Economics of Education Review 107 (2025) 102661

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

The demand for data analytical skills by gender: Evidence from a field experiment

Menghan Shen [a], Xiangrui Zheng [b,*], Tong Wang [c], Xiaoyang Ye [d]

[a] Center for Chinese Public Adminstration, School of Government, Sun Yat-sen University, Guangzhou, 510275, China
[b] Graduate School of Education, Peking University, Beijing 100871, China
[c] College of International Relations, Ritsumeikan University Kinugasa Campus, 56-1 Toji-in Kitamachi, Kita-ku, Kyoto 603-8577, Japan
[d] Annenberg Institute, Brown University, Providence, Rhode Island 02912, United States

ARTICLE INFO

JEL Classification:
A2
I23
J01
J16

Keywords:
Audit experiment
Return to education
Data skill
Gender

ABSTRACT

This paper examines the return to advanced data analysis skills among job applicants from economics undergraduate programs employing a resume audit experiment. We randomly assigned fictitious resumes with three levels of data analysis skills (basic, medium, and strong) and submitted them to online job postings. Resumes with basic data analysis skills indicated proficiency in Excel. Resumes with medium data analysis skills demonstrated proficiency in Stata and SPSS, while resumes with strong data analysis skills indicated proficiency in Python and SQL, in addition to Stata and SPSS. Compared to resumes with basic skills, those with medium and strong skills received callback rates that were 2.5 and 2.8 percentage points higher, representing increases of 19.2 % and 21.5 %, respectively. For female applicants, resumes with medium and strong skills received callback rates that were 3.4 and 5.1 percentage points higher, corresponding to increases of 29.8 % and 44.7 %, respectively. These differences in callback rates were statistically significantly different from zero for both the overall sample and female applicants. On the other hand, no statistically significant effect was observed for male applicants. Interview evidence suggests that employers demand data analysis skills as tangible skills, rather than merely considering them as signals of ability. This finding is consistent with human capital theory, as opposed to signaling theory. Moreover, we find evidence of gender discrimination among applicants with basic data analysis skills, where women received statistically significantly lower callback rate than men. However, for resumes indicating advanced data analysis skills, no significant gender differences emerged, suggesting statistical discrimination.

1. Introduction

In modern knowledge-based economies, the skills of the population play a crucial role (Hanushek and Woessmann, 2008). Among these skills, data analysis stands out as an important yet under-studied area. The availability of vast amounts of data and advancements in data collection, analysis, and prediction technologies have revolutionized companies across industries (Caplin et al., 2024). This exponential growth of data, coupled with the industry's drive to leverage it for improved business outcomes, has resulted in a rising demand for data analysis talent. A report by LinkedIn reveals that among 22 countries analyzed, 19 listed data science roles among their top jobs (LinkedIn, 2022). According to the U.S. Bureau of Labor Statistics, data science is projected to experience even greater growth than any other field by 2029 (Davenport & Patil, 2022).

With a strong demand for workers with data analysis talent, students majoring in social science disciplines such as economics may be uniquely positioned to capitalize on the increasing demand for these skills. In addition to building writing and analytical skills (Flynn & Quinn, 2010), they can also acquire data analysis skills through both within-major and out-of-major course choices to enhance their labor market outcomes. Typically, social science majors like economics require the completion of courses in mathematics, statistics, and econometrics. Moreover, there is an increasing trend among social scientists to incorporate applied statistics in applied economics (Ricca et al., 2019; Keniston, 2022), enabling students to address empirical

* Corresponding author at: Xiangrui Zheng, Graduate School of Education, Peking University, Beijing, China.
E-mail addresses: shenmenghan@gmail.com (M. Shen), ray.zhengxr@gmail.com (X. Zheng), tongw@fc.ritsumei.ac.jp (T. Wang), xiaoyang.ye26@gmail.com (X. Ye).

https://doi.org/10.1016/j.econedurev.2025.102661
Received 17 October 2024; Received in revised form 27 April 2025; Accepted 5 May 2025
Available online 17 May 2025
0272-7757/

questions using statistical methods. Furthermore, students in social science majors have the advantage of taking data and programming courses outside their major, which are increasingly available to non-STEM students within their institutions and beyond (Light & Wertz, 2022).

However, despite the strong interest and emphasis on teaching and learning data analytics, there is currently limited causal evidence regarding the demand for data analysis skills among students in social science disciplines. Traditionally, the human capital required for data science has been closely associated with degrees in statistics or programming (Ahmad et al., 2022). Market demand has primarily focused on hiring technical specialists proficient in coding, data processing, and information generation. This specialized approach aligns with the theory of specialization, where job repetition contributes to higher productivity levels, a characteristic commonly observed in information technologies. However, as the field of data science is experiencing a growing level of automation, with the assistance of large language models that can generate high-quality code, coding skills may become less important as job requirements compared to the past. Instead, candidates with an understanding of statistics along with expertise in specific application domains are increasingly deemed suitable for data analysis positions. Tambe (2021) argues that modern algorithmic technologies, including algorithms, AI, and data science, differ fundamentally from previous technical skills in the workforce. Informational complementarities arise when workers possess a deep understanding of algorithmic technologies and their application domains. Using job post data, he finds evidence that employers now seek algorithmic literacy across a broad range of occupations, deviating from the requirements of past technical skills (Tambe, 2021).

This paper examines the return on advanced data analysis skills for undergraduate students majoring in economics using a resume audit experiment. While the labor market generally recognizes that students with economics major have the capacity to acquire data analysis skills, there could be significant variation in the extent of skills they actually acquire even within the same school. Depending on the intensity of quantitative methods offered in the program, economics majors can be considered either as STEM majors or non-STEM majors (Keniston, 2022). Thus, both our control and treatment group resumes would appear credible to human resource departments.

We sent 4,035 fictitious resumes in response to online job postings in China, randomly varying the indicated level of data analysis skills—basic or advanced—and measured employer responses through interview callbacks. The resume with basic data analysis skills indicated the ability to analyze data using Excel. The resume with medium data analysis skills demonstrated proficiency in Stata and SPSS, while the resume with strong data analysis skills indicated proficiency in Python and SQL, in addition to Stata and SPSS. For the purposes of our study, both medium and strong skill levels are considered indicative of advanced data analysis skills.

