we augment equation (1) by including various interaction terms between different resume attributes, different internship characteristics, as well as resume and internship characteristics. Equation (3) is estimated as a linear probability model. The standard errors associated with estimates from all regression models (except when otherwise noted) presented in Section 6 are clustered at the name-major-university-semester level.

1.2.4 Determinants of Internship Opportunities

Because we regress positive employer responses on a large number of applicant and internship characteristics, we present the estimates from equation (3) in a series of different tables: Tables 7a-7d. The estimates presented in these tables are from the same regression model. In Table 7a, we examine the impact of part-time and unpaid status on positive employer responses. The overall estimated positive response gap between part- and full-time internships is negative but quantitatively small. The small overall estimate is driven by a 1.4 percentage point positive differential for unpaid internships (column 2) and a 1.7 percentage point negative differential for paid internships (column 3). The estimated difference in the positive response gap between part- and full-time status for paid internships is statistically significant at the five-percent level.

Employers are more likely to respond to an applicant for an unpaid internship. In column (1), we show that applicants who apply to unpaid internships are 3.4 percentage points (57 percent in terms of probability) more likely to receive positive responses than those who apply to paid internships. These findings are robust for both part- and full-time internships.

[31]Note that occupation-level fixed effects comprise a full set of dummy variables for the six-digit SOC code ascribed to the internships.

The percentage-point differential is larger for part-time internships than it is for full-time internships, however, and the difference between part- and full-time internships is statistically different from zero.

Table 1.8: The Effects of Part-Time and Unpaid Status on Positive Response Rates

	All Internships (1)	Unpaid Internships (2)	Paid Internships (3)	Part-Time Internships (4)	Full-Time Internships (5)
Part Time	-0.0006	0.0136	-0.0170**	–	–
	(0.0052)	(0.0088)	(0.0070)	–	–
Unpaid	0.0340***	–	–	0.0474***	0.0204**
	(0.0054)	–	–	(0.0079)	(0.0088)
R2	0.1645	0.2271	0.1869	0.2323	0.1977
N	11,520	5,471	5,583	5,539	5,515

In Table 7b, we present estimates the impact of major fields of study and academic ability (proxied by different grade point averages) in Panels A and B, respectively. In columns 1 and 2, we find that the majors are jointly statistically significant, implying statistical differences in positive-response rates between the different majors. The joint test is statistically statistical significance at the 10-percent level, although majors are jointly not statistically different from zero for the subsamples presented in columns 3, 4 and 5.[32] Overall, we find little empirical evidence for a relationship between majors and internship opportunities.

When examining the full sample (column 1), a subsample of unpaid internships (column 2) and a subsample of full-time internships (column 3), applicants who report a 3.8 or 4.0 GPA

[32]We conducted a number of robustness checks regarding the null effects (in general) detected for the particular majors. In one sensitivity check, we combined the three business-related majors into one category and the non-business-related majors into another category. Such a grouping also leads to the conclusion support no relationship between business degrees and internship opportunities. This finding holds when the sample is restricted to include only internships belonging to SOCs in which business majors represent the largest share (relative to the other majors in the experiment) of workers. It is the case, however, that we are unable to estimate more detailed comparisons (involving interaction effects) due to small sample sizes.

are more likely than applicants who report a 3.0 or 3.2 GPA in an economic and statistical sense. The estimated difference between applicants who report 3.4 or 3.6 GPAs and applicants who report 3.0 or 3.2 GPAs is positive, but the estimated differences are not statistically significant. Moreover, the F-test for joint exclusion of the GPA indicators indicates statistical significance for all internships (column 1) and full-time internships (column 5), but not unpaid (column 2), paid (column 3) or part-time internships (column 4). In sum, there is some evidence that suggesting having a very high GPA improves internship opportunities, but the overall evidence suggests that GPA is not an overly important determinant of internship access.

Table 1.9: The Effects of Major Field of Study and GPA on Positive Responses

	All Internships (1)	Unpaid Internships (2)	Paid Internships (3)	Part-Time Internships (4)	Full-Time Internships (5)
Panel A: Major Field of Study					
Economics	-0.0061	0.0041	-0.0187**	-0.0027	-0.0090
	(0.0079)	(0.0126)	(0.0091)	(0.0108)	(0.0111)
Marketing	0.0001	0.0062	-0.0011	-0.0081	0.0060
	(0.0083)	(0.0125)	(0.0105)	(0.0115)	(0.0113)
Psychology	-0.0085	-0.0060	-0.0088	-0.0104	-0.0047
	(0.0080)	(0.0127)	(0.0095)	(0.0104)	(0.0118)
Business Administration	0.0095	0.0199	0.0010	0.0026	0.0181
	(0.0084)	(0.0138)	(0.0098)	(0.0108)	(0.0120)
English	0.0112	0.0332**	-0.0053	0.0111	0.0098
	(0.0085)	(0.0141)	(0.0094)	(0.0119)	(0.0121)
P-value for F-test	0.0808*	0.0678*	0.1782	0.4769	0.1452
Panel B: Grade Point Average					
High GPA (3.8 or 4.0)	0.0130**	0.0164*	0.0101	0.0108	0.0176**
	(0.0056)	(0.0095)	(0.0067)	(0.0079)	(0.0081)
Medium GPA (3.4 or 3.6)	0.0082	0.0123	0.0016	0.0074	0.0088
	(0.0058)	(0.0094)	(0.0069)	(0.0081)	(0.0080)
P-value for F-Test	0.066*	0.1951	0.2807	0.3731	0.0981*
R2	0.1527	0.2122	0.1687	0.2191	0.1810
Observations	11,520	5,688	5,832	5,786	5,734

Table 1.10: The Effects of Past Internships, Volunteer Work, and Work Experience on Positive Responses

	All Internships (1)	Unpaid Internships (2)	Paid Internships (3)	Part-Time Internships (4)	Full-Time Internships (5)
Panel A: Volunteer Experience					
Volunteer Job #1	0.0039	0.0010	0.0073	0.0087	0.0034
	(0.0071)	(0.0114)	(0.0083)	(0.0102)	(0.0098)
Volunteer Job #2	-0.0048	-0.0108	-0.0037	-0.0089	-0.0030
	(0.0067)	(0.0112)	(0.0077)	(0.0092)	(0.0090)
Volunteer Job #3	0.0026	0.0005	0.0059	-0.0020	0.0096
	(0.0064)	(0.0103)	(0.0082)	(0.0087)	(0.0096)
p-value for F-test	0.7339	0.7727	0.6490	0.5123	0.7102
Panel B: Work Experience During College					
Outside Sales	0.0062	0.0116	-0.0034	0.0132	0.0008
	(0.0087)	(0.0136)	(0.0105)	(0.0112)	(0.0123)
Retail Clothing	-0.0031	-0.0090	-0.0010	0.0011	-0.0066
	(0.0076)	(0.0121)	(0.0094)	(0.0107)	(0.0112)
Server in Restaurant	0.0059	0.0112	-0.0065	0.0147	-0.0042
	(0.0084)	(0.0133)	(0.0104)	(0.0124)	(0.0115)
Campus Recreation	0.0111	0.0175	-0.0020	0.0292**	-0.0025
	(0.0084)	(0.0131)	(0.0098)	(0.0114)	(0.0116)
University Dining Services	-0.0036	-0.0013	-0.0103	0.0149	-0.0226**
	(0.0079)	(0.0125)	(0.0097)	(0.0111)	(0.0114)
p-value for F-test	0.3683	0.3250	0.8873	0.1063	0.3235
Panel C: Past Internship Experience					
Research Internship	0.0128*	0.0152	0.0087	0.0203**	0.0043
	(0.0070)	(0.0114)	(0.0076)	(0.0097)	(0.0097)
Finance Internship	0.0136**	0.0209**	0.0100	0.0304***	0.0051
	(0.0069)	(0.0105)	(0.0083)	(0.0094)	(0.0093)
Insurance Internship	0.0126*	0.0208*	0.0083	0.0288***	0.0026
	(0.0068)	(0.0113)	(0.0078)	(0.0097)	(0.0089)
p-value for F-test	0.0538*	0.0912*	0.4260	0.0008***	0.9322
R2	0.1527	0.2122	0.1687	0.2191	0.1810
Observations	11,520	5,688	5,832	5,786	5,734

In Table 7c, we examine the impact of volunteer experience (Panel A), work experience during college (Panel B) and past internship experience (Panel C) on positive responses from employers. For volunteer experience (Panel A), the comparison group is applicants with no volunteer experience. The estimates for different volunteer jobs are not statistically different from zero on an individual basis as well as jointly. The same is true work experience obtained while working toward ones degree (Panel B). The particular college jobs are, in general, not statistically significant on an individual basis, and the different jobs are not jointly statistically significant either. Past internship experience (Panel C) is statistically significant predictor of internship opportunities for particular types of internships: unpaid and part time. The estimated differences between applicants with past internship experience and those without past internship experience in column 2 and column 4 drive the overall estimates in column 1. Past internship experience is not a strong predictor of internship access for paid (column 3) and full-time (column 5) internships. In sum, volunteer or work experience obtaining during college are not important predictors of internship opportunities. However, interning prior to applying for a subsequent internship aids in internship opportunities for those that are unpaid or part time.

Per our experimental design, applicants were assigned white- and black-sounding first and last/family names. In Table 7d, we examine how race (Panel A) and distance from an employer (Panel B) affects positive-response rates for all internships (column 1), unpaid internships (column 2), paid internships (column 3), part-time internships (column 4) and full-time internships (column 5). From Panel A, we find strong evidence that black-named applicants receive fewer positive response than white-named applicants. The estimated black-white differences in positive-response rates range approximately from 20-30 percent in terms of probability, depending the type of internships. The estimated differentials are highly statistically significant in most cases, except that for full-time internships (column 5).

In Panel B, we examine the impact of an applicants distance from an employer on positive-response rates. The base group for the estimates presented in columns 1-5 is applicants who live less than 250 miles from an employer. A general pattern emerges across the different types of internships: On average, intern employers are less likely to respond positively to an applicant who lives (relatively) far away from the companys location.

Table 1.11: The Effects of Race and Distance from Intern Employer on Positive Response

	All Internships (1)	Unpaid Internships (2)	Paid Internships (3)	Part-Time Internships (4)	Full-Time Internships (5)
Panel A: Race					
Black	-0.0161***	-0.0217***	-0.0124**	-0.0224***	-0.0119*
	(0.0047)	(0.0078)	(0.0055)	(0.0064)	(0.0065)
Panel B: Distance					
250-500 Miles Away from Employer	-0.0121	-0.0057	-0.0146	-0.0013	-0.0156
	(0.0096)	(0.0155)	(0.0114)	(0.0133)	(0.0135)
500 or More Miles from Employer	-0.0295***	-0.0277**	-0.0291***	-0.0226**	-0.0353***
	(0.0077)	(0.0130)	(0.0096)	(0.0101)	(0.0118)
R2	0.1527	0.2122	0.1687	0.2191	0.1810
Observations	11,520	5,688	5,832	5,786	5,734

In Table 8, we examine more closely racial differences in positive-response rates between applicants using different dependent variables. The findings for the black-white difference indicate two noteworthy results: overall discrimination, using positive responses as the dependent variable, is driven by black-named applicants receiving far fewer interview opportunities. Interestingly, black-named applicants were more likely to receive a location inquiry from intern employers. In terms of probability, black applicants are over 50 percent more likely to be asked about their location than white applicants.

In Panel B, we conduct head-to-head comparisons between white- and black-named applicants living particular distances from employers. What becomes clear from the estimates is that racial discrimination is greater for applicants living 500 or more miles away from employers. In fact, the greater discrimination against black applicants living relatively far away from employers is a primary driver of the overall findings (compare to estimates in Panel A).

Table 1.12: Racial Discrimination, Distance from Employer, and Various Response Rates

	Dependent Variable			
	Positive Response	Employer Expresses Interest	Employer Offers Interview	Employer Inquires About Location
	(1)	(2)	(3)	(4)
Panel A: Baseline				
Black	-0.0160***	-0.0062*	-0.0099***	0.0046***
	(0.0047)	(0.0033)	(0.0034)	(0.0017)
Panel B: Interactions with Distance				
Racial Gap for Applicants Less than 500 Miles from Employer	-0.0107	-0.0030	-0.0078	0.0039* (0.0023)
	(0.0071)	(0.0050)	(0.0053)	(0.0023)
Racial Gap for Applicants 500 Miles or Further Away from Employer	-0.0208***	-0.0090**	-0.0118***	0.0052**
	(0.0060)	(0.0041)	(0.0045)	(0.0024)
Observations	11,520	11,520	11,520	11,520

Overall, we find that intern employers discriminate more extensively when applicants 500 or more miles from the companies. While there is no definitive test for the different types of discrimination, we view these findings in the following ways. The greater discriminatory behavior against black applicants residing relatively further away from the companys location is suggestive of statistical discrimination. A possible way through the statistical discrimination occurs is that employers may view race as an indicator of socioeconomic status, and discriminate on the basis of socioeconomic status in lieu of race. Such an interpretation supports the notion that discrimination is rooted in incomplete information (Aigner and Cain 1977; Arrow 1973; Phelps 1972), rather than animus (Becker 2010).

Implicit bias is a widely cited channel for discrimination to occur (Bertrand, Chugh and Mullainathan (2005). Although Price and Wolfers and Rooth (2010) are notable exceptions, resume audits have struggled to deflect concerns that the estimates are driven by unconscious discrimination in lieu of conscious discrimination. For similar reasoning, the greater discrimination against black applicants living further away from the firm is indicative of a conscious decision, as employers are tying race and distance from internship and discriminating to a greater extent.

