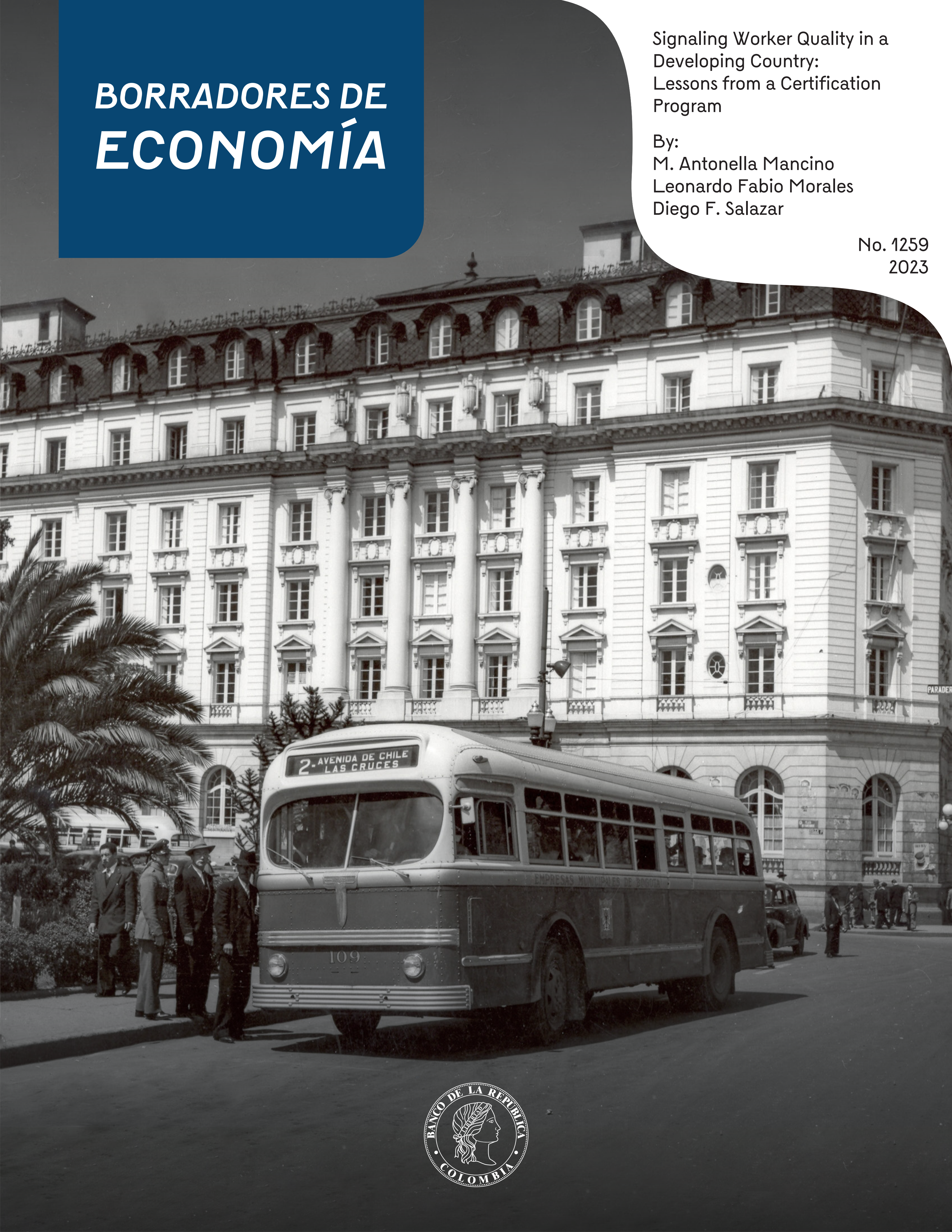
BORRADORES DE ECONOMÍA

Signaling Worker Quality in a Developing Country: Lessons from a Certification Program

By:
M. Antonella Mancino
Leonardo Fabio Morales
Diego F. Salazar

No. 1259
2023



BANCO DE LA REPUBLICA
COLOMBIA

Signaling Worker Quality in a Developing Country: Lessons from a Certification Program [1]

M. Antonella Mancino, Leonardo Fabio Morales, Diego F. Salazar

Abstract

We evaluate the returns to signaling occupation-specific skills using unique administrative data from a nationwide certification program in Colombia. The program certifies skills and issues three certificates: basic, intermediate, and advanced. We use regression discontinuity methods to compare workers' earnings around certificate-assignment thresholds. Signaling advanced occupation-specific skills yields significant returns: 9.7% on average within two years of certification. Instead, we find no effects from signaling basic or intermediate occupation-specific skills. Two mechanisms drive our findings. First, the increase in earnings for salaried workers comes from promotions within a firm. Second, the certificate facilitates transitions from self-employment to salaried work.

Keywords: Signaling, Labor demand, Wages

JEL Codes: J01, J31, J44

1. M. Antonella Mancino: Wilfrid Laurier University amancino@wlu.ca . Leonardo Fabio Morales: Banco de la República lmoralzu@banrep.gov.co. Diego F. Salazar: National University of Singapore diegosalazar.net@gmail.com. We want to thank Lance Lochner, Juan Morales, Xiaoxiao Li, Hong Long, Jungho Lee, Sunha Myong, Christine Ho, Siddharth George, and Louis-Pierre Lepage for helpful comments and suggestions, as well as seminar participants at Wilfrid Laurier University, Singapore Management University, National University of Singapore, Central Bank of Colombia, the 2022 LAMES-LACEA Annual Meeting, the 2023 Midwest Economics Association Annual Meeting, the 2023 RIDGE Forum Labor Workshop, the 2023 SOLE Annual Meeting, AMES 2023, ESAM 2023, and the 2023 EEA-ESEM Congress. For invaluable help with the data, we are grateful to William Orozco Daza and Mario Javier Rincón (SENA). For excellent research assistance, we thank Linda Nataly Corredor, Felipe Gaviria, and Simon Yarce. All errors are ours.

Señalizando la calidad de los trabajadores en un país en desarrollo: Lecciones de un programa de certificación [2]

M. Antonella Mancino, Leonardo Fabio Morales, Diego F. Salazar

Abstract

En este trabajo evaluamos los beneficios de señalar habilidades específicas de una ocupación, utilizando para dicho fin datos administrativos de un programa de certificación a nivel nacional en Colombia diseñado y ejecutado por el SENA. El programa certifica habilidades en tres niveles: básico, intermedio y avanzado. A través de métodos de regresión discontinua comparamos los ingresos de los trabajadores alrededor de los umbrales de asignación de certificados. Se muestra evidencia de que certificar habilidades avanzadas, genera retornos significativos en el ingreso de los trabajadores: 9,7% en promedio dentro de los dos años posteriores a la certificación. No encontramos efectos de certificar habilidades básicas o intermedias. Dos mecanismos explican estos hallazgos. En primer lugar, el aumento de los ingresos de los trabajadores asalariados proviene de ascensos dentro de las empresas. En segundo lugar, el certificado facilita la transición del trabajo por cuenta propia al trabajo asalariado.

Keywords: Señalización, Demanda Laboral, Salarios

JEL Codes: J01, J31, J44

---

2. M. Antonella Mancino: Wilfrid Laurier University amancino@wlu.ca . Leonardo Fabio Morales: Banco de la República lmoralzu@banrep.gov.co. Diego F. Salazar: National University of Singapore diegosalazar.net@gmail.com. Agradecemos a Lance Lochner, Juan Morales, Xiaoxiao Li, Hong Long, Jungho Lee, Sunha Myong, Christine Ho, Siddharth George, y Louis-Pierre Lepage por sus comentarios y sugerencias; tambien a los participantes del seminario de Laurier University, Singapore Management University, National University of Singapore, Banco de la República, 2022 LAMES-LACEA, 2023 Midwest Economics Association, 2023 RIDGE Forum Labor Workshop, 2023 SOLE, AMES 2023, ESAM 2023, y el congreso de 2023 EEA-ESEM. Agradecemos a William Orozco Daza and Mario Javier Rincón (SENA) por su ayuda con la información. Agradecemos por su asistencia de investigacióna a Linda Nataly Corredor, Felipe Gaviria, and Simon Yarce.

1 Introduction

One of the fundamental drivers of low productivity in developing countries is the mismatch between workers and firms.[3] This issue arises because there is limited information about the skills of potential workers, and local labor markets are unable to gather and transfer this information effectively (Chiplunkar and Banerjee 2023). It has been shown that, in developed countries, academic credentials, diplomas, and college reputation can help to mitigate these information challenges.[4] These devices provide job seekers with indicators of their skills and offer firms valuable tools for screening and comparing candidates. However, the same devices may not be equally effective in the developing world. First, workers usually lack formal education and training. Second, for certain industries and sectors, there might not be enough variation in schooling among workers to accurately infer productivity. Third, traditional measures of academic ability may not reflect productivity in occupations that require mainly skills accumulated on the job and that are not taught in formal academic institutions.[5]

Because of informational frictions and the imperfect role of education in forecasting occupation-specific productivity, there are reasons to think that policies aimed at providing accurate information about occupation-specific skills and productivity can generate significant efficiency gains. On the one hand, occupation-specific skills provide means for finding employment and long-run earnings growth for low-educated workers (Bandiera et al. 2020). Therefore, reducing information frictions likely has a positive impact on income and reduces income differentials among comparable workers.[6] On the other hand, such policies also have the potential to rectify the presence of employer information disparities regarding occupation-specific productivity and induce the efficient allocation of workers (Greenwald 1986; Pinkston 2009; Schönberg 2007).

3. See, for example, Bloom et al. (2010), Hall and Jones (1999), and McKenzie (2017).

4. See, for example, Altonji and Pierret (2001), Arcidiacono et al. (2010), Bedard (2001), and Clark and Martorell (2014).

5. Regarding the characteristics of the labor markets in developing countries, see, for example, Behrman (1999) and Rosenzweig (1988).

6. Recent literature has shown that providing credible information on workers' skills or recent job performance can largely improve workers' labor market outcomes (Abebe et al. 2021; Abel et al. 2020; Bassi and Nansamba 2022; Carranza et al. 2022).

In this paper, we identify the worker's returns to signaling occupation-specific skills in the context of a developing country. To estimate these returns within a diverse population of workers, we employ a sharp regression discontinuity design using unique administrative data. Our study focuses on a particular signaling device: a certificate issued by the Colombian National Training Service (SENA), a governmental organization responsible for evaluating and certifying workers' skills in Colombia. Starting in 2004, SENA implemented a rigorous evaluation procedure to assess occupation-specific skills acquired as a byproduct of work experience.[7] Program participants are assigned to one of four mutually exclusive and exhaustive categories on the basis on their performance on the certification exam: no certificate, basic, intermediate, or advanced. Each category is defined by sharp thresholds corresponding to different exam-score intervals.

In an ideal scenario, where certificates are randomly assigned, the wage variation across these predefined categories would capture the signaling value of obtaining a certificate. However, wage differentials not only reflect the signaling value but also encompass productivity disparities among workers and match-quality factors pertaining to the interaction between workers and firms. To address potential confounders, we leverage the sharp differences in scores required to obtain a particular certificate as a proxy for random assignment. We first focus on individuals who barely pass the exam and those who barely fail, assuming they are similar in all other dimensions that matter for productivity. Under certain conditions, passing status constitutes a valid approximation for random assignment for individuals with scores close to the passing cutoff (Cattaneo et al. 2020; Cattaneo and Titiunik 2022; Lee and Lemieux 2010). We use this insight to estimate the unbiased signaling returns of the basic certificate by comparing the earnings of the two groups within the two years after certification. We then extend this approach to estimate the returns of obtaining either an intermediate or an advanced certificate, which allows us to explore the distributional effects associated with the content of the signal. Our unique dataset also provides comprehensive information regarding employment status (i.e.,

7. The certificates under analysis are not legally required to practice the corresponding occupations.

salaried work, self-employed, unemployed) and the transitions between employers. We use this information to identify the mechanisms driving income growth and income differentials in response to obtaining the certificate.

Several features of our study set it apart from previous literature. First, we estimate the returns to signaling occupation-specific skills, whereas previous literature has examined mostly the effects of signaling academic aptitude as a proxy for productivity.[8] There are strong arguments suggesting occupation-specific skills are valuable and are key to explaining post-schooling wage growth. Nonetheless, limited attention has been devoted to analyzing the consequences of providing reliable information about them.[9] Some exceptions can be found in the emerging literature studying the returns to occupation certification, predominantly in the context of developed economies, with estimates that span from no effect to a 13% increase in income.[10] However, for the most part, this literature has used imprecise (self-reported) certification measures, and only a few studies have national-level data, preventing them from examining the effects across various segments of the economy. In addition, most papers rely on observable characteristics to estimate the effects of certification, and therefore, it is unclear whether self-selection into certification leads to biased estimates.

Our study design is able to isolate the returns to signaling occupation-

8. Relevant papers in this literature are Bedard (2001), Clark and Martorell (2014), Macleod et al. (2017), and Tyler et al. (2000). As we mentioned before, more recently, a few studies have focused on the value of signaling noncognitive skills (Bassi and Nansamba 2022), general skills at the hiring stage (Abel et al. 2020; Carranza et al. 2022), or field-specific skills (Busso et al. 2023)

9. Regarding the theoretical importance of task-specific skills, or more generally occupation-specific skills, see, for example, Becker (2009) and Gibbons and Waldman (1999, 2004). For empirical papers discussing the value of occupation-specific skills, see, for example, Kambourov and Manovskii (2009), Neal (1995), Parent (2000), Poletaev and Robinson (2008), and Sanders and Taber (2012). Finally, regarding the importance of skills for post-schooling wage growth, see Rubinstein and Weiss (2006) and Sanders and Taber (2012).

10. This literature further concludes that the effects of licensing on labor market outcomes are larger than the effects of certification. Unlike certification, occupational licensing mandates that individuals can engage in a particular occupation only if they satisfy specific predetermined criteria for competence. See, for example, Albert (2017), Kleiner and Krueger (2013), Kleiner and Vorotnikov (2017), and Xia (2021). In addition, the literature finds that the effects of certification are larger among less-educated individuals (Baird et al. 2021), suggesting the potential role of certification as a signal of skills in situations when general measures of human capital are not available.

specific skills from confounding channels, such as education and schooling. In our analysis, individuals are primarily full-time workers who have completed their formal education. The skills they signal are acquired mainly on the job, as a byproduct of experience rather than through formal education. In addition, the program tests skills without facilitating investment in human capital through lectures or training.[11] Importantly, we provide evidence spanning a wide group of occupations from a national sample of program participants. Therefore, our conclusions are not restricted to particular firms or sectors.

A second contribution to the literature is the possibility of studying the effects on experienced workers. In contrast, most of the literature evaluates the returns to signaling skills when workers first enter the labor market.[12] The effects on experienced workers have been consistently understudied, and signaling can provide this population group with significant benefits, particularly in contexts where productivity evolves heterogeneously (Kahn and Lange 2014). In fact, our findings suggest that experienced workers capture significant returns. In such a way, we provide novel evidence about the importance of information frictions on post-schooling income growth.

Lastly, in the literature, limited attention has been given to signals that convey information about different skill levels. Previous studies have focused mainly on dichotomous signals or whether the worker possesses an occupational certificate at all.[13] We are among the first to provide direct evidence of the distributional effects of displaying signals with different

11. It is possible that participants in the certification program are still investing in human capital by studying for the exam. Our empirical methodology allows us to deal credibly with this issue.

12. Regarding the value of academic credentials, see, for example, Arcidiacono et al. (2010), Bedard (2001), Clark and Martorell (2014), Jepsen et al. (2016), Machin et al. (2020), and Tyler et al. (2000). As we have mentioned, more recent literature has focused on the value of signaling general skills Abel et al. (2020), noncognitive skills Bassi and Nansamba (2022), and field-specific skills Busso et al. (2023). In all cases, these papers focus on young workers entering the labor market.

13. Examples of papers analyzing dichotomous signals are Clark and Martorell (2014), Jepsen et al. (2016), Machin et al. (2020), and Tyler et al. (2000). The literature studying the economic importance of certificates is relatively new. The following papers provide important contributions in this area: Albert (2017), Kleiner and Krueger (2013), and Kleiner and Vorotnikov (2017).

content.[14] Our context enables us to do so because the signaling device categorizes workers' skills into four mutually exclusive and exhaustive categories, each defined by a sharp threshold.

Our results reveal no effects from signaling basic or intermediate occupation-specific skills. Notably, the absence of positive returns remains consistent across education levels, tenure, and firm sizes. The evidence suggests that obtaining basic or intermediate certificates does not enhance the chances of workers transitioning to other firms or achieving a different employment status. Specifically, we find no evidence of job-to-job transitions, leaving self-employment, or moving from salaried work to other types of employment. We argue that two explanations (not necessarily mutually exclusive) are compatible with such a pattern. First, it is possible that the basic and intermediate certificates do not reveal new information to current employers. Second, signaling basic or intermediate skills may not be valued by potential employers. Because employers are unwilling to pay more for average workers, transitions are limited, and those with the certificate at hand capture no returns.

Instead, signaling advanced occupation-specific skills yields large and significant returns. Our estimates show that obtaining an advanced certificate generates an average increase in income of 9.7% during the two years following certification. Importantly, we find that returns depend on the employment status at the time of certification. Salaried workers capture the returns in the first quarter after certification, and they remain constant during the next two years. Such effects are larger for more-educated workers. We further find that raises for salaried workers typically come from within the firm rather than from switching employers and are concentrated among tenured workers working at large firms. In such a way, the certificate serves as a screening device that facilitates promotions and justifies income increases within organizations.

The situation is different for self-employed participants with advanced certificates. They capture significant returns but only after nine months of

14. Another paper that provides an analysis of signals with different contents is Bassi and Nansamba (2022).

certification, coinciding with transitions to salaried work. In this case, the certificate provides a valuable signal to new employers, facilitating transitions and generating an average increase in income of 15% the second year after certification. Instead, workers who are unemployed at the time of certification do not capture any positive returns. For this group, the negative signal from unemployment seems to overwrite important information about occupation-specific skills. To rationalize the basic transition patterns, in the last part of the paper, we present a stylized search theoretic model with asymmetric information.

Overall, our results suggest that alleviating information frictions benefits only a subgroup of workers and that other frictions in the labor market make wage adjustments unlikely, even after relevant information has been revealed to current and potential employers. In the end, certifying skills may have a more limited role in earnings growth and job transitions than previous findings suggest (Bandiera et al. 2020). Additionally, our findings cast doubts about the ability of wages to aggregate information about productivity in developing countries.

The rest of the paper is structured as follows. In Section 2, we provide a detailed description of the program's key features. In Section 3, we describe the sample of program participants and elaborate on the procedure for obtaining information on labor market outcomes at and after applying for the certificate. We outline the empirical strategy employed to estimate the causal effects of obtaining a certificate in Section 4. In Section 5, we present the core findings, and in Section 6, we examine the heterogeneity of returns. In Section 7, we delve into the mechanisms behind our main results. Finally, we conclude in Section 8.

2 Program Description

Since 2004, the Colombian National Training Service (SENA), a government agency in Colombia, has been responsible for implementing a nationwide certification program.[15] In Colombia, technical norms define the tasks and activities specific to different occupations and the quality standards governing the production and provision of goods and services within those occupations. These norms are drafted by industry skill councils and are approved by the government. Based on the criteria defined by such norms, the certification program aims to assess and certify the knowledge and skills that workers acquire through their work experience. In doing so, the program is tailored to occupations where knowledge is mostly acquired outside of formal education institutions, indirectly targeting lower-educated individuals.

From its inception, the Colombian government has recognized this program as a pivotal policy to enhance firms' productivity and bolster their competitiveness. During the past decade, the program has gone through a significant expansion resulting in SENA being entrusted with continually expanding participation, free of charge, across the country. To date, SENA issues certificates in 912 technical norms, which have been developed by 74 industry skill councils representing the major sectors of the economy.[16] During 2019, SENA certified approximately 243,000 workers across 117 different locations.[17]

The policy's underlying objectives are to reduce the costs associated with identifying productive workers, streamline personnel selection processes, and minimize potential mismatches between firms and workers. Furthermore, the policy provides workers with means, i.e., certificates, to publicly demonstrate their skills and knowledge, facilitating transitions to higher-paying jobs and reducing unemployment among participants.

15. The two primary legal dispositions governing this program are Decreto 933/2003 and Decreto 4108/2011.

16. The six most popular technical norms, accounting for 24% of certifications between 2017 and 2019, were: serving customers following service procedures, handling food in compliance with current regulations, controlling access to restricted areas based on service characteristics and regulations, operating forklifts following technical manuals, and promoting safe and healthy practices in work environments.

17. In general, candidates can take the exam for all norms in any municipality. The primary constraint is the waiting time, as they may have to wait for an instructor from a different location to visit and assist with the practical test. However, SENA does not consider this issue a significant barrier for participants to attain certification.

Participants enter the program hoping to be certified in a particular technical norm. To obtain the certificate, they must provide evidence of proficiency in executing the task and work activities defined by such a norm, as well as their knowledge of the quality standards. As a result, the certificate contains valuable information about the skills and knowledge required to perform a specific occupation. We refer to such skills and knowledge as occupation-specific skills.[18] To initiate the certification process, participants must demonstrate at least six months of experience in a given occupation. Typically, the process is completed within four weeks.

To obtain the certification, individuals must take a two-part exam. The first part, known as the competence exam, entails participants performing a series of tasks and work activities under the observation of a panel of SENA officials. Evaluators assign a pass/fail grade based on the participants' performance. Since all participants must showcase relevant work experience to start the certification process, most of them successfully pass this stage (see Table 1). The second part of the examination is the knowledge test, when participants undergo a multiple-choice exam. The exam evaluates participants' understanding of the various concepts related to the occupation and the quality standards prescribed by the technical norm. SENA administers the exam, which is designed using a randomly selected set of predefined questions.[19] Therefore, the exam's difficulty is constant across participants. The exam is graded by a computer on a scale of 0 to 100 points. Using computerized grading ensures that SENA officials cannot manipulate the results. The score determines the level of certification conferred by SENA. Individuals who score below 30 points do not get a certificate, even if they pass the first part. Participants scoring between

18. For instance, the technical norms pertaining to plumbing primarily outline the tasks and work activities related to installing and repairing piping fixtures and systems. These norms also define proper network installation, functionality, and durability standards. As a result, the certification program evaluates if individuals can perform occupation-specific activities efficiently while producing outcomes of higher quality and durability.

19. While the technical norm itself is publicly available, the question bank is not accessible. According to SENA, the question bank contains more than 100,000 questions for all the different technical norms. This question bank is constantly enhanced and revised. It is important to add that the number of questions in the competence exam ranges from 18 to 44. The specific number of questions for a particular norm depends on the number of tasks and activities described by the norm.

30 and 59.9 points receive a basic certificate, while those scoring between 60 and 89.9 points are granted an intermediate certificate. Lastly, participants who score 90 or higher obtain an advanced certificate. Participants are only informed of the certification level attained, and the exact grade remains confidential.

According to SENA's guidelines, participants can improve their certification level by undertaking the second part of the examination in the subsequent fiscal year, leading to infrequent exam retakes among individuals. In fact, our data show that only 0.5% of participants retake the knowledge test. Furthermore, the lack of information about how close participants are to the certification cutoff likely discourages retaking, even among those who scored just below a given cutoff.

A typical certificate (see Figure OA1 in online Appendix A) includes the participant's name, identifier number, the certificate level awarded, and the specific technical norm for which the participant has been certified. Employers can access the certificate information by entering the workers' identification number into SENA's web portal. Therefore, SENA provides the same information to workers and employers, making it unlikely that participants hide the certificate when unsatisfied with the outcome.