We report two major findings. First, resumes indicating advanced data analysis skills received higher callback rate than those with basic skills. More specifically, compared to resumes with basic skills, those with medium data analysis skills received callback rates that were 2.5 percentage points higher, representing a 19.2 % increase. This coefficient is statistically significant at the 10 % level. Resumes with strong data analysis skills received callback rates that were 2.8 percentage points higher, representing a 21.5 % increase. The coefficient is statistically significant at the 5 % level. For female applicants, having medium data analysis skills increased callbacks by 3.4 percentage points, corresponding to a 29.8 % increase. This coefficient is statistically significant at the 10 % level. Resumes with strong data analysis skills received callback rates that were 5.1 percentage points higher, corresponding to a 44.7 % increase. This coefficient is statistically significant at the 1 % level. Among male applicants, compared to resumes with basic skills, those with medium and strong skills received callback rates that were 1.7 and 0.8 percentage points higher, respectively. However, these differences are not statistically significant. Our interview evidence suggests that employers demand data analysis skills as tangible skills, rather than merely considering them as signals of ability. This result is consistent with prior empirical literature that distinguish human capital effects versus signaling effects (Kane & Rouse, 1995; Ishikawa & Ryan, 2002; Lange, 2007).

Second, for resumes with basic data analysis skills, the callback rate for women is 2.9 percentage points (20 %) lower than for men, with statistical significance at the 10 percent level. However, for resumes with advanced data analysis skills, there is no longer a statistically significant difference in callback rate by gender. This finding aligns with the theory of statistical discrimination because providing additional information eliminated discrimination. In statistical discrimination, when uncertainty is reduced, beliefs about group statistics would play a smaller role in assessing quality (Bertrand & Duflo, 2016; Bohren et al., 2019). In contrast, for taste-based discrimination, providing information would not reduce the difference in callback rate, and discrimination would persist even if quality were perfectly observable.

This paper contributes to four sets of literature. First, it adds to the robust literature estimating the return to cognitive and non-cognitive skills (Deming & Kahn, 2018; Autor, 2014; Borghans et al., 2014; Weinberger, 2014; Deming, 2017; Hanushek, et al., 2015; Heckman et al., 2006; Bowles et al., 2001). In the realm of cognitive skills, researchers have found that literacy, numeracy, and problem-solving skills are associated with higher earnings (Hanushek, et al., 2015; Autor, 2014). In the non-cognitive skills domain, researchers have identified leadership and interpersonal skills (Weinberger, 2014; Bowles et al., 2001; Heckman et al., 2006), as well as traits such as self-control, time preference, and sociability as likely causes of labor market success (Heckman et al., 2006). However, these studies generally have not isolated the causal effect of cognitive or non-cognitive skills. Only a few studies have overcome this problem. Notably, Koedel & Tyhurst (2012) examine the link between math skills and labor-market outcomes using a resume-based field experiment. Our paper extends the literature in two ways: first, by focusing on data analysis skills, an important but under-studied cognitive skill in the modern knowledge-based economy; and second, by using a field experiment to isolate the causal effect of a subcomponent of cognitive skills (Koedel & Tyhurst, 2012).

Second, our paper contributes to the existing body of research that investigates labor market outcome differences across and within academic majors (Light & Wertz, 2022; Light & Rama, 2019; Leighton & Speer, 2020; Bridet & Leighton, 2015). In particular, we build upon the work of Light & Wertz (2022), who identified that taking occupationally-specific, non-disciplinary courses can contribute to within-major earnings variation. Our study finds that data analysis skills, which can be acquired within or outside of the major, can also be an important source of variation in labor market outcomes within the same major.

Third, our study is connected to the existing literature on the economic returns of STEM majors (Light & Rama, 2019; Melguizo & Wolniak, 2012; Olitsky, 2014). Previous studies have commonly adopted a dichotomous view when examining STEM majors. Light & Rama (2019) depart from this approach by introducing the concept of "STEM-intensity" and measuring it within the context of college coursework. Building upon this line of research, we also extend beyond the conventional STEM/non-STEM dichotomy and instead focus on exploring the economic implications of a specific STEM skill. While the traditional approach predicts that "STEM pays", we find an economic return to a particular STEM skill. This divergence from the traditional conceptualization of a STEM major can be useful in recognizing the potential supply of STEM professionals in economies experiencing a shortage of STEM graduates.

Fourth, this paper adds to the literature on gender discrimination in the labor market (Petit, 2007; Bertrand & Duflo, 2016; Neumark, 2018; He et al., 2023). Within this body of work, a smaller subset distinguishes between statistical discrimination and taste-based discrimination

(Bohren et al., 2019; Coffman et al., 2021; Feld et al., 2022; Bohren et al., 2023). Specifically, Feld et al. (2022) found that employers tend to believe female programmers perform worse than their male counterparts, despite there being no significant gender differences in performance. However, this form of statistical discrimination can be reduced by providing evaluators with additional information about applicants' aptitude or personality. Our study contributes to this body of research by differentiating the sources of discrimination in two ways. First, we explore it in the field of data analysis skills. Like math skills (Bohren et al., 2019; Coffman et al., 2021) and computer aptitude (Feld et al., 2022), data analysis skill is a critical cognitive skill and an area where gender-based discrimination may also be prevalent. Furthermore, we extend this work to the context of China, where gender discrimination remains widespread, but there are few studies with credible designs that differentiate the sources of discrimination.

2. Resume audit Experiment Design

We conducted our resume audit experiment between May and July 2022, submitting resumes to job vacancies in Beijing via one of China's largest job search websites: Zhaopin (zhaopin.com). In China, college students typically graduate in mid-July. In recent years, many students wait for the results of civil service examinations before applying for jobs in either the private or public sectors. Consequently, companies have adjusted their recruitment cycles for fresh graduates to align with this new job search timeline.

We were able to conduct the resume audit experiment in China because it was relatively uncommon for employers to use social media or online presence as a basis for making callback decisions. In China, the use of name-based social media platforms was very limited; individuals typically engaged under nicknames or pseudonyms, making it difficult for employers to verify identities through these channels.

2.1. Randomization

We used a between-subjects experimental design. Table 1 details the experiment design. In our experiment, each hiring firm received one fictitious resume, with one of three levels of data analysis skills randomly assigned to the applicant. Other variables, such as demographic attributes, educational background, resume length, additional skills, and salary expectation were held constant. Our research design relies on the balance of company characteristics across different treatment groups.

Table 1
Experiment design.