Table 1.13: Expected Earnings, Paid and Unpaid Internships, and Positive Response Rates

	(1)	(2)
Panel A: All Internships		
Between Top Quartile and Middle of Distribution	-0.0109	-0.0192
	(0.0109)	(0.0193)
Between Bottom Quartile and Middle of Distribution	0.0162*	0.0187*
	(0.0092)	(0.0110)
Between Top Quartile versus Bottom Quartile	-0.0271**	-0.0378**
	(0.0116)	(0.0193)
Panel B: Unpaid versus Paid Internships		
Within Top Quartile	0.0152	0.0165
	(0.0116)	(0.0118)
Within Middle of Distribution	0.0352***	0.0354***
	(0.0065)	(0.0065)
Within Bottom Quartile	0.0507***	0.0489***
	(0.0117)	(0.0121)
SOC-Level Fixed Effects:		
Major (2 digit)	Yes	No
Minor (3 digit)	No	Yes

Compared to internships related to occupations in the middle of the earnings distribution, applicants who apply to internships ascribed to occupations in the top quartile of the earnings distribution are less likely to receive positive responses, but the estimated differentials are not statistically significant. By contrast, applicants who apply to internships ascribed to occupations in the bottom quartile of the earnings distribution are more likely to receive positive responses, and the estimated differential is statistically significant at the 10-percent level. Lastly, when comparing internships ascribed to occupations in the top and bottom quartiles of the earnings distribution, applicants applying to those in the top quartile are about 50 percent less likely to receive a positive response than those applying to those in the bottom quartile.

For internships that map to occupations in the top quartile of the earnings distribution, applicants who apply to unpaid internships (relative to paid ones) are more likely to receive positive responses. The estimated differential is, however, not statistically significant. By contrast, for internships that correspond to occupations in the middle of the earnings distribution, applicants who apply to unpaid internships are about 3.5 percentage points more likely to receive positive responses, which translates into an over 60 percent higher positive response rate. Likewise, for internships that correspond to occupations in the bottom quartile of the earnings distribution, applicants who apply to unpaid internships are about five percentage points more likely to receive positive responses, which translates into an over 85 percent higher positive response rate.

1.2.5 External Validity

Because there is very little data on internships in the U.S., it is difficult to determine the extent to which our experiment is valid in an external sense. However, a recent survey shows that employers generally find interns through open applications rather than through university placement offices[33]. Combining data from the American Community Survey (ACS) on employment outcomes with our experimental data provides a useful way to assess external validity. First, linking the experimental and nationally-representative survey data sources allows us to determine what share of young men are working in the occupation categories that correspond to the internships in our sample. If large shares of young men are working in occupations that correspond to our internships, such a finding is suggestive of a high degree of external validity, given that employers prefer to hire workers who obtain relevant experience and internships offer opportunities for college students to obtain such work experience (Cappelli 2014).

We also determine whether workers who majored in the fields of study used in our experiment constitute relatively large or small shares of the workers employed in a particular occupation category. Documenting whether the majors used in our experiment are commonly

[33] See https://www.naceweb.org/uploadedfiles/content/static-assets/downloads/executive-summary/2016-internship-co-op-survey-executive-summary.pdf (last seen 30 October 2016).

held by workers employed in those occupations lends support to the external validity of the experimental data.

Table 1.14: Percentages Employed by Major Field of Study and Occupational Category

		Major Fields of Study Used in Experiment					
	All Majors	Business	Biology	Economics	English	Marketing	Psychology
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Management (11)	11.3%	5.9%	2.4%	4.9%	1.8%	5.6%	2.2%
Business and Finance (13)	11.4%	7.2%	1.6%	7.9%	1.2%	3.7%	1.9%
Computer and Mathematical (15)	10.1%	2.2%	1.3%	3.4%	1.1%	1.3%	1.1%
Community and Social Services (21)	1.8%	2.0%	2.2%	1.5%	2.1%	1.3%	12.0%
Arts, Design, Entertainment,Sports and Media (27)	4.9%	1.9%	0.9%	1.4%	4.7%	1.9%	1.4%
Sales (41)	12.6%	8.4%	2.6%	4.2%	2.1%	7.8%	2.9%
Office and Administration (43)	9.4%	6.3%	2.8%	4.2%	3.6%	3.7%	3.7%
Other	38.5%	3.4%	6.7%	1.8%	2.0%	1.6%	3.1%

In Table 10, column (1), we show the share of employment in the 2011-2014 ACS for men aged 21-29 with Bachelors degrees across seven major SOC codes of the internships in the experiment, which represent around 94 percent of our sample. Employment among young, college-educated men in these occupational categories represents approximately 62 percent of total employment for this sample of workers[34]. In columns (2) through (7) of Table 10, we also present the share of workers with the six majors used in our experiment who are employed in each major SOC category. Among the majors used in our experiment, business, economics and marketing, which could be grouped as business-type majors, are disproportionately represented in the following major SOCs: management, business and finance, sales, and office and administration. Workers who majored in psychology are employed at greater rates in the community-and-social-services category, and workers who majored in English are disproportionately employed in the arts, design, entertainment, sports and media category. In the computer-and-mathematical category, economics majors are the most likely to be employed (relative to the other majors in our experiment). Taken together, the data compiled from the

[34]The notes at the bottom of Table 2 provide details on how the sample is constructed. We tend to find similar patterns in the data when alternative samples are used (e.g., when women are included).

ACS suggests large shares of young men work in the occupations that correspond to our internships and the majors assigned to our fictive applicants represent nontrivial shares of the workers employed in the occupation categories.

1.3 Discussion and Conclusion

To our knowledge, this is the first paper to examine the market for internships in the United States. We estimate the relationship between internship characteristics, individual characteristics, and positive responses to internship applications. By linking internships to data external to our experiment, we are able to more fully explore the relationship between (a) different internship characteristics (b) internship characteristics and employer responses, (c) resume attributes and employer responses and (d) interaction effects between internship characteristics and resume attributes.

Internships with downstream occupations that have short periods of training are more likely to be paid. Holding all else equal, when internships are paid, employers are less likely to contact the applicants.However, the differential we observe in response rates between paid and unpaid internships is driven by downstream occupational wage differences, which are tangentially related to internship remuneration.

We find that the majority of internships advertised are unpaid and that there is a strong positive relationship between paid status and and full time status. The availability of internships in general and paid/unpaid internships in particular varies widely across states, with more populated states such as California and New York having larger number of internships. However, these internships are disproportionately unpaid, while less populated states such as Utah and Vermont tend to have more paid than unpaid internships.

In terms of the econometric evidence based on the advertisement data, we find that a one-unit increase in the unemployment rate triggers a nine percent reduction in the probability that an internship is paid. The relationship between paid status and months of on-the-job training (OJT) required is convex in its shape; that is, the probability that an internship is paid falls at low levels of OJT but begins to increase around 17 months of required OJT. We find no statistical evidence linking particular skills demands, such as social skills, routine tasks or service skills, to the paid status of internships. The exception is nonroutine analytical skills, which is associated with a 35 percent higher probability of paid status. The evidence for expected earnings indicates a positive relationship: the probability of paid status is around one percentage point higher when expected earnings rises by five percent.

The econometric evidence from the resume audit reveals a number of noteworthy findings. Unpaid internships have higher positive employer response rates than paid internships, suggesting that firms may need to cast a wider net to fill unpaid internships. However, the employer response gap between paid and unpaid internships is eliminated for internships that correspond to high-paying occupations. Part-time internships carry similar response rates to full-time internships for unpaid internships, but they have lower response rates for paid internships. We tend to find no effects for major field of study, volunteer experience, and college work experience. We find limited statistical evidence that having a high grade point average or working as an intern in the past improves internship opportunities. The only resume attributes that have robust effects on employer responses are race and distance from employers. Black applicants receive approximately 27 percent fewer positive responses than their white counterparts. Applicants living 500 miles or more away from employers are about 50 percent less likely to receive positive responses. The overall racial difference in positive responses rates is, however, driven entirely by greater discrimination against black-named applicants living 500 or more miles away from employers. This finding is supported by an analysis of the text from employer responses in which black-named candidates are more likely to receive a request for information concerning their current location.

This paper provides the first initial view of the internship market within the United States. However, there are areas of research within this market that have yet to be explored. Limitations in the data set do not allows us to follow the outcome of applicants post-interview, and our measure of downstream wages is a proxy for actual wage. More descriptive longitudinal survey data following subjects over time including questions on internship experience will help characterize the causal impact of internships on future wages.

Chapter 2

Skill Demand for Virtual Internships

2.1 Introduction

As the labor market for skilled jobs becomes more competitive, students are often put in a position to have more specific rather than general labor market experience (Finch, et al. 2013; Cappelli 2014; Nunley, et al. 2016a; Nunley, et al. 2016b). For college graduates, this specific work experience has historically been provided via internships in a respective area of interest. The National Association of Colleges and Employers defines an internship as " a form of experiential learning that integrates knowledge and theory learned in the classroom with practical application and skills development in a professional setting. Internships give students the opportunity to gain valuable applied experience and make connections in professional fields they are considering for career paths; and give employers the opportunity to guide and evaluate talent." [1]

The 1992 lindquist-Endicott report's survey of 258 corporations indicated that 17% of incoming entry level workers had previously been interns. In addition, 37% of interns were offered jobs upon completion with 49% of interns accepting the invitation. In contrast to today's labor market, a 2016 survey of 271 corporations by the National Association of Colleges and Employers indicates that over 60% of students have taken part in an internship, with 72.7% being offered full time employment. Of these 72.7%, over 85% had accepted the invitation. This survey data seems to indicate that the internship is increasingly becoming the pathway to smooth the school-to-work transition.

Despite the increasing completion rate of internships in undergraduate, the literature on the subject remains thin. Jaeger et. al (2017) studies the determinants of internship opportunities and the skill demands for a wide variety of internship opportunities using experimental

[1] See more at: $http://www.naceweb.org/about-us/advocacy/position-statements/position-statement-us-internships$

data and information from job advertisements scraped from a large online job board. However, experimental studies do not isolate the causal impact of an internship on future worker outcomes. Studies using observational data on summer and gap year internships from German Universities indicate that completion of an internship leads to nearly a six percent return in future wages on average (Saniter and Siedler 2014). This is in conflict with studies analyzing post-graduation internships which indicate a negative return on future monthly earning (Harms 2015). However, such studies which focus on internships have exclusively concentrated on location based internships (i.e. internships in which you must be physically present). This leaves a large gap in the literature with respect to virtual internships (i.e. internships in which you must *not* be physically present).

Research on virtual internships is of growing importance as telecommuting in computer and science related positions has increased by 69% from 2000-2010 (Mateyka et al 2012), with data from the American Community Survey indicating increases of over 100% from 2005-2014 across all occupations[2]. The increases in telecommuting positions are not limited to those in science and technology. Mateyka (2012) also find that nearly one quarter of all telecommuting positions were in the fields of finance, business, and management.

As the workforce of the United States continues to transition to a higher percentage of workers who telecommute, it necessary to document the characteristics of these positions and the underlying skills which are demanded from employers. This paper describes the current demand for virtual interns and the associated skill sets. Using a large database of web scraped internship job advertisements, I follow an analysis similar to what is performed by Deming and Kahn (2016) to study the heterogeneity in skills demands for virtual internships. Parsing of the text data from the job advertisements allows the isolation of several different skill types, including: cognitive skills, social skills, management skills, customer service skills, writing skills, financial knowledge, coding skills, general computer skills, and key phrases which demonstrate sound character. In addition, I link the job advertisements to CBSA level unemployment rates using the date and location information from the job posting. Exploiting the

[2]refer to http://globalworkplaceanalytics.com/telecommuting-statistics for detailed analysis of ACS telecommuting data

variation in unemployment rates over time, I study the the impact of unemployment on the paid status of virtual internships.

First, I categorize the virtual internships by standard occupation classification codes (SOC) to explore the distribution of occupations within the sample. I find that a majority of virtual internships within the sample belong to the Arts, Design, Entertainment, Sports and Media occupation category, as defined by the standard occupation classification system. These internships are largely unpaid, with only 13% receiving compensation. This is in contrast to fields such as office, administration, and sales which have greater than 50% of virtual internships receiving compensation. Virtual internships categorized as including business, finance, computer, and math have little over one quarter of internships receiving compensation.

Second, I explore the geographic distribution of virtual internships by matching the individual ads to their respective metropolitan statistical areas. I find large variation in the metropolitan statistical areas and states that offer virtual internships in comparison to standard internships.

Third, I find that the paid status of virtual internships is positively associated with coding, financial, and customer service skills and negatively associated with writing skills and keywords that demonstrate strong character. Contrary to other labor markets, I find that social skills are not correlated with the paid status of virtual internships. In addition, I find that there is a negative relationship between the unemployment rate rate and the paid status of virtual internships at the metropolitan statistical area level. A discussion of this relationship appears in section (2.3.2). Next, I find part-time virtual internships to require a higher degree of writing and project management skills. In contrast to paid internships, part time internships present a negative association with customer service skills. Finally, I find that that cognitive and financial skills are statistically significant and are positively correlated with downstream wages. A job that is associated with a with these skills is likely to be more highly paid. This in contrast to job which strong character traits, which is statistically significant and negatively associated with downstream wages.

This paper is structured as follows. Section 2.2 describes the data and descriptive statistics. This is followed by section 2.3 which describes the empirical analysis of the text data, including the skill demand for virtual internships and the relationship between the unemployment rate and paid status of virtual internships in section. Finally, the conclusion and discussion of the results is located in section 2.4.

2.2 Data

Using data scraped from a large online job board in the fall of 2016 and the spring of 2017, I collect all job advertisements listed on the site. This job board has specialized in providing internship opportunities for students rather than entry level positions or positions that require more experience. However, it does include entry level positions for applicants that have a special designation as a “job” on the advertisement page.

From the fall data source I am able to extract 173,152 employment advertisements. I then restrict the database to job advertisements which exclude the “job” indicator signaling an ad for entry level employment. The remaining database includes all internship advertisements listed on the site, which constitutes roughly 25% of the original sample. Further, I then restrict our sample to internships which are tagged as “virtual” internships in which the student performs telework for the firm. This leaves a final fall sample of 13,882 virtual internship advertisements as the sample of interest. I repeat the same process with the data scraped in the spring of 2017. After accounting for duplicate internships, I extract an additional 1,635 virtual internship advertisements that had been posted since the last scrape. This brings the final sample to 15,517 virtual internships to be used in the analysis. Of these 15,517 observations I am limited to 12,428 which are located in major metropolitan statistical areas. The remaining observations could not be matched to an unemployment rate and consists of virtual internships located in minor metropolitan statistical areas or rural areas outside of the areas in which data is collected[3]. Finally, I am able to match 12,242 observations to proxy wage data from the Occupation Employment Statistics Survey Data[4].

[3]Deming and Kahn (2016) and Jaeger et. al (2017 also limit their analysis to job ads within metropolitan statistical areas.

