Do online certifications improve job market outcomes? Evidence from an IT skills certification platform in India

Nandana Sengupta[1,*], Neaketa Chawla[1,†], Anna Agarwal [‡]; James Evans [§]

Abstract

We estimate the returns to online skills certification for engineering graduates in India using a regression discontinuity design. Individuals in our dataset take a computer programming exam, and are required to score above a threshold to receive a widely used software engineer IT-services certificate provided by an online job skills credentialing platform in India. We find that certified candidates have approximately 0.25 higher probability of finding employment following the exam. Our results indicate that skill certification in this context is a strong one-time signal of quality. Certification cannot replace education, however, and fails to predict longer-term job market outcomes as demonstrated by the lack of estimated causal impact on current employment status or income level. Our findings suggest the promise and limits of certification in an educational market where university training is highly variable.

Keywords: Regression Discontinuity; Certification; Skilling; Signalling; Computerised assessments

Conflict of interest Statement: The authors declare that they have no competing interests that could influence the work reported in this article.

[*] [Corresponding author] School of Public Policy, Indian Institute of Technology Delhi, New Delhi `Nandana.Sengupta@sopp.iitd.ac.in`

[†] Brij Disa Centre of Data Science and Artificial Intelligence, Indian Institute of Management, Ahmedabad `neaketac@iima.ac.in`

[‡] Center for Policy Research, New Delhi `annaagarwal@gmail.com`

[§] University of Chicago and Santa Fe Institute `jevans@uchicago.edu`

Funding Source: The research carried out for this article was generously supported by a grant from the University of Chicago Delhi Center.

1 Introduction

Engineering graduates in India find it difficult to land lucrative jobs. Recruiters at top IT firms often bring up the low employability of the country's engineering graduates[1]. Although many private engineering colleges have opened in the past decade, a majority are plagued with outdated curricula and poor infrastructure that produce low-quality engineers [Blom and Saeki, 2011]. By some estimates, only 50% of the engineering graduates from private colleges could achieve campus placement in 2017, compared to a combined 77% for engineering graduates from top government-funded institutions [2]. In such a context, external certifications based on aptitude tests can be a useful tool in the hands of prospective employers to screen fresh graduates from the myriad of educational institutions below the prestigious top tier [Spence, 1978]. Our study examines whether digital certifications provided by online platforms can improve job market outcomes for engineering graduates from non-premier engineering institutions.

We consider skills certification provided by a popular digital platform in India. Candidates appear for a computerised assessment and are awarded a certificate if they score above a pre-established cut-off. We use a regression discontinuity analysis [Thistlethwaite and Campbell, 1960] to study if the employment outcomes for certified candidates are significantly different from those who do not receive a certificate. Our identification strategy exploits the fact that candidates have to clear specific cutoffs in order to obtain certification. Our results show that certification increases the probability of finding a job by approximately 0.25. Certification cannot replace the value of intrinsic quality or education, however, and we do not find evidence of a sustained impact of certification on other future job market outcomes as measured by the candidate's current employment status and salary level. Results suggest therefore that certifications are a strong proximate signal of quality facilitating a candidate to obtain an initial match with an employer.

Recruiters depend on various sources of information to ascertain the suitability of an applicant for a job. These include a candidate's grades and diplomas, perceived quality of their educational institution, psychometric surveys, and even general certifications, which function as signalling devices to indicate student quality and contextual fit. Such proxy measurements are not enough, however, to enable effective signaling and efficient hiring. College degrees based on traditional assessments in Indian engineering colleges often focus on knowledge recall and pencil-and-paper problem-solving skills, which are of little value to employers [Varghese, 2015]. On the other hand, employability tests seek to directly measure skills and capabilties employers consider most relevant for common tasks [Weber, 2015]. Technological advancements and the proliferation of the internet have also provided opportunities to create innovative and cost-effective tools to aid in job search [Denzer et al., 2018]. Certification based on computerized assessments can help candidates broadcast their skills to prospective employers and increases employers' confidence to hire. LinkedIn, the largest job search and networking website for professionals, provides "skill-badges" based on "skill-quizes" for individuals to advertise their profile. Ed-tech companies like *Unacademy*, *AMCAT* and *Co-Cubes* provide resume-building and interview preparation services to job seekers. In India, where the formal labour market is inundated with engineering graduates from private colleges whose quality of education is unreliable [Tilak and Choudhury, 2021], these tools are often advertised as a means to increase employability. However, there is very limited research that studies the effectiveness of such tools both as search and matching instruments, as also predictors of employment outcomes.

A range of empirical literature, primarily based in developed countries, studies the impact of adult vocational certification. Evidence from studies conducted in the UK suggests that returns to additional certification amongst adults are very limited [Silles, 2007; Blanden et al., 2012,Booth and Bryan, 2005; Wolf et al., 2006 and Wolf 2011]. Findings show that returns to early life education are higher [Silles, 2007], and "life-long learning" among middle-aged adults only affects earnings for women [Blanden et al., 2012]. Studies on the General Education Development or GED test, which is conducted in Canada and the US reach similar conclusions [Heckman et al., 2011 and Jepsen et al., 2017]. The GED program offers a credential based on a battery of seven-hour subject tests considered equivalent to a high school diploma. The authors find that obtaining a GED has little to no impact on labour market outcomes, suggesting that even if the GED measures cognitive skills, it fails to signal non-cognitive qualities such as persistence, reliability and motivation required to initially finish high school. More recent work by [Bratsberg et al., 2020] documents the effect of external certification of skills, acquired through apprenticeship, on job market outcomes, finding that individuals without upper secondary education experience a positive effect of certification on future earning.

Research on digital skills certification remains in its infancy. A small set of studies have attempted to capture the causal impact of e-learning and e-certification on job market outcomes. Kässi and Lehdonvirta [2019] examine the impact of voluntary skills certifications on freelance earnings in the context of online labour market platforms and find that certified candidates win more work contracts leading to an increase in their earnings. Castaño-Muñoz and Rodrigues [2021] investigate the causal impact of enrolment in Massive Open Online Course (MOOCs) on labour market outcomes. The study finds that enrolment in MOOCs improves workers' employment retention but has no effect on their wages. Most recently Kässi and Lehdonvirta [2022] show that obtaining a microcredential in an online freelancing setting increases workers' earnings. Their analysis suggests however that online microcredentialling is only a partial substitute of verified past work given that the effects are smaller for more experienced workers. While a substantial proportion of their sample consists of individuals residing in India, Bangladesh Pakistan and Phillipines, we are not aware of any other such study in the context of India or other developing countries in the global south.

Most studies within the Indian context have focused on vocational training and its impact on labour market outcomes [Adhvaryu et al., 2018 and Kumar et al., 2019] with the population of interest typically involving samples of informal workers. To the best of our knowledge, our paper is among the first to estimate the causal impact of digital skill certification on labour market outcomes for a set of workers within the software industry. In doing so, we add to the nascent literature on digital skill certification globally.

Our work is most closely related to Kässi and Lehdonvirta [2019] and Kässi and Lehdonvirta [2022] who find a positive impact from digital skills certification in the context of an online labour market platform. In line with findings from studies on the GED tests and related certifications in the UK, [Heckman et al., 2011, Silles, 2007; Blanden et al., 2012,Booth and Bryan, 2005; Wolf et al., 2006 and Wolf 2011] we find no long-term impact of certification on labour market outcomes. While we hypothesize that this is largely because once hired, the impact of certification is not expected to add to the benefits of on-the-job experience, we acknowledge an important caveat here. Our estimates of the causal impact of certification on long-term labor market outcomes are not significantly different from zero, however they are positive. This leaves open the possibility that our estimates are indicative of noisy positive impacts of certification which could not be captured due to the limited sample size of the study. Nevertheless, we show that certification dramatically improves the odds of finding a job, making it a credible and effective one-time search and matching tool. Our analysis also demonstrate that an individual's logical reasoning abilities and caste identity are significantly associated with long-term job market success. Our results have important implications for the Indian labour market, particularly around questions of employability and labour supply in the highly competitive and growing IT sector in India.

The remainder of the paper is organized as follows. We describe our data in section 2 along with an explanation of our methodology that lays down our identification and empirical strategy. We present our results in section 4 and section 5 concludes.