		College selectivity	Data analysis skills		
			Strong	Medium	Basic
Gender	Male	Selective	Strong, Male, Selective	Medium, Male, Selective	Basic, Male, Selective
		Less-selective	Strong, Male, Less-selective	Medium, Male, Less-selective	Basic, Male, Less-selective
	Female	Selective	Strong, Female, Selective	Medium, Female, Selective	Basic, Female, Selective
		Less-selective	Strong, Female, Less-selective	Medium, Female, Less-selective	Basic, Female, Less-selective
Key Characteristics Constant across Jobs			*Positions: Marketing, administration and product operation *Expected Salary: 5000-10,000 CNY *Hukou: Beijing *Major: Economics *Internship: Middle-scale private company *Self-evaluation: Pleasant personality, team spirit, good social skills, stress resistance, etc. *Certificates: English certificate (CET-6) and Computer grade II certificate *Other: Moderate student work experiences and social activities		

To explore potential heterogeneity in the effect of data analysis skills on the likelihood of receiving a job interview invitation, we also varied gender and college reputation in the fictitious resumes. Gender was indicated on the resume and via a photo in the upper-right corner, two photos differ by gender generated using AI portrait fusion technology to ensure consistent physical appeal level by gender (See Appendix C). To minimize bias due to appearance, we used photos of average-looking individuals. Additionally, we varied the resumes by the reputation of the applicant's college, categorizing them as "selective" or "less-selective." Beijing Jiaotong University (BJTU) represented the selective university (ranking in China No. 41), while the Capital University of International Economics and Business (CUIEB) represented the less-selective institution (ranking in China No. 139). We fully cross-randomized the treatment and control conditions, resulting in a 3 × 2 × 2 factorial design with 12 total resume templates. Each template included an email address and a phone number registered in Beijing to receive callbacks. To ensure realism, we invited HR professionals from both public and private firms in China to review the resumes.

2.2. Creating Fictitious Resumes

To create the fictitious resumes, we developed realistic yet fictional templates. First, we gathered 50 resumes from the same job search site to identify key resume features. We performed a statistical analysis of common characteristics among economics graduates applying for jobs in Beijing, considering factors such as expected positions, salaries, internship experience, student work experience, English proficiency, and certifications.

Based on these attributes, we created a profile of a fictitious new college graduate applying for marketing, administration, and product operation positions in Beijing, with a monthly salary expectation of 5,000-10,000 CNY (707-1415 USD). This applicant holds a Beijing hukou[1], is about to graduate with a bachelor's degree in economics, has completed an internship at a mid-sized private company, possesses some student work experience and social activities, and holds both an English CET-6 certificate and a Computer Grade II certificate. Additionally, the resume includes a "self-evaluation" paragraph—a unique feature in Chinese resumes—that highlights traits such as a pleasant personality, team spirit, good social skills, stress resistance, execution ability, and openness, which are commonly expressed in many resumes we reviewed.

For each fictitious resume, one of three levels of data analysis skills was randomly assigned: basic skills (Excel only), medium skills (Excel + Stata + SPSS), and strong skills (Excel + Stata + SPSS + Python + SQL). We categorize medium and strong skills as advanced data analysis skills. These categories represent different levels of data analysis proficiency. Excel is widely used for office work and data analysis; Stata and SPSS are statistical software commonly taught and used in social science in China; Python and SQL are advanced tools for more complex data analysis, particular in big data processing and machine learning applications. The self-evaluation section elaborate on the practical application of these skills in internships, work experiences, and academic competitions. More specifically, for basic skills (Excel only) resumes, in the self-description section, they stated that "*I am able to use Excel, PPT, and Word in Microsoft Office proficiently, have a strong problem analysis ability.*" In the internship experience, the resume stated that "*assist in maintaining the local online community, perform daily product scheduling*

[1] The Beijing hukou, or household registration, is a system in China that grants residents specific social benefits and access to services like education and healthcare in Beijing.

and sharing with my team, and pay real-time attention to the market and competing products using Excel." In the project experience (which usually refers to student competitions or research project), the resume stated that "*analyzing and visualizing the questionnaire responses using Excel.*"

For the medium skills (Excel + Stata + SPSS) resumes, in the self-description section, the resumes stated that "*I have studied professional data analysis software such as Stata and SPSS in school and I am able to analyze business, sales, and user data, and perform data visualization.*" In the internship experience and project experience, the resume stated similar experiences as the control group, but added the data analysis program to SPSS and Stata, and to be realistic, slightly increased the difficulty of the task accomplished. For example, in the task accomplished section, the resume wrote "*such as assist in the weekly summary and visualization of sales data by SPSS/Excel to optimize the selection of recommended products in order to meet the needs of users in different channels.*"

For the strong skills (Excel + Stata + SPSS + Python + SQL) resumes, in the self-description section, the resumes stated that "*I am proficient in using certain programming methods for data analysis. I have studied professional data analysis software such as Stata and SPSS in school; moreover, I am familiar with Python and SQL languages and am able to use crawler technology to collect data, analyze business, sales, and user data, and perform data visualization. Able to explore automation methods and improve office efficiency by programming.*" In the internship experience and project experience, the resume stated similar experiences as the control group, but added the data analysis program to SPSS and Stata in addition to Python and SQL. Similar to the medium group, we slightly increased the difficulty of the task accomplished. For example, in the task accomplished section, we wrote "*perform sales data management and analysis by Python/SQL; assist in the weekly summary and visualization of sales data by Python to optimize the selection of recommended products in order to meet the needs of users in different channels*." The resume templates can be found in Appendix B

2.3. Selecting Job Vacancies and the Application Procedure

We searched for all possible positions that college graduates from humanities and social sciences in China might apply for at zhaopin.com. This includes roles in operations, product management, market promotion, sales, procurement, finance and insurance, human resources, administrative management, data analysis, and others.

Between May and July 2022, we collected information on job vacancies in Beijing requiring a bachelor's or associate bachelor's degree in these fields. We restricted the vacancies to jobs for fresh graduates, defined as positions requiring one year or less of work experience. Each day, we used crawler technology to collect hundreds of real job advertisements, cataloging the application link and all related information. During this period, we collected data on 4,035 vacancies, covering a wide range of company and industry types.

To send out resumes, we recruited four college student research assistants. In April 2022, we provided field training on campus. All research assistants were instructed on how to create and send out assigned resumes on the job search website, respond to employers, and record responses. In addition to the four research assistants, a project manager oversaw the team, randomly checked for recording errors, and asked the research assistant team to cross-check the recorded results every week.