In general, no legal barriers prevent workers from continuing their current occupation, even if they do not obtain certification in the relevant technical norm. However, an exception exists for specific technical norms that apply to regulated occupations. Regulated occupations involve situations where workers face exceptional hazards, or where failure to comply with the technical standards could lead to unacceptable risks for consumers and workers.[20] For such occupations, certification is mandatory and granted only upon achieving a minimum score of 90 points on the knowledge test. Hence, workers who score below 90 are not certified and should not en-

20. Some examples of technical norms regarding occupations involving situations where workers face exceptional hazards include those related to tasks performed at elevated heights or the evaluation of equipment utilizing natural gas as an energy source. Examples of activities in which failure to comply with the technical standards could lead to unacceptable risks include installing and maintaining home networks for natural gas distribution and water purification procedures.

gage in such occupations. Within our dataset, technical norms related to regulated occupations comprise less than 10% of the observations, and we choose to exclude them from the analysis. This exclusion is mainly motivated by the fact that the 90-point threshold leads to distinct and incomparable outcomes for technical norms associated with regulated occupations and those associated with non-regulated occupations. While participants scoring above 90 receive an advanced certificate, regardless of the underlying occupation, participants who score below this threshold would not obtain a certificate in the case of technical norms associated with regulated occupations, and they would receive an intermediate certificate in the case of technical norms associated with non-regulated occupations.

Our analysis of the certification program relies on administrative data provided by SENA, which covers all participants seeking certification in technical norms linked to non-regulated occupations from January 2017 to December 2019. The causal analysis leverages the discontinuity observed in the certification levels (i.e., basic, intermediate, advanced) that arise from variations in underlying scores near the three respective cutoffs (i.e., 30, 60, and 90). SENA has meticulously documented participants' demographic characteristics and the specific technical norms for which they sought certification. Over the course of this period, the program issued approximately 627,000 distinct certificates to more than 470,000 participants.[21] It is important to highlight that the institutional context remained consistent throughout the study period. Factors such as the number of evaluators, program coverage, and exam format remained unchanged since 2017. Section 3.1 provides additional information regarding SENA's data.

3 Data

Our analysis relies on two sources of information. First, we use SENA's novel administrative data with information about all participants in the certification program between 2017 and 2019. Second, given that our primary data on program participants do not contain information on labor market

21. The number of certificates exceeds the number of participants due to individuals being eligible for certification in multiple technical norms.

outcomes after certification, we use administrative records from contributions to the social security system to obtain income and employment status information.

3.1 SENA

We obtained data from SENA on all individuals who started the certification process between January 2017 and December 2019. The data contain information on the technical norm individuals applied to be certified on, the scores on the two-part exam, the test date, employment status at the time of certification, and socio-demographic information, such as educational achievement, age, and geographic location. In total, the data set contains information on 627,340 applications for certification. Table 1 presents descriptive statistics for the complete dataset.

Individuals in the sample are predominantly male. In our analysis, we focus on the sample of men for two main reasons. First, responses from women may be obscured by changes in home production activities that are not accounted for in our data. Second, the sample of men is more homogeneous and performs better on the balanced tests for different covariates and thresholds (see Section 4.2), proving better support to our empirical strategy. However, as discussed in Section 5.3, estimates using the full sample of men and women are compatible with our core results. As shown in the second column of Table 1, men in the sample have, on average, 38.5 years old. Consistent with the type of occupations targeted by the certification program, there is a large share of low-educated individuals: the proportion with some college education is only 32%.

As previously mentioned, the first part of the exam is generally considered a pass for nearly all participants and therefore does not actually determine their eligibility for the certificate. In fact, the pass rate for the sample of men stands at 98.9%. The mean score for the second part of the exam among the sample of men is 82 points, indicating that, on average, participants obtain an intermediate certificate. It is worth noting that the objective of the second part is to assess participants' comprehension of concepts and standards outlined in the technical norm. Consequently, despite participants' familiarity with the tasks and work activities due to their work experience, there exists variability regarding the knowledge required to execute these tasks in accordance with the quality standards. For instance, while all plumbers likely possess the capability to install and repair drinking water and drainage networks, only the most skilled individuals are familiar with the design principles of the network that allows them to determine the minimum pipe diameter required for the different devices in the restroom.[22] Given that such knowledge enhances the production and quality of goods and services associated with specific occupations, we interpret the results from the second part of the exam as indicative of participants' occupation-specific skills.

3.2 PILA and Estimation Sample

To obtain the labor market histories, we use employer-employee-linked administrative data from the Unified Social Security Contributions Form (PILA). By law, all workers and firms in the formal sector must report to PILA their contributions to the social security system. PILA provides monthly information on wages, payroll-tax payments, employment type (salaried work or self-employment), and firm and job characteristics. We also observe workers' transitions between employers and in and out of PILA. However, we lack information on individuals working in the informal sector.

We can only match personal identifier numbers between PILA and SENA data for the subsample of program participants who reported to PILA at any point during 2010.[23] This implies that for individuals who did not report to PILA in 2010, we cannot observe their labor market outcomes at any point in time. In the third column of Table 1, we report descriptive statistics for the matched sample of men. Our estimation sample is the matched sample of men. Section 5.3 presents additional results for the matched sample of men and women.

22. For example, according to the technical norm, a Toilet with a flush tank requires a minimum pipe diameter of 1/2 inch, while a toilet with a Toilet flush valve requires 1 inch, and a urinal with a flushometer requires 3/4 inch.

23. Law 1581 of 2012 is the general legal framework applicable to managing and protecting personal data. Because of the restrictions imposed by the law, individual identification numbers were part of PILA only in 2010.

The matched sample contains 39% of individuals from the entire sample of men. There are significant differences between the matched and unmatched samples.[24] Nevertheless, the magnitude of such differences is, in most cases, subtle. First, the matched sample is older than the unmatched one. This difference is not surprising since matched individuals reported to PILA in 2010. Hence, younger individuals, who are less likely to have worked in 2010 (seven to nine years before certification), are less likely to be matched. Second, unemployment is less prevalent in the matched sample. This fact is also expected since the matched sample contains individuals already employed at a younger age. Third, individuals are equally likely to obtain a basic, intermediate or advanced certificate. Importantly, all 912 technical norms are present in the matched sample.[25]

We aggregate PILA information at the quarterly level in the following way. First, our measure of income is the average monthly reported income. Second, we classify an individual as employed if he appeared in PILA at least one month in the quarter. Third, if an individual does not report to PILA in any month during the quarter, we classify him as not being employed. Given our data on labor market outcomes, we cannot distinguish unemployment from employment in the informal sector, in which reporting to PILA is not mandatory. Nevertheless, taking advantage of the self-reported data on employment in SENA and the employment data on PILA at the time of certification, we can infer the relevance of the informal sector in our sample of participants. While the measure of employment in SENA data is likely more comprehensive than the one from PILA, the employment rate in both samples is remarkably similar. In both data sets, the overall employment rate at the time of certification is around 91%, suggesting that the informal sector is not as relevant for our sample as it may be in the general population.

Each month, workers must classify their occupational status using PILA's

24. Descriptive statistics for the matched and unmatched sample of men are displayed in Table OA1 in online Appendix A.

25. Table OA2 in online Appendix A displays the top 10 technical norms in the matched and unmatched samples. There is a fair degree of overlap in the more prevalent norms between the two samples, with only four norms in the top 10 for the matched sample and not appearing in the top 10 for the unmatched sample.

categories. We classify individuals as salaried workers if they are categorized as dependents or any other category in which their employers make contributions to the social security system on their behalf. Conversely, individuals are classified as self-employed if they report being independent workers or belonging to any other category in which they have to pay their entire contribution to the social security system.

Workers are allowed to be certified in multiple norms and 23% of them have more than one certification. In our preferred specification, we focus on the returns to the first certificate. As we show in Section 5.3, our conclusions remain if we restrict the sample to include those with only one certificate between 2017 and 2019. Lastly, we look at outcomes up to two years after certification, when the certificates remain valid.[26]

In our panel, an observation is a worker-quarter pair. The last panel of Table 1 reports summary statistics on labor market outcomes for the estimation sample within two years of certification. The overall employment rate in this two-year period is 86%, with 77% of workers being salaried. The monthly average income is 1,153 thousand Colombian pesos (approximately USD 427 in 2018 dollars).[27]

26. SENA certificates are valid for three years. However, we choose to look at outcomes up to two years out because we lack data beyond the third quarter of 2021 and because we want to get a balanced sample of individuals applying for a certificate between 2017 and 2019.

27. The monthly average income in the estimation sample is above the average minimum wage of 826 thousand Colombian pesos between 2017 and 2021.

Table 1: Descriptive Statistics

	SENA data		Estimation Sample
	Full Sample	Men Only	(Men)
	A. Demographic Characteristics (SENA)		
Demographic Characteristics (Mean)			
Male	0.70		
Age	38.24	38.53	41.99
Less Than High School	0.19	0.22	0.20
High School	0.41	0.46	0.46
Some College	0.37	0.30	0.30
More Than College	0.04	0.02	0.04
Employment Status (Mean)			
Salaried Worker	0.78	0.80	0.87
Self-Employed	0.05	0.05	0.05
Unemployed	0.17	0.15	0.09
	B. Certification Program (SENA)		
Skills Certified			
Technical Norms	912	912	912
Industry Skill Councils	74	74	74
Certification Level (Mean)			
No Certification	0.01	0.01	0.01
Basic	0.13	0.13	0.13
Intermediate	0.40	0.39	0.39
Advanced	0.47	0.48	0.47
Certification Two-Part Exam (Mean)			
Knowledge	81.97	82.11	82.02
Competence	99.15	98.91	99.02
Individuals	627,340	429,272	181,395
	C. Post Certification Labor Market Outcomes (PILA)		
Employment Status (Mean)			
Salaried Work	...	...	0.77
Self-Employment	...	...	0.09
Income (Mean)	...	...	
Income	...	...	1,153,149
Ln of Income - Salaried Worker	...	...	13.97
Ln of Income - Self-Employed	...	...	13.90
Observations	...	...	1,434,061

Notes: This table reports descriptive statistics for the full sample of men and women applying to get SENA certificates between 2017 and 2019 (first column), the restricted sample of men (second column), and the matched (estimation) sample of men. The estimation sample corresponds to the subsample of workers we could match with PILA. The first panel reports demographic characteristics and employment information, calculated using SENA data only. The second panel reports information regarding the certification program. The last panel shows descriptive statistics two years after certification, using PILA data. The income variable contains zeros in periods when individuals are not salaried workers or self-employed.

4 Empirical Strategy

4.1 Research Design

This section describes the empirical strategy used to estimate the causal effect on income of obtaining a basic, intermediate, or advanced certificate. Given the nature of the SENA certification program, we use a sharp regression discontinuity (RD) design (Cattaneo et al. 2020; Lee and Lemieux 2010). In a typical RD design, all units receive a score, and the treatment is assigned to units with a score above a known cutoff. The key feature of the RD design is that, given the score, the probability of receiving treatment changes discontinuously at the cutoff. As long as units cannot sort around the known cutoff, which can be verified empirically, the abrupt change in the probability of receiving treatment is as good as random. Therefore, it can be used to learn about the local causal effect of the treatment.

Let $T_{it}^{c} = \mathbf{1}(score_{it} > c)$ be an indicator variable that takes the value of 1 if individual i, taking the exam in year t, obtains a certification score, $score_{it}$, above the threshold c. As noted in Section 3.1, we consider three thresholds: 30, 60, and 90, which correspond to obtaining a basic, intermediate, or advanced certificate, respectively. The standard local linear estimator of the RD treatment is implemented by running the following weighted least squares regression:

$$Y_{is} = \alpha + \beta score_{it} + \delta_{RD}^{c} T_{it}^{c} + \tau score_{it} \times T_{it}^{c} + \gamma Z_i' + \varepsilon_{is}, \qquad (1)$$

where, Z_i are predetermined covariates and Y_{is} represents the labor market outcome of interest $s > 0$ quarters after certification. Equation (1) is estimated with only individuals with scores within a chosen bandwidth h, such that $score_{it} \in [c-h, c+h]$, and with weights applied according to some kernel function. The main parameter of interest, δ_{RD}^{c}, is estimated as

$$\delta^c_{RD} = \lim_{score_{it} \downarrow c} E[Y_{is}|score_{it}, Z'_i] - \lim_{score_{it} \uparrow c} E[Y_{is}|score_{it}, Z'_i]. \qquad (2)$$

Our primary outcome of interest is the natural logarithm of income, which includes earnings from salaried work and self-employment.[28] It is important to note that PILA data does not capture earnings in the informal sector. Thus, our findings should be interpreted within the context of returns to the certificate in the formal sector.

As mentioned in Section 3.2, we look at outcomes up to eight quarters after certification (that is, $s \in [1, 8]$). The predetermined covariates, Z_i, include age and education dummies. We also include industry skill councils' fixed effects and year-of-certification fixed effects.[29] Therefore, we estimate the returns to the certificate by exploiting variation within age, education, year of certification, and industry groups. Following Cattaneo et al. (2020), we use a triangular kernel, a first-order polynomial, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator.[30] We report RD point estimators with robust bias-corrected confidence intervals (Cattaneo et al. 2020). Lastly, standard errors are clustered at the technical-norm level to adjust for the correlations induced by industry- and occupation-specific unobserved components.

28. By employing log income as the outcome measure, we exclude individuals reporting zero earnings. This approach helps to avoid potential confounding effects arising from the impact of signaling on employment. For example, if signaling negatively affects overall employment, we would observe an increase in zero-income cases, potentially attenuating the estimated effect on income. A detailed discussion of the effects on employment can be found in Section 7.

29. Ideally, we would like to include technical norm fixed effects. However, given the large number of technical norms (912), we instead choose to include industry skill councils' fixed effects (74), which can be regarded as industry fixed effects. In addition, while it would be interesting to add firm fixed effects, which would allow us to estimate the returns to the certificate within the firm, there is not enough variation in the sample to perform such an exercise. As a matter of fact, the median proportion of workers participating in the certification program between 2017 and 2019 is 4% of firms' total workers.

30. In Section 5.3, we show that our results are robust to the inclusion of additional controls (year and location fixed effects or no controls at all), using alternative methodologies to choose the bandwidth, using fixed bandwidths, using non-bias-corrected RD estimates, and not adjusting for the presence of mass points during estimation.

4.2 Validity of the Design

The main threat to identification in an RD design is the possibility that program participants actively manipulate their score around the threshold of interest, implying that individuals just above the threshold are systematically different from individuals just below it. To mitigate this concern, we perform two falsification tests that support the validity of the RD design. First, we examine the density of the running variable, $score_{it}$, around each threshold. Second, we investigate whether treated individuals are similar around each threshold. The intuition for these two falsification tests is that if individuals cannot manipulate their score, the number of observations just above the threshold should be similar to the number of observations just below the threshold and there should be no systematic differences across groups.

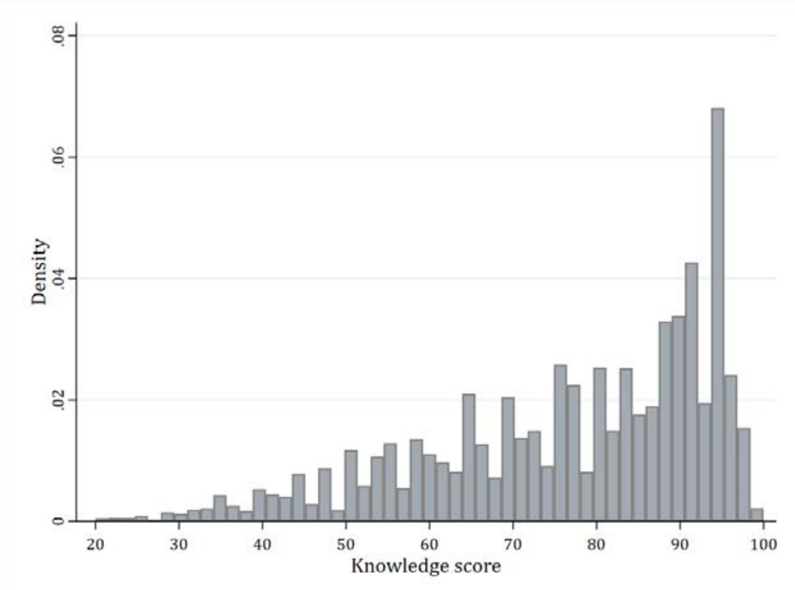
Manipulation of the score seems unlikely in our context for several reasons. As discussed in Section 2, the test format is multiple choice, and the grading is performed (by a computer) in a location different from the testing location, reducing the chances of manipulation. Furthermore, the underlying score is not revealed to participants or employers, who only get to see the certification level. Figure 1 displays the distribution of the scores.[31] Visual inspection of the histogram shows no apparent discontinuities in the scores around the thresholds of interest: 30, 60, and 90. We formally test for the presence of manipulation of the score around each threshold using the test proposed by Frandsen (2017). Frandsen (2017) is the adequate manipulation test in our context since the running variable has mass points at integer values. The critical parameter in the manipulation test is k, which dictates the maximum degree of nonlinearity in the probability mass function that is still considered compatible with the absence of manipulation.[32] The results of the manipulation test around all three thresholds lead us to

31. In the histogram in Figure 1, we exclude the highest score, 100, since it represents a significant mass point, denying a straightforward exploration of continuity. In Figure OA2 in online Appendix A, we display the complete histogram.

32. A smaller k means even tiny deviations from linearity will lead the test to reject the null of no manipulation with high probability (Frandsen 2017). We choose k using the entire distribution of the running variable, not just around the thresholds. Given our sample, the maximum suggested value for k is 0.001.

fail to reject the null hypothesis of absence of manipulation (p-value = 1).

Figure 1: Distribution of Scores



Notes: This figure displays the distribution of scores in the second part of the certification exam (knowledge test) for the matched sample of men. The histogram excludes the highest score, 100, since it represents a significant mass point, denying a straightforward exploration of continuity. In Figure OA2 in online Appendix A, we display the complete histogram.

To perform the falsification analysis on predetermined characteristics, we estimate Equation (1), using each characteristic as the outcome variable. We analyze the set of predetermined covariates used in the primary analysis, Z_i, and income at the time of certification (that is, $s = 0$). The results are presented in Table 2. Our analysis shows that, at the moment of certification, individuals just below the threshold for receiving a basic certificate are not statistically different from individuals just above the threshold in terms of age or schooling. In addition, there are no observed differences regarding the reported income at the time of certification. We reach the same conclusion for individuals around the intermediate and advanced thresholds. In all, our tests show a smooth evolution through the different thresholds, confirming that participants just above and below the respective cutoffs are very similar.

Table 2: Covariate Balance Check

	(1) Threshold 30	(2) Threshold 60	(3) Threshold 90
Age	1.204	-0.142	-0.105
	(0.959)	(0.342)	(0.734)
	[48.255]	[45.369]	[44.739]
High School	0.052	-0.049	-0.050
	(0.054)	(0.019)	(0.032)
	[0.475]	[0.506]	[0.483]
Some College	-0.014	-0.001	-0.031
	(0.061)	(0.023)	(0.056)
	[0.200]	[0.265]	[0.304]
More Than College	0.017	0.006	0.001
	(0.007)	(0.004)	(0.014)
	[0.015]	[0.035]	[0.042]
Income at Certification (1000s)	36.989	-31.001	91.652
	(50.301)	(29.250)	(79.819)
	[863.888]	[1,024.739]	[1,121.895]
Number of Observations	181,395	181,395	181,395
Effective # of Control Observations	774	11,141	8,819
Effective # of Treatment Observations	2,894	17,158	19,950
Bandwidth	8.600	9.900	3.300

Notes: Standard errors are reported below the point estimates in parentheses. Standard errors are clustered at the technical norm level. The sample mean for the control group is displayed below the standard error in squared brackets. For each threshold, the analysis uses a fixed bandwidth that is the average of the optimal bandwidths in Table 3. Bandwidths are displayed below the effective number of observations.

5 Results

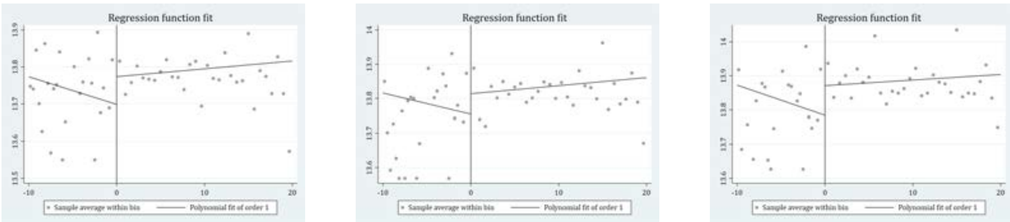
In this section, we present the core results. We use Equation (1) to estimate the returns from obtaining a given certification level. Our primary outcome of interest is the natural logarithm of income. We begin by examining the returns from the basic and intermediate certificates, separately. Subsequently, we present the estimates for the advanced certificate. These results allow us to directly investigate the distributional effects associated with the content of the signal.

5.1 Effects of Obtaining a Basic or Intermediate Certificate

With regard to the basic certificate, the first panel of Table 3 displays the results on the log of income for quarters one to eight after certification.

Figure 2 presents a visual depiction of the effects on log income one, four, and eight quarters after certification. Following Equation (2), the effect of obtaining a basic certificate is measured by the discontinuity observed between individuals who score just below 30 points and those who score just above 30 points. In general, our estimates reveal no discernible effect on income within the first two years following certification. Specifically, for individuals with a basic certificate, income remains unchanged even after eight quarters, relative to marginal individuals without a certificate.

Figure 2: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate on Log of Income



(a) Quarter One (b) Quarter Four (c) Quarter Eight

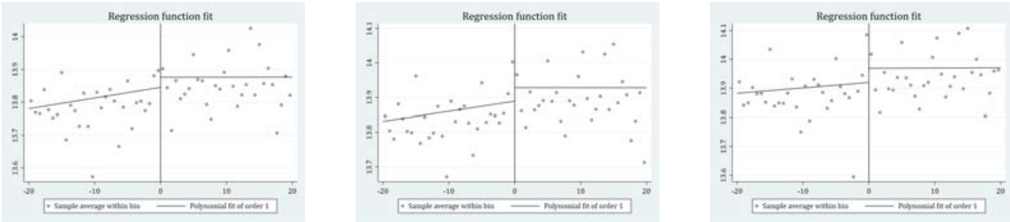
Notes: The three figures summarize the estimated results of Equation (1), one, four, and eight quarters after certification, using the log of income as the main outcome. The running variable is the exam score, and the discontinuity threshold is 30. The regressions include the controls described in Section 4.

Turning to the effects of acquiring an intermediate certificate, the second panel of Table 3 presents the results on the log of income one to eight quarters after certification. The effect is measured by the discontinuity observed between individuals who score just below 60 points and those who score just above 60 points. Figure 3 presents a visual depiction of the effects one, four, and eight quarters after certification. Our analysis indicates no effect on income for individuals obtaining an intermediate certificate relative to marginal individuals who obtain a basic certificate.

Two plausible explanations can account for the absence of returns from obtaining either a basic or an intermediate certificate. First, it is possible that the basic and intermediate certificates do not provide new information about workers' productivity, thus leading to minimal revisions in employers' priors and, consequently, earnings.[33] Second, the basic or intermediate cer-

33. One may be worried that employers pay lower wages when the certificate is not

Figure 3: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate on Log of Income



(a) Quarter One (b) Quarter Four (c) Quarter Eight

Notes: The three figures summarize the estimated results of Equation (1), one, four, and eight quarters after certification, using the log of income as the main outcome. The running variable is the exam score, and the discontinuity threshold is 60. The regressions include the controls described in Section 4.

tificates may not significantly enhance workers' prospects for transitioning to other firms (for salaried workers, unemployed, or self-employed individuals). Potential employers likely do not want to incur the costs of poaching average workers and, consequently, there is no incentive for current employers to increase wages in response.

as expected (e.g., basic certificate). However, this is likely not the case given that, in the Colombian labor market, wages are characterized as downward rigid (Agudelo and Sala 2017). Likewise, it is unlikely that workers would be fired for not getting a given certification level, as such a decision would violate the principles of cause termination. Additionally, as discussed in Section 7, analysis of the transitions does not reveal evidence compatible with workers being fired for performing below expectations in the certification exam.