[4]There are several reasons why all of these observations can not be matched. First, the OES data may not include all of the standard occupation classification (SOC) codes listed in the original scraped data set. Second, the OES data derives the SOC codes from an occupation classification (OCC), which is then cross-walked to a corresponding SOC code. In some cases, the occupation classification is not detailed enough to provide a 6-digit

From the job advertisements I am able to capture many key data features, including: job title, firm name, firm location, application deadline, posting date of the job ad, job time frame, job description, job requirements, and job skills needed to complete the internship. Other information on whether the internship is full time or part time, and paid or unpaid is also provided.

I then follow Jaeger et. al (2017) in the use of a machine learning algorithm [5] to match the virtual internships advertisement to the most closely associated 8-digit SOC (Standard Occupation Classification) code. The O*net-SOC Autocoder uses the job title and parses the text in the job description, requirements, and skills section from the internship advertisement, then matches it to a corresponding SOC code based on analyst weighted terms. Use of job titles in this case is important, as job titles explain more than 90% of wage variance (Marinescu and Wolthoff 2016). In addition, the autocoder provides a match score[6] to the classification in the range of 0-100. As an example, a score of 75 would indicate that the autocoder matches the correct occupation code 75% of the time. Figure (2.1) below estimates the distribution of match scores for the occupation codes from the O*net-SOC Autocoder. [7]

---

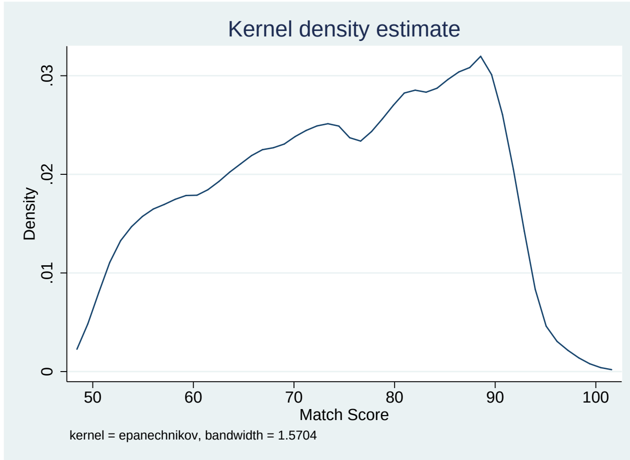
level SOC code. These observations are then rolled up to a three or four digit code, which is then matched with the three or four digit SOC code available in the original scraped data source.

[5]http://www.onetsocautocoder.com/plus/onetmatch

[6]from the FAQ section on the autocoder website: "*The match score is the weighted average of the various match methods. If a code is the top selection of all methods (and meets expected match thresholds and separation from other codes), then the overall match score will be near 100, the maximum possible. Scores above 70 accurately predict the correct code at least 70% of the time; scores above 80 accurately predict the correct code at least 80% of the time; and scores above 90 accurately predict the correct code at least 90% of the time. Accuracy rates drop off rapidly when scores are in the 60s or lower.*"

[7]Career Builder, a large online job board, recently used the same algorithm to label job advertisements to be used in the classification of job titles for the creation of a large training data set (Javed et. al 2015). Career Builder's algorithm, Carotene, marks a slight improvement on the Autocoder application with respect to classifying past work histories of job seekers , but it is inconclusive on whether the algorithm outperforms Autocoder on classifying jobs ads . Also of importance is that the taxonomy used in the Carotene algorithm is of finer granularity, and not consistent with the current O*net taxonomy. Thus, I concluded that use of the Autocoder algorithm was the most appropriate choice in the classification of the internship jobs ads from the experiment.

Figure 2.1: Distribution of Match Scores



The average match score for the algorithm is 74.5, indicating that the algorithm was able to correctly classify the virtual internships 74.5% of the time. As can be seen in the figure (2.1) above, this distribution is skewed left. This indicates that the majority of virtual job advertisements had match score greater than the average, providing evidence that the O*net-SOC Autocoder algorithm coded the advertisements properly.

Given the high rate of transition from internship to full-time offers of employment[8] , it is likely that interns end up in occupations very similar to their given virtual internship. Because the data set does not include information on salary, I measure occupational wages by linking the internship advertisement by SOC code to data from the Occupation Employment Statistics (OES) from the Bureau of Labor Statistics. The OES data is superior to alternative sources for many reasons. First, the survey produces employment and wage estimates for over 800 occupations. Second, the estimates are available at the metropolitan statistical area (MSA) level. This provides a unique data set in which there is variation in average wages within occupations and across Metropolitan Statistical Areas. Using cross-sectional data from the OES for 2012-2016, I measure the average earnings of employees by occupation code and MSA. This data is then merged with our advertisement data to provide a proxy for a given wage in the downstream occupation of the virtual intern.

Table 2.1 provides the percentage of virtual internships which fall into a specific occupation category. The occupation category is obtained from the first two digits of the six digit

[8]See National Association of Colleges and Employers 2016 Internship and Co-op Survey

SOC code. Column (1) of this table includes all internships. Columns (2) and (3) break out the virtual internships into unpaid and paid, respectively. Columns (4) and (5) repeat this process for full-time and part-time internships. Finally, columns (6) and (7) provide the share of paid and unpaid internships within a respective category.

Table 2.1 offers insight into the different occupation categories which have paid, unpaid, full-time, and part-time virtual internship opportunities. For example, most occupation categories have a higher ratio of unpaid virtual internships to paid virtual internships. However, certain categories of occupations run contrary to this. Education, Training and Library (25), Healthcare Practitioner (29), Healthcare Support (31), Food Preparation and Serving-Related (35), Sales (41), Office and Administration (43), Construction and Extraction (47), and Transportation and Moving (53) all provide a higher ration of paid virtual internships than unpaid virtual internships. Arts, Design, Entertainment, Sports and Media (27) represent a disproportionate amount of the virtual internships captured in the scrape, and consist mainly of unpaid virtual internships. Most surprisingly, occupations including Business and Finance (13) and Computer and Math (15) consists of virtual internships that are primarily unpaid.

The table also breaks down the proportion of virtual internships that are unpaid, paid, full-time, and part-time across the respective industries. Over 40% of the sample consists of virtual internships which are categorized as being in the Arts, Design, Entertainment, Sports and Media (27) major SOC. Other large sectors represented in the sample include:

Table 2.1: Percentages of Different Types of Internships by Occupation Category

Major SOC Category	Total Ads	Percentage of Ads within Each Internship Category						
		All (1)	Unpaid (2)	Paid (3)	Full Time (4)	Part Time (5)	Share Paid (6)	Share Unpaid (7)
Management (11)	548	4.09%	2.6%	1.47%	0.86%	3.22%	36%	64%
Business and Finance (13)	2,563	19.11	13.9%	5.20%	2.63%	16.48%	27%	73%
Computer and Math (15)	1,192	8.89	6.5%	2.36%	2.13%	6.76%	27%	73%
Architecture and Engineering (17)	41	0.31	0.2%	0.10%	0.10%	0.21%	32%	68%
Life, Physical and Social Sciences (19)	36	0.27	0.2%	0.04%	0.04%	0.22%	17%	83%
Community and Social (21)	152	1.13	1.0%	0.14%	0.12%	1.01%	13%	88%
Legal (23)	23	0.17	0.1%	0.03%	0.04%	0.13%	17%	83%
Education, Training and Library (25)	220	1.64	0.6%	1.01%	0.61%	1.03%	61%	39%
Arts, Design, Entertainment, Sports and Media (27)	6,373	47.52	41.4%	6.08%	3.76%	43.76%	13%	87%
Healthcare Practitioner (29)	27	0.20	0.1%	0.13%	0.10%	0.10%	67%	33%
Healthcare Support (31)	27	0.20	0.1%	0.15%	0.09%	0.11%	74%	26%
Protective (33)	10	0.07	0.1%	0.01%	0.01%	0.07%	10%	90%
Food Preparation and Serving-Related (35)	12	0.09	0.0%	0.05%	0.04%	0.05%	58%	42%
Building/Grounds Cleaning and Maintenance (37)	2	0.01	0.0%	0.01%	0.01%	0.01%	50%	50%
Personal Care (39)	36	0.27	0.1%	0.12%	0.07%	0.20%	44%	56%
Sales (41)	758	5.65	2.1%	3.53%	1.34%	4.31%	63%	37%
Office and Administration (43)	1,338	9.98	2.6%	7.35%	4.82%	5.15%	74%	26%
Farm, Fish and Forestry (45)	1	0.01	0.0%	0.00%	0.01%	0.00%	0%	100%
Construction and Extraction (47)	3	0.02	0.0%	0.01%	0.01%	0.01%	67%	33%
Install and Repair (49)	6	0.04	0.0%	0.01%	0.01%	0.04%	33%	67%
Production (51)	35	0.26	0.2%	0.05%	0.03%	0.23%	20%	80%
Transportation and Moving (53)	7	0.05	0.0%	0.04%	0.01%	0.04%	71%	29%

unpaid Business and Finance (13) at 13.9%, unpaid Computer and Math (15) at 6.5%, paid Business and Finance (13) at 5.2%, paid Arts, Design, Entertainment, Sports and Media (27) at 6.08%, and Office and Administration (43) at 7.35%. In addition, many of the virtual internships within the sample are part time, with Office and Administration (43) having the largest number of internships that a full time as a fraction of the entire sample.

2.2.1 Virtual Internship Skills

Following Deming and Kahn (2016) I use categories of job skills that are determined from the text of the job description, job requirements, and job skills section of the individual advertisement. They choose the cognitive and social skill sets to match and compare to "non-routine analytical" job tasks that are used in Autor, Levy and Murnane (2003). They also choose the words communication, teamwork, and collaboration for the "social skills" to closely match the key words in Deming (2016). Finally, they follow the literature on soft skills to define the "character" skill, including such words as "detail-oriented" and "time management". This allows a more direct comparison to previous literature describing non-routine analytical and social skills. The other skills are determined by the words and phrases that commonly show up within their large set of job ads.

Table 2.2: Description of Job Skills (From Deming and Kahn 2016)

Job Skills	**Key Words and Phrases**
Cognitive	Problem Solving, Research, Analytical, Critical Thinking, Math, Statistics
Social	Communication, Teamwork, Collaboration, Negotiation, Presentation
Character	Organized, Detail-Oriented, Multi-Tasking, Time Management, Meeting Deadlines, Energetic
Writing	Writing
Customer Service	Customer, Sales, Client, Patient
Project Management	Project Management
People Management	Supervisory, Leadership, Management, Mentoring, Staff
Financial	Budgeting, Accounting, Finance, Cost
Software	Java, SQL, Python, SAS, Stata, C#, C+, C++
Computer	Computer, Spreadsheets, Common Software (e.g. Microsoft Excel, Powerpoint)

The individual job skills are coded as a binary variable if they include at least one of the key words of phrases listed in table 2.2. It is possible that a job ad may include more than one of the words of phrases listed. Table 2.3 provides the summary statistics for each of the skill sets listed above. This table shows that approximately 41% of all job ads include a key word or phrase signifying the demand for cognitive skills, while roughly 55% of ads demand some sort of social skills. As expected, job advertisements explicitly stating the need for financial and coding skills is much smaller at 13% and 3%, respectively. Though the demand for specific coding skills is rather limited in the virtual internship market, the demand for general computer skills is much greater. 54% of all job advertisements included terms such as computer, spreadsheets,and other common software portals such as Microsoft word and excel.

Table 2.3: Summary Statistics for Firm Level Demands

	(1)		(2)		(3)		(4)		(5)	
	All Internships		Unpaid Internships		Paid Internships		Part-Time Internships		Full-Time Internships	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Cognitive	0.413	0.492	0.428	0.495	0.376	0.485	0.416	0.493	0.398	0.490
Social	0.548	0.498	0.568	0.495	0.495	0.500	0.553	0.497	0.519	0.500
Character	0.252	0.434	0.264	0.441	0.219	0.414	0.256	0.436	0.232	0.422
People Management	0.351	0.477	0.332	0.471	0.399	0.490	0.337	0.473	0.422	0.494
Financial	0.129	0.336	0.077	0.266	0.266	0.442	0.107	0.309	0.242	0.428
Customer Service	0.482	0.500	0.423	0.494	0.634	0.482	0.449	0.497	0.643	0.479
Writing	0.548	0.498	0.619	0.486	0.367	0.482	0.585	0.493	0.368	0.482
Project Management	0.037	0.188	0.039	0.193	0.032	0.175	0.038	0.192	0.028	0.166
Coding (specific)	0.034	0.181	0.033	0.179	0.036	0.186	0.031	0.172	0.050	0.219
Computers (general)	0.534	0.499	0.504	0.500	0.613	0.487	0.513	0.500	0.636	0.481
Observations	13411		9668		3743		11153		2258	

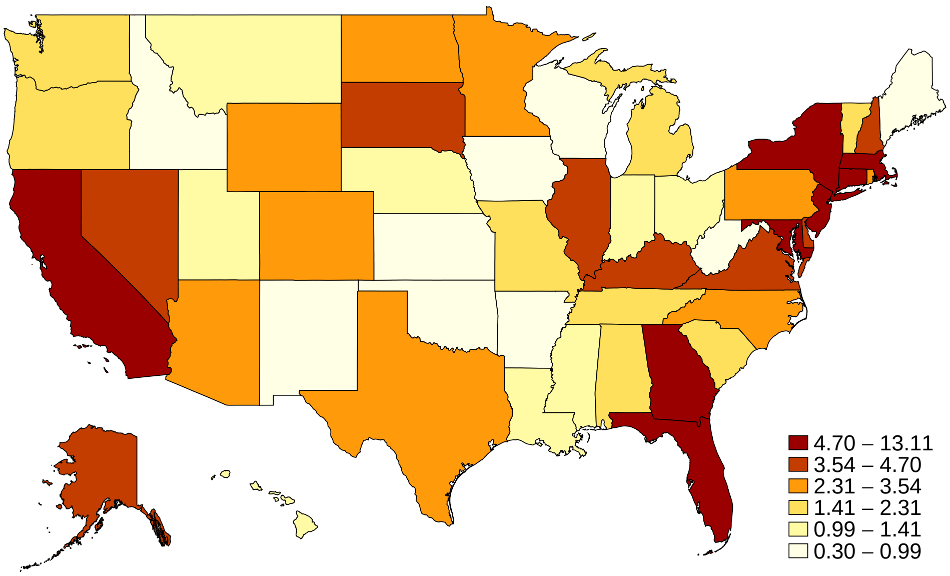
Table 2.3 then breaks down the demand of skills further by breaking out the sub-categories of internships. Paid internships have a relatively similar mean with unpaid internships with regards to cognitive, social, character, people management, project management, and coding skills. However, paid internships deviate from unpaid internships by having a higher demand on average for financial, customer service, and general computer skills. Unpaid internships have nearly double the number of advertisements which mention writing skills when compared to paid internships.