Each research assistant was assigned a set of three resumes that were identical in all respects except for the level of data analysis skills, which were randomly assigned to represent three distinct levels: basic, medium, and strong. Thus, each set of three otherwise identical resumes included the control group (basic level of data analysis skills) and the treatment groups (medium and strong levels of data analysis skills). We randomized research assistants to sets of resumes by resume gender. The advantage of this approach was that it minimized confusion and errors, as research assistants were responsible for a consistent set of three resumes.

Our experiment took place on an online job search site (zhaopin.com) that allows employers and applicants to engage in instant chats for more detailed information before deciding on an official interview. This feature, common on Chinese job search sites in recent years, means that employer responses to our fictitious applications could come via traditional email, mobile phone, or messages from the job search site. Quick conversations between our research assistants and employers could potentially contaminate our standardization and objectivity. The characteristics of the research assistants may have influenced their interactions with HR, and consequently, the callback rate. To mitigate this, we conducted extensive training to standardize the procedure and process. We trained our research team to chat with employers in a unified and standardized manner using a "script" that summarized answers to each type of question, based on the pilot experiment. We registered the final response of an invitation or rejection regardless of the number of conversation rounds. In line with previous resume audit experiments, we defined a positive callback as a situation where the applicant received: a clear invitation to a job interview via email, mobile phone, or job site message; a proposal to exchange contact information and further chat on WeChat (a widely used social platform in China, similar to Facebook or WhatsApp); an inquiry to provide the employer with more information about the applicant's traits and competence. The translated "script" can be found in the Appendix D. We trained our RA for a two weeks during the pilot period in April 2022.

As each research assistant was assigned a set of three resumes, each person managed three accounts. The accounts linked corresponding emails and mobile phone numbers to receive callbacks and register the responses. Upon receiving a positive response, the research assistants promptly replied to employers to decline the offer, minimizing inconvenience for the employers. Responses received within 30 days of the application were registered. The outcome variable in our analysis is set to 1 if the applicant received an explicit invitation to a job interview, and 0 otherwise.

Each day, after collecting newly posted job advertisements, we randomly assigned one of the twelve resume templates to each job vacancy and assigned a research assistant to apply for the position. All resumes were submitted via the job search site's application channel, each containing an active email address and mobile phone number linked to the respective fictitious applicant through which employers could contact us. To maximize visibility and overall callback rate, all applications were sent between 9 AM and 2 PM on workdays. The audit experiment ran from May 31 to July 21, 2022.

2.4. Data Collection

Our primary outcome of interest was whether employers would call back a job applicant based on their data analysis skills. We also collected detailed information about the vacancies, including the offered salary, degree requirement, work experience requirement, recruitment channel (social/campus/internship), and the number of workers wanted. Additionally, we linked to the Tianyancha database, a well-known business search platform, to gather enterprise characteristics such as company scale, property nature, establishment year, stock market listing status, registered capital, industry type, and business scope.

3. Experiment results

3.1. Covariate balance

Table 2 presents the summary statistics for our study. Panel A describes the key variables of the individuals in our fictitious resumes. The gender proportions of job seekers in the basic, medium, and strong skill groups closely resemble those of the overall sample, with no significant differences observed ($p = .443$). However, the basic skilled group is slightly more likely to have resumes with selective colleges ($p = .065$). In

Table 2
Balance checks.

	(1) Total	(2) C (Basic)	(3) T (Medium)	(4) T (Strong)	ANOVA *P*-value
Panel A. *Individual characteristics*					
Female	0.486	0.489	0.473	0.497	0.443
	(0.500)	(0.500)	(0.499)	(0.500)	
Selective school	0.515	0.535	0.518	0.490	0.065*
	(0.500)	(0.499)	(0.500)	(0.500)	
Panel B. *Company characteristics*					
Company scale (large=1)	0.641	0.636	0.649	0.638	0.748
	(0.480)	(0.481)	(0.477)	(0.481)	
Foreign-owned	0.109	0.108	0.109	0.109	0.999
	(0.311)	(0.311)	(0.312)	(0.311)	
Listed Company	0.117	0.123	0.115	0.113	0.676
	(0.322)	(0.329)	(0.319)	(0.317)	
Registered capital (CNY)	38678.1	39243.1	34718.1	42176.1	0.631
	(220689.2)	(252708.0)	(196719.3)	(207658.1)	
Data-driven company	0.199	0.198	0.195	0.203	0.893
	(0.399)	(0.399)	(0.397)	(0.402)	
Agriculture & manufacture industry	0.071	0.076	0.061	0.077	0.196
	(0.257)	(0.264)	(0.240)	(0.267)	
Finance industry	0.035	0.036	0.036	0.034	0.966
	(0.185)	(0.187)	(0.185)	(0.182)	
Internet industry	0.073	0.077	0.066	0.077	0.440
	(0.260)	(0.267)	(0.248)	(0.266)	
Technology industry	0.384	0.382	0.375	0.396	0.504
	(0.486)	(0.486)	(0.484)	(0.489)	
Core district	0.456	0.460	0.445	0.464	0.579
	(0.498)	(0.499)	(0.497)	(0.499)	
Panel C. *Job post characteristics*					
Annual wage (CNY)	91314.1	88471.9	92033.5	93565.1	0.170
	(73113.5)	(65816.9)	(68134.8)	(84466.6)	
Working experience needed (1 year or less=1)	0.358	0.360	0.371	0.344	0.334
	(0.480)	(0.480)	(0.483)	(0.475)	
Bachelor degree needed	0.543	0.553	0.528	0.549	0.356
	(0.498)	(0.497)	(0.499)	(0.498)	
Number of recruiting workers	6.838	7.692	6.581	6.203	0.773
	(52.819)	(60.799)	(48.240)	(48.125)	
Full-time job	0.687	0.667	0.698	0.696	0.159
	(0.464)	(0.471)	(0.459)	(0.460)	
Office skills needed	0.109	0.105	0.120	0.103	0.314
	(0.312)	(0.307)	(0.325)	(0.303)	
Computer major needed	0.130	0.119	0.139	0.132	0.278
	(0.336)	(0.324)	(0.346)	(0.339)	
Data skills needed	0.169	0.168	0.155	0.184	0.125
	(0.375)	(0.374)	(0.362)	(0.388)	
Product & operation position	0.138	0.138	0.142	0.135	0.856
	(0.345)	(0.345)	(0.349)	(0.341)	
Procurement position	0.023	0.020	0.021	0.027	0.457
	(0.148)	(0.139)	(0.145)	(0.161)	
Finance & insurance related position	0.059	0.054	0.054	0.070	0.178
	(0.236)	(0.227)	(0.226)	(0.255)	
Human Resources position	0.128	0.120	0.135	0.130	0.486
	(0.334)	(0.325)	(0.342)	(0.337)	
Administrative position	0.313	0.314	0.298	0.327	0.268
	(0.464)	(0.464)	(0.458)	(0.469)	
Business analysis position	0.043	0.047	0.037	0.045	0.363
	(0.203)	(0.212)	(0.189)	(0.208)	
Marketing & sales position	0.336	0.331	0.338	0.339	0.900
	(0.472)	(0.471)	(0.473)	(0.474)	
Accounting position	0.027	0.031	0.030	0.020	0.100
	(0.163)	(0.174)	(0.172)	(0.140)	
Post length	880.144	882.761	891.808	865.331	0.567
	(619.050)	(571.042)	(600.884)	(683.042)	
N	4035	1377	1351	1307	