Table 3: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic, Intermediate, and Advanced Certificate on Log of Income

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
Basic Certificate	0.019	0.028	0.002	-0.009	0.027	0.049	0.039	0.062
	(0.030)	(0.037)	(0.037)	(0.040)	(0.034)	(0.043)	(0.031)	(0.033)
# of Observations	164,242	162,247	160,912	159,819	158,344	156,895	155,954	140,679
Eff. # of Control Obs.	756	633	698	713	617	680	611	641
Eff. # of Treatment Obs.	3,617	2,237	2,743	3,510	2,188	2,648	2,139	2,182
Bandwidth	10.14	7.738	9.994	10.47	7.661	9.038	7.203	8.306
Mean	13.726	13.738	13.751	13.777	13.775	13.777	13.805	13.814
Intermediate Certificate	-0.010	-0.002	-0.001	-0.013	-0.006	0.009	0.017	0.008
	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.018)	(0.018)
# of Observations	164,242	162,247	160,912	159,819	158,344	156,895	155,954	140,679
Eff. # of Control Obs.	9,611	9,921	9,808	9,319	11,296	11,248	11,393	10,147
Eff. # of Treatment Obs.	13,679	15,196	14,987	13,323	17,176	17,302	17,182	15,226
Bandwidth	8.798	9.714	9.357	8.753	10.30	10.79	11.00	10.41
Mean	13.829	13.840	13.863	13.873	13.878	13.892	13.909	13.908
Advanced Certificate	0.087	0.087	0.082	0.127	0.104	0.094	0.099	0.099
	(0.028)	(0.031)	(0.034)	(0.029)	(0.029)	(0.027)	(0.032)	(0.032)
# of Observations	164,242	162,247	160,912	159,819	158,344	156,895	155,954	140,679
Eff. # of Control Obs.	8,322	8,009	6,843	7,986	7,901	10,235	7,619	6,929
Eff. # of Treatment Obs.	18,098	17,852	14,410	17,669	17,400	23,044	16,776	14,754
Bandwidth	3.955	3.548	2.981	3.684	3.625	4.113	3.082	3.013
Mean	13.885	13.901	13.913	13.927	13.937	13.956	13.970	13.971

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification, for each discontinuity threshold. The outcome is the log of income. The running variable is the exam score, and the three discontinuity thresholds are 30 (first panel), 60 (second panel), and 90 (third panel). All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. The total and effective number of observations change across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019. Robust bias-corrected standard errors are reported below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. The sample mean for the control group is displayed below the optimal bandwidth.

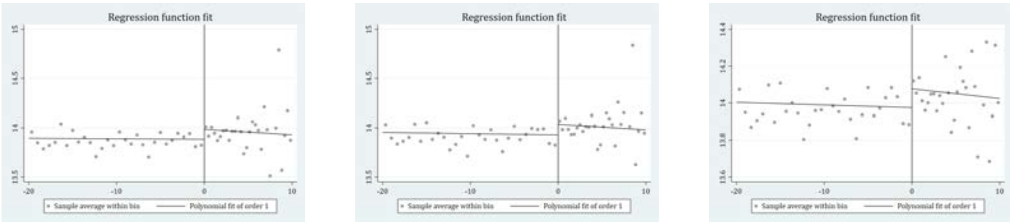
5.2 Effects of Obtaining an Advanced Certificate

For the advanced certificate, the results for all eight quarters are displayed in the third panel of Table 3. Figure 4 presents a visual depiction of the effects on the log of income one, four, and eight quarters after certification. For all quarters, obtaining an advanced certificate has a substantial effect on income. The estimated effect is relatively stable over time, ranging between 8.7% and 12.7%.

The sizable returns associated with acquiring an advanced certificate suggest that it provides new and reliable information to incumbent employers. By obtaining an advanced certificate, individuals can effectively signal their advanced occupation-specific skills, thereby conveying valuable information about their productivity, which employers reward by increasing wages. In this sense, the certificate serves as a screening mechanism, enabling firms to identify and appropriately compensate the most productive workers. Moreover, to the extent that the information conveyed by the certificate is trusted within the labor market, an advanced certificate also enables workers to signal their skills to potential employers. Therefore, income adjustments may arise from new employers seeking to attract the most productive workers or current employers aiming to retain their talented workforce. Obtaining an advanced certificate can therefore lead to wage growth within the firm, even when the current employer knows accurately the worker's productivity. In Section 7, we further discuss these possibilities. Lastly, the stability of returns over time further indicates that employers do not typically update or revise their expectations about productivity after certification.

There are three additional considerations to highlight. First, our returns for the advanced certificate are within the range of estimated returns to certification in developed economies (Albert 2017; Kleiner and Krueger 2013; Kleiner and Vorotnikov 2017). Second, like previous estimates in the literature, we provide evidence that returns to signaling are captured only by workers with advanced skills (Bassi and Nansamba 2022). Second, our estimated returns to signaling advanced occupation-specific skills are comparable to spending an additional year at school in Colombia (García-

Figure 4: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate on Log of Income



(a) Quarter One (b) Quarter Four (c) Quarter Eight

Notes: The three figures summarize the estimated results of Equation (1), one, four, and eight quarters after certification, using the log of income as the main outcome. The running variable is the exam score, and the discontinuity threshold is 90. The regressions include the controls described in Section 4.

Suaza et al. 2014; Herrera-Idárraga et al. 2015; Morales et al. 2021). In that sense, the certificate seems to transfer information about a worker's productivity that is as valuable as that coming from schooling. Such a conclusion is important for individuals with characteristics similar to our estimation sample—namely, less educated workers in their late thirties who have finished school and likely have limited opportunities for wage growth.

5.3 Robustness Checks

We consider several alternative specifications to evaluate the robustness of our findings. The complete set of results is presented in online Appendix B. Upon examination of the estimates, it becomes evident that our main conclusions remain robust across various specifications and samples.

In Figure OA3 in online Appendix B, we show that our main conclusions are robust to using different specifications to evaluate the effects of the certificate.[34] For instance, our core findings are robust to including year and location fixed effects, excluding controls, using an optimal bandwidth that minimizes the coverage error, using a fixed bandwidth, employing both a fixed bandwidth and a sample reporting earnings in all eight periods con-

34. The complete set of results, with detailed information on standard errors, number of observations, and bandwidth, are presented in Tables OA4 to OA6 in online Appendix B.

sidered (leading to a fixed number of total and effective observations over time), using non-bias-corrected RD estimates (Calonico et al. 2014), and not adjusting for the presence of mass points during estimation. The magnitude and significance of the returns to signaling occupation-specific skills remain in all specifications. The most notable deviation arises when employing a fixed bandwidth to evaluate the effects of obtaining an advanced certificate. However, even in this case, the estimates consistently confirm the existence of positive and significant returns. On average, these returns amount to 80% of the baseline estimates.

In addition, in Tables OA7 to OA9 in online Appendix B, we show that our main conclusions are robust to using different samples.[35] First, our primary findings remain largely unchanged when utilizing a sample comprising both men and women. In this scenario, acquiring either a basic or intermediate certificate yields no returns, whereas obtaining an advanced certificate leads to positive, significant, and permanent returns. Compared to the baseline results for the advanced certificate, the returns in the full sample are consistently larger and more precisely estimated for all quarters. This observation implies that women face larger returns to signaling occupation-specific skills, which aligns with traditional models of statistical discrimination. According to these models, employers possess prior beliefs that women have lower productivity than men, resulting in differential compensation (Aigner and Cain 1977; Lang and Spitzer 2020). The presence of an advanced certificate likely corrects these prior beliefs by providing evidence that both men and women possess equivalent occupation-specific skills. Consequently, the proportional increase in returns for women should be greater.

Second, our main conclusions remain mostly unchanged when we restrict the sample to include only men who applied for just one certificate in 2017, 2018, and 2019. For the advanced certificate, the results are slightly smaller in magnitude and less precisely estimated, suggesting that the marginal value of signaling more than one occupation-specific skill within an industry

35. The complete set of results for alternative samples, with detailed information on standard errors, number of observations, and bandwidth, are presented in Tables OA7 to OA9 in online Appendix B.

is somewhat important. Therefore, our results may be partly driven by the presence of multiple certificates or the existence of complementarities between certificates.

Third, to account for the fact that our measure of income may be subject to underreporting, we analyze two subsamples to evaluate the robustness of our results.[36] First, we estimate Equation (1) excluding self-employed individuals, who are more likely to underreport earnings. Second, we use a more restricted sample consisting of those individuals who only worked as salaried employees within the two years following certification. Overall, the conclusions for all three certificates remain robust, suggesting that underreporting does not significantly influence our main results.

To further validate our results, we perform a falsification test to look at placebo thresholds, that is, thresholds other than the real thresholds determining the treatment assignment (Cattaneo et al. 2020). The intuition behind this falsification test is that the probability of receiving treatment should change abruptly only at the true thresholds. Hence, we should not observe abrupt changes in log income at artificial thresholds. We perform the falsification tests by estimating Equation (1) using six placebo thresholds: 25, 35, 55, 65, 85, and 95. The results are summarized in Figure OA5 in online Appendix B.[37] We find no effect on log income for any of the alternative thresholds, which further validates our research design.

The evidence in this section suggests that our dataset satisfies the critical assumptions for sharp RD estimation. Furthermore, our core findings are not driven by a specific bandwidth choice or specific controls and are not affected by underreported income. Moreover, we provide evidence that our key insights are not sample-specific.

36. In Colombia, the monthly contribution to social security includes three categories: pension, health insurance, and insurance to cover occupational hazards. Since the coverage does not depend on the contribution for health and occupational hazards, individuals have incentives to underreport earnings obtained from self-employment or working in the informal sector. Such an incentive does not exist for salaried workers since the employer makes the payments, which count as labor expenses toward tax returns.

37. The complete set of results for placebo thresholds, with detailed information on standard errors, number of observations, and bandwidth, are presented in Table OA10 in online Appendix B.

6 Heterogeneity

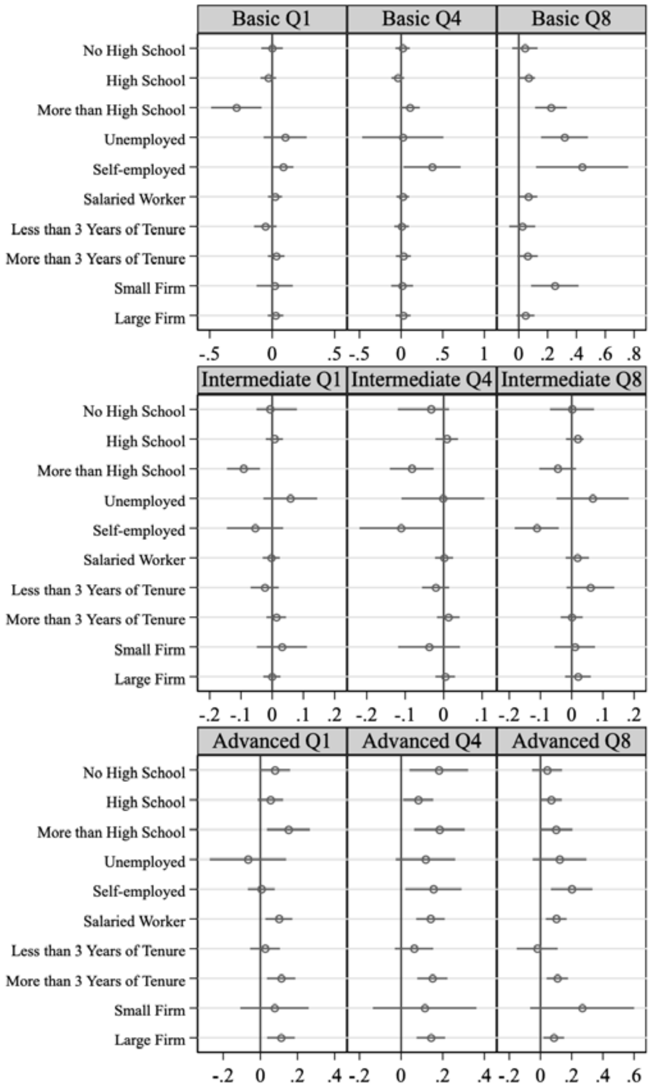
The estimates from Section 5 provide compelling evidence that signaling advanced occupation-specific skills yields significant returns. However, we find no effects from signaling basic or intermediate occupation-specific skills. A key question is whether these returns differ across various groups of workers. In this section, we explore how different the returns are based on education and initial employment status.[38] Furthermore, for salaried workers at the time of certification, we analyze heterogeneous returns by tenure and firm size.

Besides studying whether the results are concentrated in certain population groups, the heterogeneity analysis provides insights regarding how general our results are. As explained in Section 3, because of data constraints, our analysis focuses on a subsample of program participants who reported to PILA at any point during 2010. Given the small but significant differences between the matched and unmatched samples regarding education and initial employment status, exploring how estimates change with these characteristics gives us a better understanding of how the results would look if data were available to all participants.

The heterogeneity results are summarized in Figure 5. Each panel presents results for each type of certificate and different subgroups. The first panel shows the estimated effects for the basic certificate, while the second and third panels display the effects for the intermediate and advanced certificates, respectively. We focus on quarters 1, 4, and 8 after certification. Complete results for all eight quarters are displayed in Tables OA11 to OA13 in online Appendix C. Although our estimation sample is big, dividing it into different subgroups significantly reduces the number of observations near the threshold to the point that, in some cases, the sample is too thin to support strong claims. This issue is particularly salient for the basic certificate.

38. We also explore heterogeneity by age and find no evidence of heterogeneous effects for the basic and intermediate certificates. Returns for the advanced certificate do not seem to vary with age either.

Figure 5: Sharp Regression Discontinuity Estimates of the Effects on Log of Income: Heterogeneity Analysis



Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one, four, and eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 30 for the basic certificate, 60 for the intermediate certificate, and 90 for the advanced certificate. The first panel corresponds to the basic certificate, the second for the intermediate certificate, and the third for the advanced certificate. We show results by initial education level (rows 1 to 3) and initial employment status (rows 4 to 6). For salaried workers (at certification), we also present the results by tenure (rows 7 and 8) and firm size (rows 9 and 10). All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. Standard errors used to compute the 95% confidence interval are clustered at the technical norm level. Detailed results for all eight quarters are displayed in Tables OA11 to OA13 in online Appendix C.

6.1 Returns by Education Level

The first three rows of Figure 5 present RD estimates for the returns to signaling for various subgroups of individuals who differ in their education level at the time of certification: less than high school, high school, and more than high school. Like our core results, we do not find evidence of significant returns for workers with or without a high school diploma who obtained a basic certificate. Workers with at least some college education display a range of returns from negative to null, transitioning to positive, throughout quarters one to eight. However, the number of observations around the threshold is too small to support strong claims. We find similar results for workers with or without a high school diploma who obtained an intermediate certificate. Nevertheless, we estimate a negative effect on income for individuals with some college education or more, which decreases in magnitude and significance towards quarter 8 (-6.8%, on average).

Unlike the basic and intermediate certificates, we observe a gradient of positive and significant returns for the advanced certificate. For all quarters, the largest effect is estimated for workers with at least some college education (an average of 14.1% over quarters one to eight). In some cases, the returns for the more educated are two times larger than those for the least educated. Such a finding indicates that information about schooling reinforces the revelation of ability provided by the certificate, as suggested by Arcidiacono et al. (2010). In that sense, employers seem to complement the signal from the certificate with information from formal education, update their priors about productivity, and pay accordingly.

6.2 Returns by Employment Status at the Time of Certification

Rows 4 to 6 of Figure 5 present RD estimates for subgroups of participants that differ in their employment status at the time of certification: unemployed, self-employed, and salaried. For the basic certificate, the sample size for the first two categories is too small to draw general inferences. We also find no evidence of heterogeneous effects regarding initial employment status for the intermediate certificate, with the exception of some negative effects for self-employed workers.

For the advanced certificate, our analysis reveals positive returns for all workers except those initially unemployed. In this sense, the certificate does not seem to facilitate transitions from unemployment to higher-paid jobs. While we estimate positive and significant effects for workers who were initially self-employed, the returns take time to materialize, becoming evident after one year of certification. The delay in responses may be partially explained by the fact that some self-employed individuals eventually transition into salaried work, where higher-paid jobs are found. We provide evidence of these mechanisms in Section 7. Lastly, the effect on income for individuals who are salaried workers at the time of certification closely resembles our core estimates.

6.3 Returns by Tenure and Firm Size

The results in the previous section indicate that the estimated effects of the advanced certificate are mostly driven by individuals who are either self-employed or salaried workers at the time of certification. To have a better understanding of these results and the potential underlying mechanisms, in this section, we focus on salaried workers and explore heterogeneity by tenure at the time of certification and firm size. The results are displayed in the last four rows of Figure 5.

We start by dividing the sample of salaried workers into two groups: less than three years of tenure and three or more years.[39] Our estimates reveal that the returns associated with the advanced certificate are primarily observed among salaried workers with a minimum of three years of tenure.[40] One potential explanation to reconcile these findings is that, while employers revise their priors about workers' productivity upon observing the certificate, promotion or wage adjustments are usually available for tenured workers, for whom we estimate large and significant effects. In other words,

39. The choice of three years to define tenure groups is motivated by the findings in Lange (2007), who estimates that expectation errors regarding worker's productivity at the beginning of the match are reduced by half within three years of tenure.

40. For the basic and intermediate certificates, we find no effects on income in any of the tenure groups.

the certificate helps sort high-skilled tenured individuals who are eligible for promotions, acting as a screening device. In this sense, employers can rely on the certificate to choose among the most productive workers to fill out the few available positions. First, the certificate provides a mechanism to determine relative performance among the top workers, which is known as a key determinant for promotions (DeVaro 2006). Second, the certificate can be used to break ties among tenured workers, providing firms with an objective criterion for awarding promotions.[41] Importantly, the existence of limited positions within firms makes wage adjustments unlikely even after relevant information has been revealed to current and potential employers. In addition, our results suggest that the certificate does not necessarily facilitate job-to-job transitions and that most movements occur within the firm. Otherwise, we would observe effects for salaried workers regardless of their tenure. We evaluate these mechanisms in Section 7.

To further support the notion that the advanced certificate is a screening device facilitating promotions, we explore the heterogeneity of returns by firm size. In principle, larger firms offer better career progression opportunities than smaller firms (Oi and Idson 1999). Subsequently, if the advanced certificate acts as a screening device, the effects among salaried workers should be mostly driven by those in large firms. We corroborate this possibility by dividing our sample of salaried workers at the time of certification into two groups: those working at small firms (i.e., 1 to 50 employees) and those working at large firms (i.e., more than 50 employees). Consistent with our hypothesis, we find no effects on earnings for salaried workers in small firms and large and significant effects for individuals working in large firms at the time of certification.[42]

6.4 Heterogeneity Remarks

The previous results suggest that gains from signaling exhibit some degree of variation among different groups of workers. On the one hand, we do

41. Benson et al. (2019) discusses the importance of establishing criteria for promotion that do not lead to perceptions of favoritism or unfairness or the impression that effort in one's job goes unrewarded.

42. For the basic and intermediate certificates, we find no effects on income at any firm size.

not find strong evidence of heterogeneity regarding education or initial employment status for basic and intermediate certificates. On the other hand, for the advanced certificate, the returns increase with formal education and are positive only for individuals who are initially salaried or self-employed.

As previously mentioned, our findings can be used to gauge the generalizability of our results. In Section 3, we showed that individuals in the estimation sample are less likely to be unemployed and slightly more educated. Given the absence of evidence of heterogeneity in terms of education and initial employment status for the basic and intermediate certificates, our findings should be largely similar to those we would obtain if we had PILA data for all program participants. Conversely, our results likely provide an upper bound for the effects of obtaining an advanced certificate since salaried and more educated workers are overrepresented in our sample, and they captured the largest returns.

Moreover, the heterogeneity analysis hints at the forces behind income adjustments. First, the certificate seems to convey important information about workers' productivity to potential employers. For instance, self-employed individuals experience an increase in income a few quarters after being certified, only after transitions into salaried work occurred, which is compatible with traditional signaling models in the context of a frictional labor market. Second, the certificate also appears to bring additional information to incumbent employers regarding workers' productivity. Nevertheless, the evidence indicates that promotions and wage adjustments are potentially limited to tenured workers in large firms. Therefore, the certificate acts as a screening device, helping sort high-skilled tenured individuals.[43]

43. It can be argued that in some cases, the certificate enables workers to perform tasks they were not allowed before, and consequently, it helps boost earnings within the same firm. Nevertheless, such technical norms (i.e., regulated) are not considered in the analysis, and therefore, we do not contemplate such a mechanism.

7 Mechanisms

Our baseline estimates suggest that signaling basic or intermediate occupation-specific skills does not affect income, whereas signaling advanced skills has a large positive effect. In addition, the heterogeneity results indicate that changes in income may result from moves across employment sectors (for example, from self-employment to salaried work) or within employment sectors (for example, promotions within firms). In light of these findings, in this section, we formally discuss the potential mechanisms driving our core results. Motivated by our empirical findings, we develop a search model of salaried work and self-employment with asymmetric information on an individual's occupation-specific skills. We summarize the model below, and the model's detailed description is relegated to Appendix A.

To match the characteristics of the labor market under analysis, we assume the labor market comprises two sectors: salaried work and self-employment. Individuals have either no skills, low skills, or high skills to perform a particular occupation.[44] Firms have one of three types of productivity: low, intermediate, and high. Our measure of ability is equal to the occupation-specific skill individuals possess. For simplicity, we assume ability is fixed. Individuals also differ in self-employment productivity. The surplus generated by each firm type depends on the match quality between the firm and the workers. For instance, high-skilled workers generate larger output when matched to a high-type firm.

In the model, an individual chooses to be a salaried worker, self-employed, or unemployed. Each period, unemployed and self-employed individuals face a chance of receiving a job offer from one type of firm. Self-employment is readily available to all workers. We assume asymmetric information on workers' skills such that the probability of receiving an offer depends on the expected productivity of the worker at the matched firm.

44. The model does not distinguish between individuals with low and intermediate skills and, consequently, cannot be used to understand the effects of obtaining an intermediate certificate. We choose to work with a simpler model since the results for intermediate certificates indicate no effect on earnings relative to individuals obtaining a basic certificate.

Since production at a given firm depends on the worker type, firms try to identify the worker's type through a screening mechanism (for example, a job interview) that generates a signal. We interpret the SENA certificates as a credible device that increases the precision of the signal. Therefore, the certificate helps reveal the worker's productivity to the firm and changes the expected value of the match with a firm. Given the introduction of SENA certificates, the model predicts the following testable implications for labor market outcomes. First, regarding income, we derive Proposition 1.

Proposition 1 (Income). Conditional on being a salaried worker, the following is true:

- *(i) Income for high-skilled workers with advanced certificates increases since they are more likely to be employed at high-productivity firms.*
- *(ii) The effect on income is ambiguous for individuals with low skills who get a basic certificate. The direction of the effect is a function of productivity differences across worker types and the former precision of the signal.*
- *(iii) The effect on income is also ambiguous for individuals without skills who do not get a certificate.*

Second, we derive Proposition 2 regarding the employment sector of choice.

Proposition 2 (Employment). The following is true for workers after receiving the certificate:

- *(i) High-skilled workers reallocate to (or continue to work for) high-productivity firms and are less likely to be employed at type-m or type-l firms. Self-employed high-skilled individuals are also more likely to receive (and accept) offers from type-H firms, and, at the same time, some individuals (for example, those formerly employed in type-l firms) move to self-employment. Therefore, the overall effects on salaried work and self-employment are ambiguous.*
- *(ii) Low-skilled workers are less likely to be employed at high-productivity*

firms and more likely to reallocate to (or continue to work for) medium-productivity firms. They are also more likely to reject offers from low-productivity firms. As low-skilled workers are equally productive in self-employment and working at medium-productivity firms, the overall effects on salaried work and self-employment are ambiguous.

(iii) *The effects for workers without skills who receive no certificate are equivalent to those for low-skilled workers.*

The theoretical predictions of the model regarding income largely map the effects documented in Sections 5.1 and 5.2. In what follows, we empirically test the theoretical implications regarding the reallocation of workers. We start by exploring the effects on salaried work and self-employment for the three thresholds of interest. Motivated by the findings in Section 6, we discuss implications by employment status at the time of certification: unemployed, self-employed, and salaried.