2 Data

This study is conducted in collaboration with an Indian ed-tech company that provides online skill assessment and certification across a variety of job types with the main aim of quantifying employability of job seekers through automated assessments for which it charges candidates a nominal fee. Candidates are allowed to register for this exam as many times as they want; however, they can attempt it only after a minimum of 45 days from their previous attempt. These assessments run on multiple question banks and typically have auto-proctoring.[3] Corporate employers or small businesses can register with the platform to access its database of candidates and to find suitable talent for job requirements.

Candidates who register for assessments are typically from medium-tier Indian private engineering colleges with limited opportunities for campus placement. For our study we focus on a software engineering IT-services certificate and its impact on job market outcomes. The platform conducts computer adaptive test that assesses the computer programming skills of a candidate. The test consists of four sections: (i) Quantitative Aptitude, (ii) Logical Ability , (iii) English, and (iv) Computer Programming. To obtain the IT-services certificate, a candidate needs to score 450 or more in logical ability, 950 or more in logical ability plus quantitative aptitude, 480 or more in English and 400 or more in computer programming.

There are two plausible ways through which a candidate can secure a job after taking the test. Firstly, candidates can independently apply for jobs and utilise the certification to signal the relevant skill set on their resumes and other job search portals. Alternatively, candidates can sign up for selection for an interview by companies registered with the platform.

Obtaining an IT-certification in general does not lead to the platform mechanically matching test takers to companies registered on the platform and a test taker who does not obtain the cutoff score may still be matched with such companies. This is because companies on the platform establish their own distinct cut-off requirements for various combinations of subjects. For example, a firm can create a hiring criteria based on individual scores in English Proficiency, Quantitative Ability and Computer Programming. Moreover, these cut-off scores are typically not the same as those needed to acquire a specific certification. In simple terms, companies could establish varying cut-off thresholds, higher or lower, for different subjects based on the specific job roles they are recruiting for. Candidates who meet these company and role specific cut-off requirements are then shortlisted for interviews.

Our population of interest is the set of all individual test takers who appeared for the exam between 2015-2018. We do not consider candidates still enrolled in college at the time of the test as they have not yet fully entered the labour market. [4] A total of 105,984 exam-takers fit the criteria. The dataset used in the study is a combination of the information collected using an online survey of test takers and data provided by the platform. The platform provided section-wise scores for these individuals, and through the survey we collected self-reported information regarding each individual's employment status and salary along with other demographic information.

The treatment and control groups in our study consist of those individuals who obtain the IT-service certificate and those who do not, respectively. As mentioned earlier, to obtain the IT-services certificate, a candidate needs to clear a range of cutoffs, including a score of 400 or more in computer programming. We select computer programming to create comparison groups because it is the most decisive subject to determine award of the certificate as majority of students failed to clear the computer programming threshold in spite of passing the remaining three. We therefore use the computer programming score as the running variable in the regression discontinuity design, implying that individuals in the treatment group clear all cutoffs whereas individuals in the control group only miss certification from obtaining less than 400 in computer programming. Individuals in our sample scored between $285 - 515$ in computer programming.

Sample size calculations for the study were carried out using the R package `rdpower` [Cattaneo et al., 2022] and updated based on response rates to a pilot survey. To calculate the appropriate sample size, we ran simulations of the following potential outcomes of interest: annual income levels and employment status. Note that the platform does not track these outcomes which is what necessitated the use of simulations. Simulation outcomes were either randomly distributed across levels or distributed as a function of the running variable i.e. computer programming score. We considered standardized effect sizes of $\tau = (0.25, 0.5, 0.75)$.

Of these only the larger effect sizes of $\tau = (0.5, 0.75)$ were found to be feasible given our budget constraints and produced a set of sample sizes ranging from 264 to 511 with a maximum bandwidth of 96 for the running variable. We then conducted pilot surveys amongst a set of randomly sampled individuals (who were dropped from the subsequent e-mail survey) and based on these response rates our target bandwidth was updated to 115.

Of the $105,984$ test-takers in our population of interest, we sent the final e-mail survey to a total of $17,752$ randomly sampled individuals ($11,444$ and 6308 individuals in the treatment and control groups respectively) and received 358 valid responses (241 and 114 individuals in the treatment and control groups respectively). The resulting sample has two main drawbacks. First, individuals in our sample self-select into the survey. We describe in the next section the appropriate weighting schemes we utilize to address the potential biases that are likely to arise from such self selection. Second, our budget constraints limit this study's ability to estimate smaller effect sizes. The conclusions of our analysis thus correspond to large effect sizes only. In other words, estimated null effects in our estimations may in fact correspond to smaller non-zero effects which our estimations are unable to capture.

We consider 3 primary measures of job market outcomes collected via the e-mail surveys. First, we measure the respondent's success in gaining entry into the labour market via a survey question that captures whether the respondent was employed for ANY period of time after taking the exam. Second, we measure sustained employment over a longer time period by ascertaining whether the respondent is currently employed. Third, we capture the respondent's income levels in the longer term via the survey, where respondents select one of the following options to report their income-(i) no income, (ii) less than 1 lakh (i.e., 100,000 rupees) per annum, (iii) between 1 and 3 lakh per annum, (iv) between 3 and 5 lakh per annum, and (v) more than 5 lakh per annum. These options are coded as income levels of $0, 1, 2, 3, 4$ respectively. [5]

We control for several other factors that may be correlated with job market outcomes in our analysis. This includes scores on other tests, i.e., logical reasoning, quantitative ability and English language as well as demographic factors such as respondents' age, religion, caste, number of years of education and location of undergraduate institution. The information on scores in other tests, respondents' age and location of undergraduate institution are provided by the platform and information on respondents' years of education, religion and caste is collected via the online survey.

Statistic	N	Mean	St. Dev.	Min	Max
Outcomes of Interest					
Employment Status Post Exam	349	0.57	0.50	0.00	1.00
Current Employment Status	351	0.66	0.47	0.00	1.00
Income Level	347	2.35	1.41	0.00	4.00
Subject Scores					
Computer Programming Score	355	426.72	59.66	285	515
Quantitative Score	355	577.59	91.97	385	885
Logical Ability Score	355	540.61	47.97	450	725
English Score	355	593.58	73.92	490	865
Other Demographic Variables					
Age	353	23.52	1.11	21.00	29.00
Female	352	0.25	0.43	0.00	1.00
Years of Education	352	16.18	0.73	12.00	18.00
Hindu	353	0.76	0.42	0.00	1.00
Forward Caste	352	0.69	0.46	0.00	1.00
Location of College (Metro/Non-Metro)	355	0.50	0.50	0	1
Class 12 Percentage	355	83.10	10.15	52	99
Class 10 Percentage	355	87.97	6.72	61	100

Table 1: **Summary Statistics**

Table 1 provides information regarding the distribution of the key variables in our dataset. Our sample demonstrates a fairly symmetric distribution of respondents across the five income levels. While one-fifth of the people reported no income, 47% of the respondents earned an income between 3 and 5 lakh per annum. Additionally, 18% of our respondents earned an income of more than 5 lakh per annum, while 15% reported other income levels. Many estimates put the average salary of recently graduated engineering graduates from non-elite schools between INR 30-40 thousand per month (i.e 3.6 to 4.8 lakhs per annum) [Tilak and Choudhury, 2021], which approximates the median range from our sample. The average age of a respondent in our sample is 23 years and 33% of our respondents are female. Approximately 88% have completed 16 years of education, which includes 12 years of schooling and an average of four years of higher education. Approximately 75% of candidates are Hindu and half of the respondents belong to the forward (nondisadvantaged) castes [6]. The college location for respondents is equally divided between metro and non-metro cities. We also have access to the percentage marks scored by each respondent in Grade X and Grade XII, along with their year of graduation from high school.

3 Methods

3.1 Empirical Strategy

We use a regression discontinuity design (RDD) to identify the causal impact of certification on job-market outcomes. In general, comparing respondents who obtain the certificate (the treatment group) and those who do not (the control group) would not lead to an unbiased estimate of the causal impact of the certificate because individuals in the two groups may manifest other unmeasured differences. However, in this case, a test taker has to score more than 400 in computer programming to obtain the certificate, besides clearing a set of predetermined cut-offs in other subjects. This leads to a sharp discontinuity in the computer programming score as demonstrated in Figure 1. Insofar as respondents do not have control over where the cutoff is set, comparing individuals just below and above the cut-off leads to an unbiased estimate of the causal impact for certification.