Note: This table presents the result of balance checks across all control variables. The last column reports p-values from ANOVA analyses that compare the differences in each variable across the control and two treatment groups. Standard deviations are reported in parentheses. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The *p*-value of a joint *F*-test from a multinomial logit model that predicts treatment assignment using all the variables is 0.348.

Panel B, we find no significant differences in firm characteristics across the three groups. This includes the proportions of large-scale firms, foreign-owned companies, listed companies, various industries, and whether the company is in the core district in Beijing. Additionally, the assignment to different skill-level groups shows no significant differences in the registered capital of the companies. Panel C reveals no significant differences in the average salary, the average number of recruiting workers per advertisement, various job requirements, and the distribution of different types of positions by the skill-level groups. Only one out of the 29 ANOVA tests is significant at the 10 % level ($p = .065$). A test of joint significance across all variables fails to reject the null hypothesis of equality between the treatment and control groups ($p =$

.348).

Given that we submitted only one resume per position to mitigate the risk of detection, a potential concern is that this approach could introduce bias if certain types of firms systematically received a particular treatment. However, as shown in Table 2, the assignment of treatments is well balanced across a range of firm and job characteristics. This is consistent with a successful randomization.

3.2. Demand for data analysis skills

Table 3 presents results for the full sample and then separately for male and female resumes. Table 3, Column 1 provides the results of the OLS regression analysis for the baseline model without controlling for individual, company, or position characteristics. Column 2 includes individual characteristics only, while Column 3 includes both individual and firm characteristics, and Column 4 includes individual, firm, and job post characteristics.

Panel A presents regression results for all resumes. The average callback rate for the basic data analysis skill group is 13 %. Both the medium and strong data analysis skill groups exhibit higher callback rate compared to the basic data analysis skill group. In Column 1, with no controls, the coefficient for the medium data analysis skill group is 3 percentage points, statistically significant at the 5 % level. In Column 2, with control variables for individuals, the coefficient is 2.9 percentage points. In Column 3, with control variables for individuals and companies, the coefficient is 2.8 percentage points. In Column 4, with control variables for individuals, companies, and positions, the coefficient is 2.5 percentage points, indicating an increase of approximately 19.2 % from the baseline. This coefficient is statistically significant at the 10 % level. These results indicate that the inclusion of controls does not qualitatively alter the results. Thus, we focus on discussing the coefficients with all controls in Column 4.

The coefficient associated with the strong data analysis skill group is 2.8 percentage points in Column 4, statistically significant at the 5 percent level. This implies that the strong data analysis skill group has a 2.8 percentage point (or 21.5 %) higher callback rate than the basic data analysis skill group. There is no significant difference in callback rate between the medium and strong skill groups, indicating that knowing Python and SQL does not generate additional benefits in the labor market beyond other data analysis skills.

In Panels B and C, we break down the results by gender. Among women, having advanced data analysis skills leads to a substantial increase in the callback rate. Having medium data analysis skills increased callbacks by 3.4 percentage points, corresponding to a 29.8 % increase. This coefficient is statistically significant at the 10 % level. Having strong data analysis skills increased callbacks by 5.1 percentage points, corresponding to a 44.7 % increase. This coefficient is statistically significant at the 1 % level. For men, compared to resumes with basic skills, those with medium and strong data analysis skills received callback rates that were 1.7 and 0.8 percentage points higher, respectively. However, these differences are not statistically significant.

We conducted robustness checks using logistic regression, as presented in Table A.1, and found similar results. In Appendix Table A.2, Column 1, we excluded a subset of samples where there was any interaction between HR personnel and our research assistants beyond direct interview invitations. This subset consisted of only 29 samples, and removing them did not alter the results. In Appendix Table A.2, Column 2, we restricted the sample to those for which the resume was downloaded (since HR can view partial resume information without downloading it). In Appendix Table A.3, Column 3, we included controls for post length, as well as district (within Beijing) and posting month fixed effects, and again obtained consistent results. Finally, in Appendix Table A.2, Column 4, we added fixed effects at the research assistant level and found that the results remained unchanged with this additional control.

We also explored the heterogeneous effects by job characteristics. In subgroup analyses based on Tables A.3 to A.5, positive returns were observed in the service industry and higher-wage jobs, but not in the agriculture and manufacturing sectors or lower-wage jobs. The service industry relies heavily on data analysis for improving customer experience, optimizing operations, and data-driven decision-making. In contrast, the agriculture and manufacturing sectors traditionally rely less on data analysis techniques. Higher-wage jobs typically involve complex tasks requiring advanced data analysis skills, whereas lower-wage jobs tend to involve routine tasks with less emphasis on data analysis. Positive returns of advanced data analysis skills were observed in finance, technology, and internet industry, but not among other industries, potentially due to the high demand for data-driven innovation and strategy in these fields.

Moreover, positive returns were noted in listed companies, in medium-sized or larger companies, but not in smaller companies. Larger companies typically handle greater data volumes, necessitating skilled data analysts to extract insights and drive decisions. These firms also have more resources to invest in data analysis tools and infrastructure. Smaller companies, with limited data resources, may prioritize data analysis less.

Our study also indicates positive returns to advanced data analysis skills in both data-driven and non-data-driven companies, underscoring the broad value of these skills. Data analysis is increasingly recognized for its potential to drive insights and informed decision-making across various organizational contexts.