We start by discussing the effects of obtaining a basic certificate. The results are displayed in Figure 6.[45] We find no evidence of increases in salaried work or self-employment for individuals who are unemployed at the time of certification. We also find suggestive evidence indicating that individuals move from self-employment to salaried work, resulting in positive income effects. However, the sample size is too small to make any strong claims. Consistent with the model, we find that signaling basic occupation-specific skills does not motivate marginal salaried workers to move out of salaried work.[46] The absence of changes in salaried work further suggests that workers are not fired upon displaying a basic certificate. Intuitively, if individuals were fired upon showing their (not-so-encouraging) certificate, we may observe a reduction in salaried work accompanied by (involuntary) increases in self-employment or unemployment.

The results for advanced-certificate holders, by employment status at the

45. Detailed results for the basic certificate, by employment status at the time of certification, are shown in Table OA14 in online Appendix C.

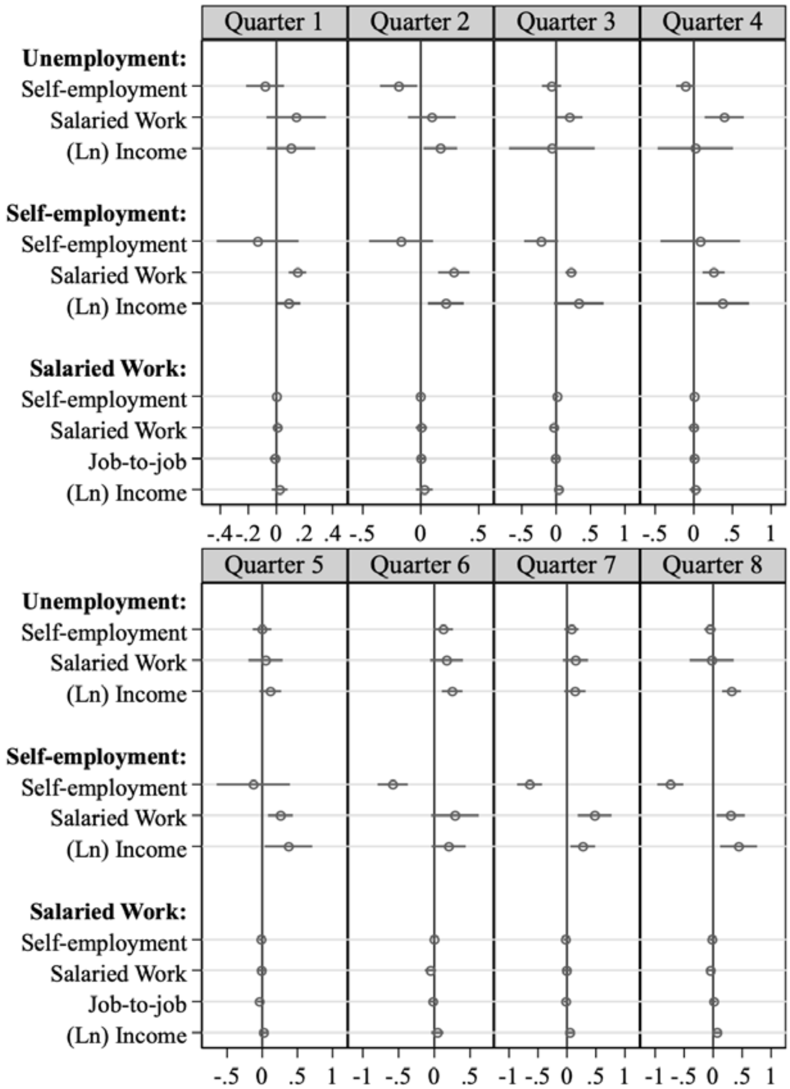
46. The results for the intermediate certificate also suggest no movements across employment sectors. These results for the intermediate certificate, by initial employment status, are summarized in Figure A1 in Appendix B. Detailed results are displayed in Table OA15 in online Appendix C.

time of certification, are displayed in Figure 7.[47] Our results strongly support the model's prediction that individuals initially self-employed transitioned to salaried work after obtaining the certificate, resulting in positive income effects. The effect on salaried work appears three quarters after certification, consistent with a frictional labor market (Lain 2019; Narita 2020). In opposition, we do not find evidence of increases in salaried work for those initially unemployed. Such a pattern also explains why the advanced certificate has no positive returns. Together, these findings suggest that employers complement the signal from the certificate with information from current employment status. In that sense, obtaining an advanced certificate is not enough to compensate for the negative signal of being unemployed.

As predicted by the model, salaried workers at the time of certification, for whom we estimate a sizable effect on income, continue with the same status upon getting the advanced certificate. Although not explicitly accounted for in the model, this sizable effect on income may result from promotions within the same firm or moves to other firms. In other words, moving to a more productive firm in the model could be regarded as moving to a better-suited job within the same firm in our data. In the last two rows of Figure 7, we document that job-to-job transitions significantly increase by 4.2% in the first quarter after certification and by smaller and insignificant percentages after that. Moreover, in Table 4, we show that the change in income one and two years after certification exists only for individuals who did not undertake job-to-job transitions. Therefore, returns are only coming from within-firm adjustments. Although we cannot disentangle if adjustments are due to promotions or wage increases without moving up the organizational ladder, the results show that firms are willing to compensate tenured workers with advanced certificates.

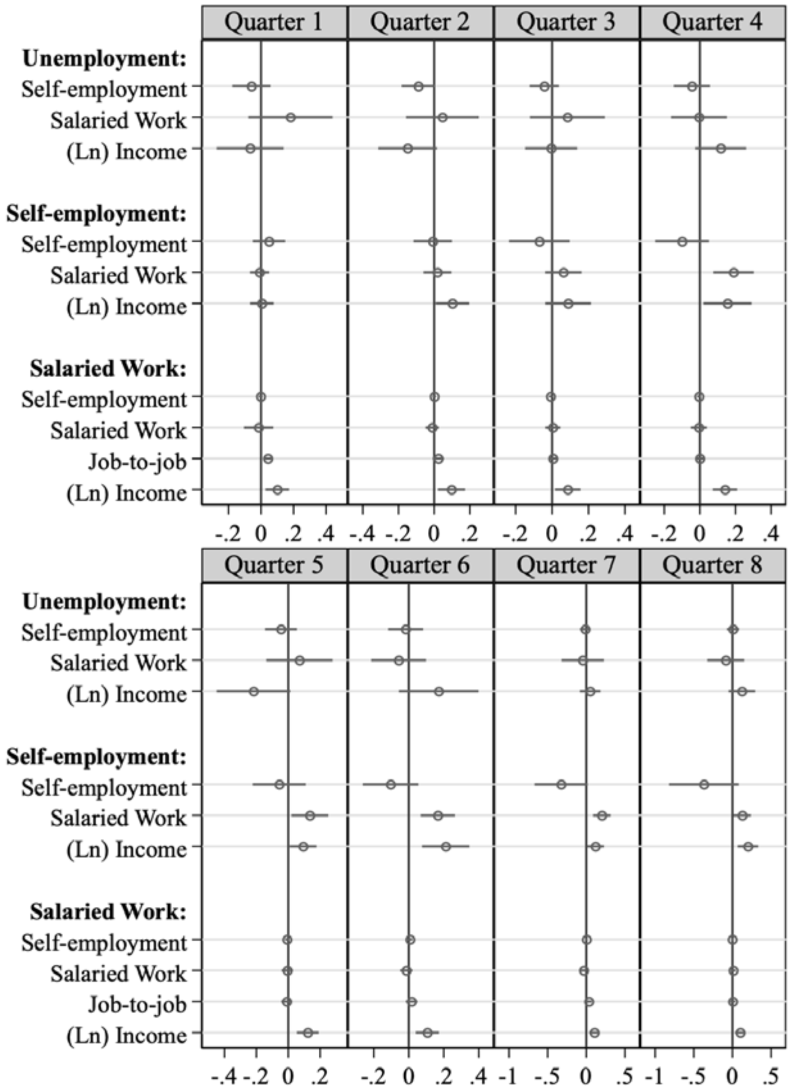
47. Detailed results for the advanced certificate, by initial employment status, are shown in Table OA16 in online Appendix C.

Figure 6: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate by Initial Employment Status: Additional Outcomes



Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one to eight quarters after certification, for four outcomes: salaried work, self-employment, log of income, and job-to-job transitions (for salaried workers only). The running variable is the exam score, and the discontinuity threshold is 30. All regressions include controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. Standard errors used to compute the 95% confidence intervals are clustered at the technical-norm level. Detailed results are displayed in Table OA14 in online Appendix C.

Figure 7: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate by Initial Employment Status: Additional Outcomes



Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one to eight quarters after certification, for four outcomes: salaried work, self-employment, log of income, and job-to-job transitions (for salaried workers only). The running variable is the exam score, and the discontinuity threshold is 90. All regressions include controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. Standard errors used to compute the 95% confidence intervals are clustered at the technical-norm level. Detailed results are displayed in Table OA16 in online Appendix C.

Table 4: Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Advanced Certificate: Income and Job-to-Job Transitions

Outcome	**(1) Quarter 4 - J2J Transition**	**(2) Quarter 8 - J2J Transition**	**(3) Quarter 4 - No J2J Transition**	**(4) Quarter 8 - No J2J Transition**
Ln (Income)	0.084	0.086	0.136	0.107
	(0.094)	(0.061)	(0.033)	(0.036)
# of Observations	23,599	32,661	116,295	90,188
Eff. # of Control Obs.	912	1,548	6,074	4,701
Eff. # of Treatment Obs.	2,246	3,573	12,755	9,616
Bandwidth	2.448	3.026	3.599	3.393

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1). In columns 1 and 2, the outcomes are the log of income one and two years after certification for individuals who did not undertake a job-to-job (J2J) transition in the corresponding periods. Columns 3 and 4 report similar results for individuals who undertook a job-to-job transition after one and two years of certification, respectively. The running variable is the exam score, and the discontinuity threshold is 90. All regressions include controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level.

In Section 6, we found that salaried workers with advanced certificates and less than three years of tenure do not capture any returns. These workers also decide to remain with the same employer. Likely, transitioning into a new firm after a short spell can be interpreted as a negative signal (e.g., Benson and Lepage 2023; Pries and Rogerson 2022; Schönberg 2007). Additionally, if promotions are restricted to tenured workers, movers would not improve their prospects since they lose seniority and still have to accumulate the time necessary for promotion.[48]

Furthermore, our findings suggest that traditional models assuming that wages adequately capture new information about productivity may understate the importance of internal market frictions.[49] In such a way, these results cast further doubts about the ability of wages to aggregate information about productivity in developing countries.

48. These patterns are also compatible with the well-documented fact that firms prefer to train and promote their own workers (Demougin and Siow 1994).

49. The relevant literature and examples of this family of models can be found in Rubinstein and Weiss (2006).

Overall, our results for the advanced certificate can be explained mainly by two mechanisms. First, for self-employed individuals, increases in income result from their ability to signal advanced occupation-specific skills to prospective employers. Second, for salaried workers, increases in income result from the possibility of demonstrating advanced skills to incumbent and potential employers. This can lead to adjustments by employers who wish to retain their talented workforce and use the certificate as a screening device.

8 Conclusion

In this paper, we provide causal evidence of the effect of signaling occupation-specific skills on income. We take advantage of a novel program of certifications in Colombia that assesses workers' abilities in particular tasks and employ a regression discontinuity design to estimate the certificate's returns. Our study design is unique in that it allows us to evaluate the effect of the signal's content, as the certification program offers three levels of certification: basic, intermediate, and advanced. We estimate returns on earnings up to two years after certification and find that the effects vary significantly with the signal content.

On the one hand, workers with a basic or intermediate certificate have no effects on earnings within two years. On the other hand, there is a sizable effect on average earnings (9.7%) for individuals with advanced certificates. We further explore whether returns differ across various groups of workers. We observe a gradient of positive and significant returns for the advanced certificate based on schooling. In addition, we find positive returns for all workers displaying an advanced certificate, except for those initially unemployed.

Overall, our findings suggest that obtaining an advanced certificate significantly impacts earnings through two main channels. First, it helps incumbent employers to update their knowledge about workers' productivity, leading to higher compensation. However, such increases are limited by opportunities for promotion and wage adjustment, which are more common in large firms and among tenured individuals. Second, the advanced certificate allows individuals to effectively signal their occupation-specific skills to potential employers, increasing their likelihood of receiving new job offers. Nevertheless, the differential effect observed for unemployed and self-employed individuals suggests that employers complement the signal from the certificate with information regarding current employment status. In such a way, obtaining the certificate only partially alleviates inefficiencies due to information frictions, even among the workers with advanced skills. Our evidence points out to other sources of frictions important enough to prevent particular groups of workers to captures the returns. Such a finding has important policy implications since it suggest that the provision of information alone has limited capacity to correct wages below productivity.

Our research provides compelling evidence that skill certification programs have some potential to impact earnings substantially among high-skilled individuals. Nevertheless, it is important to acknowledge one caveat when interpreting our findings. Our sample is confined to Colombia, which implies that our results primarily pertain to labor markets characterized by significant information frictions and wage rigidities. In this sense, our findings hold broader implications for developing countries sharing similar labor market characteristics.

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.

Abel, Martin, Rulof Burger, and Patrizio Piraino. 2020. "The Value of Reference Letters: Experimental Evidence from South Africa." *American Economic Journal: Applied Economics* 12 (3): 40–71.

Agudelo, Sonia A, and Hector Sala. 2017. "Wage Rigidities in Colombia: Measurement, Causes, and Policy Implications." *Journal of Policy Modeling* 39 (3): 547–567.

Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30 (2): 175–187.

Albert, Kyle. 2017. "The Certification Earnings Premium: An Examination of Young Workers." *Social Science Research* 63:138–149.

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The Effects of Labour Market Policies in an Economy with an Informal Sector." *The Economic Journal* 119 (539): 1105–1129.

Altonji, J. G., and C. R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *The 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.

Baird, Matthew D, Robert Bozick, and Melanie A Zaber. 2021. "Beyond Traditional Academic Degrees: The Labor Market Returns to Occupational Credentials in the United States." *IZA Journal of Labor Economics* 11 (1).

Bandiera, Oriana, Livia Alfonsi, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2020. "Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda." *Econometrica* 88 (6): 2369–2414.

Bassi, Vittorio, and Aisha Nansamba. 2022. "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda." *The Economic Journal* 132 (642): 471–511.

Becker, Gary S. 2009. *Human capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* University of Chicago press.

Bedard, Kelly. 2001. "Human Capital Versus Signaling Models: University Access and High School Dropouts." *Journal of Political Economy* 109 (4): 749–775.

Behrman, Jere R. 1999. “Chapter 43 Labor Markets in Developing Countries,” 3:2859–2939. Handbook of Labor Economics. Elsevier.

Benson, Alan, and Louis-Pierre Lepage. 2023. “Learning to Discriminate on the Job.” Working Paper, March.

Benson, Alan, Danielle Li, and Kelly Shue. 2019. “Promotions and the Peter Principle.” *The Quarterly Journal of Economics* 134 (4): 2085–2134.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts. 2010. “Why Do Firms in Developing Countries Have Low Productivity?” *American Economic Review* 100 (2): 619–23.

Busso, Matias, Sebastian Montaño, and Juan S. Muñoz-Morales. 2023. “Signaling Specific Skills and the Labor Market of College Graduates.” IZA Discussion Paper 16449.

Calonico, Sebastian, Matias D Cattaneo, and Rocío Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82 (6): 2295–2326.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin. 2022. “Job Search and Hiring with Limited Information about Workseekers’ Skills.” *American Economic Review* 112, no. 11 (November): 3547–3583.

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, Matias D., and Rocío Titiunik. 2022. “Regression Discontinuity Designs.” Working Paper, February.

Chiplunkar, Gaurav, and Abhijit V Banerjee. 2023. “How Important are Matching Frictions in the Labor Market? Experimental & Non-Experimental Evidence From a Large Indian Firm.” *Experimental & Non-Experimental Evidence From a Large Indian Firm.*

Clark, Damon, and Paco Martorell. 2014. "The Signaling Value of a High School Diploma." *Journal of Political Economy* 122 (2): 282–318.

Demougin, Dominique, and Aloysius Siow. 1994. "Careers in Ongoing Hierarchies." *The American Economic Review* 84 (5): 1261–1277.

DeVaro, Jed. 2006. "Internal Promotion Competitions in Firms." *The RAND Journal of Economics* 37 (3): 521–542.

Frandsen, Brigham R. 2017. "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete." In *Regression Discontinuity Designs,* 281–315. Emerald Publishing Limited, May.

Garcıa-Suaza, Andrés Felipe, Juan Carlos Guataquı, José Alberto Guerra, and Darıo Maldonado. 2014. "Beyond the Mincer Equation: The Internal Rate of Return to Higher Education in Colombia." *Education Economics* 22 (3): 328–344.

Gibbons, Robert, and Michael Waldman. 1999. "A Theory of Wage and Promotion Dynamics Inside Firms." *The Quarterly Journal of Economics* 114 (4): 1321–1358.

———. 2004. "Task-Specific Human Capital." *The American Economic Review* 94 (2): 203–207.

Greenwald, Bruce C. 1986. "Adverse Selection in the Labour Market." *The Review of Economic Studies* 53 (3): 325–347.

Hall, Robert E, and Charles I Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *The Quarterly Journal of Economics* 114 (1): 83–116.

Herrera-Idárraga, Paula, Enrique López-Bazo, and Elisabet Motellón. 2015. "Double Penalty in Returns to Education: Informality and Educational Mismatch in the Colombian Labour Market." *The Journal of Development Studies* 51 (12): 1683–1701.

Jepsen, Christopher, Peter Mueser, and Kenneth Troske. 2016. "Labor Market Returns to the GED Using Regression Discontinuity Analysis." *Journal of Political Economy* 124 (3): 621–649.

Kahn, Lisa B., and Fabian Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *The Review of Economic Studies* 81 (4): 1575–1613.

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "Occupational Specificity of Human Capital." *International Economic Review* 50 (1): 63–115.

Kleiner, Morris M, and Alan B Krueger. 2013. "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market." *Journal of Labor Economics* 31 (S1): S173–S202.

Kleiner, Morris M, and Evgeny Vorotnikov. 2017. "Analyzing Occupational Licensing among the States." *Journal of Regulatory Economics* 52:132–158.

Lain, Jonathan. 2019. "Discrimination in a Search and Matching Model with Self-Employment." *IZA Journal of Development and Migration* 9 (1): 6.

Lang, Kevin, and Ariella Kahn-Lang Spitzer. 2020. "Race Discrimination: An Economic Perspective." *The Journal of Economic Perspectives* 34 (2): 68–89.

Lange, Fabian. 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1): 1–35.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

Machin, Stephen, Sandra McNally, and Jenifer Ruiz-Valenzuela. 2020. "Entry Through the Narrow Door: The Costs of Just Failing High Stakes Exams." *Journal of Public Economics* 190:104224.

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): 223–61.

McCall, John Joseph. 1970. “Economics of Information and Job Search.” *The Quarterly Journal of Economics,* 113–126.

McKenzie, David. 2017. “How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence.” *Labour Economics* 83 (J62): C93.

Morales, Leonardo Fabio, Christian Posso, and Luz A Flórez. 2021. “Heterogeneity in the Returns to Tertiary Education for the Disadvantage Youth: Quality vs. Quantity Analysis.” *Borradores de Economıa; No. 1150.*

Narita, Renata. 2020. “Self-Employment in Developing Countries: A Search-Equilibrium Approach.” *Review of Economic Dynamics* 35:1–34.

Neal, Derek. 1995. “Industry-Specific Human Capital: Evidence from Displaced Workers.” *Journal of Labor Economics* 13 (4): 653–677.

Oi, Walter Y, and Todd L Idson. 1999. “Firm Size and Wages.” *Handbook of Labor Economics* 3:2165–2214.

Parent, Daniel. 2000. “Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics.” *Journal of Labor Economics* 18 (2): 306–323.

Pinkston, Joshua C. 2009. “A Model of Asymmetric Employer Learning with Testable Implications.” *The Review of Economic Studies* 76 (1): 367–394.

Poletaev, Maxim, and Chris Robinson. 2008. “Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000.” *Journal of Labor Economics* 26 (3): 387–420.

Pries, Michael J., and Richard Rogerson. 2022. "Declining Worker Turnover: The Role of Short-Duration Employment Spells." *American Economic Journal: Macroeconomics* 14, no. 1 (January): 260–300.

Rosenzweig, Mark R. 1988. "Chapter 15 Labor Markets in Low-Income Countries," 1:713–762. Handbook of Development Economics. Elsevier.

Rubinstein, Yona, and Yoram Weiss. 2006. "Post schooling Wage Growth: Investment, Search and Learning." *Handbook of the Economics of Education* 1:1–67.

Sanders, Carl, and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annual Review of Economics* 4:399–425.

Schönberg, Uta. 2007. "Testing for Asymmetric Employer Learning." *Journal of Labor Economics* 25 (4): 651–691.

Tyler, John H, Richard J Murnane, and John B Willett. 2000. "Estimating the Labor Market Signaling Value of the GED." *The Quarterly Journal of Economics* 115 (2): 431–468.

Xia, Xing. 2021. "Barrier to Entry or Signal of Quality? The Effects of Occupational Licensing on Minority Dental Assistants." *Labour Economics* 71:102027.

Appendix A Model of Salaried Work and Self-Employment

Setup

The following model is based on Bassi and Nansamba (2022), and it is extended to account for self-employment and to better adjust to the treatment explored in our paper. It builds on the framework of McCall (1970) and the literature on on-the-job search models with heterogeneous firms and self-employment (Albrecht et al. 2009; Narita 2020).

The labor market is composed of two sectors: salaried work and self-employment. There are three types of individuals, which we refer to as N, L, and H, for individuals with no, low, and high skills, respectively.[50] There are also three types of firms, which we refer to as L, M, and H, for low-, medium- and, high-productive firms, respectively. Both firms and workers have mass one, so there is always a possible match between them. Individuals differ in their ability to perform an occupation. A fraction ϕ_n of workers have no skills, a fraction ϕ_l has low skills, and the rest of the workers have high skills.

Firms differ in productivity. The surplus generated by each firm type depends on the match quality between the firm and the workers, as defined below. Each firm only hires one worker, who inelastically supplies one labor unit. As a result, firms only differ in productivity, which we allow to be match specific. A firm of type H produces a surplus equal to $y_H^h = a$ when matched to a type-h worker and a negative surplus equal to $y_H^l = -b$ when matched to a worker of type l or n. A type M firm produces a surplus equal to $y_M^h = c$ when matched to a type-h or type-l worker. The surplus is negative $y_M^l = -d$ when matched to a type-n worker. Lastly, workers are equally productive at type-N firms, generating a surplus equal to e, such that $a > c > e > 0$. Self-employment productivity increases in the type of worker.

Search model with two sectors: Individuals in the model can be salaried workers, unemployed, or self-employed. Each period unemployed and self-employed individuals face a chance of receiving a job offer from either type of firm. Self-employment is readily available to all workers. We assume asymmetric information on workers' skills, such that the probability of receiving an offer depends on the expected productivity of the worker at the matched firm. If an offer is made, the offered wage corresponds to half of the expected surplus generated by the match. We further assume there are no separations.

50. The model does not distinguish between individuals with low and intermediate occupation-specific skills and, consequently, cannot be used to understand the local effects of obtaining an intermediate certificate. We choose to work with a simpler model since the results for intermediate certificates indicate no effect on earnings, relative to individuals obtaining a basic certificate.

The surplus from a match with a type-L firm does not depend on the worker's skills, and the probability of meeting and receiving a wage offer from these firms is the same for all workers. The per-period wage offered by firms of type L is half the surplus: $w_N = e/2$.

