Title

Does signaling college-level human capital matter? An experimental study in India

Authors
Batheja, Deepshikha
Hirshleifer, Sarojini
Kaur, Opinder

Publication Date
2024-10-19

DOI
10.26085/C3PW26

Series Name: WPS
Paper No: 245
Issue Date: 10/19/2024

Does signaling college-level human capital matter? An experimental study in India

Deepshikha Batheja, Sarojini Hirshleifer, and Opinder Kaur

CEGA
Center for Effective Global Action

Working Paper Series
Center for Effective Global Action
University of California

Does signaling college-level human capital matter? An experimental study in India*

Deepshikha Batheja[†], Sarojini Hirshleifer[‡], Opinder Kaur[§]

November 8, 2024

Abstract

We measure the impact of two main signals of tertiary-level human capital accumulation, college quality and certification, on hiring in India. Using a correspondence experiment, we send 16,944 resumes to 1412 job postings for recent engineering graduates at small and medium firms. In precisely estimated results, we find that these employers do not respond to signals of tertiary education quality. Specifically, there is no impact on callbacks of having graduated from a mid-tier college ranked in the top 300 relative to an unranked college outside of the top 1000, despite significant government investment in college rankings. There is also no impact of scoring in the highest as opposed to the lowest quartile of a post-tertiary certification test that has been taken by millions of graduating students. There is evidence that women modestly benefit in the first stage of hiring in this market, with this effect concentrated in some regions.

Keywords: labor markets, hiring, tertiary education, college quality, development, audit study, correspondence experiment
JEL Codes: I23, I25, I26, J23, M51, O15

*We gratefully acknowledge J-PAL's Post-Primary Education (PPE) Initiative (grant ID: PPE-1646), funded through the United Kingdom Government's Department for International Development (DFID). We appreciate the excellent research support that Rashi Maheshwari and Aman Haider provided in the field. We thank Anil Deolalikar, Vittorio Bassi and seminar participants at the CEGA Global Workshop for their helpful comments. The RCT was pre-registered at the AEA registry with ID AEARCTR-0007940.

[†]One Health Trust (email: deepshikha@onehealthtrust.org)
[‡]University of California, Riverside and CEGA (email: sarojini.hirshleifer@ucr.edu)
[§]University at Buffalo, SUNY (email: opinderk@buffalo.edu)

1 Introduction

Limited information about jobseekers' skills has long been recognized as a source of inefficiency in the hiring process (Spence, 1973; Altonji and Pierret, 2001). One potential approach to addressing information asymmetries is through tertiary education degrees, which are a potentially important signal of skills to employers (Arcidiacono, Bayer and Hizmo, 2010). Poor quality degrees, however, are unlikely to provide effective signals (Hanushek and Woessmann, 2010). Thus, a profusion of such degrees is likely to create frictions in the labor market, which can in turn contribute to aggregate unemployment (Mortensen and Pissarides, 1999).

Such labor market information frictions are particularly a concern in India, which has seen massive growth in the number of tertiary education institutions over the past two decades, many of which are of uncertain quality (University Grants Commission, 2021). This has coincided with an unemployment rate for college graduates that is more than double the rate for those with less than a college degree.[1] These circumstances have led to the widespread use of certification exams, offered by large testing companies, which are taken by many recent graduates entering the labor market in order to signal quality (Aggarwal, 2022). Thus, two first-order questions in this setting are how employers view candidates with degrees from colleges of uncertain quality, and whether certification exams can mitigate the impact of having graduated from such a college. Furthermore, the ability to signal quality to employers may be particularly relevant for job seekers who are potentially subject to discrimination (Bertrand and Duflo, 2017).

This correspondence experiment examines the effect of two signals of tertiary-level human capital accumulation, college quality, and certification scores, on the initial stages of hiring for graduating engineering students in India. These signals are likely to be particularly salient for this type of applicant since they do not have a meaningful work history. We also examine the gender of applicants to better understand the potential discrimination in a STEM field, as well as how signals of tertiary education quality may mitigate potential discrimination. Our main outcome measure is callbacks from employers in response to the resumes in the experiment. We sent 16,944 resumes to 1412 relevant job openings in eight major cities across India. Thus, we were able to include a majority of the relevant available information technology (IT) jobs in those cities that were listed on major online job search websites during the study period. We focused on openings at small and medium local companies, given their relevance to employment in developing countries (World Bank, 2012).

Our two treatments are college quality and scores from a large-scale certification exam. According

[1] The unemployment rate was 14.9% for those with college degrees and 6.3% for those with higher secondary education in 2021-22 (Government of India, 2023).

to a long-established government ranking system in India, as well as popular perception, the vast majority of the approximately 6,000 engineering colleges in India are relatively low-quality tier-3 colleges, with a few hundred at the top of the distribution considered to be tier-1 or tier-2 (All India Council for Technical Education, 2020; British Council, 2014). Thus, college ranking systems have been the significant focus of policy, and the government has invested in developing three separate ranking systems (Gohain, 2019; KPMG, 2023). We focus on the most recent and transparent government ranking system to generate the college quality treatments we use in this study, which compare tier-2 to tier-3 colleges.[2] There are large differences in quality between these two types of colleges; tier-2 colleges are ranked between 200 and 300, while tier-3 colleges are ranked outside of the top 1000. We confirm that there are also substantial differences across the tier-2 and tier-3 colleges according to two longer established government ranking systems.[3]

Given this large number of tertiary institutions of uncertain quality, it is not surprising that graduating students often take certification exams. For the certification scores on resumes, we use one of the most widely recognized employability assessments in India, with millions of tests conducted each year (Aggarwal, 2022). Many large companies require this test, and it is common for colleges to have a relationship with the testing company. In this experiment, we compare the average impact of having a certification score relative to resumes without certification. We further examine whether the information about the quality of the applicant conveyed by the certification score matters by comparing callback rates for scores in the lowest quartile to those in the highest quartile. Thus, it is plausible that even the small- and medium-sized firms that are the focus of our sample would respond to such a large discrepancy in scores from a widely-used exam.

The main findings of our experiment are that we rule out modest effects of the two signals of tertiary human capital accumulation, college quality and certification, on callbacks. Specifically, as discussed below, we can rule out effects that are much smaller than those in related literature from other settings, and thus these null results meaningfully advance our knowledge of the role of such signals in developing countries. First, with regards to college quality, the coefficient on having graduated from a tier-3 as opposed to tier-2 college is just -0.22 percentage points (pp), and not significantly different from zero. Our precise estimates allow us to rule out effects below -0.53pp with 95% confidence. In addition, there is no impact on having a certification score on a resume. Specifically, the coefficient on the certification score treatment is -0.11pp, and we can rule out effects below -0.44pp and above 0.22pp with 95% confidence. One explanation for this result is that high certification scores have a positive effect on callbacks, while low certification scores have a negative effect. We find no effects of having a certification score in the lowest or

[2] These rankings are from the National Institutional Ranking Framework (NIRF).

[3] National Board of Accreditation (NBA) and the National Assessment and Accreditation Council (NAAC).

highest quartiles, however. As with the rest of our results, these are close to zero and precisely estimated. We also merge a subsample of firms in the experiment with publicly available data on firm characteristics, and we do not find evidence of heterogeneity along those measures.

Next, we consider potential interactions between these two signals, especially since the prevalence of certification tests in India may be designed to mitigate the impact of colleges of low or uncertain quality. Given that we do not find a negative impact of attending a tier-3 college, it may not be surprising that we do not find a mitigating effect of high certification scores. We do find a modest negative interaction effect of a low certification score and low college quality, perhaps suggesting that negative signals are particularly salient.

Our next set of findings focuses on the role of gender on callback rates for engineering jobs. In our study, however, we can rule out discrimination against female candidates at least in this early stage in the hiring process. In fact, we find a modest positive and significant effect of being female on callbacks, although there is evidence that the effect varies meaningfully by region. We do not find evidence, however, that there are any interactions between gender and college quality or certification.

To better understand how employer beliefs may affect their response to the signals we tested in the correspondence experiment, we conducted a small complementary survey of hiring managers. The suggestive findings from this survey are generally aligned with our main results, since only a minority of firms in the survey indicate that they use the two main signals of quality during resume screening. Another minority of employers indicate that they only use these signals at a later stage of hiring, which is difficult to explain given that such signals are most likely to be salient at the stage in which there is little additional information about applicants available to employers. We also find that many employers who do not use the measures of college quality and certification in the study indicate they are aware of them. So, barriers to using such signals effectively may be more complex.