Table 2.3 also breaks down the demand for skills for part-time vs. full-time virtual internships. When compared to part-time internships, full-time internships have a similar demand cognitive, social, character, people management, project management, and coding skills. Similar to paid virtual internships, full time virtual internships have a higher demand for financial, customer service, and general computer skills when compared to their counterpart. The only skills which has a significant higher demand for part-time virtual internships is writing.

2.2.2 Distribution of Virtual Internships

Although this paper is primarily focused on the demand for different job skills, it is important to show how these virtual internships are distritbuted throughout the United States. Figure 2.2 provides a visual representation of the distribution of virtual internships by state per 100,000 individuals using data from the firm location in the job advertisement.

Figure 2.2: Virtual Internships Per 100,000 Individuals

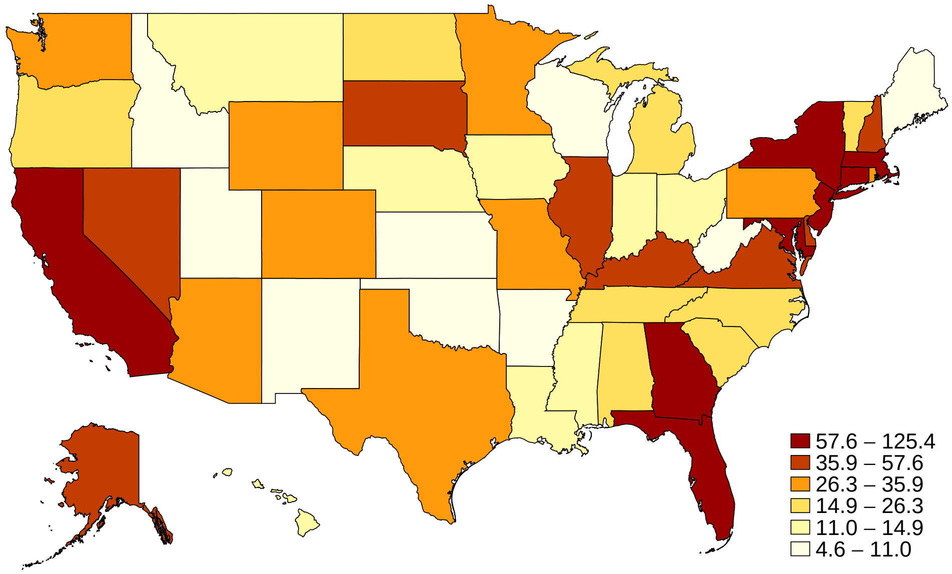


Virtual internships are concentrated in the states of California, Georgia, Florida, and much of the Northeast with roughly 5-13 virtual internships per 100,000 individuals.[9] This is in contrast to the Midwest, Heartland, Mountain, and Southwest regions which have a lower per capita volume of internships, measuring less than three virtual internships per 100,000 individuals on average.

Internships are most often accepted by college undergraduates and not the older populace within a state. Thus, it is also of interest to study the distribution of virtual internships based on the population of 19-25 year olds. Figure 2.3 provides a visual representation of the distribution of virtual internships by state per 100,000 19-25 year olds. Consistent with the previous map, the states of California, Georgia, Florida, and much of the Northeast have the highest concentration of virtual internships, while the Midwest, Heartland, Mountain, and Southwest regions have lower concentrations on average.

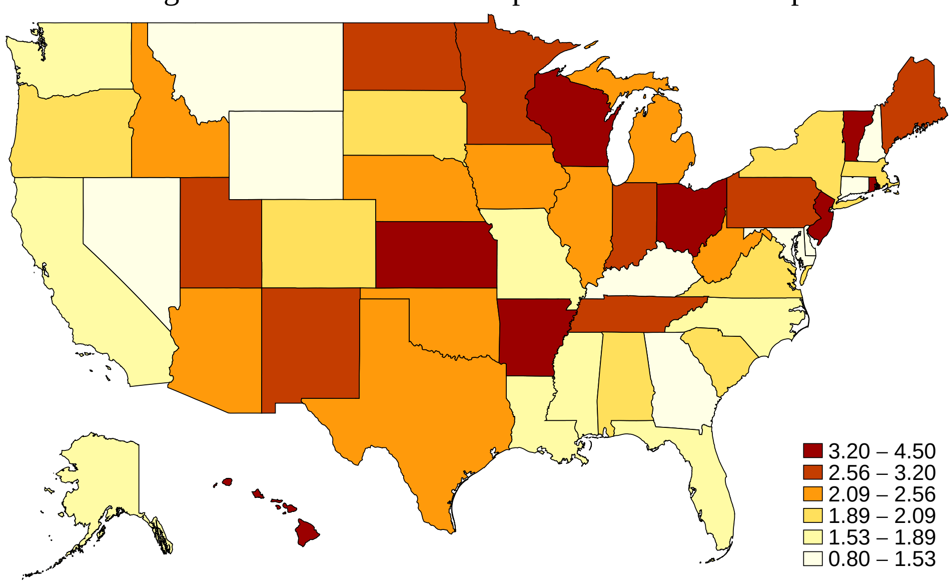
[9]measure of 5-13 internships per 100,000 individuals is state dependent.

Figure 2.3: Virtual Internships Per 100,000 19-25 Year olds



As can be seen in figures 2.2 and 2.3, there is substantial variation in the location of the internships throughout the United States. When these internship locations are linked to unemployment data at the metropolitan statistical area level, this offers a unique way to identify the impact of the unemployment rate on the paid status of virtual internships.

Figure 2.4: Ratio of Internships to Virtual Internships



The distribution of virtual internships in relation to more traditional location based internships is located in the final figure. Figure 2.4 provides the ratio of internships to virtual internships by state within the U.S. As seen above, the Midwest, Heartland, Mountain, and

Southwest regions have a higher ratio of standard location based internships when compared to the West, Southeast, and Northeast Regions.

2.3 Skill Demand

To study the demand for skills I look at the correlation between the skills variables that are extracted from the text of the job advertisement and three different features of the internships. First, I discuss the relationship between skills, unemployment, and the paid status of a virtual internship. Next, I examine the relationship between these skills, unemployment and the part time/full time status of the internship. Finally, I examine the correlation between the skills and the wages of the downstream occupations that these internships lead to. It is important to note that the wages for these downstream occupations are a proxy taken from data in the American Community Survey, based on the associated standard occupation classification given by the O*net-SOC autocoder algorithm.

Table 2.4: Correlations between Skills and Internship Attributes

Skills	Paid Status	Part Status	Log Wage
cognitive	-0.0371	-0.0005	0.0489
social	-0.0588	0.0172	0.1047
character	-0.0761	0.0225	-0.0580
people	0.0215	-0.0429	0.0573
financial	0.1975	-0.1034	0.0071
customer	0.1727	-0.1410	0.1544
writing	-0.2027	0.1498	-0.0159
project management	-0.0342	0.0318	0.0446
coding	0.0121	-0.0432	0.1216

Table 2.4 provides the bivariate correlations between the paid status of an internship, part time status of an internship, and the log proxy wage with the skills associated with each internship advertisement. This table presents a negative correlation between cognitive, social, writing, and character skills with paid status and a positive correlation with people, financial,

customer, and coding skills. This pattern is reversed when we look at the bivariate correlation between these skills and part time status. Finally, the table indicates that there is a positive correlation between cognitive, social, people, financial, customer, project management, and coding skills with the downstream wage associated with the internship. While character and writing skills seem to present a negative correlation. However, a simple analysis of the bivariate correlations does not take into consideration the impact of the unemployment rate and other omitted variables that vary within occupations and metropolitan statistical areas which may effect the paid status of virtual internships. Thus, I proceed in this section by analyzing the correlation between paid status, part time status, the wage proxy and the skills while controlling for time invariant unobservables and the unemployment rate.

2.3.1 Skill Demand for Paid vs. Unpaid Virtual Internships

The decision on whether or not internships should be paid or unpaid remains a contested topic. Recent lawsuits have seen the exploitation of student labor within for-profit firms [10].The Department of Labor (DOL) provides fact sheet #71 to offer guidance on whether or not an internship should be paid or unpaid[11]. This sheet provides a test for unpaid interns to determine if providing zero wages is appropriate. To this extent, the test provides six criteria for for-profit employers:

1. The internship, even though it includes actual operation of the facilities of the employer, is similar to training which would be given in an educational environment;
2. The internship experience is for the benefit of the intern;
3. The intern does not displace regular employees, but works under close supervision of existing staff;
4. The employer that provides the training derives no immediate advantage from the activities of the intern; and on occasion its operations may actually be impeded;

[10] i.e. Glatt vs. Fox Searchlight Pictures Inc.

[11] refer to the following web address for more details on fact sheet #71: $https : //www.dol.gov/whd/regs/compliance/whdfs71.htm$

5. The intern is not necessarily entitled to a job at the conclusion of the internship; and

6. The employer and the intern understand that the intern is not entitled to wages for the time spent in the internship.

Most recently, the second circuit court provided an amended test on whether or not an intern should be paid by looking at a three part test as to whom is the "primary beneficiary" of the interns labor. The three provided features are as follows[12]:

1. First, it focuses on what the intern receives in exchange for his work

2. Second, it also accords courts the flexibility to examine the economic reality as it exists between the intern and the employer

3. Third, it acknowledges that the intern-employer relationship should not be analyzed in the same manner as the standard employer-employee relationship because the intern enters into the relationship with the expectation of receiving educational or vocational benefits that are not necessarily expected with all forms of employment (though such benefits may be a product of experience on the job

While these six criteria from the FLSA and three criteria from the second circuit court offer guidance on whether or not an internship should be paid, it is unknown which skills are associated with paid internships that do not meet this criteria. In order to examine the relationship between skills, unemployment, and paid status I estimate a linear probability model of the following form:

$$paid_{o,m,t,j} = \alpha + \beta unemployment_m + X_j\gamma' + \phi_o + \psi_m + \epsilon_{o,m,t,j}$$

where the index o is occupation (SOC), m is the metropolitan statistical area, t is time, j is an individual virtual internship advertisement. X is a set of controls taken from the job advertisement and include information such as part time/full time status of the internship and the various skills derived from the job advertisement text. ϕ is occupation fixed effects and ψ

[12]Taken from case documents Nos. 13-4478-cv, 13-4481-cv in the Glatt vs. Fox Searchlight Pictures Inc. Second Circuit court ruling

is metropolitan area fixed effects which control for time invariant unobservables and help to eliminate omitted variable bias within the model. β is the coefficient of interest and describes the impact of the unemployment rate on the paid status of virtual internships[13]. γ is also of interest, and represent a vector of coefficients which describe the association between different skills and paid status.

In estimating the linear probability model above, it is likely that observations are related to one another within the occupation groups and the metropolitan statistical areas (MSA). Standard practice is to cluster the errors at the highest level of aggregation in order to control for the correlation within clusters. However, because of the non-nested structure between occupation categories and the metropolitan areas it is not immediately clear at which level the errors should be clustered. Recommendations are to use fixed effects in one dimension and cluster the standard errors in the other [14]. Alternatively, one may follow Cameron, Gelbach and Miller (2006) to estimate the model with multi-way clustering.

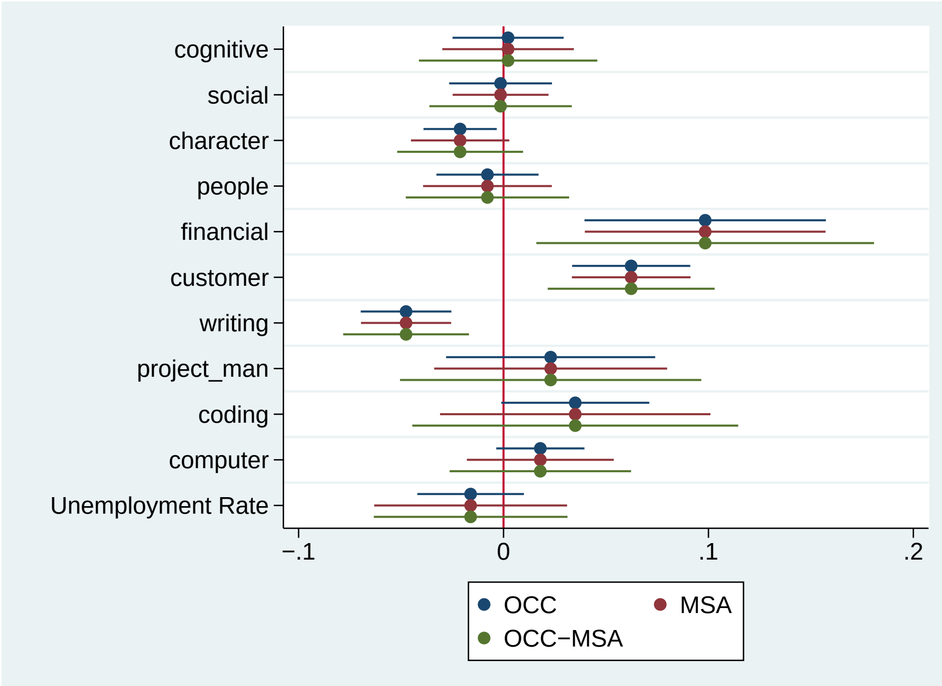
Due to the uncertainty in the level at which the errors should be clustered I present the results in three different tables. Table B.1 provides the point estimates for β and γ with errors clustered at the metropolitan statistical area level. Column (1) of table B.1 is a baseline model excluding the unemployment rate and any fixed effects. Column (2) presents the baseline model including the unemployment rate but excluding any fixed effects. Column (3) is the same as column (2), but includes area fixed effects. Column (4) continues to add area and occupation fixed effects, while column (5) presents the results of the fully specified model which includes the unemployment rate, area, occupation, and time fixed effects. Table B.2 presents the same models, but clusters the errors at the occupation level rather than the metropolitan statistical area level. Table B.3 follows Cameron, Gelbach, and Miller(2006) to estimate the model with multi-way clustering at the occupation-metropolitan statistical area level. Figure 2.5 compiles the point estimates and confidence intervals across each table for the fully specified model.[15].