Production at type-M and type-H firms depends on the worker type; hence, these firms try to identify the worker's type through a job interview, which generates a signal σ. The signal can be null (N), basic (B), or advanced (A). The probability that the signal is correct, q, is more informative the closer it is to 1. Therefore, we assume that signals are somewhat informative, such that $q = P(A|h) = P(B|l) = P(N|n) \in (0.5, 1)$, and that $P(A|l) = P(A|n) = P(B|h) = P(B|n) = P(N|h) = P(N|l) = (1-q)/2$. Lastly, we assume that the share of each worker type is 1/3.

Given a signal, firms compute the posterior probability of the worker type using Bayes' rule: $P(h|A)$, $P(l|B)$, $P(n|N)$, etc. Based on the assumption that the share of each worker type is 1/3, it follows that q is also the posterior probability. Namely, $q = P(h|A) = P(l|B) = P(n|N)$. Firms then compute the expected productivity of the worker. The expected productivity of the worker at type-H firms, given signal σ, is,

$$E[y_H|\sigma] = \begin{cases} qa + (1-q)(-b) & \text{if } \sigma = A, \\ (\frac{1-q}{2})a + (q + \frac{1-q}{2})(-b) & \text{if } \sigma = B, N. \end{cases}$$

Similarly, the expected productivity of the worker at type-M firms, given signal σ, is,

$$E[y_M|\sigma] = \begin{cases} (q + \frac{1-q}{2})c + (\frac{1-q}{2})(-d) & \text{if } \sigma = A, B \\ (1-q)c + q(-d) & \text{if } \sigma = N. \end{cases}$$

We assume that the expected output from the match for a type-H firm is negative in the case of no signal or a basic signal, so they do not make offers in those cases. In this case, the only wage they offer for an advanced signal is:

$$w_H = \frac{qa + (1-q)(-b)}{2}$$

Likewise, the expected output from the match for a type-M firm is negative in the case of no signal, so they do not make offers. Therefore, the only wage they offer for advanced or basic signals is

$$w_M = \frac{(q + \frac{1-q}{2})c + (\frac{1-q}{2})(-d)}{2}.$$

The per-period compensation in self-employment for type-n workers equals $w^n_{SE} = e/2$, such that $w^n_{SE} = w_N$. Furthermore, the per-period compensation in self-employment for type-l and type-h workers is equal to $w^l_{SE} = c/2$.

Worker's Problem

The value function of an unemployed worker of type k, for $k \in \{n, l, h\}$ is given by,

$$\begin{aligned} V^U_k &= \beta[p^L_k V^L_k p^M_k V^M_k + p^H_k V^H_k + \\ &\quad (1 - p^L - p^M_k - p^H_k) \max(V^U_k, V^{SE}_k)]. \end{aligned}$$

An unemployed individual earns zero in the current period. Next period, with probabilities p_L, p_M, and p_H, he receives an offer from a firm of type L, M, and H, which are valued at V^L_k, V^M_k, V^H_k, respectively. If he does not receive an offer from any firm, he chooses between unemployment and self-employment, valued at V^{SE}_k. The value function of self-employment, V^{SE}_k, is defined as,

$$
\begin{aligned}
V_k^{SE} &= w_{SEk} + \beta[p_k^L V_k^L + p_k^M V_k^M + p_k^H V_k^H \\
&\quad +(1 - p^L - p_k^M - p_k^H)\max(V_k^U, V_k^{SE})]
\end{aligned}
$$

Lastly, the value function of employment at a firm of type X, V_k^X, for $X \in \{L, M, H\}$, is defined as,

$$
V_k^X = \max\left[\frac{w_X + \theta}{1-\beta}, V_k^U, V_k^{SE}\right],
$$

where θ is a match-specific taste shock for the firm. Upon facing an offer, individuals compare the utility of being employed at that firm forever to unemployment and self-employment, where the individual continues to search for a job.

Decisions in the model depend on the probabilities of getting job offers from the three types of firms, wages conditional on being a salaried worker, and self-employment earnings. Probabilities, wages, and self-employment earnings depend on q. The following subsection discusses the implications of changing the value of q.

Implications of Certificates

We interpret the SENA certificates as a credible increase in the precision of the signal, q. In particular, by assuming that $q = 1$, the posterior probabilities are also more accurate (e.g., $P(H|A) \to 1$). We now discuss the implications of SENA certificates on earnings, salaried work, and self-employment.

- **Positive assortative matching**: As q increases, type-h workers are more likely to be employed at type-H firms, type-l workers are more likely to be employed at type-M firms, and type-n workers are more likely to be employed at type-L firms. Note that type-M firms still make offers to type-h workers, but they are more likely to accept offers from type-H firms. Similarly, type-L firms still make offers to type-l workers, but they are more likely to accept offers from type-M firms.

- **Unconditional wages**: As q increases, type-H firms pay higher wages due to the reduced probability of hiring type-l and type-n workers.[51] At type-M firms, as q increases, w_M increases as there is a lower probability of hiring type-n workers whose productivity is low. There is no change in the average wages paid at type-L firms, as everybody is equally productive in these firms.

- **Conditional wages**: Conditional on being a salaried worker, wages for type-h individuals increase since they are more likely to be employed at type-H firms, which pay higher salaries. Type-l workers are more likely to be employed at type-M firms but less likely to be employed at type-H firms, which pay higher salaries. The effect on wages, conditional on salaried work for type-l workers, is ambiguous. Type-n workers are more likely to be employed at type-L firms, which pay lower wages, but less likely to be employed at type-H or type-M firms, which pay higher salaries. Therefore, the effect on wages for type-n workers, conditional on being salaried workers, is ambiguous.

- **Salaried work and self-employment**: Type-h workers are more likely to receive (and accept) offers from type-H firms. They are also more likely to reject offers from type-N and type-M firms and reallocate to self-employment since the average self-employment compensation is now larger than the wage offered by type-N firms and equal to the wage offered by type-M firms.[52] Self-employed individuals are also more likely to receive (and accept) offers from type-H firms. Therefore, the overall effects on salaried work and self-employment are ambiguous.

  Type-l workers are less likely to be employed at type-H firms and more likely to be employed at type-M firms. They are also more likely to

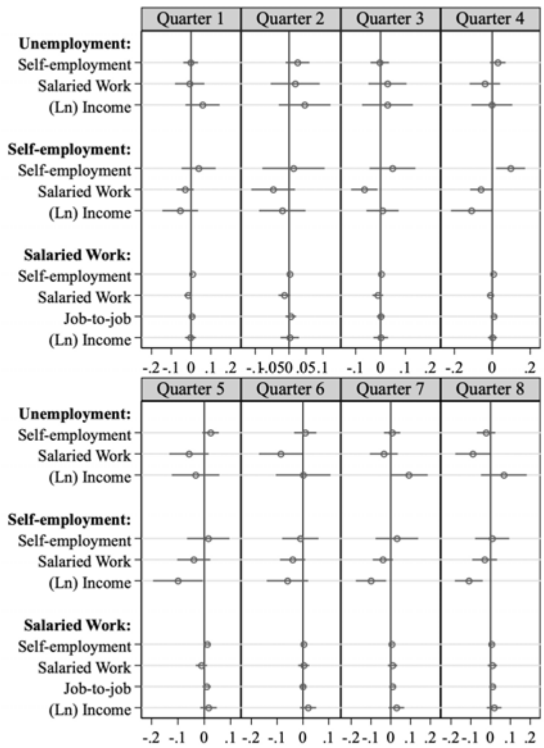
51. The new compensation offered by type-H firms increases to $\mathrm{w}_H = a/2$.

52. The new wage offered by type-M firms is $w_M = c/2$, which is equal to the per-period self-employment compensation for type-h workers.

reject offers from type-L firms. Since type-l workers are indifferent between working at a type-M firm and being self-employed, the overall effects on salaried work and self-employment are ambiguous. The impact on salaried work and self-employment are similar for type-n workers.

Appendix B Additional Figures

Figure A1: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate by Initial Employment Status: Additional Outcomes



Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one to eight quarters after certification, for four outcomes: salaried work, self-employment, log of income, and job-to-job transitions (for salaried workers only). The running variable is the exam score, and the discontinuity threshold is 60. All regressions include controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. Standard errors used to compute the 95% confidence intervals are clustered at the technical-norm level. Detailed results are displayed in Table OA15 in online Appendix C.

Online Appendix

Online Appendix A Additional Figures and Tables

Figure OA1: Example of an Advanced Certificate


Libertad y orden
RÉPUBLICA DE COLOMBIA

El Servicio Nacional de Aprendizaje SENA

En Cumplimiento del Decreto 933 de 2003
otorga
Certificado de Competencia Laboral a

[redacted]

Con CÉDULA DE CIUDADANÍA No. [redacted]

Quien demostró su Competencia Laboral en la Norma

LEVEL OF CERTIFICATE

TECHNICAL NORM

NIVEL AVANZADO **Transportar usuarios en los vehículos automotores de servicio especial de acuerdo a normatividad vigente.**

Código: 280601013

En testimonio de lo anterior, se firma el presente en CHÍA, A los veintiuno (21) días del mes de Junio de dos mil diecisiete (2017)

DATE

Firmado Digitalmente por

[redacted]

86097 - 21/06/2017
No Y FECHA REGISTRO

VIGENCIA
21 de Junio de 2020

La autenticidad de este documento puede ser verificada en el registro electrónico que se encuentra en la página web http://certificados.sena.edu.co, bajo el número 951300280601013CC11435671C.

Table OA1: Descriptive Statistics - Matched and Unmatched Sample of Men

	Unmatched sample (men)	Estimation sample (men)	P-value (Matched vs. Unmatched)
	A. Demographic Characteristics (SENA)		
Demographic Characteristics (Mean)			
Male			
Age	35.98	41.99	0.00
Less than High School	0.23	0.20	0.00
High School	0.46	0.46	0.07
Some College	0.29	0.30	0.00
More than College	0.01	0.04	0.00
Employment Status (Mean)			
Salaried Worker	0.74	0.87	0.00
Self-Employed	0.06	0.05	0.00
Unemployed	0.19	0.09	0.00
	B. Certification Program (SENA)		
Skills Certified			
Technical Norms	912	912	
Expert Tables	74	74	
Certification Level (Mean)			
No certification	0.01	0.01	0.00
Basic	0.13	0.13	0.02
Intermediate	0.38	0.39	0.00
Advanced	0.48	0.47	0.00
Certification Two-Part Exam (Mean)			
Knowledge	82.18	82.02	0.00
Competence	98.83	99.02	0.00
Individuals	247,581	181,395	

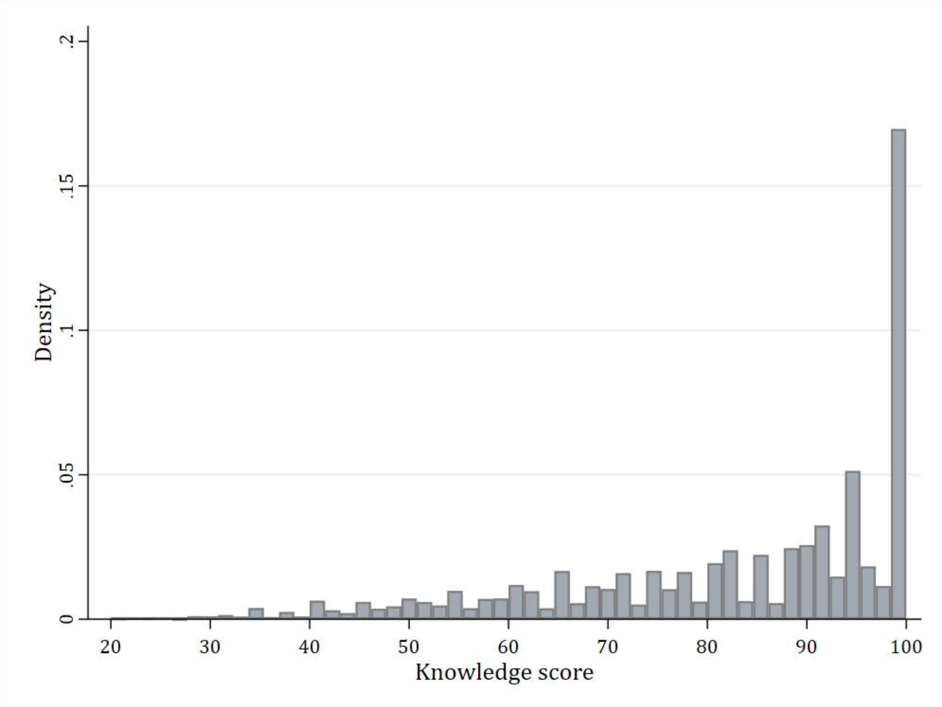
Notes: This table reports descriptive statistics for the the unmatched (first column) and matched (second column) samples of men. The estimation sample corresponds to the subsample of workers we could match with PILA. The third column provides the p-value of a difference-in-means t-test for the matched (estimation) sample and the unmatched sample of men. The first panel reports demographic characteristics and employment information. The second panel reports information regarding the certification program.

Table OA2: Top 10 Technical Norms: Matched and Unmatched Samples

Matched Sample			
Ranking in Matched	Name	Share of Participants	Ranking in Unmatched
1	Serve customers in accordance with service procedures and regulations.	0.10	1
2	To guide classroom training in accordance with technical procedures and regulations.	0.05	3
3	Control access in accordance with private security regulations.	0.04	5
4	Operate forklifts according to the technical manual.	0.04	23
5	Handle food according to current standards.	0.02	2
6	Operate the hydraulic excavator according to the technical manual.	0.02	17
7	Drive light vehicles according to technical procedures and traffic and transportation regulations.	0.02	7
8	Prevent security and surveillance incidents by technical regulations.	0.02	15
9	Prepare light vehicles in accordance with legal and technical regulations.	0.02	9
10	To drive inter-municipal or special passenger service motor vehicles, category c2, by the regulations in force.	0.02	30
Un-Matched Sample			
Ranking in Unmacthed	Name	Share of Participants	Ranking in Matched
1	Serve customers in accordance with service procedures and regulations.	0.21	1
2	Handle food according to current standards.	0.07	5
3	To guide classroom training in accordance with technical procedures and regulations.	0.07	2
4	Administer immunobiological according to delegation and health regulations.	0.05	76
5	Control access in accordance with private security regulations.	0.04	3
6	Orient people according to health standards.	0.03	65
7	Drive light vehicles according to technical procedures and traffic and transportation regulations.	0.03	7
8	Collect potentially recyclable solid waste according to established procedures and current regulations.	0.02	110
9	Prepare light vehicles in accordance with legal and technical regulations.	0.02	9
10	Transfer users in accordance with coexistence, transit, and land transportation regulations.	0.02	12

Notes: This table displays the top 10 technical norms in the matched (estimation) and unmatched samples of men.

Figure OA2: Distribution of Scores



Notes: This figure displays the distribution of scores in the second part of the certification exam (knowledge test) for the matched sample of men.

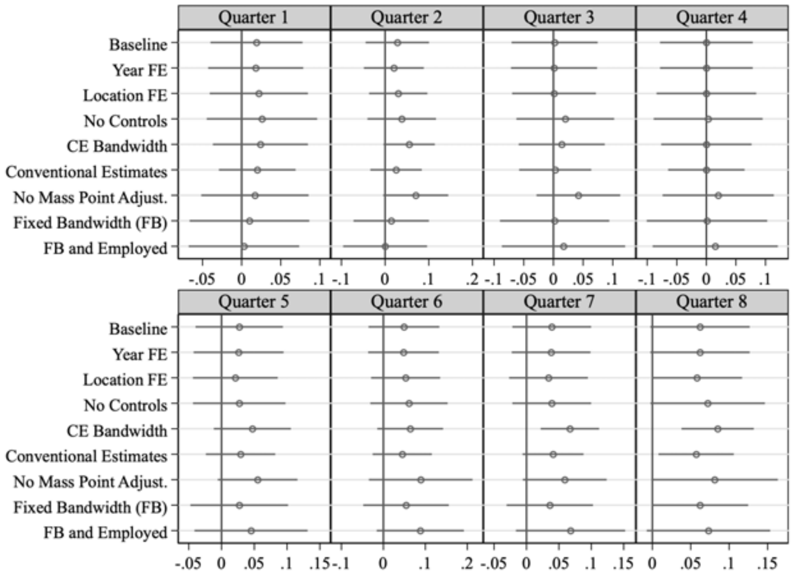
Table OA3: Descriptive Statistics: Estimation Sample of Men and Women

Variable	**# Observations**	**Mean**	**Std. dev.**
Employed	2,043,639	0.88	0.32
Income	2,043,639	1,186,950	983,688
Salaried Work	2,043,639	0.74	0.44
Self-employment	2,043,639	0.12	0.32
Ln of Income	1,800,266	13.97	0.56
Ln of Income - Salaried Worker	1,503,237	13.99	0.55
Ln of Income - Self-Employed	237,918	13.90	0.35

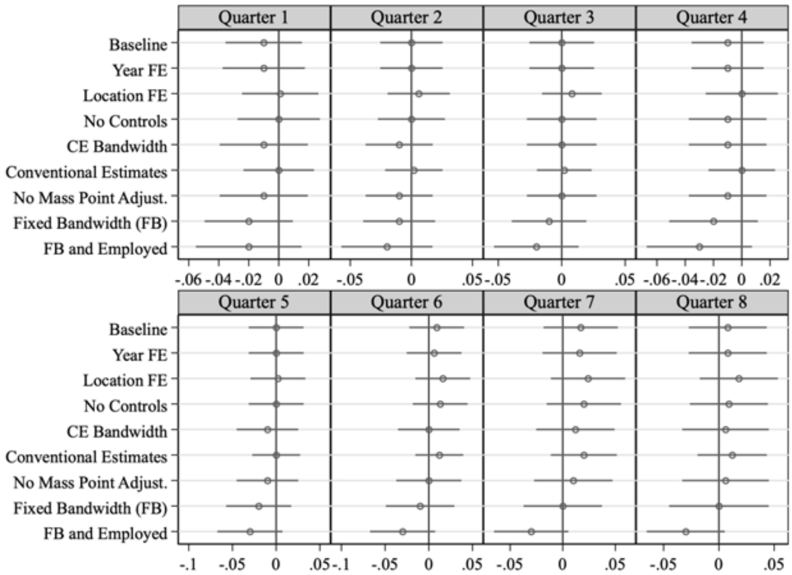
Notes: The table shows descriptive statistics for the panel of men and women two years after certification. Employment and income information are calculated using PILA data. Individuals are classified as employed if they are categorized as salaried workers, self-employed, or if they are assigned to any other category in which employers make contributions to the social security system on their behalf. Therefore, the employment rate is larger than the sum of self-employment and salaried work. The income variable contains zeros in periods when individuals are not employed.

Online Appendix B Robustness Checks

Figure OA3: Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Certificate on Log of Income: Robustness Checks

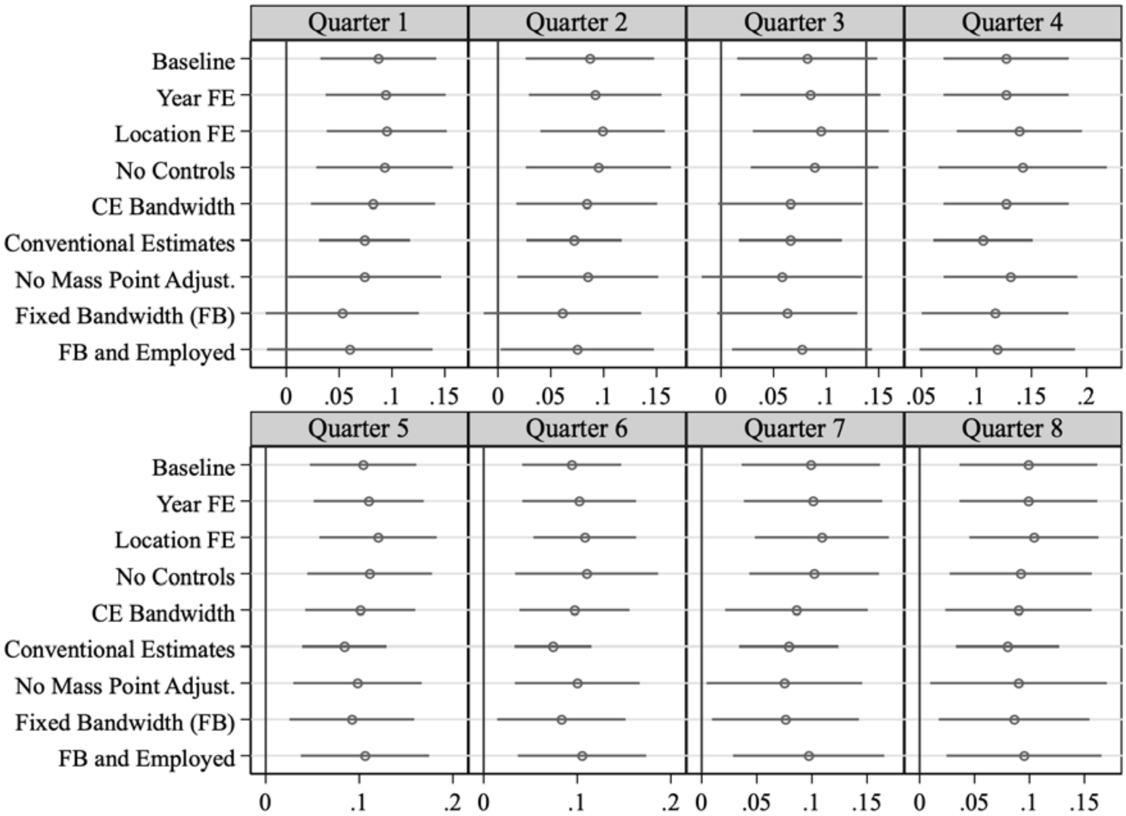


(a) Basic



(b) Intermediate

Figure OA3 Continued. Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Certificate on Log of Income: Robustness Checks



(c) Advanced

Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 30 for the basic certificate (panel a), 60 for the intermediate certificate (panel b), and 90 for the advanced certificate (panel c). In each plot, row 1 displays the baseline results from Table 3. Rows 2 and 3 show the results when we add year and location fixed effects (FE), respectively. Row 4 excludes controls. Row 5 uses the bandwidth that minimizes the coverage error (CE) of the local polynomial regression discontinuity estimator. Row 6 reports non-bias-corrected estimates. Row 7 does not adjust for the presence of mass points. Row 8 uses a fixed bandwidth. Row 9 uses a fixed bandwidth and restricts attention to individuals who are employed in all eight quarters. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator, except for rows 5, 8, and 9. Standard errors used to compute 95% confidence intervals are clustered at the technical-norm level. Detailed results for all three certificates are displayed in Tables OA4, OA5, and OA6.

Table OA4: Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Basic Certificate on Log of Income: Robustness Checks

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
1. Baseline	0.019	0.028	0.002	-0.009	0.027	0.049	0.039	0.062
	(0.030)	(0.037)	(0.037)	(0.040)	(0.034)	(0.043)	(0.031)	(0.033)
Eff. # of Control Obs.	756	633	698	713	617	680	611	641
Eff. # of Treatment Obs.	3,617	2,237	2,743	3,510	2,188	2,648	2,139	2,182
Bandwidth	10.140	7.738	9.994	10.470	7.661	9.038	7.203	8.306
2. Year FE	0.018	0.020	0.001	-0.009	0.026	0.048	0.038	0.062
	(0.031)	(0.035)	(0.037)	(0.040)	(0.035)	(0.043)	(0.031)	(0.033)
Eff. # of Control Obs.	715	704	739	713	616	681	610	641
Eff. # of Treatment Obs.	2,821	2,760	3,511	3,510	2,174	2,660	2,126	2,182
Bandwidth	9.937	9.359	10.200	10.470	7.432	9.223	7.138	8.306
3. Location FE	0.022	0.030	0.001	-0.003	0.021	0.053	0.034	0.058
	(0.032)	(0.034)	(0.036)	(0.043)	(0.033)	(0.042)	(0.031)	(0.030)
Eff. # of Control Obs.	644	676	672	648	615	680	612	600
Eff. # of Treatment Obs.	2,294	2,523	2,510	2,504	2,173	2,456	2,140	1,935
Bandwidth	7.584	8.109	8.188	8.086	7.214	9.005	7.296	7.853
4. No Controls	0.026	0.038	0.020	0.003	0.027	0.061	0.039	0.072
	(0.036)	(0.040)	(0.042)	(0.047)	(0.036)	(0.047)	(0.031)	(0.038)
Eff. # of Control Obs.	645	574	628	648	682	657	656	599
Eff. # of Treatment Obs.	2,295	1,934	2,224	2,504	2,708	2,439	2,453	1,918
Bandwidth	7.643	6.795	7.591	8.200	9.447	8.133	8.919	7.245
5. CE Bandwidth	0.024	0.055	0.014	-0.001	0.047	0.064	0.067	0.085
	(0.031)	(0.030)	(0.037)	(0.039)	(0.030)	(0.040)	(0.023)	(0.024)
Eff. # of Control Obs.	643	484	621	604	472	546	470	462
Eff. # of Treatment Obs.	2,277	1,758	2,191	2,198	1,709	1,854	1,658	1,503
Bandwidth	7.186	5.484	7.081	7.416	5.431	6.405	5.105	5.895

Continues in the next page...