Finally, our findings show positive returns to advanced data analysis skills in job posts that explicitly require such skills as well as in those that do not. The rising importance of advanced data analysis across industries drives demand for candidates with these capabilities. Data analysis skills, being highly transferable, contribute significantly to organizational effectiveness in various roles and sectors. However, no return was observed in job posts explicitly requiring office skills or computer science majors, likely because these roles focus on administrative tasks or technical programming skills rather than advanced data

Table 3
OLS regression results for predicting callbacks.

	(1)	(2)	(3)	(4)
Panel A. All				
Control means	0.130	0.130	0.130	0.130
T (Medium)	0.030**	0.029**	0.028**	0.025*
	(0.013)	(0.013)	(0.013)	(0.013)
T (Strong)	0.031**	0.030**	0.033**	0.028**
	(0.014)	(0.014)	(0.013)	(0.013)
P-value for T(Medium) = T (Strong)	0.956	0.971	0.728	0.831
N	4,035	4,035	4,035	4,035
Panel B. Female				
Control means	0.114	0.114	0.114	0.114
T (Medium)	0.042**	0.042**	0.038**	0.034*
	(0.019)	(0.019)	(0.019)	(0.018)
T (Strong)	0.052***	0.051***	0.053***	0.051***
	(0.019)	(0.019)	(0.019)	(0.019)
P-value for T(Medium) = T (Strong)	0.638	0.680	0.450	0.389
N	1,962	1,962	1,962	1,962
Panel C. Male				
Control means	0.145	0.145	0.145	0.145
T (Medium)	0.018	0.017	0.018	0.017
	(0.019)	(0.019)	(0.019)	(0.019)
T (Strong)	0.010	0.010	0.013	0.008
	(0.019)	(0.019)	(0.019)	(0.019)
P-value for T (Medium) = T (Strong)	0.698	0.701	0.815	0.640
N	2,073	2,073	2,073	2,073

Note: The dependent variable is an indicator for receiving a personalized response from a potential employer. Column 2 only includes control variables for individual. Column 3 includes control variables for individuals and companies. Column 4 includes all control variables for individuals, companies, and positions. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

analysis.

It is important to point out that this study may have underestimated the effect of data analysis skills since the estimated coefficient reflects both the return to skills and the expected effect on the wage. Since firms employ a person with high skills, they must expect to pay a higher wage than they would for someone with lower skills. All else equal, the expectation of a higher wage would have a negative effect on the callback probability.

3.3. Human capital effects or signaling effects

While our research design provides a high degree of confidence in the causal interpretation of our results, our experiment does not allow us to differentiate the effects of signaling versus human capital in the impact of data analysis skills. Signaling effects refer to the idea that certain skills, credentials, or qualifications serve as proxies for unobservable traits such as ability, work ethic, or productivity, enabling employers to infer an individual's potential value to the firm (Spence, 1973). Human capital effects, on the other hand, refer to the direct productivity-enhancing value of skills and knowledge that individuals acquire, which contribute to their ability to perform job-related tasks more effectively (Becker, 1964).

It is possible that data analysis skills serve as a signal to employers, indicating a candidate's capacity to quickly learn, adapt, or handle complex tasks. However, data analysis skills also represent a tangible, productivity-enhancing asset, improving an individual's ability to extract meaningful insights from data and make informed decisions—skills that are valued in today's data-driven labor market. It is important to differentiate signaling effects from human capital effects because this distinction has critical implications for understanding the mechanisms driving labor market returns to skills, as well as for designing effective education and training policies.

To understand whether the labor market return is due to signaling effects versus human capital effects, we conducted interviews with ten HR professionals from ten companies across various industries, including clothing, real estate, car manufacturing, technology, education, internet, and games. The interview questions and firm information are provided in Appendix E and Appendix F. We first asked HR professionals whether candidates with advanced data analysis skills receive higher callback rate at their firms. All agreed that such candidates were more likely to be called back, as data analysis was considered a crucial competency for understanding business trends and making informed decisions. We also inquired about the relative value of using Stata/SPSS compared to Python and SQL. Some HR professionals noted that their companies did not have specific software requirements; instead, they prioritized candidates' analytical abilities and insight generation. They believed that specific software skills can be learned on the job, emphasizing the importance of advanced analytical skills over familiarity with particular tools.

We also explicitly asked HR professionals if the labor market return to having advanced data analysis skills was due to signaling or human capital effects. None of the HR professionals we interviewed attributed the higher return to data analysis skills to signaling effects. From their perspective, data analysis is regarded as a concrete and practical skillset highly valued by firms across various industries, extending to almost all positions within organizations, regardless of whether job postings explicitly mention data analysis as a requirement. This focus on the direct utility of data analysis skills suggests that the observed labor market return is rooted in human capital effects—the productivity-enhancing value of these skills—rather than signaling.

3.4. Taste-based discrimination versus statistical discrimination

In our experiment, we randomly assigned different levels of data analysis skills and gender to fictitious resumes. This randomization design allows us to examine potential gender discrimination and differentiate between taste-based and statistical discrimination.

Taste-based discrimination occurs when individuals or organizations discriminate against a specific group due to disutility from interacting with its members (Becker, 1971). In contrast, statistical discrimination arises when decisions are based on group averages or stereotypes, rather than individual characteristics (Phelps, 1972; Arrow, 1973). This type of discrimination is typically driven by incomplete information or reliance on generalizations to predict behavior, skills, or productivity.

To differentiate between taste-based and statistical discrimination, previous literature pointed out that statistical discrimination can be mitigated by providing more detailed information to employers about individual candidates, while taste-based discrimination would not respond to additional information (Bertrand & Duflo, 2016). Understanding whether discrimination is statistical or taste-based is crucial for developing effective policies to reduce discrimination.

Our analysis in Table 4 reveals evidence of gender discrimination for applicants with basic data analysis skills. Specifically, the callback rate for male applicants in the basic skill group is 14.5 percent. The rate for female applicants is 2.9 percentage points (or 20 %) lower compared to men. The difference is statistically significant difference at the 10 percent level. Given the random assignment of fictitious CVs to employers, this difference suggests the presence of gender discrimination against economics graduates with basic data analysis skills.

However, we did not find gender difference in the group with advanced data analysis skills. More specifically, in the medium skill group, the rate for females is only 1.3 percentage points lower than males and this difference is no longer statistically significant. In the strong skill group, the difference in callback rate by gender is also not statistically significant. When we combine the medium and strong skill groups, the gender differences in callback rate are also not statistically significant. This indicates that there is no significant difference in callback rate between males and females in the medium and strong skill groups.