[13]unemployment rate is given at the metropolitan statistical area level.

[14]see $http://econweb.umd.edu/\ sarzosa/teach/2/Disc2Clusterhandout.pdf$ for details

[15]column 5 in tables 2.4 and 2.5, column 3 in table 2.6

Figure 2.5: Paid Point Estimate Comparison



The first thing of note is that there is a positive and statistically significant relationship between financial and customer skills with paid status when clustering the errors at the occupation, MSA, or occupation-MSA level. In addition, there is a statistically significant and negative relationship between writing and paid status across all three specifications of the error structure. Words that are representative of strong character are statistically significant and negatively associated when errors are clustered at the MSA or occupation level.

These results contrast with Deming and Kahn (2016), in which they find a positive correlation between cognitive skills and log average wages across occupations and MSA's. Also in contrast to their paper, I find that paid positions do not require more social skills, but that there is a positive and statistically significant relationships between paid status and coding, financial, and customer skills[16]. When taken into context, these results do not necessarily contradict Deming and Kahn's main finding. Virtual internships are a small subset of internships and are not included in their analysis of the general labor market. In addition, virtual internships leaves the intern isolated from the rest of the firm's labor force except for planned virtual or

[16]customer skills are statistically significant at the .01 level for occupation and MSA clustered errors and statistically significant at the .05 level for OCC-MSA clustered errors. Financial skills are not statistically significant when clustering the errors at the OCC-MSA level.

in-person meetings. It is likely that these types of internships require minimal social skills. As an example, a location based sales internship will require a significant amount of social interaction while a software engineering internship requiring coding expertise can likely be achieved virtually with limited face-to-face contact and interaction.

Tables B.1-B.3 suggest that the skills variables can account for a substantial proportion of the variation in the paid status of the internships. This is substantiated by the p-values for the F-test for joint significance of all of the skills listed. Given the F-test values, and the p-values listed in tables B.1-B.3, we can conclude in this section that the joint impact of these skills on paid status is statistically different from zero.[17]

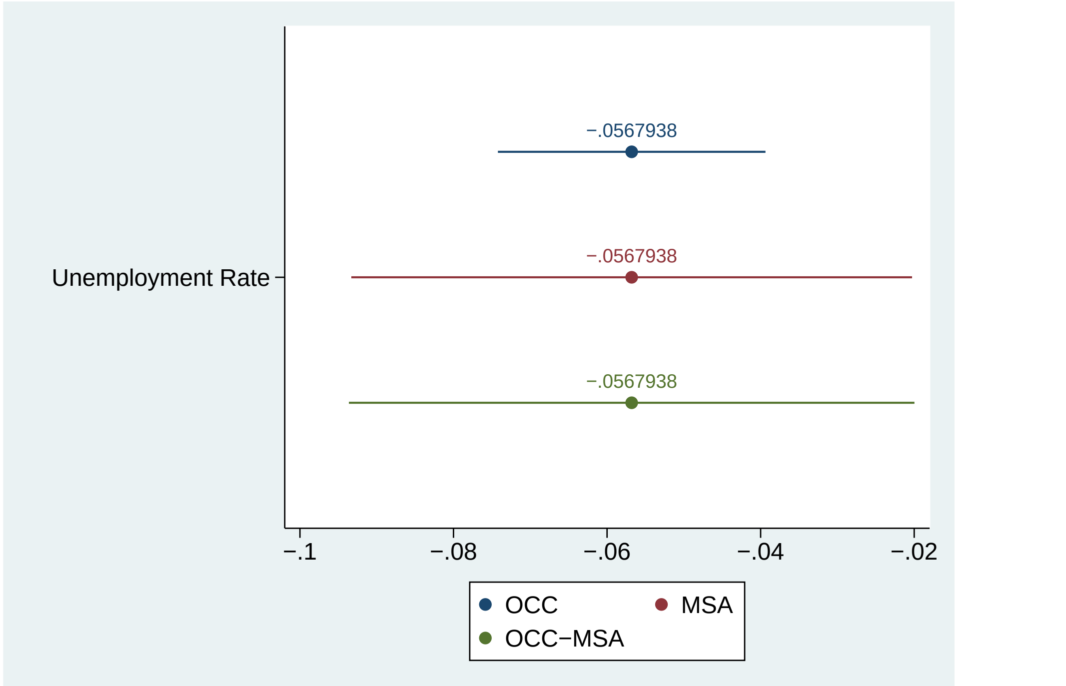
2.3.2 The Impact of Unemployment on Paid Status

Next, I focus on the impact of unemployment on the paid status of internships by linking each internships location and posting date to its corresponding metropolitan statistical area unemployment rate. Figure 2.6 compares the point estimates and confidence intervals across each table for the models including area and occupation fixed effects, but excluding time fixed effects[18]. In these models, I use state-wide variation in the unemployment rates at the MSA level in order to identify the impact of unemployment on the paid status of internships. The exploitation of this variation allows me to identify the impact of the unemployment rate independent of any nationwide trends in the paid status for virtual internships. In addition, the inclusion of MSA and occupation level fixed effects helps to control for time-invariant unobservables that may have an impact on the paid status within different metropolitan statistical areas and occupations.

[17] F-test values for joint significance are listed in along with the p-values in the appendix for chapter two.

[18] Time fixed effects were excluded due to collinearity between the unemployment rate and year dummies.

Figure 2.6: Unemployment Point Estimate Comparison



As can be seen in figure 2.6, the impact of the unemployment rate on the paid status of virtual internships is rather robust across all three specifications of the model. On average, a one unit increase in the unemployment rate leads to a 5.7% decrease in the probability that the virtual internship is paid. This indicates that there is a negative relationship between the unemployment rate and the paid status of virtual internships.

The explanation for this negative relationship is not as immediately clear as it would be with standard location based internships. For location based internships, employers must begin to offer payment for an intern as the labor market begins to tighten and quality interns become more scarce. In contrast, employers may be able to take advantage of loose labor market conditions and provide zero pay to interns trying to boost the quality of their resume. This explanation weakens when applied to virtual internships. One would expect that as the labor market tightens in one state, a firm may take advantage of loose labor market conditions in another. Hence, there would be no relationship between virtual internships paid status and the unemployment rate.

An alternative explanation for this relationship would be that the firm does care about the distance between an applicant and the firm's location. Using experimental data, Jaeger et. al

(2017) shows that there is a strong negative relationship between how far away an applicant is located and the firm's response to an application. Though one can only speculate on the exact mechanism without a field experiment, it is likely that the firm's local labor market conditions impact the paid status for virtual internships for the same reason they impact standard location based internships.

It is important to note that the data does not extend through an entire business cycle. Thus, the interpretation of this results must be taken under careful consideration. Nationally, from January 2012 to January 2017, the U6 unemployment rate declined from 8.3% to 4.8%. This is a total decline of 3.5%. It is possible that the identified impact is picking up on a decreasing trend in the unemployment during the period. In order to properly identify the impact of the unemployment rate on the paid status, more data will be needed over a complete business cycle.

2.3.3 Skill Demand for Part-Time vs. Full-Time Virtual Internships

In order to examine the relationship between skills, unemployment, and part time status I estimate a linear probability model of the following form:

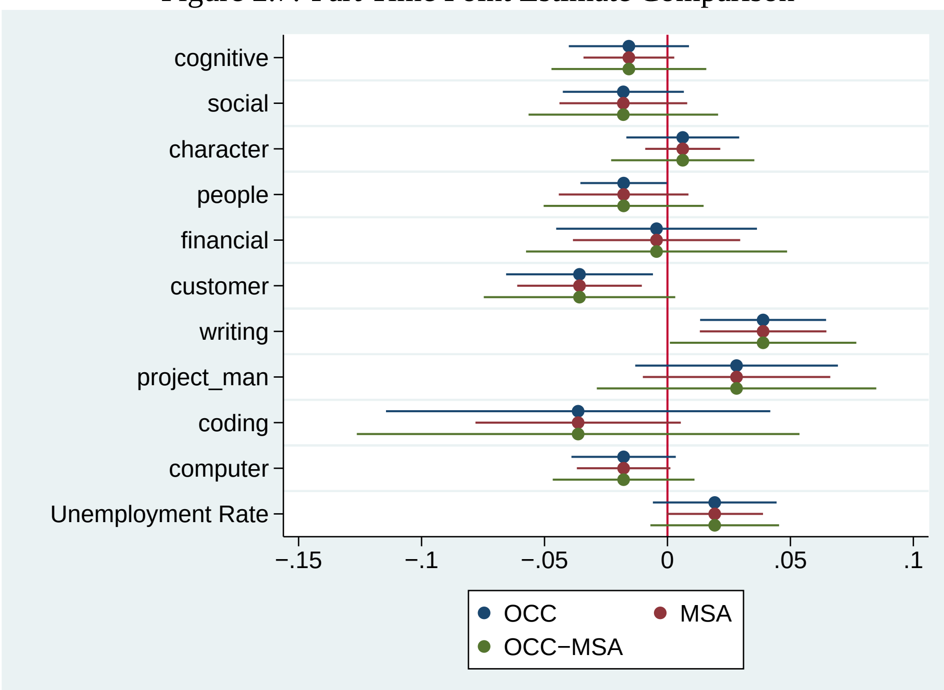
$$part_{o,m,t,j} = \alpha + skill_j\beta' + X_j\gamma' + \phi_o + \psi_m + \epsilon_{o,m,t,j}$$

where the index o is occupation (SOC), m is the metropolitan statistical area, t is time, j is an individual virtual internship advertisement. X is a set of controls taken from the job advertisement and include information such as part time/full time status of the internship. ϕ is occupation fixed effects and ψ is metropolitan area fixed effects which control for time invariant unobservables and help to eliminate omitted variable bias within the model. β is the coefficient of interest and represents a vector of coefficients which describe the association between different skills and part time status.

As with estimating the relationship between skills and paid status, it is not immediately clear at which level the errors should be clustered. To this extent, I estimate models with the errors clustered at the occupation, metropolitan statistical area, and OCC-MSA level. Table B.4 provides the point estimates for β with errors clustered at the metropolitan statistical area level. Column (1) of table B.4 is a fully specified model including occupation and area fixed effects but excluding the unemployment rate. Column(2) presents the baseline model including the unemployment rate but excluding any fixed effects. Column(3) is the same as column(2), but includes area fixed effects. Column (4) continues to add area and occupation fixed effects, while column(5) presents the results of the fully specified model which includes the unemployment rate, area, occupation, and time fixed effects. Table B.5 presents the same models, but clusters the errors at the occupation level rather than the metropolitan statistical area level. Table B.6 follows Cameron, Gelbach, and Miller(2006) to estimate the model with multi-way clustering at the occupation-metropolitan statistical area level.

Figure (2.7) compiles the point estimates and confidence intervals across each table for the fully specified model[19].This figure provides a visual representation of the point estimates and confidence intervals across the three different model types which include the most detailed level of controls.

Figure 2.7: Part Time Point Estimate Comparison



From tables B.4-B.6 and figure 2.7, writing skills are the only skills that are statistically significant and positively correlated with part time status. In contrast to paid positions, part time positions require less customer skills.

[19]column 5 in tables B.4 , B.5,and B.6

Despite writing and customer service skills being the only skills listed that are statistally significant, tables B.4-B.6 suggest that the skills variables as a whole can account for a substantial proportion of the variation in the part time status of the internships. This is substantiated by the p-values for the F-test for joint significance of all of the skills listed. Given the F-test values, and the p-values listed in tables B.4-B.6, we can conclude in this section that the joint impact of these skills on part-time status is statistically different from zero.[20]

2.3.4 Skills and Firm Pay for Virtual Internships

Finally, I look at the relationship between the average wage for each downstream occupation and the skills derived from the job advertisement text. In order to examine the relationship between skills and the proxy wage I estimate a linear probability model of the following form:

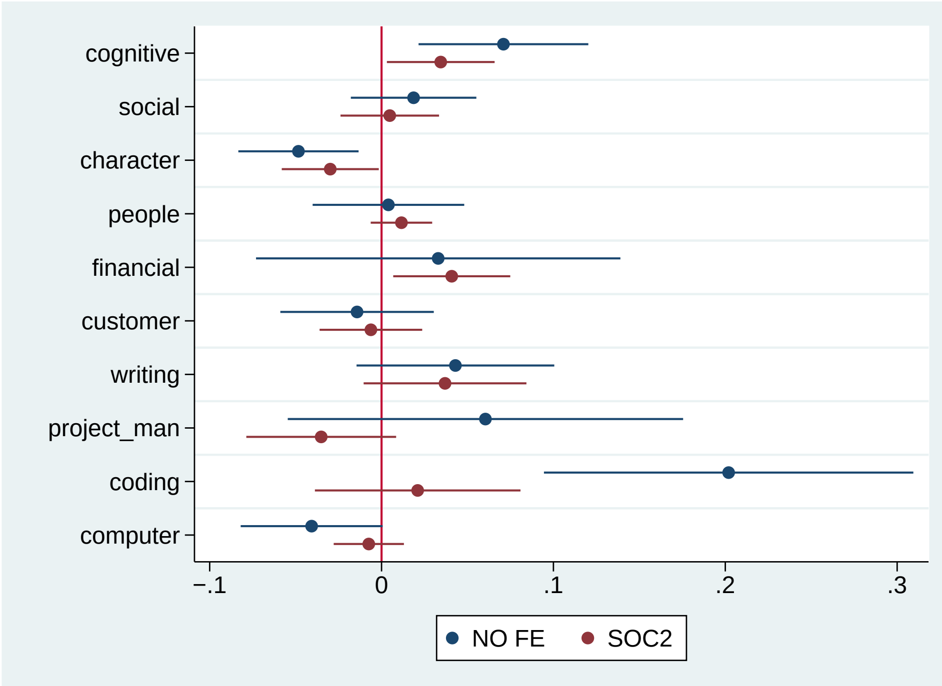
$$log(wage)_{o,m,t,j} = \alpha + skill_j\beta' + X_j\gamma' + \phi_o + \epsilon_{o,m,t,j}$$

where the index o is occupation (SOC), m is the metropolitan statistical area, t is time, j is an individual virtual internship advertisement. X is a set of controls taken from the job advertisement and include information such as part time/full time status of the internship. ϕ is occupation fixed effects which control for time invariant unobservables and help to eliminate omitted variable bias within the model[21]. β is the coefficient of interest and represents a vector of coefficients which describe the association between different skills and the logarithm of the proxy wage.