Table OA4 Continued. Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Basic Certificate on Log of Income: Robustness Checks

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
6. Conventional Estimates	0.020	0.025	0.003	-0.007	0.029	0.045	0.041	0.057
	(0.025)	(0.030)	(0.031)	(0.033)	(0.027)	(0.036)	(0.024)	(0.025)
Eff. # of Control Obs.	756	633	698	713	617	680	611	641
Eff. # of Treatment Obs.	3,617	2,237	2,743	3,510	2,188	2,648	2,139	2,182
Bandwidth	10.140	7.738	9.994	10.470	7.661	9.038	7.203	8.306
7. No Mass Points Adjustment	0.017	0.070	0.042	0.020	0.055	0.089	0.059	0.081
	(0.035)	(0.038)	(0.036)	(0.048)	(0.031)	(0.063)	(0.033)	(0.042)
Eff. # of Control Obs.	326	321	410	298	413	311	309	302
Eff. # of Treatment Obs.	800	777	949	745	936	743	743	659
Bandwidth	3.689	3.514	4.077	3.029	4.398	3.704	3.618	3.060
8. Fixed Bandwidth	0.010	0.014	0.002	0.001	0.027	0.054	0.036	0.062
	(0.039)	(0.044)	(0.047)	(0.052)	(0.038)	(0.052)	(0.034)	(0.032)
Eff. # of Control Obs.	690	679	675	650	660	660	656	641
Eff. # of Treatment Obs.	2,592	2,527	2,515	2,508	2,483	2,444	2,441	2,187
Bandwidth	8.600	8.600	8.600	8.600	8.600	8.600	8.600	8.600
9. Fixed Bandwidth and	0.003	-0.001	0.017	0.015	0.045	0.088	0.068	0.073
Always Employed	(0.036)	(0.049)	(0.053)	(0.054)	(0.044)	(0.053)	(0.043)	(0.041)
Eff. # of Control Obs.	461	461	461	461	461	461	461	461
Eff. # of Treatment Obs.	1,557	1,557	1,557	1,557	1,557	1,557	1,557	1,557
Bandwidth	8.600	8.600	8.600	8.600	8.600	8.600	8.600	8.600

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 30. Row 1 displays the baseline results from Table 3. Rows 2 and 3 show the results when we add year and location fixed effects (FE), respectively. Row 4 excludes controls. Row 5 uses the bandwidth that minimizes the coverage error (CE) of the local polynomial regression discontinuity estimator. Row 6 reports non-bias-corrected estimates. Row 7 does not adjust for the presence of mass points. Row 8 uses a fixed bandwidth. Row 9 also uses a fixed bandwidth and restricts attention to individuals that are employed in all eight quarters. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator, except for rows 5, 8, and 9. In all specifications, except for row 7, we report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA5: Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Intermediate Certificate on Log of Income: Robustness Checks

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
1. Baseline	-0.010	-0.002	-0.001	-0.013	-0.006	0.009	0.017	0.008
	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.018)	(0.018)
Eff. # of Control Obs.	9,611	9,921	9,808	9,319	11,296	11,248	11,393	10,147
Eff. # of Treatment Obs.	13,679	15,196	14,987	13,323	17,176	17,302	17,182	15,226
Bandwidth	8.798	9.714	9.357	8.753	10.300	10.790	11.000	10.410
2. Year FE	-0.012	-0.006	-0.003	-0.013	-0.008	0.006	0.016	0.008
	(0.014)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.018)	(0.018)
Eff. # of Control Obs.	9,608	9,913	9,808	9,319	11,295	11,248	11,128	10,147
Eff. # of Treatment Obs.	13,645	15,163	14,983	13,323	17,176	17,071	17,179	15,226
Bandwidth	8.653	9.544	9.272	8.753	10.200	10.540	10.910	10.410
3. Location FE	0.001	0.006	0.008	-0.005	0.002	0.016	0.024	0.018
	(0.013)	(0.013)	(0.012)	(0.013)	(0.016)	(0.016)	(0.018)	(0.018)
Eff. # of Control Obs.	10,058	11,673	9,817	9,727	11,296	11,518	11,393	10,147
Eff. # of Treatment Obs.	15,270	17,688	15,031	14,856	17,197	19,505	19,365	15,452
Bandwidth	9.412	10.210	9.647	9.183	10.450	11.020	11.030	10.680
4. No Controls	-0.009	0.000	-0.000	-0.010	-0.005	0.013	0.020	0.009
	(0.014)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.018)	(0.018)
Eff. # of Control Obs.	9,611	11,673	9,808	9,727	11,296	11,518	11,128	10,147
Eff. # of Treatment Obs.	13,692	17,688	14,991	14,865	17,196	19,505	17,182	15,227
Bandwidth	8.851	10.160	9.465	9.213	10.380	11.030	10.990	10.540
5. CE Bandwidth	-0.018	-0.012	-0.009	-0.016	-0.013	-0.003	0.012	0.006
	(0.015)	(0.014)	(0.014)	(0.014)	(0.018)	(0.018)	(0.019)	(0.020)
Eff. # of Control Obs.	7,235	7,191	7,103	7,024	8,059	8,053	7,996	7,281
Eff. # of Treatment Obs.	10,890	10,911	10,675	10,595	11,819	11,751	11,680	10,510
Bandwidth	6.139	6.777	6.528	6.107	7.183	7.530	7.671	7.278

Continues in the next page...

Table OA5 Continued. Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Intermediate Certificate on Log of Income: Robustness Checks

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
6. Conventional Estimates	-0.004	0.002	0.002	-0.008	-0.003	0.012	0.020	0.012
	(0.012)	(0.012)	(0.011)	(0.012)	(0.014)	(0.014)	(0.016)	(0.016)
Eff. # of Control Obs.	9,611	9,921	9,808	9,319	11,296	11,248	11,393	10,147
Eff. # of Treatment Obs.	13,679	15,196	14,987	13,323	17,176	17,302	17,182	15,226
Bandwidth	8.798	9.714	9.357	8.753	10.300	10.790	11.000	10.410
7. No Mass Points Adjustment	-0.018	-0.012	-0.009	-0.016	-0.013	-0.005	0.010	0.006
	(0.015)	(0.014)	(0.014)	(0.014)	(0.018)	(0.019)	(0.019)	(0.020)
Eff. # of Control Obs.	7,232	7,167	7,099	7,036	6,918	8,040	7,971	7,281
Eff. # of Treatment Obs.	10,890	10,781	10,673	10,595	10,498	11,745	11,661	10,510
Bandwidth	6.028	6.328	6.428	6.204	6.531	7.222	7.202	7.207
8. Fixed Bandwidth	-0.023	-0.019	-0.014	-0.020	-0.020	-0.014	-0.001	-0.002
	(0.015)	(0.015)	(0.015)	(0.016)	(0.019)	(0.020)	(0.019)	(0.023)
Eff. # of Control Obs.	10,069	9,922	9,818	9,738	9,569	9,550	9,445	8,661
Eff. # of Treatment Obs.	15,316	15,197	15,037	14,936	14,753	14,648	14,528	13,107
Bandwidth	9.900	9.900	9.900	9.900	9.900	9.900	9.900	9.900
9. Fixed Bandwidth and	-0.022	-0.021	-0.026	-0.035	-0.033	-0.036	-0.033	-0.033
Always Employed	(0.018)	(0.019)	(0.017)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
Eff. # of Control Obs.	6,114	6,114	6,114	6,114	6,114	6,114	6,114	6,114
Eff. # of Treatment Obs.	9,244	9,244	9,244	9,244	9,244	9,244	9,244	9,244
Bandwidth	9.900	9.900	9.900	9.900	9.900	9.900	9.900	9.900

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 60. Row 1 displays the baseline results from Table 3. Rows 2 and 3 show the results when we add year and location fixed effects (FE), respectively. Row 4 excludes controls. Row 5 uses the bandwidth that minimizes the coverage error (CE) of the local polynomial regression discontinuity estimator. Row 6 reports non-bias-corrected estimates. Row 7 does not adjust for the presence of mass points. Row 8 uses a fixed bandwidth. Row 9 also uses a fixed bandwidth and restricts attention to individuals that are employed in all eight quarters. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator, except for rows 5, 8, and 9. In all specifications, except for row 7, we report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA6: Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Advanced Certificate on Log of Income: Robustness Checks

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
1. Baseline	0.087	0.087	0.082	0.127	0.104	0.094	0.099	0.099
	(0.028)	(0.031)	(0.034)	(0.029)	(0.029)	(0.027)	(0.032)	(0.032)
Eff. # of Control Obs.	8,322	8,009	6,843	7,986	7,901	10,235	7,619	6,929
Eff. # of Treatment Obs.	18,098	17,852	14,410	17,669	17,400	23,044	16,776	14,754
Bandwidth	3.955	3.548	2.981	3.684	3.625	4.113	3.082	3.013
2. Year FE	0.094	0.092	0.085	0.127	0.110	0.102	0.101	0.099
	(0.029)	(0.032)	(0.034)	(0.029)	(0.030)	(0.031)	(0.032)	(0.032)
Eff. # of Control Obs.	8,313	7,998	6,843	7,986	7,884	7,937	7,619	6,929
Eff. # of Treatment Obs.	18,087	17,852	14,410	17,669	17,400	17,330	16,776	14,754
Bandwidth	3.845	3.542	2.981	3.684	3.589	3.798	3.076	3.013
3. Location FE	0.095	0.099	0.095	0.139	0.120	0.108	0.109	0.104
	(0.029)	(0.030)	(0.033)	(0.029)	(0.032)	(0.028)	(0.031)	(0.030)
Eff. # of Control Obs.	8,122	7,997	6,843	7,881	7,782	7,820	7,619	6,034
Eff. # of Treatment Obs.	18,020	17,849	14,410	17,609	16,980	17,283	16,776	12,363
Bandwidth	3.580	3.512	2.969	3.334	3.007	3.734	3.072	2.968
4. No Controls	0.093	0.095	0.089	0.142	0.111	0.110	0.102	0.092
	(0.033)	(0.035)	(0.031)	(0.039)	(0.034)	(0.039)	(0.030)	(0.033)
Eff. # of Control Obs.	8,098	7,983	8,027	6,803	7,857	6,669	7,864	7,021
Eff. # of Treatment Obs.	17,987	17,821	17,701	14,293	17,367	14,068	17,274	15,174
Bandwidth	3.478	3.406	3.736	2.793	3.474	2.959	3.811	3.595
5. CE Bandwidth	0.082	0.084	0.066	0.127	0.101	0.09	0.086	0.090
	(0.030)	(0.034)	(0.035)	(0.029)	(0.030)	(0.030)	(0.033)	(0.034)
Eff. # of Control Obs.	6,994	6,768	6,709	6,792	6,743	6,656	6,478	5,910
Eff. # of Treatment Obs.	14,625	14,389	14,143	14,238	14,054	14,004	13,750	12,101
Bandwidth	2.744	2.462	2.068	2.556	2.515	2.854	2.138	2.094

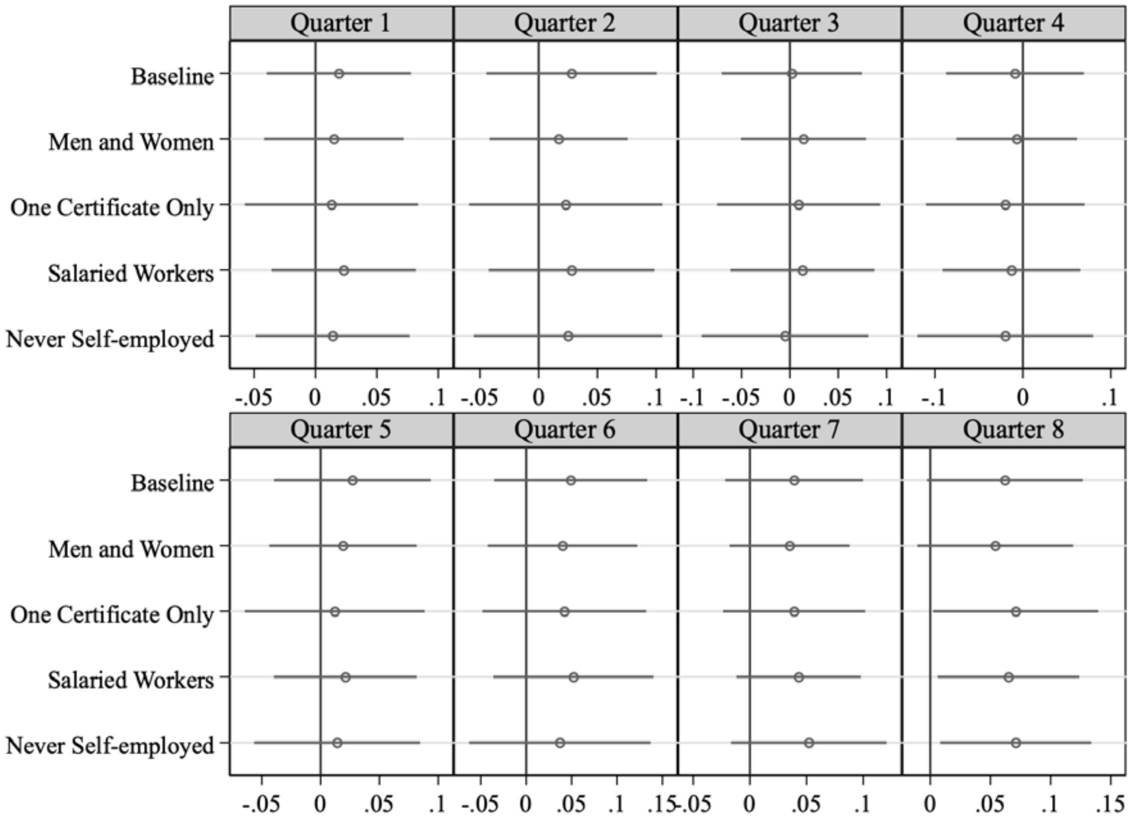
Continues in the next page...

Table OA6 Continued. Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Advanced Certificate on Log of Income: Robustness Checks

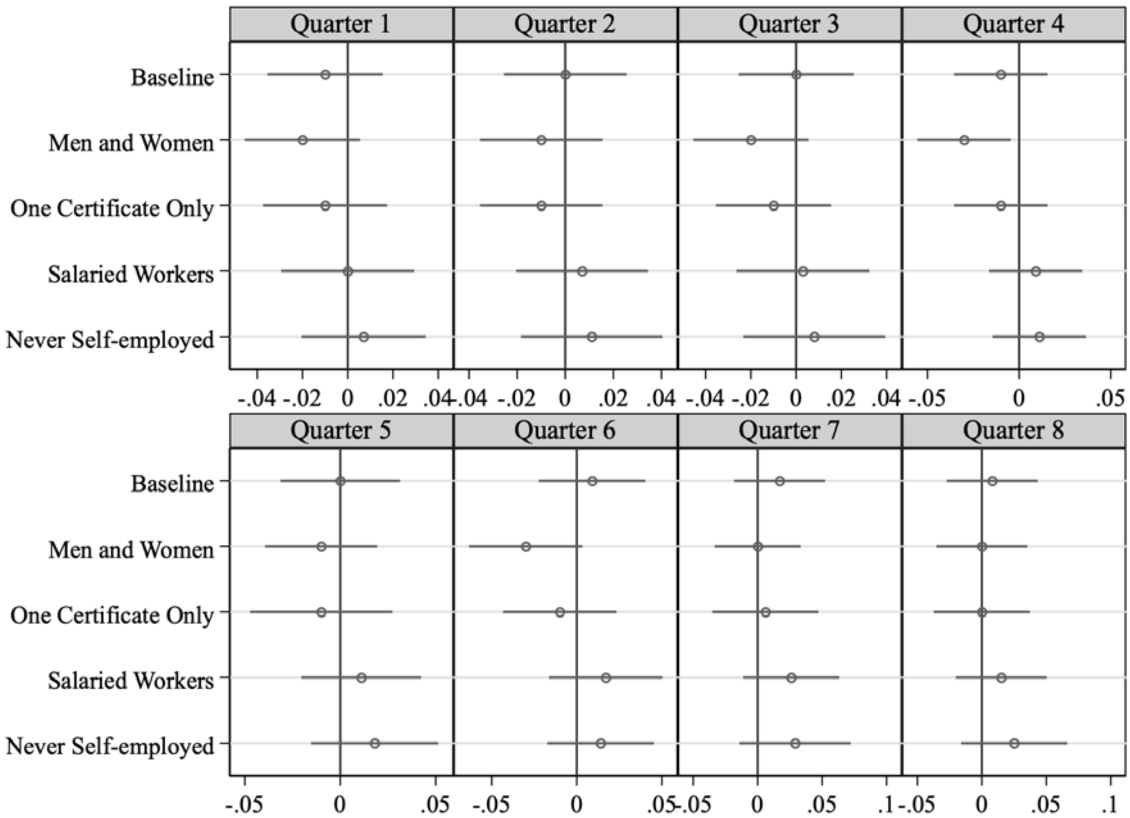
	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
6. Conventional Estimates	0.074	0.072	0.066	0.106	0.084	0.074	0.079	0.080
	(0.022)	(0.023)	(0.025)	(0.023)	(0.023)	(0.021)	(0.023)	(0.024)
Eff. # of Control Obs.	8,322	8,009	6,843	7,986	7,901	10,235	7,619	6,929
Eff. # of Treatment Obs.	18,098	17,852	14,410	17,669	17,400	23,044	16,776	14,754
Bandwidth	3.955	3.548	2.981	3.684	3.625	4.113	3.082	3.013
7. No Mass Points Adjustment	0.074	0.085	0.058	0.131	0.098	0.100	0.075	0.090
	(0.037)	(0.034)	(0.039)	(0.031)	(0.035)	(0.034)	(0.036)	(0.041)
Eff. # of Control Obs.	6,882	6,872	6,709	6,791	6,643	6,654	6,475	5,910
Eff. # of Treatment Obs.	14,495	14,479	14,143	14,238	13,927	13,989	13,750	12,101
Bandwidth	2.307	2.615	2.058	2.530	2.284	2.705	2.116	2.049
8. Fixed Bandwidth	0.053	0.061	0.063	0.117	0.092	0.083	0.076	0.086
	(0.037)	(0.038)	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.035)
Eff. # of Control Obs.	8,073	7,960	7,907	7,877	7,832	7,712	7,635	6,975
Eff. # of Treatment Obs.	17,925	17,760	17,603	17,577	17,303	17,181	17,105	15,081
Bandwidth	3.300	3.300	3.300	3.300	3.300	3.300	3.300	3.300
9. Fixed Bandwidth and	0.060	0.075	0.077	0.119	0.106	0.105	0.097	0.095
Always Employed	(0.040)	(0.037)	(0.034)	(0.036)	(0.035)	(0.035)	(0.035)	(0.036)
Eff. # of Control Obs.	5,100	5,100	5,100	5,100	5,100	5,100	5,100	5,100
Eff. # of Treatment Obs.	10,825	10,825	10,825	10,825	10,825	10,825	10,825	10,825
Bandwidth	3.300	3.300	3.300	3.300	3.300	3.300	3.300	3.300

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 90. Row 1 displays the baseline results from Table 3. Rows 2 and 3 show the results when we add year and location fixed effects (FE), respectively. Row 4 excludes controls. Row 5 uses the bandwidth that minimizes the coverage error (CE) of the local polynomial regression discontinuity estimator. Row 6 reports non-bias-corrected estimates. Row 7 does not adjust for the presence of mass points. Row 8 uses a fixed bandwidth. Row 9 also uses a fixed bandwidth and restricts attention to individuals that are employed in all eight quarters. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator, except for rows 5, 8, and 9. In all specifications, except for row 7, we report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Figure OA4: Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Certificate on Log of Income: Alternative Samples

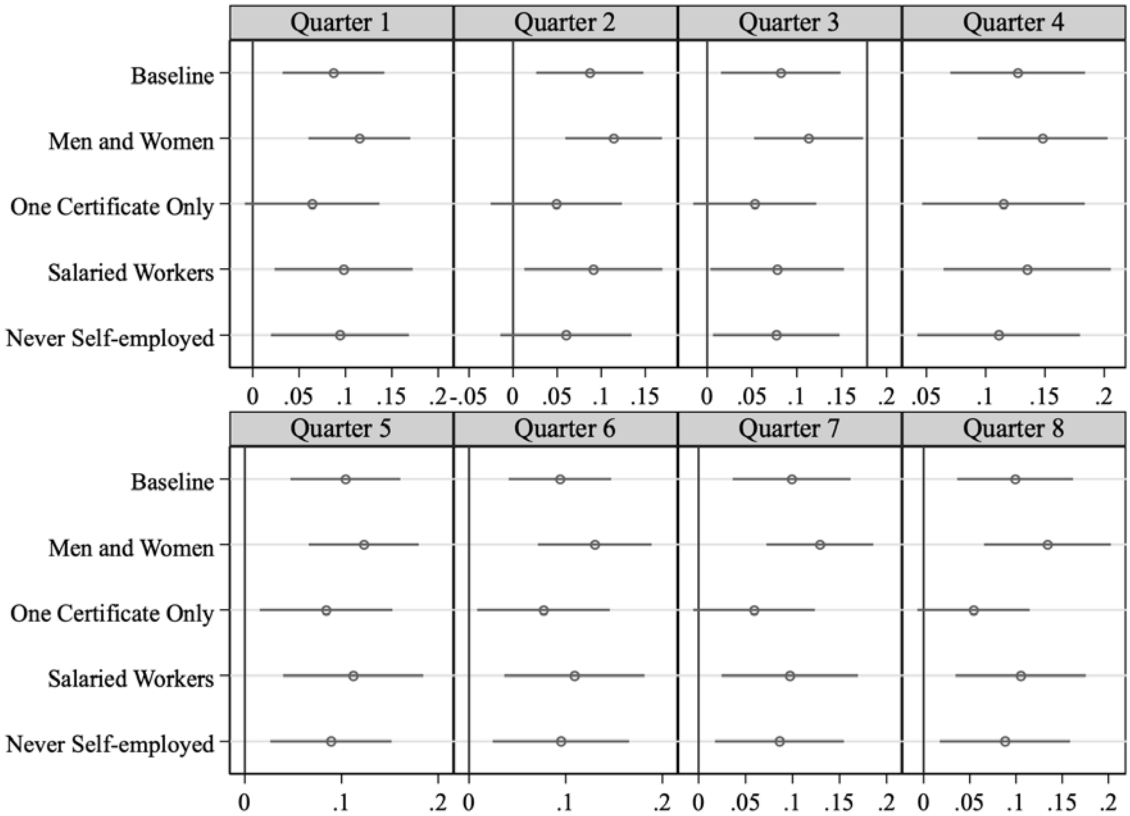


(a) Basic



(b) Intermediate

Figure OA4 Continued. Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Certificate on Log of Income: Alternative Samples



(c) Advanced

Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) and its 95% confidence interval, one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 30 for the basic certificate (panel a), 60 for the intermediate certificate (panel b), and 90 for the advanced certificate (panel c). In each plot, row 1 displays the baseline results from Table 3. Row 2 uses the full sample of men and women. Row 3 uses a sample of men who applied for only one certificate between 2017 and 2019. Row 4 uses a sample of salaried workers, while the fifth row further excludes individuals who are ever self-employed within two years of certification. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator, except for rows 5, 8, and 9. Standard errors used to compute 95% confidence intervals are clustered at the technical-norm level. Detailed results for all three certificates are displayed in Tables OA7, OA8, and OA9.

