Do Microcredentials Help New Workers Enter the Market? Evidence from an Online Labor Platform*

Otto Kässi, Vili Lehdonvirta

2022-02-14

Abstract

We investigate the effects of a voluntary microcredentialing scheme on an online freelancing labor market. Drawing on transaction-level data, we show that obtaining a microcredential increases workers' earnings. This effect is not driven by increased worker productivity but by decreased employer uncertainty. The increase in worker earnings is realized through an increase in the value of the projects won rather than an increase in the number of projects. We also find that the effect of microcredentials is lower for more experienced

---

* Otto Kässi (otto.kassi@etla.fi) is a researcher at ETLA Economic Research, a senior researcher at Turku Centre for Labour Studies, University of Turku, and a research associate at the Oxford Internet Institute, University of Oxford. Vili Lehdonvirta is Professor of Economic Sociology and Digital Social Research at the Oxford Internet Institute, University of Oxford. The authors would like to thank Fabian Braesemann, Stijn Broecke, Gretta Corporaal, Paolo Falco, Andrew Green, Marguerita Lane, Duncan MacDonald, Rafe Martyn, Andrea Salvatori, Steffen Viete, Hanna Virtanen, Alex Wood, the conference participants of the ZEW Conference on the Economics of ICT, the CESifo Economics Summer Institute, and the seminar participants in Etla, Labore, and Ifo Insitute for helpful comments and discussion. The research was supported by a grant from the European Research Council (grant number 639652), and the research project "BRIE-Etla 2019–2022 – Shaping the Future in the Era of Intelligent Tools: AI and Beyond" funded by Business Finland. Otto Kässi gratefully acknowledges a personal grant from the OECD Future of Work Fellowship. Code and instructions on how to access the data are available in https://doi.org/10.6084/m9.figshare.19161206.v1.

JEL CODES: J21, J23, J24, J31, I2

doi:10.3368/jhr.0519-10226R3

workers, which suggests that signaling by microcredentials and other forms of verified information are partial substitutes.

Keywords: signaling, human capital, skill validation, microcredentials, online freelancing, platforms, gig economy, computer-based assessment

1 Introduction

So-called “microcredentials”, “microcertificates”, or “badges” have become a prominent topic in the labor market and skills policy literature in recent years (Painter and Bamfield, 2015; OECD, 2020; European Commission, 2020; Cedefop, 2021). Microcredentials are loosely understood as digital, privately administered skill certificates, typically awarded based on voluntary computer-administered online tests. They are in many ways analogous to conventional educational qualifications and training certificates, except that they are awarded in recognition of much narrower skills or skill sets and typically certify learning outcomes rather than methods of delivery. These characteristics could make microcredentials a cheap, fast, and accessible way for job-seekers to reduce employers’ uncertainty about their abilities and thus attain better labor market outcomes for themselves. The purpose of this article is to evaluate theoretically and empirically whether obtaining such credentials indeed helps job-seekers and which job-seekers benefit the most.

In all labor markets, employers have uncertainty about candidates’ abilities especially candidates who are new to the labor market and therefore lack references from previous employers (Terviö, 2009; Pallais, 2014). This uncertainty causes labor market inefficiency, where employers must invest in costly screening procedures, and job-seekers may fail to obtain work despite possessing the necessary abilities. Employer uncertainty has been shown to have adverse effects, especially on job-seekers who are already disadvantaged in the labor market, such as immigrants (Oreopoulos, 2011), minorities (Lang and Manove, 2011), and young people (Altonji & Pierret, 2001).

Formal qualifications are a well-established labor market institution that can reduce employers' uncertainty (Spence, 1973). However, the information conveyed by formal qualifications can become quickly outdated today due to rapid changes in skill requirements in many occupations (Painter and Bamfield, 2015) and due to candidates developing important skills via online courses and other informal learning activities (Cedefop, 2020). Formal qualifications may also be less effective in transnational and remote hiring situations that have surged during the pandemic (Stephany et al., 2020; Ozimek, 2020), because foreign qualifications are more difficult for employers to evaluate (Oreopoulos, 2011).