Figure (2.8) compiles the point estimates from the fully saturated models with and without occupation level fixed effects. The first thing to note from this figure is that cognitive and financial skills are statistically significant and are positively correlated with the wage. A job that is associated with these skills is likely to be more highly paid. Second, jobs that include words associated with positive character traits have a lower return in market wage. However, The p-values on the joint significance for all of the skills together indicate they are jointly not significant in determining the wage.

[20]F-test values for joint significance are listed in along with the p-values in the appendix for chapter two.
[21]occupation fixed effects are at the 2-digit standard occupation classification level.

Figure 2.8: Wage Point Estimate Comparison



2.4 Discussion and Conclusion

Telecommuting in computer and science related positions has increased by 69% from 2000-2010 (Mateyka et al 2012),with data from the American Community Survey indicating increases of over 100% from 2005-2014 across all occupations. The increases in telecommuting positions are not limited to those in science and technology. Mateyka (2012) also find that nearly one quarter of all telecommuting positions were in the fields of finance, business, and management. As technology and communication platforms continue to become more advanced, workers in the United States will continue to transition towards more telework as a percentage of total labor hours. Simultaneously, the labor market for internships is undergoing a similar transition. More employers are offering virtual internships in which the individual can work remotely. To my knowledge, this is the first paper to address the labor market for virtual internships.

I find that a majority of virtual internships within the sample belong to the Arts, Design, Entertainment, Sports and Media occupation category, as defined by the standard occupation classification system. These internships are largely unpaid, with only 13% receiving compensation. This is in contrast to fields such as office, administration, and sales which have greater than 50% of virtual internships receiving compensation. Virtual internships categorized as including business, finance, computer, and math have little over one quarter of internships receiving compensation.

There is also large variation in the location of firms offering virtual internships throughout the United States. States with large populations tend to have a have a higher amount of virtual internships per capita. New York, California, Florida, and Georgia lead the United States in virtual internships per 100,000 19-25 year olds. As expected, less populated states have a much higher ratio of standard location based internships to virtual internships. Wisconsin, Ohio, Nebraska, and new Jersey having nearly 4 times as many location based internships vs. virtual internships.

Paid status of virtual internships is positively associated with coding, financial, and customer service skills and negatively associated with writing skills and keywords that demonstrate strong character. Contrary to other labor markets, I find that social skills are not correlated with the paid status of virtual internships. In addition, I find that there is a negative relationship between the unemployment rate rate and the paid status of virtual internships at the metropolitan statistical area level.

Part-time virtual internships require a higher degree of writing and project management skills, and present a negative association with customer service skills. When matching positions to downstream occupations, cognitive and financial skills are statistically significant and are positively correlated with wages. A job that is associated with a with these skills is likely to be more highly paid. This in contrast to job which strong character traits, which is statistically significant and negatively associated with downstream wages.

As the labor market for virtual internships continues to mature there will be many opportunities for future research. Documenting the transition of virtual internships by occupation will help to understand changing labor market conditions. In addition, future longitudinal survey data will help to assess the causal impact of virtual internships on future wages.

References

[1] Abel, Jaison R, Richard Deitz, and Yaqin Su. ”Are Recent College Graduates Finding Good Jobs?” Federal Reserve Bank of New York: Current Issues in Economics and Finance 20, no. 1 (2014): 1-8.

[2] Acemoglu, Daron, and Jorn-Steffen Pischke. ”Beyond Becker: training in imperfect labour markets.” The Economic Journal (Wiley Online Library) 109, no. 453 (1999): 112-142.

[3] Acemoglu, Daron, and Jorn-Steffen Pischke. ”The Structure of Wages and Investment in General Training.” Journal of Political Economy (JSTOR) 107, no. 3 (1999): 539-572.

[4] Acemoglu, Daron, and Jorn-Steffen Pischke. ”Why Do Firms Train? Theory and Evidence.” Quarterly Journal of Economics (JSTOR), 1998: 79-119.

[5] Aigner, Dennis J, and Glen G Cain. ”Statistical theories of discrimination in labor markets.” Industrial and Labor relations review (JSTOR), 1977: 175-187.

[6] Altonji, Joseph G, Erica Blom, and Costas Meghir. ”Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers.” Annu. Rev. Econ. (Annual Reviews) 4, no. 1 (2012): 185-223.

[7] Ameri, Mason, Lisa Schur, Meera Adya, Scott Bentley, Patrick McKay, and Douglas Kruse. ”The Disability Employment Puzzle: A Field Experiment on Employer Hiring Behavior.” Working Paper, National Bureau of Economic Research, 2015.

[8] Arrow, Kenneth J. ”The Theory of Discrimination, S. 3-33 in: Orley Ashenfelter und Albert Rees (Hg.): Discrimination in Labor Markets.” The Theory of Discrimination, S. 3-33 in: Orley Ashenfelter und Albert Rees (Hg.): Discrimination in Labor Markets. Princeton NJ.: Princeton University Press, 1973.

[9] Becker, Gary S. Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago press, 2009. The economics of discrimination. University of Chicago press, 2010.

[10] Bertrand, Marianne, and Esther Duflo. ”Field experiments on discrimination.” Tech. rep., National Bureau of Economic Research, 2016.

[11] Bertrand, Marianne, and Sendhil Mullainathan. ”Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” The American Economic Review (American Economic Association) 94, no. 4 (2004): 991-1013.

[12] Bertrand, Marianne, and Sendhil Mullainathan. ”Implicit discrimination.” The American Economic Review (JSTOR) 95, no. 2 (2005): 94-98.

[13] Cappelli, Peter. ”Skill Gaps, Skill Shortages and Skill Mismatches: Evidence for the Us.” NBER Working Paper, no. w20382 (2014).

[14] Charles, Kerwin Kofi, Jonathan Guryan, and others. ”Studying Discrimination: Fundamental Challenges and Recent Progress.” Annual Review of Economics (Annual Reviews) 3, no. 1 (2011): 479-511.

[15] Deming, David J, Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F Katz. ”The Value of Postsecondary Credentials in the Labor Market: An Experimental Study.” National Burea of Economic Research Working Paper Series No. 20528, 2014.

[16] Elbaum, Bernard. ”Why apprenticeship persisted in Britain but not in the United States.” The Journal of Economic History (Cambridge Univ Press) 49, no. 02 (1989): 337-349.

[17] Fersterer, Josef, Jorn-Steffen Pischke, and Rudolf Winter-Ebmer. ”Returns to apprenticeship training in Austria: Evidence from failed firms.” The Scandinavian Journal of Economics (Wiley Online Library) 110, no. 4 (2008): 733-753.

[18] Gaddis, S Michael, and Raj Ghoshal. ”Arab American Housing Discrimination, Ethnic Competition, and the Contact Hypothesis.” The ANNALS of the American Academy of Political and Social Science (SAGE Publications) 660, no. 1 (2015): 282-299.

[19] ”Glatt v. Fox Searchlight Pictures, Inc.” Glatt v. Fox Searchlight Pictures, Inc. Vol. 791. no. No. 13-4478-cv. Court of Appeals, 2nd Circuit, 2015. 376.

[20] ”Glatt v. Fox Searchlight Pictures, Inc.” Glatt v. Fox Searchlight Pictures, Inc. Vol. 811. no. No. 13-4478-cv. Court of Appeals, 2nd Circuit, 2016. 528.

[21] Hacker, Michael A. ”Permitted to Suffer for Experience: Second Circuit Uses” Primary Beneficiary” Test To Determine Whether Unpaid Interns are Employees Under the FLSA in Glatt v. Searchlight Pictures, Inc.” Boston College Law Review 57, no. 6 (2016): 67.

[22] Hamilton, Gillian. ”The market for Montreal apprentices: Contract length and information.” Explorations in Economic History (Elsevier) 33, no. 4 (1996): 496-523.

[23] Javed, Faizan, and Ferosh Jacob. ”Data Science and Big Data Analytics at Career Builder.” In Big-Data Analytics and Cloud Computing, 83-96. Springer, 2015.

[24] Kirkeboen, Lars, Edwin Leuven, and Magne Mogstad. ”Field of Study, Earnings, and Self-Selection*.” The Quarterly Journal of Economics, 2016.

[25] Kroft, Kory, Fabian Lange, and Matthew J Notowidigdo. ”Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment.” The Quarterly Journal of Economics (Oxford University Press) 128, no. 3 (2013): 1123-1167.

[26] Lahey, Joanna N. ”Age, Women, and Hiring: An Experimental Study.” Journal of Human Resources (University of Wisconsin Press) 43, no. 1 (2008): 30-56.

[27] Lahey, Joanna N, and Ryan A Beasley. ”Computerizing Audit Studies.” Journal of Economic Behavior and Organization (Elsevier) 70, no. 3 (2009): 508-514.

[28] Lanning, Jonathan A. ”Opportunities denied, wages diminished: Using search theory to translate audit-pair study findings into wage differentials.” The BE Journal of Economic Analysis & Policy 13, no. 2 (2013): 921-958.

[29] Mateyka, Peter J., Melanie Rapino, and Liana Christin Landivar. *Home-based workers in the United States: 2010*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau, 2012.

[30] Marinescu, Ioana, and Ronald Wolthoff. *Opening the black box of the matching function: The power of words*. No. w22508. National Bureau of Economic Research, 2016.

[31] Neumark, David, and Judith Rich. ”Do Field Experiments on Labor and Housing Markets Overstate Discrimination? Re-examination of the Evidence.” Tech. rep., National Bureau of Economic Research, 2016.

[32] Neumark, David, Ian Burn, and Patrick Button. ”Is it harder for older workers to find jobs? New and improved evidence from a field experiment.” Tech. rep., National Bureau of Economic Research, 2015.

[33] Nunley, John M, Adam Pugh, Nicholas Romero, and R Alan Seals. ”College Major, Internship Experience, and Employment Opportunities: Estimates from a Resume Audit.” Labour Economics, 2016(a).

[34] Nunley, John M, Adam Pugh, Nicholas Romero, and R Alan Seals. ”Unemployment, Underemployment, and Employment Opportunities: Results from a Correspondence Audit.” Industrial and Labor Relations Review, 2016(b).

[35] Oreopoulos, Philip. ”Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Résumés.” American Economic Journal: Economic Policy (JSTOR) 3, no. 4 (2011): 148-171.

[36] Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. ”The Short-and Long-Term Career Effects of Graduating in a Recession.” American Economic Journal: Applied Economics (American Economic Association) 4, no. 1 (2012): 1-29.

[37] Pager, Devah. ”The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future.” The Annals of the American Academy of Political and Social Science (Sage Publications) 609, no. 1 (2007): 104-133.

[38] Pallais, Amanda. ”Inefficient hiring in entry-level labor markets.” The American Economic Review (American Economic Association) 104, no. 11 (2014): 3565-3599.

[39] Phelps, Edmund S. ”The statistical theory of racism and sexism.” The American Economic Review (JSTOR) 62, no. 4 (1972): 659-661.

[40] Price, Joseph, Justin Wolfers, and others. ”Racial Discrimination Among NBA Referees.” The Quarterly Journal of Economics (Oxford University Press) 125, no. 4 (2010): 1859-1887.

[41] Riach, Peter A, and Judith Rich. ”Field Experiments of Discrimination in the Market Place.” The Economic Journal (Wiley Online Library) 112, no. 483 (2002): F480–F518.

[42] Rooth, Dan-Olof. ”Automatic associations and discrimination in hiring: Real world evidence.” Labour Economics (Elsevier) 17, no. 3 (2010): 523-534.

[43] Saniter, Nils, and Thomas Siedler. ”Door Opener or Waste of Time? The Effects of Student Internships on Labor Market Outcomes.” IZA Discussion Paper No. 8141 (IZA Discussion Paper), 2014.

[44] Wallis, Patrick. ”Apprenticeship and training in premodern England.” Journal of economic history (Cambridge Univ Press) 68, no. 03 (2008): 832-861.

Appendices

Appendix A

Chapter One Appendix

A.1 Resume Design

A.1.1 Applicant Names

Constructing a similar framework of prior correspondence studies, we randomly assign names to applicants which are unique to a particular racial group. These names include: Xavier Washington, Darius Jackson, Colin Johansson, and Wyatt Schmidt. Xavier Washington and Darius Jackson are African American names, while Colin Johansson and Wyatt Schmidt are Caucasian names. The surnames were chosen from the 2000 US Census in order to locate uniquely white and black names. For example, the 2000 US Census lists the surname Schmidt as the 88th most popular surname in the United States, with 96.5% of those self identifying as white.The surnames Jackson and Washington were the 6th and 18th most popular surnames for Blacks, with 53.0% and 89.9% self identifying as black[1]

The given names were chosen from the 1984 Social Security name database. This database was chosen because it would provide popular names which are consistent with the age of the applicants. Xavier and Darius are ranked 139 and 155, and Wyatt and Colin are ranked 124 and 197. Together with the surnames, the full names are likely to signal the race of the applicant. [2]

A.1.2 Resume Construction

Using a computerized resume generator (Lahey 2009) we randomly assign productivity characteristics to resumes while maintaining a balanced sample of covariates. Each applicant is assigned one of six college majors, including: Economics, Business Administration, Psychology, Biology, Marketing, and English. In addition, the applicant is assigned a grade

[1] Though these values seem low, they are rather high when taken into context of other popular black names

[2] Charles and Guryuan (2008) describe the importance of choosing popular names for both black and white individuals

point average from 3.0-4.0 in increments of .20 grade points. An initial job is assigned to all applicants so that a baseline level of work experience is established. Beyond this baseline, we randomly assign half of all applicants prior internship and volunteer experience. Internship experience is assigned in one of three fields: marketing, research, or business. Likewise, volunteer experience is varied among local and national well known entities. Finally, each applicant is provided a baseline skill set, with randomly assigned additional skills in statistical programming and data analysis. In what follows, we address the process of choosing these characteristics to be used in the resume construction.

A.1.3 Institutions

We chose twenty-four large public institutions throughout the United States. The public institutions had an undergraduate enrollment of at least 16,000, but did not exceed 48,000 for the 2014-2015 Academic year. The institutions varied in quality.