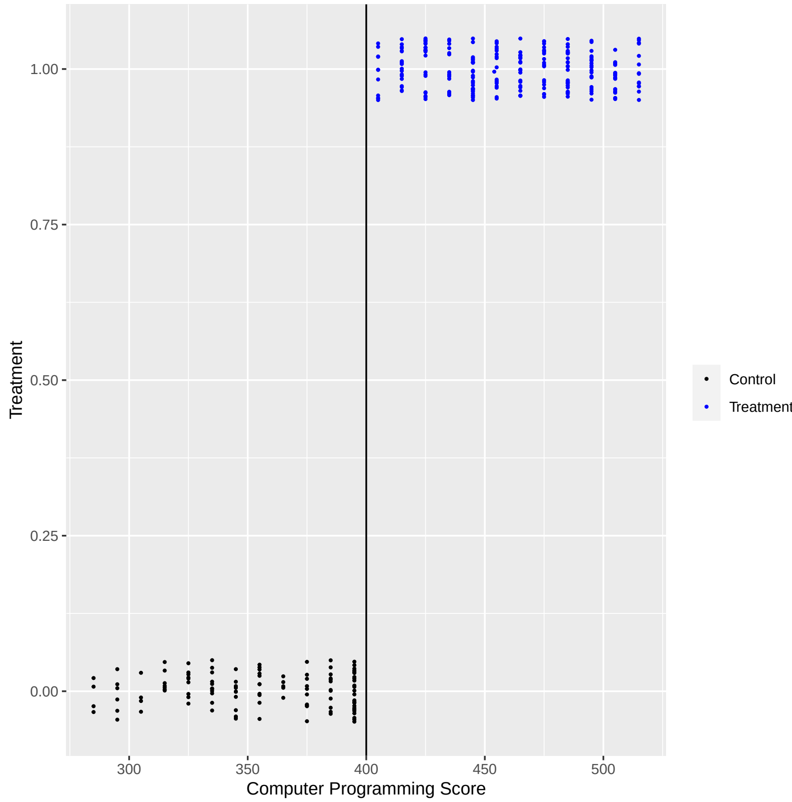

Figure 1: **A Sharp Regression Discontinuity Design**: We use a jitter plot to visualise the number of candidates who obtain a certificate. Blue dots represent certified candidates and black is for candidates without a certificate. No candidate scoring below the computer programming score cut-off receives a certificate

We employ the following regressions to estimate the causal impact of certification on labour market outcomes. First, to estimate the impact of certification on obtaining at least one job match, we run the linear probability model

$$P(EmpPostExam_i = 1) = \beta_0 + \tau D_i + \beta_1 CS_i + \beta_x X_i \tag{1}$$

where $EmpPostExam_i = 1$ if the respondent was employed for ANY period of time between the exam and the survey, and 0 otherwise. CS_i measures the percentage score in computer programming. $D_i = 1$ when an individual obtained the certificate and 0 otherwise. X_i is the vector of additional controls including demographic controls and percentage scores in sections on quantitative ability, logical ability and English.

Second, to estimate the impact of certification on sustained employment we run the linear probability model

$$P(CurrentEmp_i = 1) = \beta_0 + \tau D_i + \beta_1 CS_i + \beta_x X_i \tag{2}$$

where $CurrentEmp_i = 1$, if the respondent was employed at the time of taking the survey and 0 otherwise.

Third, to measure the effect of certification on level of current wages, we use the linear regression

$$IncomeLevel_i = \beta_0 + \beta_1 CS_i + \tau D_i + \beta_x X_i + \epsilon_i \tag{3}$$

where $IncomeLevel_i$ refers to the annual income level of respondent i. The reference income levels are as follows: (i) no income - **0**, (ii) an income of less than 1 lakh per annum- **1**, (iii) an income between 1 and 3 lakh per annum- **2**, (iv) an income between 3 and 5 lakh per annum- **3**, and (v) an income of more than 5 lakh per annum- **4**.

We note here that while the choice of a linear probability model as well as a linear regression model here is motivated by their ease of interpretation, we also run the more standard logistic regression and proportional odds models to validate our results. These models, along with other variations are described in section 3.3.

3.2 Post-stratification weights

Given that individuals in our study comprise a sample of respondents whose participation in the online survey is voluntary, it is necessary to account for potential selection bias in the sample. In order to address this concern, we use inverse probability weights for post-stratification on the models specified in eqs (1–3). The selection of specific attributes to calculate these post-stratification weights is based on the assumption that they are correlated with the probability of participation in the survey. Typically, studies show that gender can lead to systematic bias in opt-in surveys, where women are less likely to respond on topics that do not interest them [Chang and Krosnick, 2009]. Additionally, response rate may vary depending upon whether an individual's college is located in a metro city, which captures a person's access to opportunities and their general college environment. Participation may also be influenced by other characteristics, specifically the educational attributes of an individual.

We assign every respondent a weight equal to $w_i = 1/\hat{p}_i$ where $\hat{p}_i$ is the estimated probability of survey participation and corresponds to the following model.

$$p_i = logit(P(response_i = 1)) = \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_k X_k \tag{4}$$

where $response_i = 1$ if the candidate participated in the survey and 0 otherwise. $X_1, X_2, \cdots X_k$ are the relevant attributes that include marks scored in Class X, high school passing year, gender and college location. This model is estimated for the complete sample of candidates whose computer programming score fell in the required range.

Results reported in the main paper include the following versions of the models in eqs (1–3): 1) unweighted regressions with no control variables, 2) unweighted regressions with controls, 3) weighted regressions with no controls and 4) weighted regressions with controls.

3.3 Robustness checks

We perform the following robustness checks on our models. Since two of our dependent variables are binary we also specify a logistic regression of the following form.

$$logit(P(y_i = 1)) = \beta_0 + \tau D_i + \beta_1 CS_i + \beta_x X_i \tag{5}$$

where y_i is the employment status post exam and current employment status. CS_i measures the computer programming score and X_i is the vector of additional controls.Moreover, considering that current income is categorized into four levels, we utilize a proportional odds regression model as follows.

$$(\text{Income Level}_i \leq j) = \beta_0 + \beta_1 CS_i + \tau D_i + \beta_x X_i \tag{6}$$

where $IncomeLevel_i$ refers to the annual income level of respondent i and j represents the reference income levels as described in section 3.1.

We also perform a series of additional robustness checks. First, we run multiple specifications of the models specified in equations 1, 2 and 3 by adding additional educational controls. Second, we estimate 1, 2 and 3 along with a squared term corresponding to the computer programming score CS in order to account for additional non-linear effects. Third, we estimate regression models using a smaller sample of individuals, restricting the bandwidth of computer programming scores to 365-445. Smaller neighbourhoods make individuals in the treatment and control group more alike (the trade-off is a reduced sample size), providing further corroboration that any difference in outcome is due to certification alone. Results of our estimation remain qualitatively unchanged across all robustness checks. Finally, we also use local polynomial regressions to check for discontinuity at the threshold for our outcome variables. Tables and figure corresponding to our robustness checks are presented in the appendix of the paper.

3.4 Identification Evaluation

To evaluate whether our identification strategy is valid, we demonstrate that other than computer programming scores, respondents have overlapping distributions in other attributes. Figure 2 shows density plots for several attributes of candidates both in treatment and control groups. The distribution of computer programming scores for candidates in treatment and control groups are non-overlapping. However, both groups have a largely overlapping distribution of scores across the other three subjects as well as percentage scores in Grade

X and Grade XII. This suggests that candidates have comparable educational attributes except for their scores in computer programming. Table 2 presents a comparison of other characteristics for both groups. There are no significant differences in the means of the two samples across age, gender, college area, religion and caste. This strengthens our assertion that candidates below and above threshold are similar in all aspects and any difference in their employment outcomes can be attributed to certification and their scores in computer programming.