Table OA7: Sharp Regression Discontinuity Estimates of the Effects of Obtaining a Basic Certificate on Log of Income: Alternative Samples

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. Men and Women	0.015	0.017	0.014	-0.007	0.019	0.040	0.035	0.054
	(0.029)	(0.030)	(0.033)	(0.035)	(0.032)	(0.042)	(0.027)	(0.033)
# of Observations	235,346	231,969	229,967	228,652	226,496	224,094	222,619	201,123
Eff. # of Control Obs.	773	832	907	799	742	861	880	825
Eff. # of Treatment Obs.	2,666	2,988	3,627	2,939	2,517	3,288	3,548	2,951
Bandwidth	6.975	7.459	9.030	7.196	6.842	8.807	9.552	8.203
2. One Certificate Only	0.013	0.023	0.009	-0.020	0.012	0.042	0.039	0.071
	(0.036)	(0.042)	(0.043)	(0.046)	(0.039)	(0.046)	(0.032)	(0.035)
# of Observations	125,946	124,460	123,522	122,565	121,398	120,332	119,670	107,464
Eff. # of Control Obs.	500	537	527	509	482	608	483	510
Eff. # of Treatment Obs.	1,631	1,830	1,804	1,790	1,551	2,780	1,548	1,589
Bandwidth	6.810	7.333	7.649	7.161	6.651	10.26-	6.505	7.900
3. Salaried Worker	0.023	0.028	0.013	-0.013	0.021	0.052	0.043	0.065
	(0.030)	(0.036)	(0.038)	(0.040)	(0.031)	(0.045)	(0.028)	(0.030)
# of Observations	144,403	142,532	141,069	139,894	138,472	137,331	136,243	122,849
Eff. # of Control Obs.	661	615	633	653	600	591	579	535
Eff. # of Treatment Obs.	2,505	2,229	2,431	2,978	2,190	2,177	2,162	1,702
Bandwidth	9.410	8.173	9.158	10.46	8.613	8.886	8.361	7.977
4. Never Self-Employed	0.014	0.025	-0.005	-0.020	0.014	0.037	0.052	0.071
	(0.032)	(0.041)	(0.044)	(0.051)	(0.036)	(0.051)	(0.035)	(0.032)
# of Observations	125,383	123,881	123,040	122,412	121,235	120,425	119,612	107,859
Eff. # of Control Obs.	569	553	571	548	559	555	501	544
Eff. # of Treatment Obs.	1,997	1,944	2,117	2,106	2,092	2,088	1,662	1,821
Bandwidth	8.569	8.678	9.969	9.168	9.375	9.966	7.701	9.241

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 30. Row 1 uses the full sample of men and women. Row 2 uses a sample of men who apply for only one certificate between 2017 and 2019. Row 3 uses a sample of salaried workers, while row 4 further excludes individuals who are ever self-employed within two years of certification. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA8: Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Intermediate Certificate on Log of Income: Alternative Samples

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. Men and Women	-0.021	-0.016	-0.028	-0.033	-0.014	-0.030	-0.009	-0.008
	(0.013)	(0.013)	(0.013)	(0.013)	(0.015)	(0.017)	(0.017)	(0.018)
# of Observations	235,346	231,969	229,967	228,652	226,496	224,094	222,619	201,123
Eff. # of Control Obs.	12,060	11,949	9,896	9,814	12,655	9,636	11,354	10,343
Eff. # of Treatment Obs.	17,468	17,304	15,302	15,196	18,369	15,051	16,555	14,911
Bandwidth	7.209	7.768	6.584	6.326	8.095	6.828	7.739	7.667
2. One Certificate Only	-0.013	-0.010	-0.012	-0.018	-0.016	-0.011	0.006	-0.002
	(0.014)	(0.013)	(0.013)	(0.013)	(0.019)	(0.017)	(0.021)	(0.019)
# of Observations	125,946	124,460	123,522	122,565	121,398	120,332	119,670	107,464
Eff. # of Control Obs.	11,393	10,427	9,412	9,540	9,189	9,153	9,289	8,268
Eff. # of Treatment Obs.	18,774	17,868	13,990	15,897	13,897	13,631	15,501	12,310
Bandwidth	13.730	12.900	10.170	11.400	10.670	10.320	11.280	10.680
3. Salaried Worker	0.000	0.007	0.003	0.009	0.011	0.017	0.026	0.015
	(0.015)	(0.014)	(0.015)	(0.013)	(0.016)	(0.017)	(0.019)	(0.018)
# of Observations	144,403	142,532	141,069	139,894	138,472	137,331	136,243	122,849
Eff. # of Control Obs.	8,990	10,238	8,709	10,040	10,800	10,044	10,604	9,072
Eff. # of Treatment Obs.	13,060	14,730	12,787	14,361	17,906	16,330	17,678	14,624
Bandwidth	9.699	10.530	9.673	10.370	12.030	11.300	12.130	11.800
4. Never Self-Employed	0.007	0.011	0.008	0.011	0.018	0.014	0.029	0.025
	(0.014)	(0.015)	(0.016)	(0.013)	(0.017)	(0.016)	(0.022)	(0.021)
# of Observations	125,383	123,881	123,040	122,412	121,235	120,425	119,612	107,859
Eff. # of Control Obs.	8,894	9,617	8,650	8,865	10,118	8,664	9,212	8,424
Eff. # of Treatment Obs.	12,633	15,731	12,414	14,177	16,260	12,164	15,317	13,638
Bandwidth	10.730	12.160	10.930	11.430	13.020	11.000	12.800	12.320

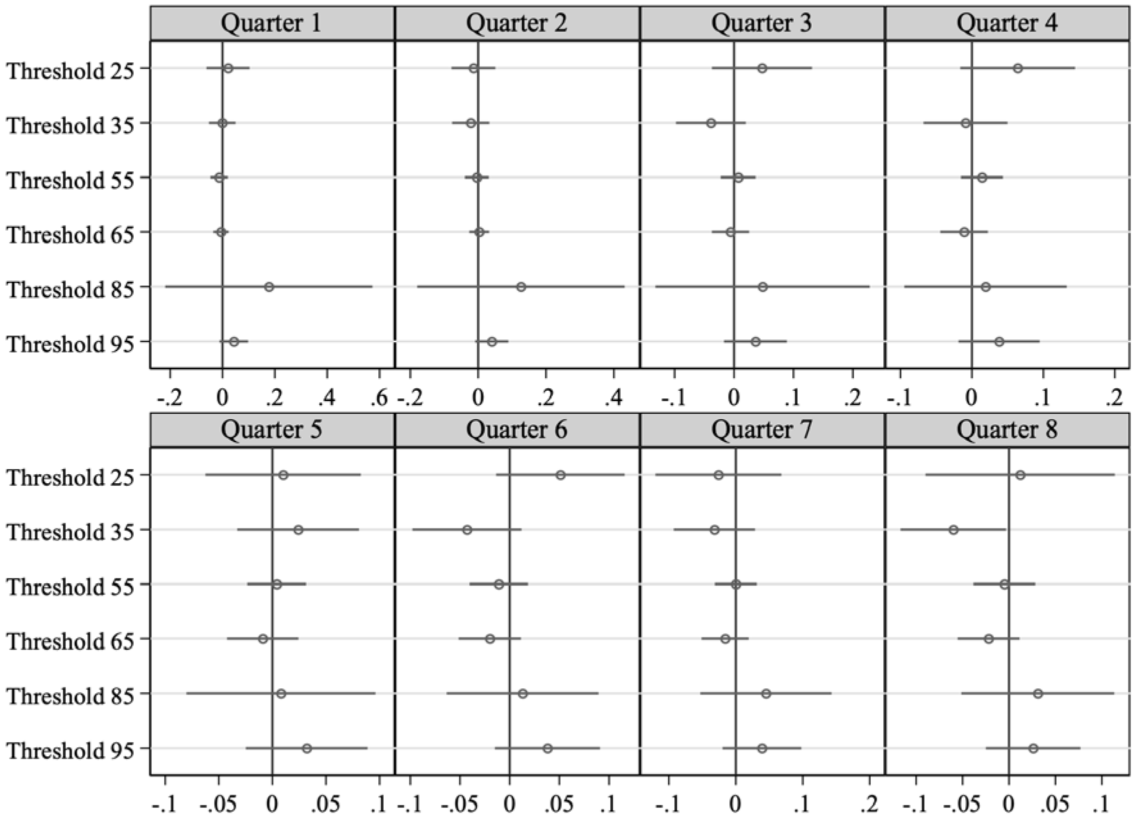
Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 60. Row 1 uses the full sample of men and women. Row 2 uses a sample of men who apply for only one certificate between 2017 and 2019. Row 3 uses a sample of salaried workers, while row 4 further excludes individuals who are ever self-employed within two years of certification. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA9: Sharp Regression Discontinuity Estimates of the Effects of Obtaining an Advanced Certificate on Log of Income: Alternative Samples

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. Men and Women	0.115	0.114	0.113	0.148	0.123	0.130	0.129	0.134
	(0.028)	(0.028)	(0.031)	(0.028)	(0.029)	(0.030)	(0.029)	(0.035)
# of Observations	235,346	231,969	229,967	228,652	226,496	224,094	222,619	201,123
Eff. # of Control Obs.	12,878	12,601	10,696	12,183	12,204	11,989	11,877	9,453
Eff. # of Treatment Obs.	24,914	24,688	20,115	24,147	24,011	23,787	23,630	17,351
Bandwidth	3.976	3.841	2.718	3.151	3.568	3.452	3.380	2.711
2. One Certificate Only	0.064	0.049	0.053	0.115	0.084	0.077	0.059	0.054
	(0.037)	(0.038)	(0.035)	(0.035)	(0.035)	(0.035)	(0.033)	(0.031)
# of Observations	125,946	124,460	123,522	122,565	121,398	120,332	119,670	107,464
Eff. # of Control Obs.	6,330	6,240	6,201	6,172	6,124	5,279	5,997	5,440
Eff. # of Treatment Obs.	13,671	13,298	13,191	13,206	12,980	10,669	12,829	11,237
Bandwidth	3.112	3.024	3.011	3.055	3.006	2.941	3.090	3.077
3. Salaried Worker	0.098	0.091	0.078	0.135	0.112	0.109	0.097	0.105
	(0.038)	(0.040)	(0.038)	(0.036)	(0.037)	(0.037)	(0.037)	(0.036)
# of Observations	144,403	142,532	141,069	139,894	138,472	137,331	136,243	122,849
Eff. # of Control Obs.	7,250	7,173	7,071	7,077	7,030	6,917	5,909	6,206
Eff. # of Treatment Obs.	15,470	15,352	15,138	15,076	14,852	15,080	12,060	12,892
Bandwidth	3.019	3.047	3.016	3.065	3.057	3.214	2.851	3.025
4. Never Self-Employed	0.094	0.060	0.077	0.111	0.089	0.095	0.086	0.088
	(0.038)	(0.038)	(0.036)	(0.035)	(0.032)	(0.036)	(0.035)	(0.036)
# of Observations	125,383	123,881	123,040	122,412	121,235	120,425	119,612	107,859
Eff. # of Control Obs.	6,337	6,241	6,216	6,219	6,257	6,087	5,191	5,463
Eff. # of Treatment Obs.	13,861	13,471	13,354	13,637	13,512	13,379	10,628	11,439
Bandwidth	3.182	3.051	3.095	3.155	3.740	3.237	2.811	3.019

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the discontinuity threshold is 90. Row 1 uses the full sample of men and women. Row 2 uses a sample of men who apply for only one certificate between 2017 and 2019. Row 3 uses a sample of salaried workers, while row 4 further excludes individuals who are ever self-employed within two years of certification. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical norm level. Bandwidths are displayed below the effective number of observations. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Figure OA5: Sharp Regression Discontinuity Estimates of the Effects on Log of Income - Placebo Thresholds



Notes: Each plot displays the estimate for the main coefficient of interest in Equation (1) its 95% confidence interval, one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score, and the (placebo) thresholds of interest are 25, 35, 55, 65, 85, and 95. All specifications include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator. Standard errors used to compute 95% confidence intervals are clustered at the technical norm level. Detailed results for all placebo thresholds are displayed in Table OA10.

Table OA10: Sharp Regression Discontinuity Estimates of the Effects on Log of Income - Placebo Thresholds

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
Threshold 25	0.021	-0.014	0.047	0.064	0.010	0.051	-0.026	0.012
	(0.042)	(0.033)	(0.043)	(0.041)	(0.037)	(0.033)	(0.048)	(0.052)
Eff. # of Control Obs.	260	358	256	272	246	311	246	236
Eff. # of Treatment Obs.	725	1,265	703	868	692	1,005	687	641
Bandwidth	5.331	8.912	5.093	6.277	5.437	7.770	5.266	5.233
Threshold 35	-0.001	-0.022	-0.039	-0.009	0.024	-0.043	-0.032	-0.060
	(0.026)	(0.028)	(0.030)	(0.030)	(0.029)	(0.028)	(0.031)	(0.029)
Eff. # of Control Obs.	1,446	1,473	1,263	1,277	1,273	1,252	1,255	1,171
Eff. # of Treatment Obs.	4,716	5,661	3,627	4,086	4,093	4,019	4,023	3,613
Bandwidth	9.549	10.060	7.265	8.724	8.097	8.285	8.347	8.911
Threshold 55	-0.013	-0.004	0.007	0.014	0.004	-0.011	-0.000	-0.005
	(0.017)	(0.018)	(0.015)	(0.015)	(0.014)	(0.015)	(0.016)	(0.017)
Eff. # of Control Obs.	9,707	9,556	10,190	8,901	10,027	8,695	9,521	7,824
Eff. # of Treatment Obs.	17,307	17,139	19,504	15,793	19,038	15,317	17,608	13,813
Bandwidth	11.950	11.740	13.490	10.990	13.380	10.230	12.420	10.250
Threshold 65	-0.006	0.003	-0.006	-0.011	-0.009	-0.020	-0.016	-0.022
	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)	(0.016)	(0.018)	(0.017)
Eff. # of Control Obs.	18,541	15,985	18,131	16,243	16,005	15,902	13,198	13,046
Eff. # of Treatment Obs.	33,475	27,920	32,882	29,086	28,727	28,301	22,386	22,321
Bandwidth	15.240	13.140	15.600	14.802	14.910	14.170	11.170	12.150

Continues in the Next Page...

Table OA10 Continued. Sharp Regression Discontinuity Estimates of the Effects on Log of Income - Placebo Thresholds

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
Threshold 85	0.177	0.126	0.048	0.019	0.008	0.013	0.045	0.031
	(0.202)	(0.156)	(0.092)	(0.058)	(0.045)	(0.039)	(0.050)	(0.042)
Eff. # of Control Obs.	4,945	5,086	5,388	7,651	7,689	7,554	7,552	6,723
Eff. # of Treatment Obs.	7,217	7,131	7,187	9,954	10,154	9,760	9,667	8,717
Bandwidth	2.126	2.376	2.651	3.050	3.280	3.065	3.121	3.163
Threshold 95	0.043	0.040	0.036	0.038	0.032	0.038	0.039	0.026
	(0.028)	(0.025)	(0.027)	(0.029)	(0.029)	(0.027)	(0.030)	(0.026)
Eff. # of Control Obs.	14,223	14,262	14,170	14,125	13,934	13,603	13,543	12,158
Eff. # of Treatment Obs.	13,838	13,735	13,595	13,541	13,417	13,268	13,192	11,454
Bandwidth	3.563	3.656	3.616	3.614	3.583	3.468	3.472	3.581

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1), one to eight quarters after certification. The outcome is the log of income. The running variable is the exam score and the (placebo) thresholds of interest are 25, 35, 55, 65, 85, and 95. All regressions include controls described in Section 4, use a triangular kernel, and choose the bandwidth to minimize the mean squared error (MSE) of the local polynomial RD estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Online Appendix C Additional Heterogeneity and Mechanisms Results

Table OA11: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate on Log of Income: Heterogeneity Analysis

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. By education level								
1a. Less than High School	-0.001	-0.028	0.010	0.018	0.063	0.034	0.026	0.042
	(0.044)	(0.046)	(0.045)	(0.044)	(0.071)	(0.039)	(0.044)	(0.045)
# of Observations	31,651	31,152	30,873	30,648	30,298	29,921	29,631	26,757
Eff. # of Control Obs.	193	164	166	175	138	176	179	193
Eff. # of Treatment Obs.	590	501	496	576	428	564	565	597
Bandwidth	7.497	6.045	6.304	7.357	5.442	7.759	7.603	8.827
1b. High School	-0.031	0.053	0.005	-0.040	0.045	0.018	0.014	0.068
	(0.032)	(0.037)	(0.046)	(0.039)	(0.034)	(0.062)	(0.033)	(0.036)
# of Observations	76,631	75,667	75,010	74,486	73,764	73,106	72,571	65,411
Eff. # of Control Obs.	314	283	310	350	231	283	276	267
Eff. # of Treatment Obs.	1,180	994	1,152	1,746	861	945	957	834
Bandwidth	7.008	6.852	7.450	10.560	5.222	6.546	6.758	6.511
1c. More than High School	-0.287	-0.041	-0.374	0.105	0.064	-0.032	-0.231	0.223
	(0.103)	(0.093)	(0.237)	(0.061)	(0.077)	(0.291)	(0.264)	(0.056)
# of Observations	55,960	55,428	55,029	54,685	54,282	53,868	53,752	48,511
Eff. # of Control Obs.	51	84	48	85	82	50	48	104
Eff. # of Treatment Obs.	163	177	121	174	172	152	114	163
Bandwidth	3.846	4.013	2.997	4.524	4.307	3.754	2.706	4.997
2. By Initial Employment Status								
2a. Unemployed	0.104	0.169	-0.062	0.020	0.115	0.249	0.137	0.317
	(0.088)	(0.074)	(0.318)	(0.249)	(0.079)	(0.074)	(0.094)	(0.083)
# of Observations	7,676	8,156	8,378	8,521	8,800	8,936	8,967	8,158
Eff. # of Control Obs.	32	33	34	38	36	39	32	25
Eff. # of Treatment Obs.	115	91	129	152	112	115	104	58
Bandwidth	7.078	5.381	7.486	8.805	6.923	6.699	5.366	4.597
2b. Self-Employed	0.088	0.216	0.330	0.372	0.373	0.200	0.273	0.439
	(0.042)	(0.079)	(0.184)	(0.175)	(0.173)	(0.122)	(0.109)	(0.163)
# of Observations	14,972	14,495	14,305	14,182	13,983	13,800	13,729	12,267
Eff. # of Control Obs.	12	11	12	10	10	35	34	11
Eff. # of Treatment Obs.	80	75	77	74	72	198	197	62
Bandwidth	3.707	3.326	3.713	3.625	3.714	7.938	8.571	3.222
2c. Salaried Worker	0.023	0.031	0.037	0.020	0.019	0.042	0.049	0.066
	(0.029)	(0.037)	(0.033)	(0.039)	(0.030)	(0.044)	(0.032)	(0.032)
# of Observations	138,151	136,228	134,891	133,826	132,261	130,891	130,005	117,239
Eff. # of Control Obs.	627	556	591	568	540	617	534	571
Eff. # of Treatment Obs.	2,385	1,861	2,112	2,102	1,823	2,734	1,784	1,980
Bandwidth	9.993	7.352	8.788	8.523	7.748	10.190	7.379	9.161

Continues in the Next Page...

Table OA11 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate on Log of Income: Heterogeneity Analysis

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
3. By Tenure in the Firm								
3a. Less than Three Years	-0.055	0.006	0.004	0.006	-0.020	-0.045	-0.057	0.024
	(0.046)	(0.043)	(0.037)	(0.045)	(0.038)	(0.112)	(0.045)	(0.046)
# of Observations	40,234	39,694	39,174	38,888	38,418	37,953	37,795	35,180
Eff. # of Control Obs.	187	206	216	148	221	195	153	194
Eff. # of Treatment Obs.	502	568	628	416	690	536	409	496
Bandwidth	6.496	7.276	8.347	5.667	9.711	7.196	5.953	7.106
3b. Three or More Years	0.032	0.014	0.047	0.027	0.025	0.065	0.066	0.061
	(0.034)	(0.036)	(0.043)	(0.046)	(0.032)	(0.040)	(0.031)	(0.036)
# of Observations	97,917	96,534	95,717	94,938	93,843	92,938	92,210	82,059
Eff. # of Control Obs.	359	376	374	359	361	358	356	350
Eff. # of Treatment Obs.	1,317	1,474	1,473	1,468	1,448	1,422	1,432	1,247
Bandwidth	7.288	8.649	8.139	8.083	8.559	8.726	8.682	8.058
4. By Firm Size								
4a. 1 to 50 Workers	0.020	-0.007	0.026	0.012	0.032	0.033	0.166	0.250
	(0.074)	(0.063)	(0.060)	(0.067)	(0.073)	(0.064)	(0.058)	(0.084)
# of Observations	20,362	20,065	19,730	19,547	19,400	19,127	18,888	17,200
Eff. # of Control Obs.	58	69	71	65	61	56	51	51
Eff. # of Treatment Obs.	226	319	315	304	263	267	209	183
Bandwidth	6.440	8.394	8.348	8.412	7.731	7.948	6.580	6.402
4b. More Than 50 Workers	0.027	0.017	0.044	0.024	0.023	0.053	0.051	0.046
	(0.032)	(0.036)	(0.038)	(0.046)	(0.038)	(0.050)	(0.030)	(0.032)
# of Observations	117,571	115,948	114,951	114,075	112,656	111,561	110,914	99,859
Eff. # of Control Obs.	531	562	486	469	422	548	362	459
Eff. # of Treatment Obs.	1,848	2,438	1,585	1,576	1,345	2,337	1,211	1,350
Bandwidth	8.591	10.260	7.852	7.408	6.437	10.910	5.362	7.469

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the basic certificate, one to eight quarters after certification. The outcome is the log of income. The first panel displays the results by initial education level. The second panel presents the results by initial employment status.The third and fourth panels present the results for salaried workers (at certification), by tenure and firm size, respectively. The running variable is the exam score and the discontinuity threshold is 30. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA12: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate on Log of Income: Heterogeneity Analysis

	(1) **Quarter 1**	(2) **Quarter 2**	(3) **Quarter 3**	(4) **Quarter 4**	(5) **Quarter 5**	(6) **Quarter 6**	(7) **Quarter 7**	(8) **Quarter 8**
1. By education level								
1a. Less than High School	-0.007	-0.001	-0.022	-0.033	-0.009	-0.032	0.015	0.001
	(0.022)	(0.027)	(0.023)	(0.024)	(0.025)	(0.034)	(0.034)	(0.036)
# of Observations	31,651	31,152	30,873	30,648	30,298	29,921	29,631	26,757
Eff. # of Control Obs.	2,672	2,794	1,933	1,914	2,361	1,545	2,210	2,013
Eff. # of Treatment Obs.	4,061	4,278	2,884	2,858	3,631	2,138	3,115	2,812
Bandwidth	12.470	13.010	9.402	9.272	11.630	7.385	10.740	10.470
1b. High School	0.007	0.000	-0.000	0.008	0.015	0.030	0.033	0.019
	(0.014)	(0.014)	(0.014)	(0.015)	(0.018)	(0.019)	(0.020)	(0.019)
# of Observations	76,631	75,667	75,010	74,486	73,764	73,106	72,571	65,411
Eff. # of Control Obs.	6,687	7,788	7,340	7,051	6,392	6,381	6,822	6,471
Eff. # of Treatment Obs.	10,548	12,765	11,851	10,849	10,171	10,120	10,559	10,286
Bandwidth	12.650	15.280	14.350	13.950	12.700	12.910	13.990	14.780
1c. More than High School	-0.092	-0.063	-0.032	-0.083	-0.091	-0.084	-0.050	-0.045
	(0.027)	(0.026)	(0.024)	(0.029)	(0.033)	(0.030)	(0.029)	(0.030)
# of Observations	55,960	55,428	55,029	54,685	54,282	53,868	53,752	48,511
Eff. # of Control Obs.	2,088	2,058	2,390	2,023	2,304	2,288	2,297	2,110
Eff. # of Treatment Obs.	4,009	3,959	4,252	3,928	4,181	4,169	4,176	3,727
Bandwidth	6.060	6.393	7.826	6.865	7.463	7.091	7.928	7.978
2. By Initial Employment Status								
2a. Unemployed	0.058	0.046	0.028	-0.002	-0.033	0.001	0.091	0.067
	(0.044)	(0.039)	(0.052)	(0.055)	(0.046)	(0.054)	(0.049)	(0.059)
# of Observations	7,676	8,156	8,378	8,521	8,800	8,936	8,967	8,158
Eff. # of Control Obs.	709	790	644	712	726	760	694	744
Eff. # of Treatment Obs.	1,139	1,330	992	1,220	1,270	1,277	1,177	1,246
Bandwidth	13.440	14.650	10.950	12.670	12.790	12.720	11.680	14.000
2b. Self-Employed	-0.055	-0.023	0.008	-0.111	-0.109	-0.062	-0.109	-0.112
	(0.046)	(0.035)	(0.033)	(0.055)	(0.048)	(0.041)	(0.039)	(0.036)
# of Observations	14,972	14,495	14,305	14,182	13,983	13,800	13,729	12,267
Eff. # of Control Obs.	1,136	1,143	771	1,061	734	772	736	680
Eff. # of Treatment Obs.	2,483	2,652	1,614	2,387	1,567	1,642	1,543	1,401
Bandwidth	10.410	12.660	8.788	10.780	8.834	9.413	8.486	8.885
2c. Salaried Worker	-0.003	0.002	0.001	0.001	0.016	0.020	0.029	0.018
	(0.014)	(0.014)	(0.015)	(0.012)	(0.016)	(0.016)	(0.020)	(0.019)
# of Observations	138,151	136,228	134,891	133,826	132,261	130,891	130,005	117,239
Eff. # of Control Obs.	8,519	9,688	9,577	8,219	10,291	10,181	10,097	8,643
Eff. # of Treatment Obs.	12,392	14,194	13,763	12,062	17,102	16,935	16,999	13,899
Bandwidth	9.467	10.950	10.060	9.698	12.080	12.160	12.550	11.970

Continues in the Next Page...