A.1.4 Academic Major

Each applicant was assigned one of the following six academic majors: Economics, Business Administration, Psychology, Biology, Marketing, and English. These majors were chosen due to their popularity amongst the student population (Carnevale et al. 2015). For example, the 2009 American Community Survey indicates business management and administration as the most popular major amongst college graduates, with 8 percent of all individuals holding a bachelor's degree choosing this major. Likewise, Psychology is ranked 5th, with 4 percent; Marketing and Marketing research ranked 7th, with 3 percent; English Language and Literature is ranked 9th, with 3 percent. These majors also have weak-labor market orientation. That is, they do not have a well-defined school to work transition when compared to majors such as nursing, engineering, etc.

A.1.5 Employment Status

To provide a baseline level of work experience, each applicant was assigned with certainty an entry level position in one of three fields. These fields include university employment, the restaurant industry, or retail industry. The field choices were the results of extracting a random sample of real resumes from the online job board. Examination of these resumes provided us with the most common out of field job experience listed on the typical bachelor degree seeking resume.

Once the baseline work experience fields had been designated, samples of work experience were pulled from the pool of real resumes in order to more accurately describe the duties and responsibilities on the job. In addition, major restaurant and retail chains were chosen. First, this mitigates the harm to any small firms. Second, it is more likely for these firms to be located in each of the college towns where the applicants list their residence.

A.1.6 In-Field, Out-of-Field Internship Experience and Volunteer Experience

Half of all applicants were randomly assigned previous internship experience in one of three fields: Marketing, Business/Finance, and Research. Assigning different fields of internships allowed us to examine the impact of in-field vs. out-of-field prior internships on the probability of receiving an internship in that field. The marketing internship and business/finance internships were with nationally located well known firms. This alleviated any biases that may be due to assigning applicants experience with firms that are only regionally active.

Similar to previous internship experience, half of all applicants were randomly assigned volunteer experience in one of three nationally recognized non-profit organizations.

A.1.7 Application Process

We used a large internet job board exclusively listing internships to locate the positions of interest. To alleviate any bias, the job category, whether the job was paid or unpaid, full time or part time, and the position of the job amongst the resulting list were all randomized. In more detail:

1. A random number determined whether we submitted resumes to an internship opening in the following job categories: marketing, research, or general business.

2. After determining the job category, a random number determined whether we applied to a full-time or a part-time internship opening.

3. After determining whether to submit the resume to the full-time or part-time internship, a random number determined whether to apply to a paid or unpaid internship.

4. Next, a random number generator determined which page the selected internship should be chosen from.

5. Finally, after #1, #2, #3, & #4 above, a random number (from 1-10) determined the particular internship to which a resume was submitted on the webpage.

We then monitored the accounts for a receipt of response. If a response was noted, we immediately replied acknowledging that we had taken another position in order to minimize any harm to the firms.

A.2 Creation of Accounts

This section describes the process of account creation for each randomly generated resume. We create twenty-five unique email addresses, associated with the twenty-five public institutions used in our study. Exploiting a tool within the email host, we are able to create twenty-four applicant accounts for each associated email address.

We use the associated public institution's assigned email address with the addition of a unique identification code to create and track the applicants account. Upon signing up for the account, we upload a PDF version of the applicant's resume, which is then transferred into a standard format by an algorithm within the job portal. Settings are defined to remove access to the original resume. This procedure removes the ability for potential firms to see the unique identification code attached to the PDF name.

Uploading the resume also removes the need to provide an email address on the web-portal generated resume. This process eliminates the need to post a contact email address or phone number. In lieu of the prior contact methods, we ask that the firm contact the applicant through a direct messaging system within the job-portal.

A.3 Standard Occupational Classification (SOC) of Internship Job Advertisements

We use a machine learning algorithm to match the individual internship job ads to their associated O*Net classification, as designated by the Standard Occupational Classification (SOC) hierarchy [3]. This algorithm, commissioned by the United States Department of Labor Employment and Training Administration, divides the text of the job ad into individual analyst weighted terms, then assigns a SOC code based on matching of these terms to the job descriptions listed in the O*Net database.

Career Builder, a large online job board, recently used the same algorithm to label job advertisements to be used in the classification of job titles for the creation of a large training data set (Javed et. al 2015). Career Builder's algorithm, Carotene, marks a slight improvement on the Autocoder application with respect to classifying past work histories of job seekers[4], but it is inconclusive on whether the algorithm outperforms Autocoder on classifying jobs ads [5]. Also of importance is that the taxonomy used in the Carotene algorithm is of finer granularity, and not consistent with the current O*net taxonomy. Thus, we concluded that use of the Autocoder algorithm was the most appropriate choice in the classification of the internship jobs ads from the experiment.

A.3.1 Matching Unemployment Values to Individual Job Ads

Each individual internship advertisement included an address for the firm and the geographic latitude/longitude coordinates. Using the opencagegeo program in stata, we utilize the latitude and longitude coordinates to create an output of the town, county, state, and zip code in which the job advertisement is located. We are able to match 100% of the county/state combinations to county level federal information processing standard codes (FIPS codes). Using a crosswalk from the NBER, we then match these county level FIPS codes to area level FIPS codes (CBSA and NECTA). We are able to match 98.2% of county level FIPS codes to the corresponding CBSA or NECTA area code. We then match these area level FIPS codes to the

[3]http://www.onetsocautocoder.com/plus/onetmatch
[4]The authors indicate a 4% higher precision.
[5]The authors are in the process of conducting a survey comparing the results of the two algorithms.

database of seasonally adjusted civilian labor force and unemployment values from the BLS. The observations for this data are at the month level. Thus, we match the individual job ads to the corresponding unemployment rate the month they are published on the job site, providing us with substantial variation in unemployment rates over time.

A.4 Sample Resumes

This section provides examples of sample resumes used in the experiment. The PDF versions of the resumes were not viewed by the firms. Upon creating a new account on the job board, a PDF version of the resume was uploaded into the site. This resume was then scanned by the job board, and a standardized resume format was created as a result. The formatting of the PDF resumes was optimized in order to minimize any errors in this process.

Figure A.1: Sample Resume One

Colin Johansson

PRESENT ADDRESS
601 R St.
Apartment #121
City, State Postal Code

EDUCATION

University of ABC

Bachelor of Science, Biology
Graduation Date: May 2017
GPA:3.2/4.0

EXPERIENCE

Student Employee	University Dining Services City, State	August 2015-Present

- Responsible for serving food to students, employees, and professors.
- Opened and closed the dining hall when scheduled.
- Organized catering events for companies and future students.

Marketing Intern	XYZ Company City, State	May 2015-August 2015

- Analyzed marketing objectives, implemented marketing plans, and modeled potential improvement to company business and advertising. Reviewed the strength of competitors in local insurance industry and created potential improvements modeled off these strengths.
- Took control of the social media outlets, resulting in over 1000 unique Facebook subscribers in one month, and drastically improved online influence.
- Utilized updated marketing techniques which engaged customers more often.

Student Volunteer	XYZ Non-Profit City, State	Jan 2015-May 2015

- Provided assistance in the construction of homes for low income families within the community.
- Helped schedule volunteers for on a weekly basis.
- Assisted with the management of materials used in the home building process.

Skills

- R Programming
-Data Analysis
-Microsoft Office
-Public Speaking
-Social Media

Figure A.2: Sample Resume Two

Darius Jackson

PRESENT ADDRESS
601 R St.
Apartment #121
City, State Postal Code

EDUCATION

University of ABC

Bachelor of Business, Marketing
Graduation Date: May 2017
GPA:3.6/4.0

EXPERIENCE

Student Employee — University Dining Services, City, State — August 2015-Present

- Responsible for serving food to students, employees, and professors.
- Opened and closed the dining hall when scheduled.
- Organized catering events for companies and future students.

Marketing Intern — XYZ Company, City, State — May 2015-August 2015

- Analyzed marketing objectives, implemented marketing plans, and modeled potential improvement to company business and advertising. Reviewed the strength of competitors in local insurance industry and created potential improvements modeled off these strengths.
- Took control of the social media outlets, resulting in over 1000 unique Facebook subscribers in one month, and drastically improved online influence.
- Utilized updated marketing techniques which engaged customers more often.

Skills

-Microsoft Office
-Public Speaking
-Social Media

Figure A.3: Sample Resume Three

Wyatt Schmidt

PRESENT ADDRESS
1925 8th Avenue
Apartment #54
City, ST ZIP

EDUCATION

University of ABC, City, State
Bachelor of Science, Psychology, 2016
GPA:3.2/4.0

EXPERIENCE

Student Liason — Recreation Center, City, ST — August 2013-Present

- Worked on several job rotations, including: assistant to management, front desk, and sanitization of equipment.
- Greeted students at the front door to provide welcoming experience.
- Provided walking tours to prospective students and athletes.

ADDITIONAL SKILLS

- Powerpoint & Prezi Presentation Software
- Social Media: Twitter, Facebook, etc.
- Market Research & Analysis
- Able to communicate with coworkers
- SPSS Statistical Analysis Software
- Hard worker who is able to work in a fast paced environment, collaborate, and meet deadlines.

A.5 Appendix Tables

Table A.1: Description of Job Skills (From Deming and Kahn 2016)

Job Skills	**Key Words and Phrases**
Cognitive	Problem Solving, Research, Analytical, Critical Thinking, Math, Statistics
Social	Communication, Teamwork, Collaboration, Negotiation, Presentation
Character	Organized, Detail-Oriented, Multi-Tasking, Time Management, Meeting Deadlines, Energetic
Writing	Writing
Customer Service	Customer, Sales, Client, Patient
Project Management	Project Management
People Management	Supervisory, Leadership, Management, Mentoring, Staff
Financial	Budgeting, Accounting, Finance, Cost
Software	Java, SQL, Python, SAS, Stata, C#, C+, C++
Computer	Computer, Spreadsheets, Common Software (e.g. Microsoft Excel, Powerpoint)

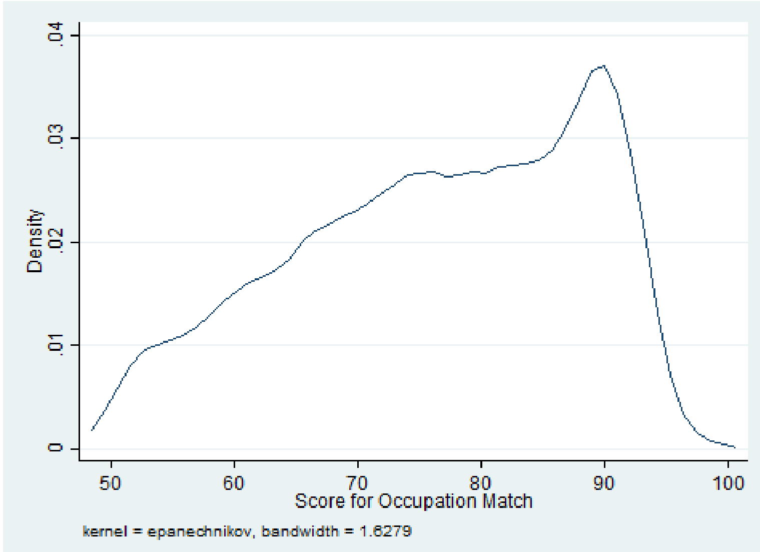
Table A.2: Summary Statistics for Resume Credentials

	All Internships	Unpaid Internships	Paid Internships	Part-Time Internships	Full-Time Internships
	(1)	(2)	(3)	(4)	(5)
Major Field of Study					
Business Administration	0.1667	0.1675	0.1658	0.1695	0.1638
Biology	0.1667	0.1653	0.1680	0.1714	0.1618
Economics	0.1667	0.1686	0.1648	0.1604	0.1730
English	0.1667	0.1633	0.1699	0.1666	0.1667
Marketing	0.1667	0.1646	0.1687	0.1649	0.1685
Psychology	0.1667	0.1707	0.1627	0.1671	0.1662
Grade Point Average (GPA)					
High GPA (3.8 or 4.0)	0.3333	0.3305	0.3361	0.3303	0.3364
Medium GPA (3.4 or 3.6)	0.3333	0.3368	0.3299	0.3320	0.3347
Low GPA (3.0 or 3.2)	0.3333	0.3326	0.3340	0.3377	0.3289
Data Skills					
Data Analysis	0.6667	0.6621	0.6711	0.6699	0.6634
Previous Internships					
Research Internship	0.1667	0.1751	0.1584	0.1659	0.1674
Finance Internship	0.1667	0.1688	0.1646	0.1649	0.1685
Insurance Internship	0.1667	0.1653	0.1680	0.1708	0.1625
Volunteer/Charitable Work					
Volunteer Experience #1	0.1667	0.1686	0.1648	0.1671	0.1662
Volunteer Experience #2	0.1667	0.1619	0.1713	0.1661	0.1672
Volunteer Experience #3	0.1667	0.1704	0.1631	0.1709	0.1624
College Work Experience					
Marketing	0.1667	0.1677	0.1656	0.1645	0.1688
Retail Clothing	0.1667	0.1656	0.1677	0.1676	0.1657
Server at Restaurant	0.1667	0.1684	0.1650	0.1640	0.1693
Campus Recreation Center	0.1667	0.1668	0.1665	0.1649	0.1685
University Dining Service	0.1667	0.1614	0.1718	0.1720	0.1613
Race					
Black	0.5000	0.4996	0.5003	0.5047	0.4953
Observations	11,520	5,688	5,832	5,786	5,734