Because the gender difference in callback rate diminishes as additional information about the applicant's data analysis skills is provided, this suggests that the gap in callback rate is sensitive to the amount of information employers have about the applicant. Under statistical discrimination, employers may rely on group-based expectations when information is limited. For example, they may assume that male applicants with basic skills are more likely to quickly acquire advanced data skills, while female applicants with basic skills are not. This leads to lower callbacks for women in the basic data analysis skill group. When advanced skills are clearly demonstrated, employers need to rely less on stereotypes, and gender differences disappears. Thus, we cautiously interpret these findings as evidence of statistical discrimination, though further research is necessary to rule out alternative explanations.

Table 4
Gender difference for predicting callbacks.

	(1) C (Basic)	(2) T (Medium)	(3) T (Strong)	(4) T (Medium)+T (Strong)
Control means (Male)	0.145	0.163	0.155	0.159
Female	-0.029*	-0.013	0.012	-0.000
	(0.018)	(0.019)	(0.020)	(0.014)
N	1,377	1,351	1,307	2,658

Note: The dependent variable is an indicator for receiving a personalized response from a potential employer. Column 1 only includes the control group samples; Column 2 only includes the medium data analysis skills group samples; Column 3 only includes the strong data analysis group samples; Column 4 includes the medium and strong data analysis skills group samples. Columns 1-4 include all control variables for individuals, companies, and positions. Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4. Survey evidence on the supply of female students with advanced data analysis skills

While the study's findings suggest that female economics graduates stand to gain substantially from acquiring data analysis skills, conversations with HR professionals and anecdotal evidence indicate that there is a relatively lower supply of women with strong data analysis skills in the labor market, despite the gender distribution in Chinese economics and STEM fields being roughly balanced (Ministry of Education, People's Republic of China, 2022).

Previous research has suggested women's underrepresentation in technology and STEM fields may be driven by differences in ability (Ceci & Williams, 2010), access to information (Li, 2018; Porter & Serra, 2020), preference (Ding et al., 2021), and confidence (Avilova & Goldin, 2024). To formulate effective policy in light of these challenges, we posed several key questions: Are female economics graduates in China less likely to acquire advanced data analysis skills or less inclined to pursue data-driven jobs? If the latter, is this due to a lack of awareness of the labor market returns, or to lower confidence and interest?

To answer these questions, we surveyed 1,781 undergraduate students in economics programs at six major Chinese universities (see Appendix G for survey design and Appendix H for methods). Survey results (see Appendix Table A.6) reveal that female and male economics majors report similar GPAs and coursework experiences involving data analysis. However, males are more likely to use data analysis skills during internships, express greater confidence in their ability to learn these skills, and show more interest in pursuing data-driven careers. Both genders recognize the value of data analysis for improving job prospects, with women estimating a 45 % increase in callback rate, compared to 40 % for men.

The internship selection process is similar to that in the United States, where students search for opportunities, submit applications, attend interviews, and compete for available positions. The observed difference in the use of data analysis skills may be due to several factors. It's possible that male and female students apply for different types of internships, receive different offers, or take different offers. Even within the same type of internship, they may be assigned different tasks or responsibilities.

Overall, these findings suggest that the gap in advanced data analysis skills between genders is not due to differences in ability or awareness of labor market returns, but rather stems from a gap in confidence and interest. This aligns with previous research indicating that while both genders possess similar math skills, males typically exhibit greater confidence, which influences career choices and skill development (Flynn & Quinn, 2010; Ellis et al., 2016).

5. Discussion and conclusion

In our audit experiment using fictitious resumes for job postings in China, we found that resumes with advanced data analysis skills is associated with higher callback rate overall. In addition, the effects are mainly concentrated among women and not among men. Interview with HR professionals suggest that the return in labor market is likely due to human capital effects and not signaling effects. In addition, we found evidence of gender discrimination against women: among resumes with basic data analysis skills, women have statistically significantly lower callback rate compared to men. For resumes with advanced data analysis skills, there is no longer a statistically significant difference in callback rate by gender. This finding is consistent with the theory of statistical discrimination.

Given the potential large benefits for women in acquiring advanced data analysis skills, to put out more concrete public policy suggestions, we conducted an additional survey and found that female students acquire advanced data analysis skills at a similar rate as their male counterparts in school but are less interested in data analysis careers. This can potentially create gender differences in the supply of advanced data analysis skills. Future research should investigate the underlying causes of gender differences in interest and develop targeted interventions accordingly. Educational initiatives, scholarships, and mentorship programs should aim to broaden exposure and cultivate both interest and confidence among women in data analysis and mathematics, thereby increasing their representation in the growing data-driven industries and reducing gender inequality in the labor force.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT4.0 in order to translate and refine the text in the appendix. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding statement

The authors were supported by the Natural Science Foundation of Guangdong Province (Grant Number 2022A1515011130) and the MOE Project of the Key Research Institute of Humanities and Social Sciences in Universities (Grant No. 22JJD880002).

Data availability statement

Data is available upon request from first author (Menghan Shen).

Ethics approval statement

The study was approved by the IRB office in the School of Government, Sun Yat-sen University (IRB No. School of Government 202202).

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Menghan Shen and Xiangrui Zheng contributed equally to this work. We thank our editor and the two anonymous referees for their constructive feedback, which greatly improved the paper. All remaining errors are our own.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2025.102661.

References

Ahmad, N., Hamid, A., & Ahmed, V. (2022). Data science: Hype and reality. *Computer, 55*, 95–101. https://doi.org/10.1109/MC.2021.3130365

Arrow, K. J. (1973). The theory of discrimination. *Discrimination in labor markets* (pp. 3–42). Princeton University Press.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent. *Science, 344*, 843–851. https://doi.org/10.1126/science.1251868

Avilova, T., & Goldin, C. (2024). Seeking the "Missing Women" of economics with the undergraduate women in economics challenge. *Journal of Economic Perspectives, 38*, 137–162. https://doi.org/10.1257/jep.38.3.137

Becker, G. S. (1971). *The economics of discrimination*. University of Chicago Press.

Becker, G.S., 1964. Human capital: A theoretical and empirical analysis with special reference to education, First Edition.

Bertrand, M., Duflo, E., 2016. Field experiments on discrimination, in: Handbook of economic field experiments. North-Holland, pp. 309–393. https://doi.org/10.1016/bs.hefe.2016.08.004.