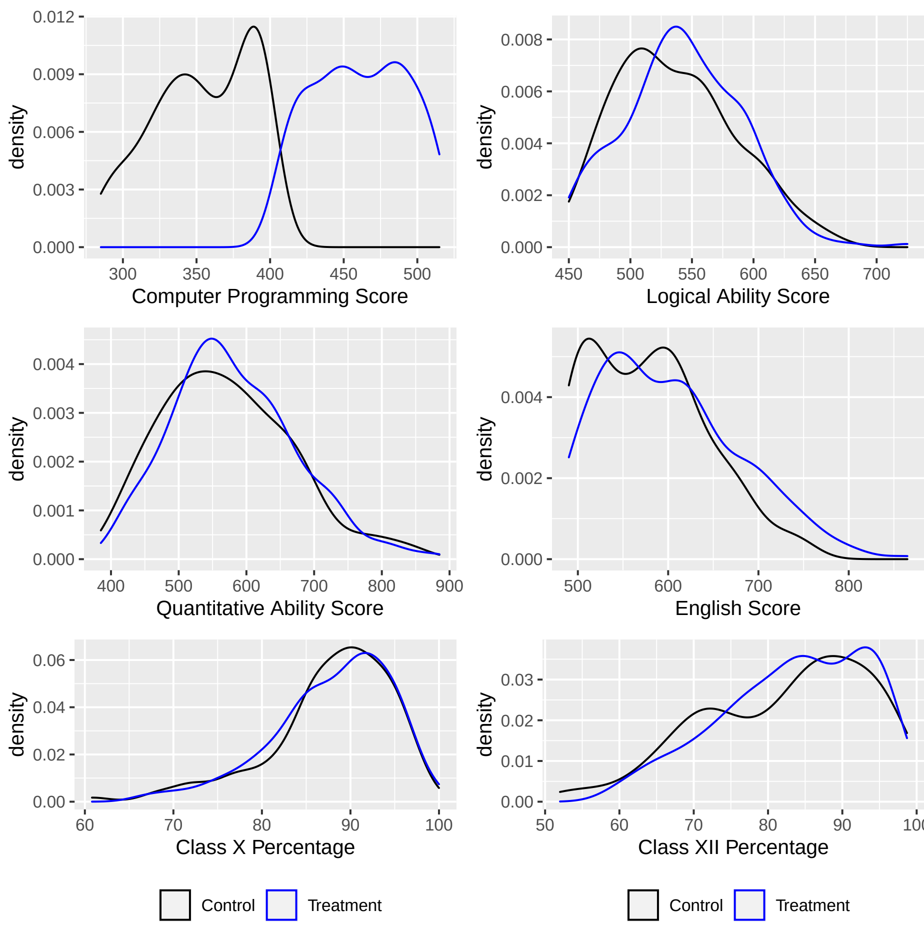

Figure 2: **Educational attributes of candidates in Treatment and Control**

	Treatment (N=241)	Control (N=114)	Difference
Mean Age (Years)	23.5	23.5	0
Gender (% Female)	26.5%	21%	5.5%
College Area (% Metro)	49%	52%	3%
Religion (% Hindu)	75.9%	76.3%	0.4%
Caste (% Forward Caste)	71%	62%	9% $^{+}$

Table 2: **Summary Statistics for Treatment and Control.**
Note: $^{+}$p<0.1; *p<0.05; **p<0.01.

4 Results

We plot the outcomes of interest against computer programming scores in Figure 3. The top panel of the figure depicts whether a respondent was employed for any period of time after taking the exam. The middle panel of Figure 3 shows the current employment status of the candidate against their computer programming scores. Finally, the bottom of Figure 3 plots the income levels of the respondents against computer programming scores on both sides of the cut-off. To perform a sensitivity analysis, we re-generate this figure using locally polynomial regression line which are presented in the appendix in Figure 4.

Employment Status Post Exam: The top panel of Figure 3 depicts whether a respondent was employed for any period of time after taking the exam. The regression line is a linear fit for the probability of being employed using computer programming score as the dependent variable. The figure shows a discontinuity in the outcome of respondents on the threshold.

Table 3 presents the unweighted and weighted regression estimates corresponding to (1) and confirm the results obtained in Figure 3 . As demonstrated by the estimates of ATE across all specifications, receiving the certificate leads to a significant improvement in the probability of finding at least one job. Compared to candidates without a certificate, whose probability of finding a job after taking the exam is about 0.4, certified candidates have an estimated probability that is higher by about 0.25 of finding employment. These estimates are significant at 95% confidence levels for the weighted regressions that correct for potential non-response bias. Thus our primary result shows that online certification in this context manifests a significant and substantial value from the perspective of matching a candidate to at least one job.

Of the control variables only 2 significant relationships emerge. First, candidates with more years of education exhibit lower chances of being employed after the exam. This result confirms findings from other studies carried out in developing countries that have observed a negative relationship between employment and education due to demand and skill mismatch or lower absorption rates of the labour markets [O'higgins, 2001; Bairagya et al., 2015]. The less studied relationship these results reveal is that a one unit increase in logical reasoning score is associated with a 0.001 increase in the probability of obtaining a job [Boissiere et al., 1985; Nikolov et al., 2020; Glewwe et al., 2022]. Albeit a correlational finding, to our knowledge this is the first evidence within the Indian context of the potential importance of logical reasoning ability in the formal labour market.

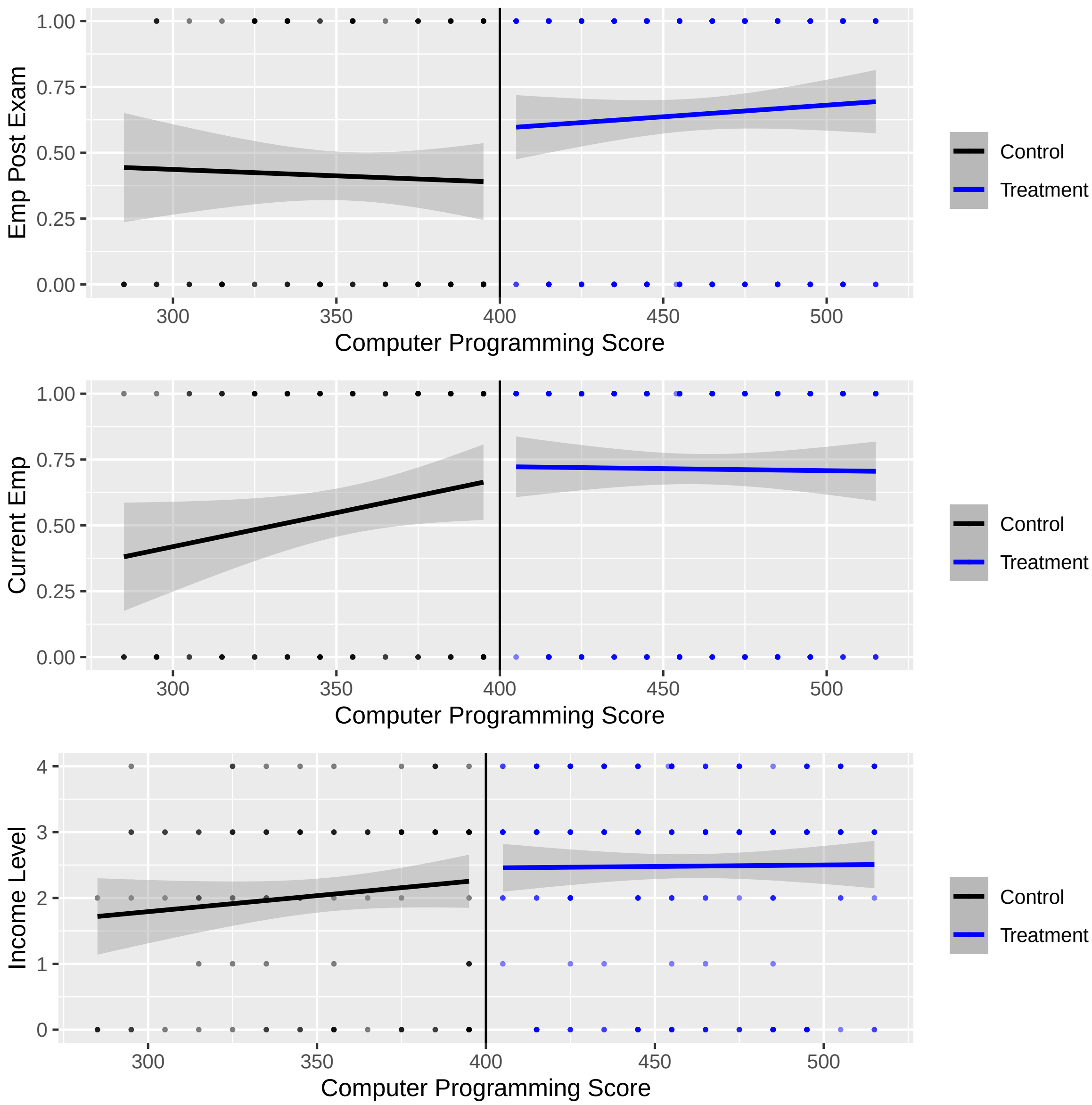

Figure 3: **Returns to certification on job market outcomes**: We plot outcomes of interest against computer programming scores. The lines represent linear regression fits using computer programming score as the running variable.