A.6 Figures

Figure A.4: Kernel Density Plot for Occupation-Match Score

Figure A.5: Kernel Density Estimate for Occupation-Match Score



Appendix B

Chapter Two Appendix

Table B.1: Correlation Between Skills and Paid Status

	(1) paid	(2) paid	(3) paid	(4) paid	(5) paid
cognitive	-0.024	-0.019	-0.016	0.003	0.002
	(0.0169)	(0.0174)	(0.0177)	(0.0163)	(0.0163)
social	-0.027*	-0.024*	-0.024*	-0.003	-0.001
	(0.0138)	(0.0138)	(0.0126)	(0.0117)	(0.0119)
character	-0.042***	-0.043***	-0.039***	-0.021*	-0.021*
	(0.0144)	(0.0149)	(0.0149)	(0.0127)	(0.0122)
people	0.025	0.021	0.020	-0.006	-0.008
	(0.0199)	(0.0206)	(0.0197)	(0.0156)	(0.0159)
financial	0.219***	0.218***	0.212***	0.099***	0.098***
	(0.0331)	(0.0332)	(0.0337)	(0.0299)	(0.0298)
customer	0.106***	0.106***	0.110***	0.062***	0.062***
	(0.0157)	(0.0156)	(0.0126)	(0.0144)	(0.0147)
writing	-0.140***	-0.140***	-0.133***	-0.047***	-0.048***
	(0.0146)	(0.0149)	(0.0139)	(0.0113)	(0.0112)
project_man	-0.002	-0.001	0.031	0.024	0.023
	(0.0370)	(0.0334)	(0.0312)	(0.0290)	(0.0289)
coding	-0.022	-0.016	-0.014	0.033	0.035
	(0.0253)	(0.0256)	(0.0254)	(0.0328)	(0.0335)
computer	0.060***	0.058***	0.058***	0.017	0.018
	(0.0191)	(0.0196)	(0.0194)	(0.0182)	(0.0182)
part_dummy	-0.344***	-0.330***	-0.323***	-0.218***	-0.217***
	(0.0335)	(0.0338)	(0.0296)	(0.0289)	(0.0289)
Unemployment Rate		-0.039***	-0.072***	-0.057***	-0.016
		(0.0108)	(0.0198)	(0.0185)	(0.0239)
Area Fixed Effects			X	X	X
Occupation Fixed Effects				X	X
Time Fixed Effects					X
F-Test P-value	0.000	0.000	0.000	0.000	0.000
Observations	12835	12428	12428	12428	12428
R^2	0.212	0.211	0.269	0.390	0.392

Standard errors in parentheses

*standard errors clustered at the MSA level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Correlation Between Skills and Paid Status

	(1) paid	(2) paid	(3) paid	(4) paid	(5) paid
cognitive	-0.025	-0.019	-0.016	0.003	0.002
	(0.0216)	(0.0225)	(0.0199)	(0.0137)	(0.0138)
social	-0.026	-0.024	-0.024	-0.003	-0.001
	(0.0193)	(0.0204)	(0.0181)	(0.0129)	(0.0127)
character	-0.042*	-0.043**	-0.039*	-0.021**	-0.021**
	(0.0214)	(0.0219)	(0.0209)	(0.0090)	(0.0091)
people	0.022	0.021	0.020	-0.006	-0.008
	(0.0202)	(0.0206)	(0.0181)	(0.0125)	(0.0127)
financial	0.226***	0.218***	0.212***	0.099***	0.098***
	(0.0329)	(0.0331)	(0.0335)	(0.0301)	(0.0299)
customer	0.107***	0.106***	0.110***	0.062***	0.062***
	(0.0212)	(0.0215)	(0.0201)	(0.0146)	(0.0146)
writing	-0.139***	-0.140***	-0.133***	-0.047***	-0.048***
	(0.0209)	(0.0209)	(0.0190)	(0.0118)	(0.0112)
project_man	-0.008	-0.001	0.031	0.024	0.023
	(0.0271)	(0.0270)	(0.0275)	(0.0260)	(0.0259)
coding	-0.027	-0.016	-0.014	0.033*	0.035*
	(0.0452)	(0.0482)	(0.0454)	(0.0183)	(0.0184)
computer	0.062***	0.058***	0.058***	0.017	0.018
	(0.0169)	(0.0177)	(0.0149)	(0.0110)	(0.0109)
part_dummy	-0.349***	-0.330***	-0.323***	-0.218***	-0.217***
	(0.0298)	(0.0329)	(0.0305)	(0.0225)	(0.0225)
Unemployment Rate		-0.039***	-0.072***	-0.057***	-0.016
		(0.0070)	(0.0116)	(0.0089)	(0.0132)
Area Fixed Effects			X	X	X
Occupation Fixed Effects				X	X
Time Fixed Effects					X
F-Test P-value	0.000	0.000	0.000	0.000	0.000
Observations	12835	12428	12428	12428	12428
R^2	0.221	0.211	0.269	0.390	0.392

Standard errors in parentheses

*standard errors clustered at the occupation level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Correlation Between Skills and Paid Status

	(1)	(2)	(3)	(4)	(5)
	paid	paid	paid	paid	paid
cognitive	-0.018	-0.019	-0.016	0.003	0.002
	(0.0248)	(0.0256)	(0.0254)	(0.0221)	(0.0222)
social	-0.025	-0.024	-0.024	-0.003	-0.001
	(0.0202)	(0.0211)	(0.0207)	(0.0177)	(0.0177)
character	-0.037	-0.043*	-0.039	-0.021	-0.021
	(0.0255)	(0.0229)	(0.0265)	(0.0159)	(0.0157)
people	0.019	0.021	0.020	-0.006	-0.008
	(0.0249)	(0.0256)	(0.0254)	(0.0200)	(0.0203)
financial	0.217***	0.218***	0.212***	0.099**	0.098**
	(0.0437)	(0.0398)	(0.0447)	(0.0424)	(0.0421)
customer	0.109***	0.106***	0.110***	0.062***	0.062***
	(0.0231)	(0.0227)	(0.0237)	(0.0208)	(0.0208)
writing	-0.131***	-0.140***	-0.133***	-0.047***	-0.048***
	(0.0212)	(0.0221)	(0.0221)	(0.0161)	(0.0157)
project_man	0.032	-0.001	0.031	0.024	0.023
	(0.0335)	(0.0303)	(0.0360)	(0.0379)	(0.0375)
coding	-0.018	-0.016	-0.014	0.033	0.035
	(0.0481)	(0.0473)	(0.0493)	(0.0391)	(0.0406)
computer	0.059***	0.058**	0.058**	0.017	0.018
	(0.0222)	(0.0229)	(0.0232)	(0.0225)	(0.0226)
part_dummy	-0.325***	-0.330***	-0.323***	-0.218***	-0.217***
	(0.0382)	(0.0409)	(0.0403)	(0.0357)	(0.0361)
Unemployment Rate		-0.039***	-0.072***	-0.057***	-0.016
		(0.0105)	(0.0189)	(0.0188)	(0.0241)
Area Fixed Effects			X	X	X
Occupation Fixed Effects				X	X
Time Fixed Effects					X
F-Test P-value	0.000	0.000	0.000	0.000	0.000
Observations	12835	12428	12428	12428	12428
R^2	0.288	0.211	0.269	0.390	0.392

Standard errors in parentheses

*standard errors clustered at the MSA-OCC level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Correlations with Part Status

	(1)	(2)	(3)	(4)	(5)
	part	part	part	part	part
cognitive	-0.015	-0.003	-0.007	-0.016	-0.016
	(0.0123)	(0.0126)	(0.0123)	(0.0125)	(0.0124)
social	-0.016	-0.011	-0.011	-0.017	-0.018
	(0.0124)	(0.0132)	(0.0118)	(0.0127)	(0.0125)
character	0.009	0.015	0.015	0.006	0.006
	(0.0115)	(0.0128)	(0.0126)	(0.0116)	(0.0116)
people	-0.019**	-0.029**	-0.029***	-0.018**	-0.018**
	(0.0089)	(0.0133)	(0.0110)	(0.0088)	(0.0089)
financial	-0.005	-0.041**	-0.036*	-0.004	-0.004
	(0.0206)	(0.0199)	(0.0196)	(0.0209)	(0.0207)
customer	-0.036**	-0.046**	-0.037**	-0.036**	-0.036**
	(0.0150)	(0.0189)	(0.0150)	(0.0152)	(0.0152)
writing	0.038***	0.075***	0.067***	0.039***	0.039***
	(0.0128)	(0.0137)	(0.0132)	(0.0132)	(0.0130)
project_man	0.028	0.021	0.031	0.028	0.028
	(0.0212)	(0.0259)	(0.0247)	(0.0208)	(0.0209)
coding	-0.031	-0.086***	-0.091***	-0.036	-0.036
	(0.0393)	(0.0267)	(0.0265)	(0.0399)	(0.0397)
computer	-0.020*	-0.039***	-0.032***	-0.017	-0.018*
	(0.0107)	(0.0125)	(0.0114)	(0.0108)	(0.0108)
paid_dummy	-0.175***	-0.243***	-0.238***	-0.174***	-0.173***
	(0.0236)	(0.0265)	(0.0247)	(0.0238)	(0.0238)
Unemployment Rate		0.003	0.015	0.008	0.019
		(0.0077)	(0.0118)	(0.0101)	(0.0128)
Area Fixed Effects	X		X	X	X
Occupation Fixed Effects	X			X	X
Time Fixed Effects	X				X
F-Test P-value	0.000	0.000	0.000	0.000	0.000
Observations	12835	12428	12428	12428	12428
R^2	0.311	0.142	0.206	0.283	0.283

Standard errors in parentheses

*standard errors clustered at the MSA level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Correlations with Part Status

	(1)	(2)	(3)	(4)	(5)
	part	part	part	part	part
cognitive	-0.015*	-0.003	-0.007	-0.016*	-0.016*
	(0.0093)	(0.0145)	(0.0139)	(0.0094)	(0.0094)
social	-0.016	-0.011	-0.011	-0.017	-0.018
	(0.0137)	(0.0134)	(0.0124)	(0.0133)	(0.0132)
character	0.009	0.015	0.015	0.006	0.006
	(0.0077)	(0.0117)	(0.0119)	(0.0077)	(0.0078)
people	-0.019	-0.029**	-0.029**	-0.018	-0.018
	(0.0131)	(0.0131)	(0.0144)	(0.0133)	(0.0134)
financial	-0.005	-0.041**	-0.036*	-0.004	-0.004
	(0.0173)	(0.0184)	(0.0188)	(0.0171)	(0.0173)
customer	-0.036***	-0.046***	-0.037***	-0.036***	-0.036***
	(0.0126)	(0.0168)	(0.0118)	(0.0127)	(0.0129)
writing	0.038***	0.075***	0.067***	0.039***	0.039***
	(0.0132)	(0.0132)	(0.0125)	(0.0130)	(0.0131)
project_man	0.028	0.021	0.031	0.028	0.028
	(0.0193)	(0.0243)	(0.0189)	(0.0196)	(0.0194)
coding	-0.031	-0.086***	-0.091***	-0.036*	-0.036*
	(0.0223)	(0.0323)	(0.0279)	(0.0216)	(0.0212)
computer	-0.020**	-0.039***	-0.032***	-0.017*	-0.018*
	(0.0097)	(0.0122)	(0.0109)	(0.0095)	(0.0097)
paid_dummy	-0.175***	-0.243***	-0.238***	-0.174***	-0.173***
	(0.0248)	(0.0299)	(0.0314)	(0.0252)	(0.0253)
Unemployment Rate		0.003	0.015**	0.008	0.019*
		(0.0053)	(0.0074)	(0.0063)	(0.0100)
Area Fixed Effects	X		X	X	X
Occupation Fixed Effects	X			X	X
Time Fixed Effects	X				X
F-Test P-value	0.000	0.000	0.000	0.000	0.000
Observations	12835	12428	12428	12428	12428
R^2	0.311	0.142	0.206	0.283	0.283

Standard errors in parentheses

*standard errors clustered at the occupation level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: Correlations with Part Status

	(1)	(2)	(3)	(4)	(5)
	part	part	part	part	part
cognitive	-0.003	-0.003	-0.007	-0.016	-0.016
	(0.0162)	(0.0163)	(0.0177)	(0.0162)	(0.0161)
social	-0.005	-0.011	-0.011	-0.017	-0.018
	(0.0160)	(0.0153)	(0.0163)	(0.0199)	(0.0197)
character	0.017	0.015	0.015	0.006	0.006
	(0.0149)	(0.0143)	(0.0165)	(0.0147)	(0.0149)
people	-0.032**	-0.029**	-0.029*	-0.018	-0.018
	(0.0147)	(0.0145)	(0.0170)	(0.0164)	(0.0166)
financial	-0.038*	-0.041*	-0.036	-0.004	-0.004
	(0.0221)	(0.0210)	(0.0243)	(0.0270)	(0.0271)
customer	-0.047**	-0.046**	-0.037**	-0.036*	-0.036*
	(0.0205)	(0.0207)	(0.0182)	(0.0199)	(0.0199)
writing	0.075***	0.075***	0.067***	0.039**	0.039**
	(0.0150)	(0.0154)	(0.0165)	(0.0195)	(0.0193)
project_man	0.023	0.021	0.031	0.028	0.028
	(0.0256)	(0.0268)	(0.0258)	(0.0293)	(0.0290)
coding	-0.084***	-0.086***	-0.091***	-0.036	-0.036
	(0.0300)	(0.0308)	(0.0340)	(0.0459)	(0.0459)
computer	-0.038***	-0.039***	-0.032**	-0.017	-0.018
	(0.0145)	(0.0141)	(0.0141)	(0.0149)	(0.0147)
paid_dummy	-0.256***	-0.243***	-0.238***	-0.174***	-0.173***
	(0.0378)	(0.0360)	(0.0380)	(0.0343)	(0.0343)
Unemployment Rate		0.003	0.015	0.008	0.019
		(0.0074)	(0.0115)	(0.0102)	(0.0134)
F-Test P-value	0.00	0.00	0.00	0.0157	0.0153
Observations	12835	12428	12428	12428	12428
R^2	0.150	0.142	0.206	0.283	0.283

Standard errors in parentheses

*standard errors clustered at the MSA-OCC level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: Correlation Between Skills and Wage

	(1)	(2)
	ln_wage	ln_wage
cognitive	0.071***	0.034**
	(0.0251)	(0.0159)
social	0.019	0.005
	(0.0185)	(0.0146)
character	-0.048***	-0.030**
	(0.0178)	(0.0143)
people	0.004	0.012
	(0.0224)	(0.0091)
financial	0.033	0.041**
	(0.0538)	(0.0173)
customer	-0.014	-0.006
	(0.0227)	(0.0152)
writing	0.043	0.037
	(0.0292)	(0.0241)
project_man	0.060	-0.035
	(0.0584)	(0.0221)
coding	0.202***	0.021
	(0.0546)	(0.0304)
computer	-0.041*	-0.007
	(0.0210)	(0.0104)
paid_dummy	-0.071**	0.007
	(0.0334)	(0.0115)
part_dummy	0.058**	0.008
	(0.0253)	(0.0128)
F-test P-Value	0.00	0.00
Observations	12242	12242
R^2	0.062	0.544

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$