Our study is the first experiment to examine the impact of college quality on the prospects of job seekers in a developing country, a question that has become increasingly relevant with the dramatic expansion of tertiary education in such countries (Arnhold and Bassett, 2021). Furthermore, these results have broad policy relevance, as they examine callbacks for resumes of engineering graduates in India, the world's most populous country, which had 806,000 engineering graduates in 2020 alone.[4] This paper contributes to a growing literature on college selectivity and the returns to higher education that has been largely focused on developed countries (Dale and Krueger, 2002; Arcidiacono, Bayer and Hizmo, 2010; Darolia et al., 2015) with the exception of MacLeod

[4]This is 12% of total number of graduates in 2020 (Government of India, 2019-20).

et al. (2017) that uses data from Colombia. Although those studies generally find positive results, our experiment can rule out small impacts of college quality. This paper is most closely related to Deming et al. (2016), a correspondence experiment, that finds a 2pp effect on callbacks in the U.S. of reporting a non-selective public institution compared to a for-profit college. In contrast, we find a coefficient of 0.22pp of graduating from a much higher quality college and we can rule out effects larger than 0.53pp. One explanation for these differences is that employers in India may have less experience interpreting college quality given the recent expansion of tertiary education there.

Our study is also the first to examine the impact of a certification program that is already at scale. Many skills are acquired through non-formal education and thus are difficult to signal to employers. Therefore, national qualification frameworks (NQFs) that certify skills are a policy priority across many developed and developing countries.[5] Certification programs have a role in developing countries, in particular, since education is often of uncertain quality. As a result, there is a nascent thread of literature studying job frictions and signaling in urban labor markets in developing countries (Abebe et al., 2021; Bassi and Nansamba, 2022; Carranza et al., 2022). These studies, however, rely on study-specific certification tests, with an emphasis on soft skills. Of these, Carranza et al. (2022) is most closely related to our study as it also implements a correspondence experiment in a developing country, South Africa. They find that certification increases callbacks by 1.6pp, while our study can rule out analogous effects larger than 0.22pp.

Our paper also contributes to a large correspondence experiment literature on discrimination in the labor market, by examining the role of gender in a STEM field in a developing country context. Gender discrimination in hiring in India is a particularly relevant question since female labor force participation remains low and gender preferences are still indicated in job postings (Jayachandran, 2015; Chowdhury et al., 2018). Furthermore, there is a strong gender disparity in STEM fields observed around the world.[6] Previous correspondence experiments from developed countries have found evidence of discrimination against women in STEM jobs (Moss-Racusin et al., 2012; Baert, De Pauw and Deschacht, 2016; Riach and Rich, 2006).[7]

More broadly, this paper contributes to the literature examining the hiring practices of firms (Oyer and Schaefer, 2016). In particular, recent work in developing countries has focused on information frictions in the labor market, and ways in which firms compensate for those frictions (Heath,

[5]Over 100 countries that have developed NQFs, which often include Recognition of Prior Learning (RPL) programs that certify human capital acquired outside of formal education (Allais, 2010).

[6]Only 29.3% of STEM researchers were women globally. This statistic is as low as 18.5% for countries in South and West Asia (UNESCO, 2019).

[7]For more on discrimination in the labor market in general, see Bertrand and Duflo (2017) for a review, and more recently, Kline, Rose and Walters (2022).

2018; Banerjee and Chiplunkar, 2024). Signaling human capital, as in this study, is a potentially important way to limit labor market frictions, but our results indicate that more work is needed for this approach to be effective.

2 Study Design

2.1 Context

India has experienced a nearly five-fold increase in tertiary enrollment in the last two decades.[8] In 2000, India had only 669 engineering institutes, which grew to around 6,166 institutes in 2019 (Agarwal 2006; All India Council for Technical Education 2020). The large growth in engineering colleges has raised concerns about the employability of engineering college graduates (NASSCOM, 2017; University Grants Commission, 2021). This is more broadly aligned with the unemployment trends for college graduates, who had an unemployment rate was 17.6% compared to 6% overall in the same year (MoSPI, 2019-20).

The Indian education system has traditionally followed a tiered system with a select group of high-performing colleges regarded as tier-1 from among a few hundred well-regarded colleges. The remainder of that group is generally considered tier-2, and then there are a much larger number of remaining colleges regarded as low tiered (British Council, 2014; Cheney, Ruzzi and Muralidharan, 2005).[9] There are three government-backed rankings that assess college quality: National Institutional Ranking Framework (NIRF) rankings, National Assessment and Accreditation Council (NAAC) accreditation, and National Board of Accreditation (NBA) accreditation (see Section SA1 for more details). The NIRF started in 2016 and thus is the most recently developed system. It also relies on the most well-defined assessment criteria, which include teaching, learning and resources, research and professional practices, graduation outcomes, outreach and inclusivity, and perception. The Ministry of Human Resource and Development (MHRD), Government of India provides a list of the top 1000 Indian colleges, out of which NIRF provides specific rankings to the top 300.

Uncertainty about the employability of college graduates creates a high screening costs to employers(Blom and Saeki 2011). Consequently, private companies in India have introduced college exit exams, such as AMCAT, E-Litmus, and Cocubes, which provide external validation of graduates' skills, potentially enhancing their employability in a competitive market. We chose one of the most

[8] The total enrollment in higher education in India increased from 8.39 million in 2000-01 to 38.27 million in 2019-20 (Government of India, 2019-20).

[9] The Indian Institutes of Technology (IITs) and National Institutes of Technology (NITs) are generally considered the most prestigious engineering colleges in the country and are classified as tier-1 institutes. Some private institutes are also regarded as tier-1 colleges. These colleges have close to 100% placement rate.

widely used employability assessments in India, with more than three million tests conducted each year across multiple countries including India. Half a million tests are purchased directly by job-seekers to gain credentials, with the remainder purchased by employers (Aggarwal, 2022). This certification exam is a computer adaptive test that focuses on topics that are likely to be relevant to employers. In particular, there are three main subject areas: English comprehension, logical reasoning and, quantitative ability as well as optional tests on job-specific domain skills.

The lack of women in science, technology, engineering, and mathematics (STEM) fields can lead to lower earnings for women, and potentially, less innovation (Dasgupta and Stout, 2014; Beede et al., 2011). In India, women are significantly underrepresented in STEM fields, but this topic remains under-researched (Choudhury and Singh, 2023). The gender gap is particularly stark in the top engineering entrance exam; only 28.8% of students who register for it are female (National Testing Agency, 2019). Furthermore, women who do manage to enter the labor market face additional barriers. Female graduates in India are twice as likely to be unemployed relative to male graduates (MoSPI, 2019-20).

2.2 Sample Selection

To ensure a broad sample, we applied to 1412 job openings across India. In particular, we aimed to apply to close to the universe of relevant IT jobs at small and medium local companies listed on major online job search websites during the study period in four global cities (Delhi, Bengaluru, Chennai, Mumbai), and we reached over 400 employers in each of these cities. We also applied to jobs in four emerging IT hubs (Mysore, Kolkata, Pune and Hyderabad) (NASSCOM, 2020). We prioritized selecting job postings from small and medium-sized companies.[10] The full correspondence experiment extended from July 2021 to January 2022. Our concerns about the potential adverse impacts of the COVID-19 pandemic on the labor market were mitigated by a reported surge in IT jobs during this period as companies went remote (NASSCOM, 2020; Economic Times, 2021).

To compile our database of job openings, we used nationally recognized online job search websites (indeed.com, naukri.com, and monster.com). Such local and global online job portals are widely used in many countries, including India (Nomura et al., 2017; Chowdhury et al., 2018). To avoid any potential suspicion from employers, we only applied to one job per company across multiple job search websites and branches.

Our focus was on job postings requiring a B. Tech degree in Computer Science or Information

[10]Furthermore, large companies are likely to prefer IT candidates from tier-1 institutions. See Section SA2 for further details on the selection of jobs into the study sample.

Technology as a minimum qualification, since a large number of job openings require this type of degree (MoSPI, 2019-20).To ensure the salience of undergraduate institutions and certification scores, we applied as graduating engineering students with no prior work experience aside from internships. To ensure the relevance of these resumes, we applied to jobs that did not require any experience. Our sample of job openings included both full-time positions and full-time paid internships, such as software engineer, junior software engineer, developer, software intern, and IT support assistant.