Current Employment Status: The middle panel of Figure 3 shows the current employment status of the candidate against their computer programming scores. The regression lines estimated on both sides of the threshold depict that there is no obvious discontinuity in the probability of being currently employed for candidates just below and above the threshold. Table 4 presents the unweighted and weighted regression results corresponding to

	Dependent variable:			
	Emp Post Exam			
	unweighted		*with post-stratification weights*	
	(1)	(2)	(3)	(4)
Treatment	0.198^{+}	0.195^{+}	0.244^{*}	0.249^{*}
Computer Programming Score	0.0004	0.0003	−0.0001	−0.0002
Quantitative		−0.00001		−0.00005
English Score		0.0003		0.0002
Logical		0.001		0.001^{*}
Hindu		0.045		0.028
Forward Caste		0.082		0.070
Years of education		-0.061^{+}		-0.074^{*}
Female		−0.048		−0.042
Control Group Mean	0.405		0.392	
Observations	349	345	349	345
R^2	0.052	0.078	0.049	0.081
F Statistic	9.567^{**}	3.163^{**}	8.984^{**}	3.270^{**}

Table 3: **Estimating Average Treatment Effect of Certification on Employment Status post exam**

Notes: Columns 1 and 2 estimate equation 1 using linear probability model. Columns 3 and 4 are weighted linear probability models using post-stratification weights to estimate equation 1. We show the outcome mean for the control group, unweighted and using post-stratification weights in column 1 and 3 respectively. $^{+}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$.

(2) where these findings are confirmed. Unlike the case of being employed at least once after certification, across all specifications the probability of being employed at the time of the survey for certified candidates is not significantly different from those not certified, indicating that the impact of certification does not extend to a long-term employed status. Reasons behind the lack of sustained impact on employment cannot be disentangled by the current analysis as a candidate's unemployed status may be driven either because they voluntarily resign or are fired.

Nevertheless, our results point to other potential predictors of sustained employment outcomes. First, as before, we find that candidates with more years of education exhibit significantly lower chances of being employed during the time of the survey [O'higgins, 2001; Bairagya et al., 2015]. Second, we find that female candidates have a lower probability of remaining employed. This too corroborates an observed trend across many labour market contexts in India [Srivastava and Srivastava, 2010; Klasen and Pieters, 2015; Mahajan and Ramaswami, 2017]. Third, we continue to find a strong and positive association between logical reasoning score and sustained employment.

Strikingly, we find a very strong and significant association between belonging to a forward (non-discriminated against) caste group and sustained employment. A candidate belonging to a forward caste group is estimated to have approximately 0.18 higher chance of being employed at the time of the survey. While the importance of caste networks for finding jobs has been widely studied in the context of rural India as well as the informal sector [Munshi and Rosenzweig, 2005; Das and Dutta, 2007; Ito, 2009], there is limited empirical evidence that caste plays the role of a mediating factor within India's formal labour market [Banerjee et al., 2009; Siddique, 2011 ; Basole et al., 2021]. Our paper adds to this literature, pointing to potential caste-based differences within India's formal labour market.

Income Levels: The bottom panel of Figure 3 plots the income levels of the respondents against computer programming scores on both sides of the cut-off. The lines are linear estimates of the average level of income as a function of the computer programming score

	Dependent variable:			
	Current Emp			
	unweighted		*with post-stratification weights*	
	(1)	(2)	(3)	(4)
Treatment	0.068	0.066	0.112	0.117
Computer Programming Score	0.001	0.001	0.001	0.001
Quantitative		−0.0001		−0.0001
English Score		−0.0001		−0.0001
Logical		0.001*		0.001*
Hindu		0.019		0.029
Forward Caste		0.170**		0.181**
Years of education		−0.113**		−0.115**
Female		−0.104+		−0.112*
Control Group Mean	0.553		0.522	
Observations	351	347	351	347
R^2	0.029	0.106	0.034	0.124
F Statistic	5.241**	4.447**	6.104**	5.306**

Table 4: **Estimating Average Treatment Effect of Certification on Current Employment Status**
Notes: Columns 1 and 2 estimate equation 2 using linear probability model. Columns 3 and 4 are weighted linear probability models using post-stratification weights to estimate equation 2. We show the outcome mean for the control group, unweighted and using post-stratification weights in column 1 and 3 respectively. +p<0.1; *p<0.05; **p<0.01.

	Dependent variable:			
	Income Level			
	unweighted		*with post-stratification weights*	
	(1)	(2)	(3)	(4)
Treatment	0.208	0.200	0.341	0.340
Computer Programming Score	0.021	0.020	0.014	0.012
Quantitative		0.008		0.007
English Score		−0.006		−0.006
Logical		0.031*		0.037*
Hindu		0.162		0.154
Forward Caste		0.093		0.173
Years of education		−0.103		−0.104
Female		−0.422*		−0.404*
Control Group Mean	2.036		1.919	
Observations	347	343	347	343
R^2	0.026	0.071	0.027	0.082
F Statistic	4.519*	2.837**	4.832**	3.318**

Table 5: **Estimating Average Treatment Effect on current income**
Notes: Columns 1 and 2 estimate equation 3 using linear probability model. Columns 3 and 4 are weighted linear probability models using post-stratification weights to estimate 3. We show the outcome mean for the control group, unweighted and using post-stratification weights in column 1 and 3 respectively. *Note*:+ p < 0.1, * p < 0.05, ** p < 0.01

for candidates on both sides of the cut-off. There appears to be no clear discontinuity in the average income of candidates just below and above the cut-off. We confirm this result in Table 5 which presents the weighted and unweighted regression results corresponding to (3). As in the case of employment status at the time of the survey, certification does not have a significant impact on a candidate's income level at the time of the survey. Therefore, there is no evidence of a causal impact of certification for long-term wage growth. This result is in line with findings from other contexts [Heckman et al., 2011, Silles, 2007; Blanden et al., 2012,Booth and Bryan, 2005; Wolf et al., 2006 and Wolf 2011]. Of the remaining covariates, the only predictor of longer-term income levels is again logical reasoning ability, as in all previously reported models. A one unit increase in the logical ability score is associated with 0.03 increase in the income level.

4.1 Robustness checks

As outlined in subsection 3.3, we conduct an array of robustness checks. These checks involve specifying logistic and odds ratio models for employment outcomes and income level, respectively. We also specify multiple regressions that include additional controls. Additionally, we account for any non linear effects by adding a squared term for computer programming score and re-evaluate our regressions. Furthermore, we also replicate our main results over a smaller sample where the bandwidth of computer programming score is restricted to 365-445. Finally, we reproduce our reduced form results presented in 3 using local polynomial regression fits. Results of our analysis remain qualitatively unchanged. Interested readers are directed to the appendix for relevant tables and figures.

4.2 Summary of findings

Taken together, our results suggest that certification leads to a strong and positive impact on the labour market outcomes of candidates, leading to nearly one-fourth higher chances of finding *some* employment. Nevertheless, given our study design, there is no measurable impact of certification on longer-term indicators such as long-term employment status or wage level. In this way, our results indicate that while certification in this context represents an effective search and matching instrument, it cannot necessarily replace sustained education and manifests limited impact for longer-term performance in the labour market. We stress once again here that while our estimates of the causal impact of certification on long-term labor market outcomes are not significantly different from zero, they are positive. This leaves open the possibility that our estimates are indicative of noisy positive impacts of certification which could not be captured due to the limited sample size of the study.

Importantly, our results point towards logical reasoning as a consistently important predictor of labour market outcomes across contexts. This is particularly interesting because other factors have more traditionally been considered crucial for long-term employment, including quantitative ability [Black et al., 2015 and Arcidiacono, 2004] and English language capacity [Azam et al., 2013 and Jain et al., 2019], which fail to post a significant impact in this context. Further, our results indicate that even within a formal sector context in India, caste and longer-term employment have a strong association. While many scholars have argued that caste networks matter beyond informal settings [Banerjee et al., 2009; Siddique, 2011] there is limited empirical evidence from India corroborating this argument, and our paper addresses this gap in the literature. However we refrain from making any causal claim with regard to both these associations [Hünermund and Louw, 2020] and only posit them as important areas of future investigation.