Bohren, J. A., Haggag, K., Imas, A., & Pope, D. G. (2023). Inaccurate statistical discrimination: An identification problem. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_01367

Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review, 109*, 3395–3436. https://doi.org/10.1257/aer.20171829

Borghans, L., Ter Weel, B., & Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *ILR Review, 67*, 287–334. https://doi.org/10.1177/001979391406700202

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature, 39*, 1137–1176. https://doi.org/10.1257/jel.39.4.1137

Bridet, L., & Leighton, M. A. (2015). The major decision : Labor market implications of the timing of specialization in college. *School of Economics and Finance, Discussion Paper Series*.

Caplin, A., Deming, D. J., Li, S., Martin, D. J., Marx, P., Weidmann, B., & Ye, K. J. (2024). The ABC's of who benefits from working with AI: Ability, beliefs, and calibration. *Working Paper Series*. https://doi.org/10.3386/w33021

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science, 19*, 275–279. https://doi.org/10.1177/0963721410383241

Coffman, K. B., Exley, C. L., & Niederle, M. (2021). The role of beliefs in driving gender discrimination. *Management Science*. https://doi.org/10.1287/mnsc.2020.3660

Davenport, T. H., & Patil, D. J. (2022). Is data scientist still the sexiest job of the 21st century? *Harvard Business Review*.

Deming, D., & Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics, 36*, S337–S369. https://doi.org/10.1086/694106

Deming, D. J. (2017). The growing importance of social skills in the labor market*. *The Quarterly Journal of Economics, 132*, 1593–1640. https://doi.org/10.1093/qje/qjx022

Ding, Y., Li, W., Li, X., Wu, Y., Yang, J., & Ye, X. (2021). Heterogeneous major preferences for extrinsic incentives: The effects of wage information on the gender gap in STEM major choice. *Research in Higher Education, 62*, 1113–1145. https://doi.org/10.1007/s11162-021-09636-w

Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *Plos One, 11*, Article e0157447. https://doi.org/10.1371/journal.pone.0157447

Feld, J., Ip, E., Leibbrandt, A., Vecci, J., 2022. Identifying and overcoming gender barriers in tech: A field experiment on inaccurate statistical discrimination. https://doi.org/10.2139/ssrn.4238277.

Flynn, P. M., & Quinn, M. A. (2010). Economics: good choice of major for future ceos. *The American Economist*. https://doi.org/10.1177/056943451005500107

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review, 73*, 103–130. https://doi.org/10.1016/j.euroecorev.2014.10.006

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*, 607–668. https://doi.org/10.1257/jel.46.3.607

He, H., Li, S. X., & Han, Y. (2023). Labor market discrimination against family responsibilities: A correspondence study with policy change in China. *Journal of Labor Economics, 41*, 361–387. https://doi.org/10.1086/719966

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*, 411–482. https://doi.org/10.1086/504455

Ishikawa, M., & Ryan, D. (2002). Schooling, basic skills and economic outcomes. *Economics of Education Review, 21*, 231–243. https://doi.org/10.1016/S0272-7757(01)00005-X

Kane, T. J., & Rouse, C. E. (1995). Labor-market returns to two- and four-year college. *The American Economic Review, 85*, 600–614.

Keniston, M., 2022. Economics reclassification as STEM concentration to benefit international students. Hamilton College. Retrieved from: http://www.hamilton.edu/news/story/economics-stem-major-international-students.

Koedel, C., & Tyhurst, E. (2012). Math skills and labor-market outcomes: Evidence from a resume-based field experiment. *Economics of Education Review, 31*, 131–140. https://doi.org/10.1016/j.econedurev.2011.09.006

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics, 25*, 1–35. https://doi.org/10.1086/508730

Leighton, M., & Speer, J. D. (2020). Labor market returns to college major specificity. *European Economic Review, 128*, Article 103489. https://doi.org/10.1016/j.euroecorev.2020.103489

Li, H.-H. (2018). Do mentoring, information, and nudge reduce the gender gap in economics majors? *Economics of Education Review, 64*, 165–183. https://doi.org/10.1016/j.econedurev.2018.04.004

Light, A., & Rama, A. (2019). Moving beyond the STEM/non-STEM dichotomy: wage benefits to increasing the STEM-intensities of college coursework and occupational requirements. *Education Economics, 27*, 358–382. https://doi.org/10.1080/09645292.2019.1616078

Light, A., & Wertz, S. S. (2022). Should English majors take computer science courses? Labor market benefits of the occupational specificity of major and nonmajor college credits. *Economics of Education Review, 88*, Article 102263. https://doi.org/10.1016/j.econedurev.2022.102263

LinkedIn, 2022. LinkedIn jobs on the rise 2022: The roles that are growing in demand Retrieved from: https://www.linkedin.com/pulse/linkedin-jobs-rise-2022-roles-growing-demand-linkedin-news-europe.

Melguizo, T., & Wolniak, G. C. (2012). The earnings benefits of majoring in STEM fields among high achieving minority students. *Research in Higher Education, 53*, 383–405. https://doi.org/10.1007/s11162-011-9238-z

Ministry of Education, People's Republic of China, 2022. The situation of students in schools at all levels and types of education Retrieved from: http://www.moe.gov.cn/jyb_sjzl/moe_560/2022/quanguo/202401/t20240110_1099535.html.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature, 56*, 799–866. https://doi.org/10.1257/jel.20161309

Olitsky, N. H. (2014). How do academic achievement and gender affect the earnings of STEM majors? A propensity score matching approach. *Research in Higher Education, 55*, 245–271. https://doi.org/10.1007/s11162-013-9310-y

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the french financial sector. *Labour Economics, 14*, 371–391. https://doi.org/10.1016/j.labeco.2006.01.006

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review, 62*, 659–661.

Porter, C., & Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics, 12*, 226–254. https://doi.org/10.1257/app.20180426

Ricca, B., Blaine, B., Donovan, K., & Geraci, A. (2019). *Changing paradigms: Faculty moving to data science from other disciplines*. Mathematical and Computing Sciences Faculty/Staff Publications.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*, 355. https://doi.org/10.2307/1882010

Tambe, P. (2021). The growing importance of algorithmic literacy. *SSRN Journal*. https://doi.org/10.2139/ssrn.3776492

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *The Review of Economics and Statistics, 96*, 849–861. https://doi.org/10.1162/REST_a_00449