2.3 Experiment Design

Our study design relies on cross-randomizing two signals of human capital accumulation at the tertiary level, namely, college quality and certification scores. To understand college quality, resumes reported a degree from either a tier-2 or tier-3 institution. We exclude colleges ranked in the top-200 in order to comprehensively exclude elite global colleges with 100% placement rates. To examine the role of certification, resumes were assigned to report high, low, and no certification scores. Additionally, we cross-randomized gender. Thus, each employer in our sample received a total of 12 resumes for a single job vacancy, with one resume from each possible combination of our treatments (Table 1).[11] In total, we sent out 16,944 realistic resumes.[12]

To ensure consistency in the classification of colleges into tier-2 and tier-3 categories, we relied on three government-approved rankings. We focus on NIRF as it is the most recently created system and it has well-defined assessment criteria. We categorized tier-2 colleges as those with 2020 NIRF rankings between 201 and 300 and tier-3 as those not included in the MHRD list of the top 1,000 colleges. We also ensured that there was a large gap in college quality between our tier-2 and tier-3 colleges in the NBA and NAAC rankings as well.[13]

There are two certification score treatments that are compared to a control, in which resumes did not have a certification score. Specifically, for the high certification score treatment we randomly selected scores from the highest quartile (75th to 95th percentile) for each of the three main subject areas as well as the domain-specific test (computer programming) that was relevant to the job postings in our study. Similarly, for the low certification score treatment, we randomly selected scores from the lowest quartile (5th to 25th percentile) for each of the tests to represent low certification scores. This wide division between high and low scores overall allowed us to clearly distinguish between these two treatments, while randomly selecting topic-specific scores within a

[11] We also randomly assigned resume bodies to treatments. See Section SA2 for further details.

[12] The pre-analysis plan indicates that we were going to send 18,000 resumes. Due to an error, pilot resumes were included in measuring the total resumes sent during the project. In addition, 12 wrong resumes were sent to one job posting. Thus, the number of resumes in the study were 16,944.

[13] See section SA1 for details on rankings.

20-percentage point spread, ensured that the scores appeared realistic.

More broadly, we ensured that all aspects of the resumes were plausible and relevant to the information that employers typically use while evaluating job applicants.[14] In particular, we selected colleges that were located in the same city as the job in order to increase the likelihood that local employers would understand the intended signal. We obtained other relevant information, such as internship experience, school and college names and grades, and applicant names, from real resumes obtained from large job portals.

3 Data and Estimation

3.1 Data

Our key outcome measure is the callback rate, which we define as a response from an employer to an applicant's resume in the form of a phone call or email. Thus, specifically, our main outcome is an indicator of whether a resume received a callback from the employer in any form. This includes callbacks for coding round, aptitude test, interview round, or to seek additional information about the applicant. In addition, to check for robustness, we created one other more restrictive version of the callback variable, that focuses only on callback for interview round (see Section SA3). It is important to note, however, that these measures do not capture actual job offers or later-stage outcomes in the job screening process.

We collected additional data to examine firm heterogeneity (see Section 4.3) and on employer beliefs about hiring (see Section 5).

3.2 Estimation

Since this is a randomized experiment, we estimate the intent-to-treat (ITT) effects for the full set of treatments using the following estimating equation:[15]

$$Y_{ib} = \alpha + \beta_j\, Treat_{ij} + \beta_k\, Treat_{ik} + \beta_{jk}\, Treat_{ij} * Treat_{ik} + \lambda_b + \epsilon_i \tag{1}$$

where Y_{ib} is employer's response to the resume for candidate i sent to the vacancy at firm b. $Treat_{ij}$ is an indicator variables equal to 1 if resume i comprises of the treatment component j, and β_j denotes the effect of a given treatment j on the callback rate. Similarly, $Treat_{ik}$ is an indicator variable equal to 1 if resume i includes treatment k, and β_k denotes the effect of a given treatment

[14] For additional details on the resumes, see Section SA1.

[15] Note that this described as the "long model" in Muralidharan, Romero and Wüthrich (2023), where they show that it is unbiased.

k on the callback rate. The interaction of the treatments j and k is denoted by β_{jk}. Since we randomized all 12 possible combinations of the condition within each firm, we include a firm fixed effect λ_b; ϵ_i denotes an unobserved error term. Before estimating the interacted treatments, however, we also examine the effect of each of the individual treatment types separately relative to just the relevant control.[16] For ease of interpretation, when we examine heterogeneous treatment effects, we interact each firm characteristic with each treatment separately.

4 Results

4.1 Main results on human capital signals

We first examine the impact of college quality on callback rates, that is, the impact of graduating from a tier-3 college relative to tier-2 college, and find a precisely estimated zero effect. The coefficient on graduating from a tier-3 college relative to a tier-2 college is −0.22pp, with a standard error of 0.0016 or 0.16pp (Table 2, column 3). Thus, we can rule out effects outside of the 95% confidence interval, which ranges from -0.5336 to 0.09pp. These effects can be compared to the mean callback rate for resumes that do not have a certification score, which is 5.58%. This rate is similar to those in other correspondence studies in developing countries, and validates our resumes as plausible in this setting.[17]

Next, we measure the impact on callback rates of including a certification score on a resume. Initially, we estimate the average effect of the certification score treatment and find that there is no significant impact on the callback rate (Table 2, column 1). Specifically, the coefficient estimate is -0.11pp, with a standard error of 0.0017 or 0.17pp. Thus, these results are precisely estimated, and the 95% confidence interval allows us to rule out effects below -0.44pp or above 0.22pp.

One reason why revealing a certification score on a resume has no impact on callbacks could be that the high certification score treatment has a positive effect, while the low certification score treatment has a negative effect. When we examine the two treatments separately, however, we find that they both have no effect. In fact, the coefficient estimate for the high (low) certification score treatment is -0.14pp (-0.07pp), both with a narrow standard error of approximately 0.0021 or 0.21pp.

One of our main initial hypotheses was that a (high) certification score may mitigate the impact

[16]Since we do this for each treatment separately, and the cross-cutting treatments are balanced for each possible combination of treatments, it should not be subject to the bias of the short model discussed by Muralidharan, Romero and Wüthrich (2023).

[17]Our average callback rate is similar to correspondence studies from India (5.23% for software jobs) (Banerjee et al., 2009), and other developing countries (Maurer-Fazio, 2012; Zhou, Zhang and Song, 2013).

of having graduated from a college of uncertain quality. Thus, apart from testing the effects of the two signaling treatments separately, we also examine whether they interact. First, we examine the interaction effect of the tier-3 college and certification treatments, and we find a small and marginally significant negative effect of −0.60pp on the callback rate (Table 2, column 4). Next, we consider the interaction effects between college quality and type of certification score (high versus low). We find that the interaction on the tier-3 college treatment and the high certification score treatment is small, negative and not significant. Thus, we do not find evidence that a high certification score can mitigate the effect of going to a tertiary institution of low or uncertain quality. We do find, however, that the combined impact of a low quality college and low certification score is modestly negative (−0.85pp) and significant at the 5% level. This suggests that there may be some value in these signals, but it is modest, and it may be easier to send negative signals as opposed to positive ones.

Our precisely estimated null results from this large-scale experiment allow us to rule out effects of the magnitude that have been found in the limited previous literature on the signaling value of certification and college quality. A correspondence experiment of a study-specific certification program in South Africa, Carranza et al. (2022), finds effects of 1.6pp on callbacks of any type. In contrast, we examine an at-scale certification program and can rule out effects of 0.22pp on a similar outcome. A unique correspondence experiment on college quality, Deming et al. (2016), compares lower quality for-profit colleges to non-selective public institutions in the U.S. They find that having a non-selective public institution on a resume increases callbacks by 2.0pp. In contrast, we can rule out effects larger than 0.53pp for the effect of a tier-2 college relative to a tier-3 college, despite the large gap in quality between the two categories. This difference in effect sizes across the two studies may be unsurprising given the study reported here is unique in examining college quality in a developing country, where employers may have less experience considering the signal of college quality.