5 Discussion

Certifications based on computerised assessments provided by online platforms can be a cost-effective tool to assess candidate aptitude, especially in markets like India that have seen unrestricted growth in private educational institutions without strong reputations and manifesting inconsistent quality. The expertise and reach of certification platforms for administering computerised tests can benefit employers by reducing the cost of finding suitable talent, but also job market applicants by providing them the means to broadcast and advertise their skills. After the COVID19 pandemic, in particular, employers may be more inclined to hire people remotely [Barrero et al., 2021], and external certifications can build trust and credibility without employers spending the time and effort to design customized tests or evaluation mechanisms.

Several online platforms have emerged in recent years within the Indian context that provide different types of certification based on standardised assessments. One of the primary objectives for such platforms is to provide a signalling mechanism for candidates that assists with job matching. Nevertheless, empirical assessment of the utility of such digital certifications remains an under-studied topic globally, particularly in the context of Global South countries like India where they may be leveraged to play an outsized role in the labour market. Our study addresses this lacunae in the literature by estimating the causal impact of IT software skills certification on job market outcomes for Indian engineering graduates. Our novel dataset, combining primary survey and proprietary information from a skills certification platform, helps us estimate the impact of certification on job market outcomes using a regression discontinuity design to identify causal impact.

Our results indicate that certification leads to significantly higher chances for an individual to find a near-term job. Nevertheless, it does not replace the role of education and cannot guarantee employment over a sustained period of time, nor does it significantly affect wage levels over time. Two plausible pathways explain these results. First, online skills certification platforms provide consultation and networking services that help candidates connect with prospective employers. Such matching services can lead to a proximate increase in employment for certified candidates, which can be attributed to the platform's mediation services. As mentioned before, registered companies shortlist candidates for an interview if they clear all pre-determined cut-offs. These may lead to reduced uncertainty among employers regarding a candidate's ability during the hiring process. Second, candidates who obtain a certificate become highly motivated to look for jobs. They may employ more search effort to seek out jobs, an influence that wanes with time. Together, these suggest that certification based on standardised assessments help to lubricate the labor market in the short term, but cannot themselves solve major structural inequities in job preparation. Effective policy must account for other factors to improve long-term employability[McKenzie, 2017; Kluve et al., 2019; Abebe et al., 2021].

Certified candidates may not experience a causal impact on longterm wage-level as certification can only measure pre-existing skills; they do not influence the accumulation of new skills that warrant higher wages [Becker, 1962]. This argument can be attributed to human capital theory, but another line of thought from the screening literature provides an alternative explanation for an absence of the difference in wages. Employers often use years of schooling or additional education to make inferences about the unobservable ability of a candidate, such as higher IQ or lower propensity to quit. From this viewpoint, certification could act as proxy or correlated signal with pre-existing indicators [Weiss, 1995].

While our primary results measure the impact of certification, our regressions also estimate two potentially important associations. First, there is a strong and positive association between logical ability scores and all job market outcomes, both for respondent's short-term and long-term employment status and income levels. This suggests that logical reasoning abilities, as captured in this test, may represent a strong signal of skills valued by employers. Standardized assessments might be designed to incorporate this aspect more directly when deciding criteria for the award of the certificate if the aim is to have an impact on larger time-scales. Second, we estimate a large and positive association between belonging to a forward caste group and a candidate's longer-term employment status. This association does not show up for the short-term employment status, providing suggestive evidence that network ties are an important aspect that affects individual employment status over the long run. Besides providing a means to certify an individual's hard skills, mediation services that aid in building connections with prospective employers can increase the chances of finding employment for those belonging to historically disadvantaged caste groups.

We acknowledge limitations associated with our study. The chief caveat of our study is that our budget constraints limit our study's sample size which in turn limits our ability to capture smaller effects. In other words, estimated null effects may in fact correspond to smaller non-zero effects which our estimations are unable to capture. Further, our measure of salary in income-levels limits our ability to calculate precise estimates of the impact of certification on income. However, self-reported income data is fraught with measurement issues [Brehm, 2009] and our approach of eliciting annual income bands was responsive to them. Initial pilot surveys suggested that respondents made fewer errors when reporting annual income levels as opposed to precise incomes. Additionally, voluntary participation of candidates in our study may lead to non-response bias [Brehm, 2009; Chang and Krosnick, 2009]. We address this by utilizing post-stratification weights using a large number of covariates. Further, we run multiple specifications of the regressions and conduct a variety of robustness checks. Our results remain qualitatively similar across all specifications considered.

Our analysis makes a number of contributions to the literature. First, the study adds to the limited literature on digital certifications globally and does so in particular for the Global South context. Specifically, our study understands the role of external certification in addressing information asymmetry present in the Indian software job market and provides evidence of proximate impact. Second, our study is amongst the first in the Indian context to shed light on the potential role of the 'soft' skill of logical reasoning ability in the labour market. Third, our study provides additional evidence of the role caste plays within India's formal labour market, adding to existing literature that has centered on informal workers.

Future research emerging from our analysis includes an impetus to study employer perspectives of external skills certification to improve long-term impact on worker careers. Our findings suggest the greater need to develop these tests in collaboration with employers after a thorough understanding of their skill requirements. Another line of inquiry involves optimizing the cut-off for certification to achieve specific impact. The potential benefit of a single indicator such as logical reasoning may also be further explored. While this paper focused solely on software certification, exploring the association between logical reasoning across a range of certifications and job types could illuminates its mechanisms of impact. Finally, our results make a case for a more thorough investigation of the role of caste networks within India's formal labor market, particularly within the high-growth IT industry.

Notes

[1]See here excerpts from an interview with the CEO of an Indian IT firm, Tech Mahindra, in 2018. https://economictimes.indiatimes.com/jobs/only-6-of-those-passing-out-of-indias-engineering-colleges-are-fit-for-a-job/articleshow/64446292.cms?from=mdr.

[2] https://indianexpress.com/article/education/how-the-indian-engineering-degree-lost-its-sheen-this-decade-6182675/.

[3]Auto-proctoring generally uses computer vision to monitor candidates during an online exam. This includes tracking movements and facial expressions, as well as analysing audio and video feeds for signs of cheating.

[4]The data was collected in March 2020, just before the nationwide COVID19 lockdown came into proper effect. Therefore, the data does not capture any effect of COVID19.

[5]Surveys that depend on voluntary participation suffer from misreporting and under-reporting of information on salary. Questions that elicit exact responses for salary are often left blank or people do not report their true income. Often, respondents mistakenly either add or miss out a digit 0 when reporting annual income. An initial pilot study conducted by us confirmed some of these common measurement issues leading us to choose this format for reporting. We specify income levels in the questionnaire for the respondents to choose from to avoid these pitfalls.

[6]The Government of India categorizes historically disadvantaged groups under 'Scheduled Castes', 'Scheduled Tribes' or 'Other Backward Classes'. The remaining population is categorized as belonging to the 'General' category. Individuals in this category are often referred to as belonging to 'Forward Castes' which are considered on average to be ahead economically and educationally.

References

G. Abebe, A. S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn. Anonymity or distance? job search and labour market exclusion in a growing african city. *The Review of Economic Studies*, 88(3):1279–1310, 2021.

A. Adhvaryu, N. Kala, and A. Nyshadham. The skills to pay the bills: Returns to on-the-job soft skills training. Technical report, National Bureau of Economic Research, 2018.

P. Arcidiacono. Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375, 2004.

M. Azam, A. Chin, and N. Prakash. The returns to english-language skills in india. *Economic Development and Cultural Change*, 61(2):335–367, 2013.

I. Bairagya et al. *Socio-economic determinants of educated unemployment in India*. 2015.

A. Banerjee, M. Bertrand, S. Datta, and S. Mullainathan. Labor market discrimination in delhi: Evidence from a field experiment. *Journal of comparative Economics*, 37(1):14–27, 2009.

J. M. Barrero, N. Bloom, and S. J. Davis. Why working from home will stick. Technical report, National Bureau of Economic Research, 2021.

A. Basole, R. Abraham, R. Lahoti, S. Kesar, M. Jha, P. Nath, R. Kapoor, and S. N. Mandela. State of working india 2021: one year of covid-19. 2021.

G. S. Becker. Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49, 1962.

S. Black, Z. He, C. Muller, and A. Spitz-Oener. On the origins of stem: The role of high school stem coursework in occupational determination and labour market success in mid life. *Unpublished Manuscript*, 2015.

J. Blanden, F. Buscha, P. Sturgis, and P. Urwin. Measuring the earnings returns to lifelong learning in the uk. *Economics of Education Review*, 31(4):501–514, 2012.

A. Blom and H. Saeki. Employability and skill set of newly graduated engineers in india.

World Bank Policy Research Working Paper, (5640), 2011.