Table OA12 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate on Log of Income: Heterogeneity Analysis

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
3. By Tenure in the Firm								
3a. Less than Three Years	-0.024	-0.011	-0.023	-0.021	0.050	0.030	0.062	0.060
	(0.023)	(0.020)	(0.019)	(0.018)	(0.037)	(0.035)	(0.039)	(0.039)
# of Observations	40,234	39,694	39,174	38,888	38,418	37,953	37,795	35,180
Eff. # of Control Obs.	3,210	3,162	3,605	3,003	2,523	3,004	2,961	2,836
Eff. # of Treatment Obs.	4,755	4,737	5,370	3,976	3,451	4,525	4,446	4,151
Bandwidth	11.860	11.660	13.140	10.160	9.383	11.570	11.140	11.960
3b. Three or More Years	0.013	0.011	0.012	0.012	0.000	0.015	0.014	0.000
	(0.016)	(0.017)	(0.017)	(0.015)	(0.016)	(0.017)	(0.019)	(0.018)
# of Observations	97,917	96,534	95,717	94,938	93,843	92,938	92,210	82,059
Eff. # of Control Obs.	6,758	6,811	6,717	6,507	6,416	6,334	6,432	5,628
Eff. # of Treatment Obs.	10,157	11,567	11,350	9,661	9,544	9,624	10,911	8,449
Bandwidth	10.980	11.550	11.390	10.080	10.350	10.750	11.090	10.750
4. By Firm Size								
4a. 1 to 50 Workers	0.031	0.022	0.012	-0.038	0.014	0.009	-0.001	0.010
	(0.041)	(0.031)	(0.031)	(0.041)	(0.034)	(0.028)	(0.030)	(0.033)
# of Observations	20,362	20,065	19,730	19,547	19,400	19,127	18,888	17,200
Eff. # of Control Obs.	1,035	1,264	1,023	852	1,046	1,196	1,007	1,045
Eff. # of Treatment Obs.	1,525	1,953	1,512	1,276	1,682	1,881	1,622	1,655
Bandwidth	8.290	10.830	8.970	7.302	9.569	10.970	9.273	10.990
4b. More Than 50 Workers	-0.001	-0.003	-0.003	0.004	0.016	0.020	0.034	0.020
	(0.014)	(0.016)	(0.016)	(0.013)	(0.016)	(0.017)	(0.021)	(0.021)
# of Observations	117,571	115,948	114,951	114,075	112,656	111,561	110,914	99,859
Eff. # of Control Obs.	10,518	8,412	8,322	8,537	8,946	8,876	9,495	8,072
Eff. # of Treatment Obs.	17,584	12,222	12,055	13,764	14,941	14,796	15,479	13,191
Bandwidth	14.190	10.910	10.750	11.860	12.660	12.710	13.470	12.480

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the intermediate certificate, one to eight quarters after certification. The outcome is the log of income. The first panel displays the results by initial education level. The second panel presents the results by initial employment status. The third and fourth panels present the results for salaried workers (at certification), by tenure and firm size, respectively. The running variable is the exam score and the discontinuity threshold is 60. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA13: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate on Log of Income: Heterogeneity Analysis

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. By education level								
1a. Less than High School	0.078	0.065	0.039	0.181	0.073	0.057	0.034	0.053
	(0.042)	(0.043)	(0.040)	(0.072)	(0.044)	(0.046)	(0.039)	(0.049)
# of Observations	31,651	31,152	30,873	30,648	30,298	29,921	29,631	26,757
Eff. # of Control Obs.	1,817	1,269	1,770	1,186	1,749	1,738	1,623	1,566
Eff. # of Treatment Obs.	4,364	3,422	4,268	3,249	4,169	4,146	3,950	3,572
Bandwidth	4.687	3.851	4.726	3.047	4.653	4.461	4.029	4.504
1b. High School	0.054	0.056	0.063	0.082	0.079	0.093	0.072	0.073
	(0.035)	(0.036)	(0.035)	(0.037)	(0.035)	(0.038)	(0.034)	(0.035)
# of Observations	76,631	75,667	75,010	74,486	73,764	73,106	72,571	65,411
Eff. # of Control Obs.	3,651	3,576	3,565	3,548	4,055	3,976	3,423	3,580
Eff. # of Treatment Obs.	6,438	6,351	6,300	6,264	7,748	7,682	6,105	6,737
Bandwidth	2.951	2.963	2.990	2.968	3.176	3.204	2.930	3.184
1c. More than High School	0.151	0.125	0.115	0.184	0.155	0.133	0.170	0.099
	(0.059)	(0.056)	(0.062)	(0.062)	(0.057)	(0.076)	(0.075)	(0.054)
# of Observations	55,960	55,428	55,029	54,685	54,282	53,868	53,752	48,511
Eff. # of Control Obs.	2,318	2,716	2,654	2,254	2,625	2,231	2,212	2,933
Eff. # of Treatment Obs.	5,301	6,495	6,384	5,211	6,294	5,100	5,129	7,742
Bandwidth	2.917	3.866	3.477	2.790	3.470	2.804	2.850	4.232
2. By Initial Employment Status								
2a. Unemployed	-0.067	-0.149	-0.005	0.117	-0.218	0.171	0.050	0.122
	(0.105)	(0.084)	(0.073)	(0.073)	(0.118)	(0.116)	(0.068)	(0.088)
# of Observations	7,676	8,156	8,378	8,521	8,800	8,936	8,967	8,158
Eff. # of Control Obs.	333	294	665	372	318	723	375	660
Eff. # of Treatment Obs.	915	841	1,643	1,038	875	1,733	1,094	1,576
Bandwidth	3.056	2.945	5.326	3.395	2.594	5.445	3.429	5.248
2b. Self-Employed	0.005	0.102	0.088	0.155	0.093	0.211	0.118	0.200
	(0.037)	(0.048)	(0.064)	(0.069)	(0.043)	(0.069)	(0.057)	(0.068)
# of Observations	14,972	14,495	14,305	14,182	13,983	13,800	13,729	12,267
Eff. # of Control Obs.	1,383	511	502	415	640	418	487	424
Eff. # of Treatment Obs.	2,924	1,595	1,591	1,392	2,008	1,351	1,525	1,345
Bandwidth	5.965	3.904	3.133	2.287	4.034	2.266	3.703	3.085
2c. Salaried Worker	0.100	0.097	0.086	0.141	0.122	0.106	0.105	0.101
	(0.037)	(0.039)	(0.036)	(0.035)	(0.035)	(0.034)	(0.034)	(0.034)
# of Observations	138,151	136,228	134,891	133,826	132,261	130,891	130,005	117,239
Eff. # of Control Obs.	6,984	6,009	5,964	6,790	6,741	6,671	6,594	6,038
Eff. # of Treatment Obs.	12,170	12,021	11,855	14,364	14,090	14,267	13,887	12,181
Bandwidth	3.003	2.978	2.990	3.011	3.037	3.182	3.056	3.071

Continues in the Next Page...

Table OA13 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate on Log of Income: Heterogeneity Analysis

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
3. By Tenure in the Firm								
3a. Less than Three Years	0.025	0.028	-0.068	0.062	0.035	0.021	-0.022	-0.021
	(0.041)	(0.046)	(0.060)	(0.047)	(0.051)	(0.053)	(0.064)	(0.067)
# of Observations	40,234	39,694	39,174	38,888	38,418	37,953	37,795	35,180
Eff. # of Control Obs.	1,788	1,734	1,703	1,733	1,701	1,645	1,643	1,534
Eff. # of Treatment Obs.	3,167	3,117	3,033	3,087	2,955	2,914	2,206	2,633
Bandwidth	2.880	2.609	2.091	2.647	2.202	2.211	2.004	2.037
3b. Three or More Years	0.112	0.103	0.120	0.150	0.139	0.107	0.114	0.108
	(0.039)	(0.038)	(0.035)	(0.037)	(0.040)	(0.033)	(0.036)	(0.035)
# of Observations	97,917	96,534	95,717	94,938	93,843	92,938	92,210	82,059
Eff. # of Control Obs.	4,962	4,957	4,875	4,835	4,122	4,883	4,695	4,274
Eff. # of Treatment Obs.	11,207	11,151	10,978	10,919	8,575	10,762	10,640	9,291
Bandwidth	3.207	3.633	3.472	3.322	2.769	3.832	3.582	3.599
4. By Firm Size								
4a. 1 to 50 Workers	0.076	0.135	0.064	0.113	0.189	0.254	0.166	0.267
	(0.094)	(0.084)	(0.140)	(0.127)	(0.120)	(0.140)	(0.104)	(0.170)
# of Observations	20,362	20,065	19,730	19,547	19,400	19,127	18,888	17,200
Eff. # of Control Obs.	724	1,580	689	832	813	797	1,133	731
Eff. # of Treatment Obs.	2,015	3,968	1,941	2,290	2,240	2,238	2,944	1,944
Bandwidth	2.582	5.251	2.966	3.178	3.108	3.538	4.638	3.254
4b. More Than 50 Workers	0.111	0.063	0.081	0.143	0.115	0.085	0.086	0.085
	(0.038)	(0.043)	(0.039)	(0.035)	(0.035)	(0.034)	(0.035)	(0.034)
# of Observations	117,571	115,948	114,951	114,075	112,656	111,561	110,914	99,859
Eff. # of Control Obs.	6,136	5,183	5,173	5,213	5,911	5,111	4,978	4,626
Eff. # of Treatment Obs.	10,125	9,867	9,774	9,825	9,730	9,570	9,457	8,419
Bandwidth	3.003	2.313	2.433	2.809	2.997	2.796	2.451	2.926

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the advanced certificate, one to eight quarters after certification. The outcome is the log of income. The first panel displays the results by initial education level. The second panel presents the results by initial employment status. The third and fourth panels present the results for salaried workers (at certification), by tenure and firm size, respectively. The running variable is the exam score and the discontinuity threshold is 90. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given variation in the number of individuals with positive earnings and the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA14: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate, by Initial Employment Status: Additional Outcomes

Outcome	**(1) Quarter 1**	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
1. Initially Unemployed								
1a. Salaried Work	0.141	0.095	0.195	0.393	0.047	0.171	0.147	-0.025
	(0.108)	(0.105)	(0.096)	(0.129)	(0.124)	(0.117)	(0.111)	(0.194)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	72	67	72	65	67	67	72	44
Eff. # of Treatment Obs.	272	237	272	185	234	237	272	101
Bandwidth	8.101	7.309	8.739	6.554	7.054	7.205	8.110	4.838
1b. Self-Employment	-0.081	-0.191	-0.067	-0.108	-0.003	0.124	0.078	-0.054
	(0.069)	(0.082)	(0.071)	(0.062)	(0.068)	(0.070)	(0.063)	(0.051)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	44	65	44	65	56	56	44	72
Eff. # of Treatment Obs.	102	184	105	184	164	162	105	221
Bandwidth	4.323	6.130	4.747	6.399	5.946	5.032	4.421	8.003
2. Initially Self-Employed								
2a. Salaried Work	0.150	0.285	0.217	0.256	0.259	0.288	0.477	0.302
	(0.032)	(0.069)	(0.042)	(0.073)	(0.090)	(0.170)	(0.149)	(0.125)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	15	13	32	32	15	13	14	15
Eff. # of Treatment Obs.	100	93	215	217	99	93	98	92
Bandwidth	4.639	3.853	5.079	5.430	4.299	3.726	4.102	4.647
2b. Self-Employment	-0.134	-0.170	-0.218	0.084	-0.127	-0.582	-0.645	-0.739
	(0.149)	(0.141)	(0.126)	(0.263)	(0.266)	(0.108)	(0.110)	(0.115)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	13	13	13	15	15	14	13	32
Eff. # of Treatment Obs.	93	89	93	99	99	98	89	92
Bandwidth	3.701	3.556	3.659	4.499	4.596	4.079	3.454	5.005

Table OA14 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining a Basic Certificate, by Initial Employment Status: Additional Outcomes

Outcome	**(1) Quarter 1**	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
3. Initially Salaried Worker								
3a. Salaried Work	0.007	0.005	-0.033	0.000	-0.012	-0.054	-0.007	-0.047
	(0.017)	(0.025)	(0.029)	(0.034)	(0.031)	(0.040)	(0.038)	(0.041)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	668	604	455	544	603	455	529	544
Eff. # of Treatment Obs.	2,531	2,039	1,587	1,750	2,039	1,577	1,743	1,553
Bandwidth	9.933	7.800	5.649	6.754	7.653	5.406	6.168	6.719
3b. Self-Employment	0.001	-0.003	0.014	0.006	-0.016	-0.002	-0.024	-0.020
	(0.003)	(0.011)	(0.011)	(0.011)	(0.015)	(0.009)	(0.018)	(0.020)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	596	530	530	530	407	544	603	647
Eff. # of Treatment Obs.	2,007	1,743	1,743	1,743	910	1,749	2,039	2,076
Bandwidth	7.052	6.186	6.204	6.274	4.835	6.596	7.758	8.866

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the basic certificate, one to eight quarters after certification, by initial employment status: unemployed (first panel), self-employed (second panel), salaried worker (third panel). The two outcomes analyzed are self-employment and salaried work. The running variable is the exam score and the discontinuity threshold is 30. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given the variation in the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA15: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate, by Initial Employment Status: Additional Outcomes

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
1. Initially Unemployed								
1a. Salaried Work	-0.007	0.018	0.028	-0.039	-0.058	-0.086	-0.035	-0.090
	(0.038)	(0.037)	(0.039)	(0.041)	(0.038)	(0.043)	(0.036)	(0.045)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	1,334	1,173	1,245	1,553	1,334	1,173	1,505	1,132
Eff. # of Treatment Obs.	2,276	2,006	2,147	2,920	2,273	2,006	2,715	1,969
Bandwidth	13.580	11.690	12.730	16.400	13.360	11.770	15.100	12.050
1b. Self-Employment	-0.002	0.025	-0.003	0.029	0.024	0.009	0.007	-0.023
	(0.019)	(0.018)	(0.019)	(0.021)	(0.016)	(0.022)	(0.021)	(0.024)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	1,165	941	1,413	899	1,407	1,326	1,334	866
Eff. # of Treatment Obs.	1,981	1,529	2,456	1,392	2,443	2,268	2,277	1,407
Bandwidth	11.150	9.499	14.590	8.830	14.110	13.040	13.670	9.176
2. Initially Self-Employed								
2a. Salaried Work	-0.030	-0.047	-0.065	-0.068	-0.045	-0.041	-0.044	-0.036
	(0.022)	(0.033)	(0.027)	(0.029)	(0.032)	(0.025)	(0.026)	(0.032)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	1,264	1,264	924	1,264	924	1,364	1,392	1,181
Eff. # of Treatment Obs.	2,744	2,744	1,986	2,719	1,986	3,103	3,220	2,614
Bandwidth	10.760	10.750	9.887	10.150	9.906	13.440	14.520	11.560
2b. Self-Employment	0.038	0.013	0.048	0.097	0.015	-0.010	0.031	0.008
	(0.044)	(0.047)	(0.047)	(0.039)	(0.041)	(0.036)	(0.055)	(0.044)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	887	1,264	1,264	916	1,278	1,314	924	1,181
Eff. # of Treatment Obs.	1,871	2,719	2,719	1,986	2,930	3,019	1,986	2,614
Bandwidth	8.349	10.290	10.150	9.232	11.040	12.380	9.990	11.440

Continues in the Next Page...

Table OA15 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Intermediate Certificate, by Initial Employment Status: Additional Outcomes

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
3. Initially Salaried Worker								
3a. Salaried Work	-0.015	-0.014	-0.011	-0.010	-0.011	0.003	0.009	0.009
	(0.010)	(0.009)	(0.011)	(0.009)	(0.011)	(0.011)	(0.012)	(0.012)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	12,916	8,628	9,038	12,462	10,477	12,452	10,793	11,848
Eff. # of Treatment Obs.	21,905	11,749	13,203	20,309	14,994	20,265	17,522	19,823
Bandwidth	14.350	8.753	9.128	13.800	10.280	13.340	11.700	14.900
3b. Self-Employment	0.007	0.002	0.003	0.006	0.011	0.003	0.005	0.005
	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)	(0.003)	(0.004)	(0.005)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	10,477	10,825	12,918	10,776	12,462	12,459	10,478	11,400
Eff. # of Treatment Obs.	14,994	17,525	21,905	17,522	20,312	20,273	15,017	18,341
Bandwidth	10.280	11.900	14.520	11.510	13.830	13.510	10.510	13.730

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the intermediate certificate, one to eight quarters after certification, by initial employment status: unemployed (first panel), self-employed (second panel), salaried worker (third panel). The two outcomes analyzed are self-employment and salaried work. The running variable is the exam score and the discontinuity threshold is 60. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given the variation in the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.

Table OA16: Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate, by Initial Employment Status: Additional Outcomes

	(1) Quarter 1	(2) Quarter 2	(3) Quarter 3	(4) Quarter 4	(5) Quarter 5	(6) Quarter 6	(7) Quarter 7	(8) Quarter 8
1. Initially Unemployed								
1a. Salaried Work	0.181	0.046	0.084	-0.005	0.070	-0.058	-0.046	-0.087
	(0.132)	(0.104)	(0.105)	(0.080)	(0.106)	(0.080)	(0.140)	(0.122)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	497	497	497	631	497	615	493	458
Eff. # of Treatment Obs.	1,515	1,519	1,518	1,851	1,514	1,837	1,499	1,314
Bandwidth	2.593	2.715	2.612	3.937	2.521	3.568	2.402	2.694
1b. Self-Employment	-0.059	-0.089	-0.043	-0.045	-0.046	-0.018	-0.013	0.007
	(0.060)	(0.048)	(0.041)	(0.052)	(0.051)	(0.051)	(0.035)	(0.040)
# of Observations	14,851	14,851	14,851	14,851	14,851	14,851	14,851	13,485
Eff. # of Control Obs.	271	494	610	606	606	497	894	784
Eff. # of Treatment Obs.	1,126	1,499	1,833	1,824	1,824	1,519	2,321	1,978
Bandwidth	1.975	2.458	3.444	3.149	3.147	2.714	4.963	4.182
2. Initially Self-Employed								
2a. Salaried Work	-0.009	0.018	0.062	0.189	0.135	0.166	0.201	0.127
	(0.030)	(0.040)	(0.051)	(0.058)	(0.059)	(0.050)	(0.058)	(0.057)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	581	581	494	223	224	503	494	513
Eff. # of Treatment Obs.	1,807	1,808	1,603	1,312	1,312	1,616	1,603	1,438
Bandwidth	3.162	3.251	2.057	1.816	1.936	2.583	2.059	3.001
2b. Self-Employment	0.049	-0.008	-0.070	-0.100	-0.057	-0.104	-0.326	-0.370
	(0.051)	(0.055)	(0.085)	(0.077)	(0.085)	(0.081)	(0.175)	(0.230)
# of Observations	16,502	16,502	16,502	16,502	16,502	16,502	16,502	14,692
Eff. # of Control Obs.	581	589	495	495	503	495	224	165
Eff. # of Treatment Obs.	1,805	1,817	1,603	1,603	1,616	1,603	1,312	1,154
Bandwidth	3.091	3.759	2.215	2.200	2.593	2.210	1.928	1.700

Continues in the Next Page...

Table OA16 Continued. Sharp Regression Discontinuity Estimates of the Short- and Medium-Term Effects of Obtaining an Advanced Certificate, by Initial Employment Status: Additional Outcomes

	(1) Quarter 1	**(2) Quarter 2**	**(3) Quarter 3**	**(4) Quarter 4**	**(5) Quarter 5**	**(6) Quarter 6**	**(7) Quarter 7**	**(8) Quarter 8**
3. Initially Salaried Worker								
3a. Salaried Work	-0.015	-0.011	0.004	-0.006	-0.006	-0.014	-0.033	0.010
	(0.046)	(0.019)	(0.021)	(0.023)	(0.018)	(0.018)	(0.021)	(0.016)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	3,745	6,325	7,377	6,327	6,419	7,393	6,429	7,051
Eff. # of Treatment Obs.	9,512	12,684	15,645	12,770	12,800	15,645	12,851	14,205
Bandwidth	1.828	2.198	3.022	2.408	2.583	3.039	2.795	3.845
3b. Self-Employment	-0.002	0.001	-0.008	-0.005	-0.008	0.007	0.003	-0.002
	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.006)	(0.006)
# of Observations	146,215	146,215	146,215	146,215	146,215	146,215	146,215	132,529
Eff. # of Control Obs.	3,302	6,325	6,325	7,453	6,432	6,424	6,430	5,918
Eff. # of Treatment Obs.	9,148	12,684	12,684	16,042	12,916	12,849	12,857	11,336
Bandwidth	1.341	2.169	2.155	3.458	2.875	2.630	2.806	2.979

Notes: Each cell displays the estimate for the main coefficient of interest in Equation (1) for the advanced certificate, one to eight quarters after certification, by initial employment status: unemployed (first panel), self-employed (second panel), salaried worker (third panel). The two outcomes analyzed are self-employment and salaried work. The running variable is the exam score and the discontinuity threshold is 90. All regressions include the controls described in Section 4, use a triangular kernel, and use the bandwidth that minimizes the mean squared error of the local polynomial regression discontinuity estimator. We report robust bias-corrected standard errors below the point estimates in parentheses. Standard errors are clustered at the technical-norm level. Optimal bandwidths are displayed below the standard errors. The effective number of observations changes across quarters given the variation in the optimal bandwidth. The total number of observations further drops in quarter 8 due to missing information in the last quarter for individuals that certified in the last quarter of 2019.