4.2 Main results on gender

Since gender is a key social and economic determinant in the Indian setting, one of our treatments examines the impact of a female name on a resume on callback rates. Our findings indicate that resumes with female names have a modestly higher callback rate of 0.36pp compared to male names; a difference that is significant at the 5% level (Table 3, column 1). This finding provides evidence for a modest preference for female job candidates in a STEM field. This result is surprising given that prior correspondence experiments focused on STEM jobs in developed country settings have found evidence of discrimination against women (Moss-Racusin et al., 2012; Baert, De Pauw and Deschacht, 2016; Riach and Rich, 2006), and that gender inequality in India is higher across a number of dimensions (Batra and Reio Jr, 2016; Jayachandran, 2015). Our results are promising for women entering entry-level STEM jobs in India, but they are only from the first round of hiring and cannot rule out gender discrimination in later stages.

Next, we consider the interaction effect of gender and the two signals of human capital accumulation that are the focus of this paper (Table 3, columns 2-4). In the presence of gender discrimination, our hypothesis was that certification or other signals of quality may mitigate those effects. We do not find any evidence, however, that these interactions are different from zero, and we can again rule out large effects.

Finally, we note that all of our main findings, in this section as well as the one above, are consistent for different versions of the outcome variable (see Section SA5.2).

4.3 Heterogeneous treatment effects on firm characteristics

Finally, we examine whether there is any evidence of treatment heterogeneity analysis on important firm characteristics. First, we examine whether there is heterogeneity in the impacts of our treatments across the North and South regions of India. This heterogeneity is particularly important for our results on gender because gender norms are more rigid in North India compared to South India (Rahman and Rao, 2004). We do find significant differences along this source of heterogeneity as expected (Table SA8).

We also consider additional firm characteristics using data we collected from Glassdoor, a website that provides data on a few basic firm characteristics for many firms. Although this data is not available for all of the firms in our correspondence experiment sample, it is important to note that since we randomize all of the treatment combinations across the 12 resumes that we sent to each firm, having a partial sample here does not impact the internal validity of the study, although it may affect power.[18] Given that, we do not find differential treatment effects for certification or college quality on callback rates in the data we could collect from Glassdoor, namely, by whether the firm is private, in the information technology (IT) industry, headquartered internationally, or has less than 51 employees (Tables SA6 and SA7).

[18] We were able to collect it for between 38% to 59% of the total sample, depending on the variable, which is around 6,840 to 10,620 resumes.

5 Mechanisms

5.1 Employer survey

To gain a deeper understanding of managers' perceptions of the importance of college quality and certification tests in hiring recent graduates, we surveyed over 250 hiring managers. We targeted both firms that were included in the experiment as well as those at other firms that hire recent engineering graduates. Social media was the most effective way to reach firms compared to phone calls or even in-person visits, thus firms in the non-study sample have a greater online presence compared to firms in the study sample. Thus, it is reassuring that findings in the larger, non-study sample are qualitatively similar to the study sample. Still, since the non-study firms likely have greater managerial capacity, it is not surprising that they report being more likely to engage on average with the signals we study.[19] Of course, these samples are opportunistic rather than representative, and thus these findings must be caveated accordingly. This survey allows us to gain insight into decision-making for some firms at least, however.

5.2 College quality

First, we consider hiring managers' perceptions of college quality in order to understand why much higher quality colleges have no impact on callbacks in our experiment (Table 4). That finding is not surprising given that only 17% (30%) of study (non-study) firms both use college quality during resume screening and indicate that they use the major rankings that we rely on to determine college quality.

The remaining firms can be grouped into three categories. First, a notable minority use alternative methods to assess college quality, such as personal knowledge of local colleges (26% of study and 27% of non-study firms). If these methods align with the study rankings, we would still expect college quality to impact hiring outcomes. Second, a minority consider college quality only later in the hiring process (23% of study and 28% of non-study firms), which is unexpected given that college quality is often most relevant during initial resume screening, before employers have any additional information about applicants. Finally, a minority (33% of study and 14% of non-study firms) do not consider college quality at all, despite the limited information available on recent graduates without work experience. Few of these firms believe all colleges are equivalent, though the rationale for disregarding college quality remains unclear. Notably, referrals do not seem to play a role in hiring by firms in this sample.

Although some firms are not aware of the study ranking systems, lack of awareness is not the

[19] We made significant efforts to reach the firms in the study sample, but there were challenges in reaching that sample of small and medium enterprises. For more on the employer survey recruitment and variables see Section SA4.

primary reason that employers do not use them. Of the firms who use an alternative approach to determine college quality, at least half indicate that they know of the study rankings (50% for study and 63% for non-study firms). In addition, even a majority of the firms that do not consider college quality at all indicate that they know of the study rankings. It is possible, however, that employers are aware of the rankings but are not sufficiently familiar with them to use them effectively.

5.3 Certification

Next, we explore possible explanations for why certification has no effect on callbacks in our sample.[20] These findings largely mirror our findings on college quality in general. Just 19% (33%) of study (non-study) firms indicate that they use study certification during resume screening. Again, we find that a large minority of employers indicate that they only use this signal at a stage in the hiring process after resume screening. In addition, we find many hiring managers indicate that they are aware of it and do not use it (29% for study and 10% for non-study firms). Thus, again, lack of knowledge is not a decisive factor.

This survey does, however, highlight the relative importance of certification in this setting, although perhaps less so for study firms. Within our sample, we actually find that more firms consider the study certification during hiring relative to the study rankings of college quality (71% v. 55% for non-study firms) suggesting the relative importance of certification in this setting. We also explore if the firms in our sample require any certification tests as part of the hiring process. A meaningful percentage of non-study firms in particular do require some type of certification test (55% and 19% of study firms). A minority of those, however, require the study certification, with many relying on their own tests.

5.4 Gender of the applicant

Our findings from the experiment indicate a significantly higher callback rate for women in the first round of hiring. We find some evidence that is consistent with this finding in the survey (Table 5). In particular, a majority of the firms in our sample indicate they have a gender diversity policy (52% of study and 74% of non-study firms). A smaller percentage of firms, however, indicate that they directly consider gender in hiring (15% of study and 47% of non-study firms).

[20] We confirmed that hiring managers trusted the certification exams during COVID-19. Many were not aware they were being held remotely (67% of study and 35% of the non-study firms), but those that were aware believed them to be fair (35% of study and 57% of the non-study firms).

6 Conclusion

We use a correspondence experiment to study the importance of two signals of tertiary education, college quality and certification, to small and medium firms in India during the first stage of hiring college graduates for IT jobs. In precisely estimated results, we find that such employers do not place importance on these two signals. In addition, we tested for gender discrimination in the male-dominated IT industry by randomizing the gender of applicants and found a significantly higher first-round callback for women. This study provides evidence of a small but significant gender-based preference in favor of female candidates in our experimental setting.

Our findings have important implications for understanding the role of tertiary human accumulation in the Indian job market. In order to utilize the full potential of its demographic dividend, improving tertiary education is a policy priority in India. Thus, there has been significant investment in creating and implementing rankings systems. Our study provides evidence, however, that hiring managers at small and medium-sized firms are neither using signals from college rankings nor are they using certification scores effectively even though the latter are frequently required by large companies. Our small complementary survey of hiring managers provides suggestive evidence that awareness of the signals is not the primary barrier to their use in resume screening. One possible explanation that warrants further consideration is that these signals are not useful in identifying quality candidates, although there is some evidence that such signals are more likely to be used by more connected or successful companies. Thus, a question for future research is how to assist smaller and less connected firms in using these signals more effectively.

References

Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. 2021. "Anonymity or distance? Job search and labour market exclusion in a growing African city." *The Review of Economic Studies*, 88(3): 1279–1310.

Agarwal, Pawan. 2006. "Higher education in India: The need for change." *Indian Council for Research on International Economic Relations (ICRIER), Working Paper No. 180.*

Aggarwal, Varun. 2022. "Meritocracy Enabled by Technology, Grounded in Science." *Making Meritocracy: Lessons from China and India, from Antiquity to the Present*, Oxford University Press.

Allais, Stephanie. 2010. *The implementation and impact of National Qualifications Frameworks: Report of a study in 16 countries.* Vol. 3, ILO Geneva.

All India Council for Technical Education, AICTE. 2020. *All India Council for Technical Education (AICTE).* Dataset accessed on 13th June 2020.

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer learning and statistical discrimination." *Quarterly Journal of Economics*, 116(1): 313–350.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond signaling and human capital: Education and the revelation of ability." *American Economic Journal: Applied Economics*, 2(4): 76–104.

Arnhold, Nina, and Roberta Malee Bassett. 2021. *Steering Tertiary Education: Toward Resilient Systems that Deliver for All.* World Bank.