M. Boissiere, J. B. Knight, and R. H. Sabot. Earnings, schooling, ability, and cognitive skills. *The American Economic Review*, 75(5):1016–1030, 1985.

A. L. Booth and M. L. Bryan. Testing some predictions of human capital theory: New training evidence from britain. *Review of Economics and statistics*, 87(2):391–394, 2005.

B. Bratsberg, T. Nyen, and O. Raaum. Economic returns to adult vocational qualifications. *Journal of Education and Work*, 33(2):99–114, 2020.

J. O. Brehm. *The phantom respondents: Opinion surveys and political representation*. University of Michigan Press, 2009.

J. Castaño-Muñoz and M. Rodrigues. Open to moocs? evidence of their impact on labour market outcomes. *Computers & Education*, 173:104289, 2021.

M. D. Cattaneo, R. Titiunik, and G. Vazquez-Bare. *rdpower: Power Calculations for RD Designs*, 2022. URL https://CRAN.R-project.org/package=rdpower. R package version 2.2.

L. Chang and J. A. Krosnick. National surveys via rdd telephone interviewing versus the internet: Comparing sample representativeness and response quality. *Public Opinion Quarterly*, 73(4):641–678, 2009.

M. B. Das and P. V. Dutta. Does caste matter for wages in the indian labor market. *Washington, DC, USA: The World Bank*, 2007.

M. Denzer, T. Schank, and R. Upward. Does the internet increase the job finding rate? evidence from a period of internet expansion. *Evidence from a Period of Internet Expansion*, 2018.

P. Glewwe, Y. Song, and X. Zou. Labor market outcomes, cognitive skills, and noncognitive skills in rural china. *Journal of Economic Behavior & Organization*, 193:294–311, 2022.

J. J. Heckman, J. E. Humphries, and N. S. Mader. The ged. *Handbook of the Economics of Education*, 3:423–483, 2011.

P. Hünermund and B. Louw. On the nuisance of control variables in regression analysis.

arXiv preprint arXiv:2005.10314, 2020.

T. Ito. Caste discrimination and transaction costs in the labor market: Evidence from rural north india. *Journal of development Economics*, 88(2):292–300, 2009.

T. Jain, P. Maitra, and S. Mani. Barriers to skill acquisition: Evidence from english training in india. *World Development*, 114:314–325, 2019.

C. Jepsen, P. Mueser, and K. Troske. Second chance for high school dropouts? a regression discontinuity analysis of postsecondary educational returns to the ged. *Journal of Labor Economics*, 35(S1):S273–S304, 2017.

O. Kässi and V. Lehdonvirta. Do digital skill certificates help new workers enter the market? evidence from an online labour platform. Technical report, CESifo Working Paper, 2019.

O. Kässi and V. Lehdonvirta. Do microcredentials help new workers enter the market? evidence from an online labor platform. *Journal of Human Resources*, 2022.

S. Klasen and J. Pieters. What explains the stagnation of female labor force participation in urban india? *The World Bank Economic Review*, 29(3):449–478, 2015.

J. Kluve, S. Puerto, D. Robalino, J. M. Romero, F. Rother, J. Stöterau, F. Weidenkaff, and M. Witte. Do youth employment programs improve labor market outcomes? a quantitative review. *World Development*, 114:237–253, 2019.

R. Kumar, S. Mandava, and V. S. Gopanapalli. Vocational training in india: determinants of participation and effect on wages. *Empirical Research in Vocational Education and Training*, 11(1):1–17, 2019.

K. Mahajan and B. Ramaswami. Caste, female labor supply, and the gender wage gap in india: Boserup revisited. *Economic Development and Cultural Change*, 65(2):339–378, 2017.

D. McKenzie. How effective are active labor market policies in developing countries? a critical review of recent evidence. *The World Bank Research Observer*, 32(2):127–154, 2017.

K. Munshi and M. Rosenzweig. Economic development and the decline of rural and urban community-based networks. *Economics of Transition*, 13(3):427–443, 2005.

P. Nikolov, N. Jimi, and J. Chang. The importance of cognitive domains and the returns to schooling in south africa: Evidence from two labor surveys. *Labour economics*, 65:101849, 2020.

N. O'higgins. Youth unemployment and employment policy: A global perspective. *O'Higgins*, (2001), 2001.

Z. Siddique. Evidence on caste based discrimination. *Labour Economics*, 18:S146–S159, 2011.

M. A. Silles. Adult education and earnings: evidence from britain. *Bulletin of Economic Research*, 59(4):313–326, 2007.

M. Spence. Job market signaling. In *Uncertainty in economics*, pages 281–306. Elsevier, 1978.

N. Srivastava and R. Srivastava. Women, work, and employment outcomes in rural india. *Economic and political weekly*, pages 49–63, 2010.

D. L. Thistlethwaite and D. T. Campbell. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6):309, 1960.

J. B. Tilak and P. K. Choudhury. Employment and employability of engineering graduates in india. *Journal of Contemporary Educational Research*, 5(3), 2021.

N. Varghese. Challenges of massification of higher education in india. *CPRHE research papers*, 1:1–52, 2015.

L. Weber. Today's personality tests raise the bar for job seekers, 2015. URL https://www.wsj.com/articles/a-personality-test-could-stand-in-the-way-of-your-next-job-1429065001. Accessed on January 24, 2023.

A. Weiss. Human capital vs. signalling explanations of wages. *Journal of Economic perspectives*, 9(4):133–154, 1995.

A. Wolf. Review of vocational education. *London: DfE*, 2011.

A. Wolf, A. Jenkins, and A. Vignoles. Certifying the workforce: economic imperative or failed social policy? *Journal of education policy*, 21(5):535–565, 2006.