Baert, Stijn, Ann-Sophie De Pauw, and Nick Deschacht. 2016. "Do employer preferences contribute to sticky floors?" *ILR Review*, 69(3): 714–736.

Banerjee, Abhijit, Marianne Bertrand, Saugato Datta, and Sendhil Mullainathan. 2009. "Labor market discrimination in Delhi: Evidence from a field experiment." *Journal of Comparative Economics*, 37(1): 14–27.

Banerjee, Abhijit V, and Gaurav Chiplunkar. 2024. "How important are matching frictions in the labor market? Experimental & non-experimental evidence from a large Indian firm." *Journal of Development Economics*, 103330.

Bassi, Vittorio, and Aisha Nansamba. 2022. "Screening and signalling non-cognitive skills: experimental evidence from Uganda." *Economic Journal*, 132(642): 471–511.

Batra, Renu, and Thomas G Reio Jr. 2016. "Gender inequality issues in India." *Advances in Developing Human Resources*, 18(1): 88–101.

Beede, David N, Tiffany A Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E Doms. 2011. "Women in STEM: A gender gap to innovation." *Economics and Statistics Administration Issue Brief*, , (04-11).

Bertrand, Marianne, and Esther Duflo. 2017. "Field experiments on discrimination." *Handbook of Economic Field Experiments*, 1: 309–393.

Blom, Andreas, and Hiroshi Saeki. 2011. "Employability and skill set of newly graduated engineers in India." *World Bank Policy Research Working Paper 5640.*

British Council, India. 2014. *Understanding India: The future of higher education and opportunities for international cooperation.* British Council. Accessed on 6th August 2024.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin. 2022. "Job Search and Hiring with Limited Information about Workseekers' Skills." *American Economic Review*, 112(11): 3547–83.

Cheney, Gretchen Rhines, Betsy Brown Ruzzi, and Karthik Muralidharan. 2005. *A profile of the Indian education system, prepared for the New Commission on the Skills of the American Workforce.* National Center on Education and the Economy.

Choudhury, Itishree, and Seema Singh. 2023. "Analysing gender differences in academic performance and labour market outcomes of engineering graduates: evidence from India." *International Journal of Manpower*, 44(8): 1622–1640.

Chowdhury, Afra Rahman, Ana Carolina Areias, Saori Imaizumi, Shinsaku Nomura, and Futoshi Yamauchi. 2018. "Reflections of employers' gender preferences in job ads in India: an analysis of online job portal data." *World Bank Policy Research Working Paper 8379*.

Dale, Stacy Berg, and Alan B Krueger. 2002. "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." *Quarterly Journal of Economics*, 117(4): 1491–1527.

Darolia, Rajeev, Cory Koedel, Paco Martorell, Katie Wilson, and Francisco Perez-Arce. 2015. "Do employers prefer workers who attend for-profit colleges? Evidence from a field experiment." *Journal of Policy Analysis and Management*, 34(4): 881–903.

Dasgupta, Nilanjana, and Jane G Stout. 2014. "Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers." *Policy Insights from the Behavioral and Brain Sciences*, 1(1): 21–29.

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

Economic Times, India. 2021. *Surge in opportunities for tech jobs pre and post-pandemic: Report.* Economic Times. Accessed on 16th August 2024.

Gohain, Manash Pratim. 2019. *Centre to pump in Rs 1,000 crore, make accreditation must for all education institutes.* Times of India. Accessed on 25th October 2024.

Government of India, GOI. 2019-20. *All India survey on higher education.* Data accessed on 9th February 2024.

Government of India, GOI. 2023. *Periodic Labour Force Survey (PLFS) Annual Report (July 2021-June 2022.* Dataset accessed on 26th July 2023.

Hanushek, Eric A, and Ludger Woessmann. 2010. "Education and economic growth." *Economics of Education*, 60: 67.

Heath, Rachel. 2018. "Why do firms hire using referrals? Evidence from Bangladeshi garment factories." *Journal of Political Economy*, 126(4): 1691–1746.

Jayachandran, Seema. 2015. "The roots of gender inequality in developing countries." *Annual Review of Economics*, 7(1): 63–88.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. "Systemic discrimination among large US employers." *Quarterly Journal of Economics*, 137(4): 1963–2036.

KPMG, India. 2023. *National Institutional Ranking Framework (NIRF) 2023: Category-specific ranking analysis.* KPMG.

MacLeod, W Bentley, Evan Riehl, Juan E Saavedra, and Miguel Urquiola. 2017. "The big sort: College reputation and labor market outcomes." *American Economic Journal: Applied Economics*, 9(3): 223–261.

Maurer-Fazio, Margaret. 2012. "Ethnic discrimination in China's internet job board labor market." *IZA Journal of Migration*, 1: 1–24.

Mortensen, Dale T, and Christopher A Pissarides. 1999. "New developments in models of search in the labor market." *Handbook of Labor Economics*, 3: 2567–2627.

MoSPI. 2019-20. *Annual Report 2019-2020.* Government of India. Data accessed on 9th February 2024.

Moss-Racusin, Corinne A, John F Dovidio, Victoria L Brescoll, Mark J Graham, and Jo Handelsman. 2012. "Science faculty's subtle gender biases favor male students." *Proceedings of the National Academy of Sciences*, 109(41): 16474–16479.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich. 2023. "Factorial designs, model selection, and (incorrect) inference in randomized experiments." *Review of Economics and Statistics*, 1–44.

NASSCOM. 2017. *Indian IT-BPM Industry Employment Skills: Navigating Through a New World.* Dataset accessed on 5th June, 2020.

NASSCOM. 2020. *Strategic Review 2020: Unlocking Opportunities, Empowering India .* Dataset accessed on 6th June 2022.

National Testing Agency, GOI. 2019. *Press Release on JEE (Main).* Accessed on 16th August 2024.

Nomura, Shinsaku, Saori Imaizumi, Ana Carolina Areias, and Futoshi Yamauchi. 2017. "Toward labor market policy 2.0: the potential for using online job-portal big data to inform labor market policies in India." *World Bank Policy Research Working Paper 7966*.

Oyer, Paul, and Scott Schaefer. 2016. "Firm/employee matching: An industry study of US lawyers." *ILR Review*, 69(2): 378–404.

Rahman, Lupin, and Vijayendra Rao. 2004. "The determinants of gender equity in India: examining Dyson and Moore's thesis with new data." *Population and Development Review*, 30(2): 239–268.

Riach, Peter A, and Judith Rich. 2006. "An experimental investigation of sexual discrimination in hiring in the English labor market." *The BE Journal of Economic Analysis & Policy*, 6(2): 0000102202153806371416.

Ribeiro, Samuel Pinto, Stefano Menghinello, and Koen De Backer. 2010. *The OECD ORBIS database: Responding to the need for firm-level micro-data in the OECD.* OECD.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 87(3): 355–374.

UNESCO. 2019. *Women 'in' Equality-Not anymore!: Gender diversity and inclusivity trends in the IT-BPM sector.* Dataset accessed on 26th June 2024.

University Grants Commission, Government of India. 2021. "Quality Mandate for Higher Education Institutions in India." (Accessed on 16th August 2024).

World Bank, Group. 2012. *World Development Report 2013: Jobs.* The World Bank.

Zhou, Xiangyi, Jie Zhang, and Xuetao Song. 2013. "Gender discrimination in hiring: Evidence from 19,130 resumes in China." *Available at SSRN 2195840*.

Table 1: Randomization design

College quality	**Certification score**	**Gender**
Tier-2 college	Highest quartile certification score Lowest quartile certification score No certification score	Female Male
Tier-3 college	Highest quartile certification score Lowest quartile certification score No certification score	Female Male

Table 2: Main results on certification and college quality

Dependent variable:	Callback rate				
	(1)	(2)	(3)	(4)	(5)
Certification score	-0.0011			0.0019	
	(0.0017)			(0.0025)	
Highest quartile certification score		-0.0014			0.0004
		(0.0020)			(0.0028)
Lowest quartile certification score		-0.0007			0.0035
		(0.0021)			(0.0029)
Tier-3 college			-0.0022	0.0017	0.0017
			(0.0016)	(0.0028)	(0.0028)
Tier-3 college*certification score				-0.0060*	
				(0.0035)	
Tier-3 college*highest quartile certification score					-0.0035
					(0.0040)
Tier-3 college*lowest quartile certification score					-0.0085**
					(0.0041)
Control mean	0.0554	0.0554	0.0558	0.0545	0.0545
N	16,944	16,944	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *No Certification* in columns (1) and (2) , *Tier-2 college* in column (3), and *Tier-2 & No certification* in columns (4) and (5).