Appendix

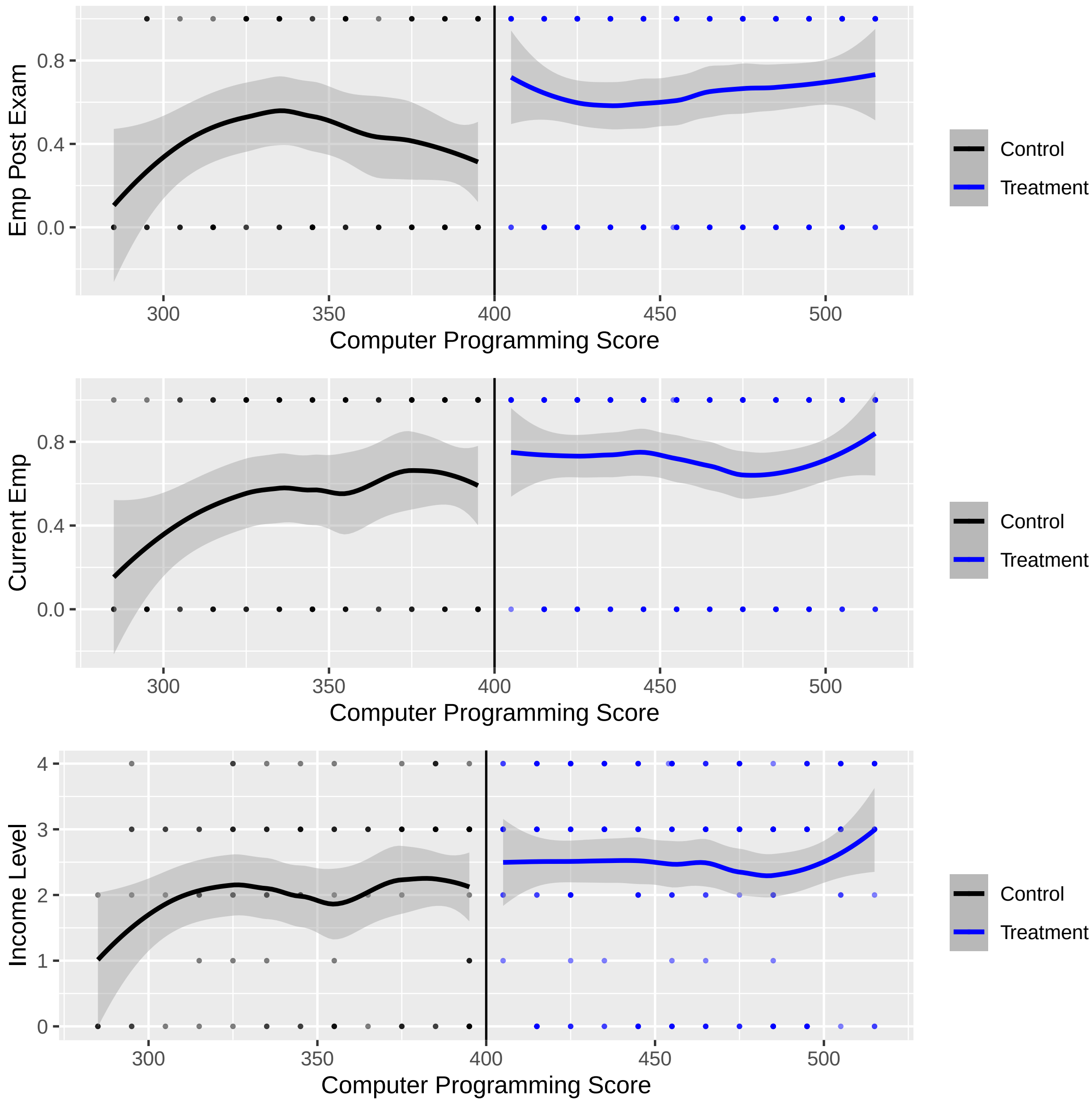

Figure 4: **Returns to certification on job market outcomes**: The figure plots the outcomes of interest against computer programming scores. The fitted lines are estimated using local polynomial regression. As shown in panel 1, compared to Figure 3 there is a stark discontinuity for employment status post exam at the computer programming cut-off. We also see some discontinuity at the threshold for current employment status, however, the effect is not significant.

	Dependent variable:		
	Emp Post Exam		
	(1)	(2)	(3)
Treatment	0.193^{+}	0.210^{*}	0.388^{*}
Computer Programming Score	0.003	-0.035	-0.025
I(Computer Programming Scoreˆ2)		0.0004	
Quantitative	0.0002	-0.0001	-0.0003
English Score	0.003	0.003	0.005
Logical	0.008	0.008	0.018^{*}
Hindu	0.045	0.047	0.007
Forward Caste	0.077	0.083	0.039
Years of education	-0.062^{+}	-0.062^{+}	-0.010
Female	-0.036	-0.049	-0.062
X10th_percentage	-0.002		
X12th_percentage	-0.001		
Observations	345	345	145
R^2	0.080	0.080	0.137
F Statistic	2.644^{**}	2.889^{**}	2.387^{*}

Table 6: **Estimating Average Treatment effect of Certification on Employment Status Post Exam using Linear Probability Model**

Notes: This table contains results from robustness checks. Column 1 estimates equation 1 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 1 with an additional squared term corresponding to the computer programming score. Column 3 estimates 1 on a smaller sample of individuals that score between 365-445 in computer programming. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

	Dependent variable:		
	Current Emp		
	(1)	(2)	(3)
Treatment	0.057	0.037	0.153
Computer Programming Score	0.008	0.079	-0.005
I(Computer Programming Score^2)		-0.001	
Quantitative	-0.0002	-0.001	-0.005
English Score	0.001	-0.001	0.002
Logical	0.011^{*}	0.010^{*}	0.017^{*}
Hindu	0.018	0.015	0.034
Forward Caste	0.149^{**}	0.168^{**}	0.113
Years of education	-0.117^{**}	-0.111^{**}	-0.031
Female	-0.060	-0.102^{+}	-0.123
X10th_percentage	-0.008^{+}		
X12th_percentage	-0.005^{+}		
Observations	347	347	146
R^2	0.139	0.111	0.095
F Statistic	4.925^{**}	4.212^{**}	1.594

Table 7: **Estimating Average Treatment effect of Certification on Current Employment Status using Linear Probability Model**:
Notes: This table contains results from Robustness Checks. Column 1 estimates equation 2 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 2 with an additional squared term corresponding to the computer programming score. Column 3 estimates estimates 2 on a smaller sample of individuals that score between 365-445 in computer programming.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Dependent variable:		
	Income Level		
	(1)	(2)	(3)
Treatment	0.200	0.163	0.432
Computer Programming Score	0.020	0.118	-0.027
I(Computer Programming Score^2)		-0.001	
Quantitative	0.008	0.008	-0.009
English Score	-0.006	-0.007	0.002
Logical	0.031*	0.031*	0.023
Hindu	0.162	0.154	0.377
Forward Caste	0.093	0.092	-0.025
Years of Education	-0.103	-0.100	0.070
Female	-0.422*	-0.419*	-0.486+
Observations	343	343	145
R^2	0.071	0.072	0.056
F Statistic	2.837**	2.587**	0.887

Table 8: **Estimating Average Treatment effect of Certification on Current Income using Linear Probability Model**

Notes: This table contains results from robustness checks. Column 1 estimates 3 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 3 with an additional squared term corresponding to the computer programming score. Column 3 estimates 3 on a smaller sample of individuals that score between 365-445 in computer programming.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Dependent variable:		
	Emp Post Exam		
	(1)	(2)	(3)
Treatment	2.249^{+}	2.414^{+}	5.944^{*}
Computer Programming Score	1.012	0.858	0.888
I(Computer Programming Score ˆ2)		1.002	
Quantitative	1.001	0.999	0.998
English Score	1.013	1.012	1.025
Logical	1.034	1.034	1.090^{*}
Hindu	1.219	1.233	1.037
Forward Caste	1.397	1.433	1.204
Years of education	0.763^{+}	0.760^{+}	0.984
Female	0.850	0.806	0.751
X10th_percentage	0.991		
X12th_percentage	0.995		
Observations	345	345	145
Akaike Inf. Crit.	467.533	465.766	199.331

Table 9: **Estimating Average Treatment effect of Certification on Employment Status Post Exam using Logistic Regression**
Notes: This table contains results from robustness checks and estimates the logistic regressions specified in 5 using Employment Status Post Exam as the dependent variable. Estimates presented correspond to odds-ratios. Column 1 estimates 5 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 5 with an additional squared term corresponding to the computer programming score. Column 3 estimates 5 on a smaller sample of individuals that score between 365-445 in computer programming. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Dependent variable:		
	Current Emp		
	(1)	(2)	(3)
Treatment	1.317	1.253	2.163
Computer Programming Score	1.040	1.426	0.981
I(Computer Programming Score ˆ2)		0.997	
Quantitative	0.999	0.994	0.970
English Score	1.004	0.996	1.010
Logical	1.059*	1.051*	1.102*
Hindu	1.072	1.058	1.144
Forward Caste	2.031**	2.208**	1.787
Years of education	0.552**	0.571**	0.850
Female	0.749	0.604^{+}	0.519
X10th_percentage	0.960^{+}		
X12th_percentage	0.975^{+}		
Observations	347	347	146
Log Likelihood	-196.675	-201.938	-81.033
Akaike Inf. Crit.	417.351	425.876	182.065

Table 10: **Estimating Average Treatment effect of Certification on Current Employment Status using Logistic Regression**:
Notes: This table contains results from robustness checks and estimates the logistic regressions specified in 5 using Current Employment Status as the dependent variable. Estimates presented correspond to odds-ratios. Column 1 estimates 5 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 5 with an additional squared term corresponding to the computer programming score. Column 3 estimates 5 on a smaller sample of individuals that score between 365-445 in computer programming.

	Dependent variable:		
	Income Level		
	(1)	(2)	(3)
Treatment	1.431	1.334	1.696
Computer Programming Score	1.031	1.228	0.991
I(Computer Programming Score^2)		0.998	
Quantitative	1.017^{+}	1.015	0.988
English Score	0.994	0.993	1.003
Logical	1.035^{+}	1.032^{+}	1.010
Hindu	1.306	1.309	1.741
Forward Caste	1.048	1.083	0.955
Years of education	1.004	1.007	1.244
Female	0.700	0.632^{+}	0.526^{+}
X10th_percentage	0.998		
X12th_percentage	0.981^{+}		
0—1	4.914	962.060**	9.998
1—2	6.131	1198.446**	12.755
2—3	11.648	2266.874**	22.144
3—4	86.973	16703.835**	166.791
Observations	343	343	145

Table 11: **Estimating Average Treatment effect of Certification on Current Income using Proportional Odds Model**
Notes: This table contains results from robustness checks and estimates the proportion odds model specified in 6 using Current Income as the dependent variable. Estimates presented correspond to odds-ratios. Column 1 estimates 6 using additional controls- Class 12 and Class 10 percentage. Column 2 estimates 6 with an additional squared term corresponding to the computer programming score. Column 3 estimates 6 on a smaller sample of individuals that score between 365-445 in computer programming.+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$