Table 3: Main results on gender, certification and college quality

Dependent variable:	Callback rate			
	(1)	(2)	(3)	(4)
Female	0.0036**	0.0036	0.0036	0.0031
	(0.0016)	(0.0028)	(0.0028)	(0.0024)
Certification score		-0.0011		
		(0.0025)		
Female*certification score		0.0000		
		(0.0035)		
Highest quartile certification score			-0.0016	
			(0.0027)	
Lowest quartile certification score			-0.0005	
			(0.0029)	
Female*highest quartile certification score			0.0005	
			(0.0040)	
Female*lowest quartile certification score			-0.0005	
			(0.0041)	
Tier-3 college				-0.0028
				(0.0024)
Tier-3 college*female				0.0011
				(0.0033)
Control mean	0.0528	0.0535	0.0535	0.0543
N	16,944	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *Male* in column (1), *Male & No certification* in columns (2) and (3), and *Tier-2 & Male* in column (4).

Table 4: Mechanisms for college quality

	Non-Study	Study
Use college quality during resume screening	0.57	0.44
	(0.50)	(0.50)
Use study ranking to determine college quality	0.30	0.17
	(0.46)	(0.38)
Use alt. college quality measure and know study rankings	0.17	0.13
	(0.38)	(0.34)
Use alt. college quality measure and do not know study rankings	0.10	0.13
	(0.30)	(0.34)
Use college quality later and know of study rankings	0.28	0.23
	(0.45)	(0.43)
Use study ranking to determine college quality	0.15	0.08
	(0.36)	(0.27)
Do not consider college quality during hiring process	0.14	0.33
	(0.35)	(0.47)
Know of study rankings	0.12	0.19
	(0.32)	(0.40)
Do not know of study rankings	0.02	0.13
	(0.15)	(0.34)
Alternative options		
Use skills listed on the resume	0.08	0.19
	(0.27)	(0.40)
Rely on recommendations	0.02	0.00
	(0.15)	(0.00)
All colleges are of the same quality	0.05	0.04
	(0.22)	(0.19)
N	206	52

Notes: Standard deviations are shown in parentheses. The remaining 1% employers from non-study firms indicate that they consider college quality later in hiring process and do not know of study rankings. Using college quality later in the hiring process includes using the college quality during the interview and technical round (after the first round of callbacks) and in making the final decision. Alternative options are not exclusive and are missing for some hiring managers. Alternative study ranking sources include local knowledge and private sector rankings.

Table 5: Mechanisms for certification

	Non-Study	Study
Use study certification during resume screening	0.33	0.19
	(0.47)	(0.40)
Use study certification later in hiring process	0.38	0.21
	(0.49)	(0.41)
Do not consider study certification during hiring process	0.29	0.60
	(0.45)	(0.50)
Know of study certification	0.10	0.29
	(0.30)	(0.46)
Do not know of study certification	0.18	0.31
	(0.39)	(0.47)
Alternative options		
Tests not in line with work done by the company	0.08	0.12
	(0.27)	(0.32)
Prefer own coding tests and technical exams	0.06	0.29
	(0.23)	(0.46)
Tests too easy or too easy to cheat	0.05	0.02
	(0.23)	(0.14)
Require any certification test as a part of the hiring process	0.55	0.19
	(0.50)	(0.40)
Require study certification test as a part of the hiring process	0.17	0.06
	(0.38)	(0.24)
Require other certification test as a part of the hiring process	0.15	0.06
	(0.35)	(0.24)
Require own test as a part of the hiring process	0.21	0.07
	(0.41)	(0.27)
Gender of the applicant		
Importance of applicant's gender on resume	0.47	0.15
	(0.50)	(0.36)
Gender diversity policy in the firm	0.74	0.52
	(0.44)	(0.50)
N	206	52

Notes: Standard deviations are shown in parentheses. Using certification quality later in the hiring process includes using the certification during the interview and technical round (after the first round of callbacks) and in making the final decision. Alternative options are not exclusive and are missing for some hiring managers.

Does signaling college-level human capital matter? An experimental study in India

Deepshikha Batheja, Sarojini R. Hirshleifer, Opinder Kaur

Supplemental Appendix for Online Publication

SA1 Resume Construction

The correspondence experiment relied on fictitious resumes. The three components of the resumes that were randomized as treatments (college name, certification score, and gender of name) are further discussed in the subsections below. In addition to those components, we included secondary school information, college GPA, tertiary extracurriculars and projects, and internships on the resumes. Those components were inspired by real resumes obtained through connections of the research team as well as a large pool posted on job portals. The specifics were modified and re-sorted to respect the individuals to whom the reference resumes belonged.

In addition to listing secondary school names on post-tertiary resumes, in India, people typically also list scores from two prominent secondary-level national exams. The names of local schools (for senior secondary and secondary education) were extracted from real resumes for each study city. Since the resumes were tested for students in low-ranked colleges, the average scores achieved by such students on secondary-level national exams were expected to lie between 65 to 69%. We then added a random number up to a single decimal point between these scores to make them appear more genuine. Similarly, for college GPA, we added a random number up to a single decimal point between 65% and 69%.

We populate our resumes with tertiary extracurriculars, projects and, internships to lend them legitimacy. We added some extra-curricular activities such as organizing college events, and participating in sports competitions. We also included plausible software certifications such as C++, Python, and Java in our resumes. Finally, we included information on project experience gained through internships, participation in undergraduate student teams, etc.

For other personal information, such as email addresses and phone numbers, we created corresponding 12 email addresses and used real 10-digit voice-over internet phone numbers (which looked similar to 10-digit Indian mobile numbers).

SA1.1 College quality

We use three official ranking sources sanctioned by the University Grants Commission (UGC) to determine college quality and to categorize colleges into tiers. The Indian education system has traditionally used the tiered system to rank institutes (Cheney, Ruzzi and Muralidharan, 2005). UGC is a central statutory body established by the Government of India, responsible for maintaining educational standards and ensuring college quality in higher education in India. The three UGC approved ranking sources are the National Institutional Ranking Framework (NIRF), the National Board of Accreditation (NBA), and the National Assessment and Accreditation Council (NAAC). We primarily rely on NIRF as it is the most recent and has well-defined assessment criteria to construct our college quality tiers. We also use the additional two official and traditional sources of NBA and NAAC to confirm our college quality tiers.

NIRF is approved by the Ministry of Human Resource and Development (MHRD), Government of India and provides rankings to educational institutions based on the overall recommendations of a core committee set up by MHRD. NIRF is the most recent accreditation system with its first ranking list published in 2016. Their ranking criteria include teaching, learning and resources, research and professional practices, graduation outcomes, outreach and inclusivity, and perception. MHRD provides a list of the top 1000 Indian colleges, out of which NIRF only provides rankings to the top 300. We define tier-1 colleges as those ranked as top 200 by NIRF, tier-2 colleges as those ranked between 201 and 300 by NIRF and, tier-3 institutions as those not in the MHRD's list of 1000 colleges.

NAAC provides accreditation to educational institutes each year using A-D categories. These rankings are based on seven assessment criteria: curriculum; teaching and learning outcomes; research, innovations and extension; infrastructure and learning resources; student support and progression; governance, leadership and management and; institutional values and best practices. We have used the accreditations from the pre-experiment year of 2021. We define tier-1 colleges as those with grade A, tier-2 colleges as those with grade B and above and tier-3 colleges as those with low or no grade/ranking by NAAC.

NBA has been established by the All India Council of Technical Education (AICTE) and provides tier-based rankings to colleges. While they do not reveal the exact criteria for their rankings, however, the rankings broadly include institutional missions and objectives, organization and governance, infrastructure facilities, quality of teaching and learning, curriculum design and review and support services (library, laboratory, instrumentation, computer facilities, etc.). NBA classifies the undergraduate degree programs for engineering into two categories of tier-1 and tier-2. We define tier-2 colleges in our study as those with either tier-1 or tier-2 of NBA ranking and tier-3 as those without any NBA rankings.

To summarize, our defined tier-2 institutions are those with NIRF rankings between 201 and 300, grade B and above NAAC ranking, and have tier-1 or tier-2 NBA ranking (See Table SA1). Similarly, our tier-3 institutions remain unranked in MHRD and NIRF rankings, have a low NAAC ranking, and have no NBA ranking. We exclude tier-1 colleges from our study as top engineering colleges have almost 100% placement rates, while students from tier-2 and tier-3 colleges struggle for placements. Furthermore, we have kept a considerable ranking difference in tier-2 and tier-3 institutions to maintain the salience of our college quality signal. Employers are likely to differentiate between tier-2 and tier-3 colleges owing to this substantial ranking difference.

SA1.2 Certification exam

The resumes which included certification scores, reported a score for each of the four main subject areas: English comprehension, logical reasoning, quantitative abilities, and computer programming. The certification test is computer adaptive, that is, the test adjusts its difficulty level based on the candidate's performance. The English comprehension exam tested the candidates on their grammar, functional vocabulary and, understanding and comprehension of texts. The quantitative exams assessed candidates on their knowledge of basic numbers, word problems, permutation-combination basic probability and, power and logarithms. The logical reasoning test included topics on deductive, inductive and abductive reasoning. Finally, the computer programming assessment contained sections on the structure and constructs of computer programs, data structures and basic algorithms and object oriented programming concepts such as data encapsulation, data abstraction and polymorphism.

The study compared two certification treatment groups of resumes to a control group. One group had high certification scores, randomly selected from the top 25% scores (75th to 95th percentile) for each afore-mentioned subject area. The other group had low certification scores, randomly selected from the bottom 25% (5th to 25th percentile). The control group had no certification scores.

SA1.3 Names and gender

To create names, we used common first and last names from the large pool of resumes to which we had access on job portals. We focused on first names from which gender easily be identified to implement our gender treatment. This process was conducted for each city/region because there are common first and last names unique to each region in India. This helped us in creating realistic names, which were representative of a region's demographic majority.

SA2 Job selection and application

We searched multiple widely used job portals (indeed.com, naukri.com and monster.com) to check for the most relevant, recent and unique job postings. We created profiles on the portals with the basic information such as name, age and contact that was aligned with each resume.

The correspondence experiment study period begins in July 2021 and ends in January 2022. We first targeted the four main metropolitan IT hubs of Bengaluru, Chennai, Delhi and Mumbai. The number of jobs applied to each week was determined by the maximum availability of the research staff. If they had extra time, they applied for jobs posted by small and medium firms in additional cities of Kolkata, Hyderabad, Mysore and Pune.

We randomly assigned resume bodies to each of the 12 treatment combinations to ensure balance across treatments, controlling for any potential effect of resume bodies on callback rates. Although there were 12 distinct treatment combinations, given the uniqueness of colleges and names specific to each city, we effectively implemented 12 treatment combinations per city. Additionally, there were 12 unique resume bodies. Resume bodies were randomized to treatment combinations for each job, but we also ensured that all resume body-treatment combinations were balanced such that they appeared an equal number of times within each city.

This randomization was conducted in batches. We searched for jobs posted in the last three days, requiring a B.Tech in Computer Science/ Information Technology with no prior work experience. Furthermore, we applied for jobs located within 30 kilometers of the city range to ensure that the employers are aware of the local city colleges.

We identified a total of 2041 job postings for fresh engineering candidates in the study period, but could apply eventually for 1412 jobs. This is because some job postings would expire in the lag between finding a job, randomization and job application. Given our intent was to apply primarily to small and medium firms, we excluded large firms from our study based on their online presence. In cases when we were unable to find a sufficient number of relevant job postings from small and medium firms in any particular week, we applied for jobs from large companies. Using the OECD classification of firm size, we applied mostly for jobs in small and medium firms with less than 200 employees (Ribeiro, Menghinello and De Backer, 2010). For those companies for which Glassdoor data is available, 82% have less than 200 employees. Overall we have Glassdoor data for 62.2% percent of our sample. So, this is likely to be a lower bound. Assuming that companies missing from Glassdoor are almost certainly smaller with less than 200 employees, then 89% of our sample constitutes of small or medium sized firms.

SA3 Callback rates

We collected information about the mode and nature of callbacks from employers for each resume sent per vacancy. Each resume provided contact information such as the phone number and email address of the job applicant to the employer. We created a total of 96 unique email addresses and 12 unique phone numbers for each of the 12 resumes sent across eight cities. We had one research analyst and three dedicated interns monitoring and handling callbacks from employers. When we received a callback, we recorded the mode of the callback (email or phone) and the reason for callback. Those reasons included: invitation for interview, request for additional information, invitation for the coding round, invitation for a certification test, invitation for any other aptitude test, rejection purposes and others.

Our main outcome is the least restrictive and includes callbacks of any kind except for rejection purposes. In addition, we created two other types of callback outcomes: one for interviews only, and one for the next round of the screening process, which includes invitations for interviews, coding, certification and aptitude tests.

SA4 Employer Survey

We conducted a post-experiment survey of 258 hiring managers, of which 52 were part of the correspondence study. The estimated completion time of the employer survey was less than twenty minutes. To incentivize participation, respondents were offered compensation in the form of a mobile recharge valued at INR 1,000.

While conducting the employer survey, multiple techniques were utilized to engage potential participants from the experiment sample. The research team used the contact numbers acquired from company websites and callbacks, but it was very challenging to reach hiring managers. There are a number of reasons for this, including that firms were not answering their phones since they were working remotely due to COVID-19 (this was verified through in-person visits in some cities). In addition, receptionists were rarely willing to connect us or pass on messages to the relevant personnel. Finally, there seemed to be significant churn in these small businesses, and we had trouble finding many of them when we returned to attempt additional surveys up to one year later.[21]

To capture a broader range of responses from firms that were not in the experiment, we used LinkedIn as a primary tool to collect survey responses. The research team focused on directly con-

[21] Employer surveys were conducted between May 2022 and January 2023. There was difficulty finding businesses (in-person or by phone) initially due to remote work as a result of the COVID-19 pandemic. It became easier over time as businesses returned to their offices.

necting with HR managers, senior and junior engineers, developers, IT support engineers, and data scientists involved in the recruitment of recent software engineering graduates.[22] This approach was relatively successful in reaching non-study firms, but it was not possible to reach study firms this way since their employees generally did not use LinkedIn. The team sent connection requests along with the survey form, and two follow-up attempts were made within two days. We also made additional follow-ups in the next couple of weeks. If someone refused to participate in the survey, no further survey request was made.

Members of the research team also shared the survey form with affinity groups on various platforms such as relevant Google, Telegram, and LinkedIn. Some of these platforms included the India HR Group, Software Tester and QA of India, India Job Networks, India HR Network, and Recruiters, Jobs & Careers. Furthermore, the study interns contacted the placement cells of their respective colleges to obtain relevant industry contacts.

SA5 Robustness Analysis

SA5.1 Alternative specifications using the main outcome

To explore the interaction effects between the various treatments, we report the results of the "long" model in the main text.[23] We also report the results for the potentially more highly powered, though also potentially biased, "short" model or the model which includes dummies for the two treatments in the regression in Tables SA2 and SA3. In addition, we report the results of specifications with separate dummies for each combination of treatments. The findings from these alternative specifications are similar to the main results.

SA5.2 Robustness to different callback outcomes

We examine whether our results are robust to an alternative version of our main outcome. Our alternative measure is the interview round, which is an indicator variable for whether the callback was for an interview (Tables SA4 and SA5). Unsurprisingly, coefficients are smaller when using these more restrictive outcome measures, but the results are consistent with those using the main version of the outcome variable. For example, the results on gender remain statistically significant.

[22]Prior to that, we tried to recruit participants through ads on LinkedIn and Facebook, but this was not successful.
[23]We report the main results in Table 2 with dummies for both treatments and their interaction in the regression. Muralidharan, Romero and Wüthrich (2023) refer to this as the long model.

Table SA1: College quality criteria

	Requirements		
	National Assessment & Accreditation Council (NAAC)	National Institutional Ranking Framework (NIRF)	National Board of Accreditation (NBA)
Tier-2 college	The college should have a NAAC grade of B and above.	The college should be ranked between 201-300 in the NIRF ranking.	The college should be in tier-1 or tier-2 of NBA Accreditation.
Tier-3 college	The college should not have a NAAC grade or have C or D grade. The University with which the college is associated, should also not have a NAAC grade or have C or D grade.	The college should not be in the MHRD list of 1000 colleges, of which NIRF is a subset (top 300 colleges).	The college should not be in tier-1 or tier-2 of the NBA Accreditation and can or cannot be an AICTE approved institution.

Table SA2: Results on certification and college quality

Dependent variable:	Callback rate		
	(1)	(2)	(3)
Certification score	-0.0011 (0.0017)		
Tier-3 college	-0.0022 (0.0016)	-0.0022 (0.0016)	
Highest quartile certification score		-0.0014 (0.0020)	
Lowest quartile certification score		-0.0007 (0.0021)	
Tier-2 college*highest quartile certification score			0.0004 (0.0028)
Tier-2 college*lowest quartile certification score			0.0035 (0.0029)
Tier-3 college*highest quartile certification score			-0.0014 (0.0028)
Tier-3 college*lowest quartile certification score			-0.0032 (0.0028)
Tier-3 college*no certification score			0.0017 (0.0028)
Control mean	0.0545	0.0545	0.0545
N	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *Tier-2 & No certification* in columns (1), (2) and (3).

Table SA3: Results on gender, certification and college quality

Dependent variable:	Callback rate				
	(1)	(2)	(3)	(4)	(5)
Female	0.0036**	0.0036**	0.0036**		
	(0.0016)	(0.0016)	(0.0016)		
Certification score	-0.0011				
	(0.0017)				
Highest quartile certification score		-0.0014			
		(0.0020)			
Lowest quartile certification score		-0.0007			
		(0.0021)			
Tier-3 college			-0.0022		
			(0.0016)		
Female*highest quartile certification score				0.0025	
				(0.0028)	
Female*lowest quartile certification score				0.0027	
				(0.0027)	
Male*highest quartile certification score				-0.0016	
				(0.0027)	
Male*lowest quartile certification score				-0.0005	
				(0.0029)	
Female*no certification				0.0036	
				(0.0028)	
Tier-2 college*female					0.0031
					(0.0024)
Tier-3 college*female					0.0014
					(0.0023)
Tier-3 college*male					-0.0028
					(0.0024)
Control mean	0.0535	0.0535	0.0543	0.0535	0.0543
N	16,944	16,944	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *Male & No certification* in columns (1), (2) and (4), *Male & Tier-2* in columns (3) and (5).

Table SA4: Robustness check for results on certification and college quality

Dependent variable:	Interview round callback rate				
	(1)	(2)	(3)	(4)	(5)
Certification score	-0.0005			0.0002	
	(0.0010)			(0.0015)	
Highest quartile certification score		-0.0008			0.0000
		(0.0012)			(0.0017)
Lowest quartile certification score		-0.0002			0.0004
		(0.0012)			(0.0017)
Tier-3 college			-0.0009	0.0000	0.0000
			(0.0009)	(0.0016)	(0.0016)
Tier-3 college*certification score				-0.0014	
				(0.0020)	
Tier-3 college*highest quartile certification score					-0.0017
					(0.0024)
Tier-3 college*lowest quartile certification score					-0.0011
					(0.0024)
Control mean	0.0166	0.0166	0.0167	0.0166	0.0166
N	16,944	16,944	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *No Certification* in columns (1) and (2) , *Tier-2 college* in column (3), and *Tier-2 & No certification* in columns (4) and (5).

Table SA5: Robustness check for results on gender, certification and college quality

Dependent variable:	Interview round callback rate			
	(1)	(2)	(3)	(4)
Female	0.0021**	0.0005	0.0005	0.0028**
	(0.0009)	(0.0016)	(0.0016)	(0.0014)
Certification score		-0.0017		
		(0.0014)		
Female*certification score		0.0025		
		(0.0020)		
Highest quartile certification score			-0.0016	
			(0.0016)	
Lowest quartile certification score			-0.0018	
			(0.0016)	
Female*highest quartile certification score			0.0015	
			(0.0024)	
Female*lowest quartile certification score			0.0034	
			(0.0024)	
Tier-3 college				-0.0002
				(0.0013)
Tier-3 college*female				-0.0014
				(0.0019)
Control mean	0.0152	0.0163	0.0163	0.0153
N	16,944	16,944	16,944	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses. The control group is *Male* in column (1), *Male & No certification* in columns (2) and (3), and *Tier-2 & Male* in column (4).

Table SA6: Heterogeneous treatment effects for certification and college quality

Dependent variable:	Callback rate				
Interaction variables:	Private Company (1)	IT Industry (2)	Foreign Headquarter (3)	Below 51 Company Size (4)	North Region (5)
Panel A: Interaction with certification					
Certification score	-0.0055	0.0014	-0.0007	-0.0022	-0.0023
	(0.0078)	(0.0042)	(0.0031)	(0.0029)	(0.0025)
Interaction variable	-0.0037	-0.0009	-0.0005	-0.0008	-0.0016
	(0.0055)	(0.0033)	(0.0027)	(0.0030)	(0.0024)
Certification score*variable	0.0055	0.0015	0.0007	0.0012	0.0025
	(0.0082)	(0.0049)	(0.0042)	(0.0045)	(0.0035)
Panel B: Interaction with Tier-3 college					
Tier-3 college	-0.0089	-0.0101**	-0.0044	-0.0052*	-0.0047*
	(0.0075)	(0.0039)	(0.0029)	(0.0030)	(0.0024)
Interaction variable	-0.0029	-0.0036	0.0014	-0.0019	-0.0025
	(0.0042)	(0.0028)	(0.0024)	(0.0024)	(0.0018)
Tier-3*variable	0.0059	0.0072	-0.0028	0.0039	0.0049
	(0.0078)	(0.0047)	(0.0041)	(0.0045)	(0.0033)
Control mean	0.0519	0.0497	0.0537	0.0495	0.0547
Interaction variable mean	0.8651	0.6599	0.1889	0.5735	0.5004
N	9,962	6,421	7,369	8,077	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses.

Table SA7: Heterogeneous treatment effects for certification

Dependent variable:	Callback rate				
Interaction variables:	Private Company (1)	IT Industry (2)	Foreign Headquarter (3)	Below 51 Company Size (4)	North Region (5)
Lowest quartile certification score	-0.0067	0.0014	-0.0010	0.0008	-0.0025
	(0.0092)	(0.0049)	(0.0036)	(0.0037)	(0.0029)
Highest quartile certification score	-0.0045	0.0014	-0.0005	-0.0052	-0.0021
	(0.0091)	(0.0047)	(0.0035)	(0.0033)	(0.0029)
Interaction variable	-0.0037	-0.0009	-0.0005	-0.0008	-0.0016
	(0.0055)	(0.0033)	(0.0027)	(0.0032)	(0.0024)
Lowest quartile certification score*variable	0.0070	0.0022	0.0010	-0.0035	0.0035
	(0.0095)	(0.0059)	(0.0052)	(0.0054)	(0.0041)
Highest quartile certification score*variable	0.0041	0.0007	0.0005	0.0058	0.0014
	(0.0094)	(0.0056)	(0.0045)	(0.0053)	(0.0040)
Control mean	0.0519	0.0497	0.0537	0.0495	0.0547
Interaction variable mean	0.8651	0.6599	0.1889	0.5735	0.5004
N	9,962	6,421	7,369	8,077	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses.

Table SA8: Heterogeneous treatment effects for gender

Dependent variable:	Callback rate				
Interaction variables:	Private Company (1)	IT Industry (2)	Foreign Headquarter (3)	Below 51 Company Size (4)	North Region (5)
Female	0.0119	-0.0027	0.0003	-0.0029	0.0066***
	(0.0075)	(0.0039)	(0.0029)	(0.0030)	(0.0024)
Interaction variable	0.0058	-0.0023	0.0023	-0.0025	0.0029
	(0.0043)	(0.0024)	(0.0022)	(0.0023)	(0.0019)
Female*variable	-0.0117	0.0046	-0.0046	0.0051	-0.0059*
	(0.0078)	(0.0047)	(0.0041)	(0.0045)	(0.0033)
Control mean	0.0519	0.0497	0.0537	0.0495	0.0547
Interaction variable mean	0.8651	0.6599	0.1889	0.5735	0.5004
N	9,962	6,421	7,369	8,077	16,944

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Callback rate is defined as a response from an employer to an applicant's resume in the form of a phone call or email. All specifications include vacancy fixed effects. Robust standard errors are shown in parentheses.