Labor market policy literature has begun to emphasize the need for new forms of skills validation that complement formal qualification systems (Cedefop, 2018; OECD, 2019). One prominent approach are microcredentialing systems administered by private entities such as labor market intermediaries, online educational providers, and even large technology companies such as Amazon. Microcredentialing systems can be designed around industry needs rather than national norms and rapidly updated to reflect changing skill requirements. They could be a flexible, low-cost, and up-to-date method for job-seekers to convey information about their abilities to employers. Indeed, the European Centre for the Development of Vocational Training describes microcredentials as a "megatrend" (Cedefop, 2021).

However, there is currently no robust empirical evidence of how well microcredentials actually work to reduce employer uncertainty. There is also very little theory on how they should work, despite the broader context of employer uncertainty and education being major topics in labor economics. In this study, we use data from an online labor market platform to examine the effects of microcredentials on labor market outcomes for workers. The platform in question hosts millions of freelance workers who bid for thousands of new projects that employers post on the platform each day. Most of the matches are transnational, and the work is performed remotely over the Internet. Workers can undertake voluntary computer-administered skill tests on the platform, which, if successfully completed, earn them microcredentials that are displayed in their profiles. Over 300 credentials are available on skills such as programming languages, graphic design techniques, and office software packages. Our data consist of the microcredentials earned and projects completed by a representative sample of 46,791 workers on the platform over a period of 9 years.

In theoretical terms, we demonstrate that obtaining microcredentials operates as a type of signal in the spirit of Spence (1973): credentials do not increase workers' productivity per se but help demonstrate their ability, leading to decreased employer uncertainty and increased worker earnings. In a standard signaling model, an agent's signaling cost only depends on their ability. However, we argue that in the context of a remote transnational labor market, the net benefit of signaling – and therefore, a worker's decision to signal – is determined by two parameters: the worker's ability and the uncertainty that prospective employers have about the worker's ability. To formalize this, we present a theoretical model that captures the idea that employer uncertainty increases the returns to signaling.

A common challenge in estimating the effects of signaling via educational qualifications is that they are confounded with increases in human capital. If we observe that education increases wages, it is impossible to tell whether this is caused by reduced uncertainty or increased productivity (Blackburn and Neumark, 1993; Chevalier et al., 2004). Our transaction-level data set has two appealing features for untangling these effects. First, these data contain a rich set of information on workers' past performance, which can be used as controls. Second, the fact that the freelancing projects are short and follow each other relatively frequently allows us to use the longitudinal dimension of the data to account for unobserved variation in worker productivity.

We implement two alternative identification strategies. First, we compare workers' earnings before and after acquiring a microcredential using a fixed effect event study design. The event study allows us to capture all time-invariant unobservable factors as fixed effects. We limit our attention to 14 days around the awarding of the microcredential. In this way, we can ensure that the return estimates are not contaminated by (uncertified) individual learning or other time-varying human capital effects. Second, we apply a conditional difference-in-differences approach and compare workers who have completed microcredentials to workers with similar observable characteristics who have not completed microcredentials.

In the event study, we find that gaining an additional microcredential results in an average earnings gain of 8.9% over the next two weeks. This effect is mainly driven by an increase in project value, which increases 9.7% following microcredential completion. Transformed into dollars, this corresponds to a $30.15 per-project return on completing a credential. Microcredential completion also leads to a 5.5% increase in the number of projects initiated within the next 14 days, but this is a relatively small effect in practical terms; the point estimate implies that workers win one new project for approximately 63 microcredentials completed. The conditional difference-in-differences strategy produces almost identical results.

We also find considerable heterogeneity in returns to microcredential completion. The increase in project value is up to 1.5 times greater for new entrants with no work history on the platform than for average workers. This suggests that microcredentials function as a partial substitute to verified past work experience. However, microcredentials' marginal effect on the number of projects initiated remains very small for new entrants. This may be because microcredentials can attest to hard skills but still leave considerable uncertainty about candidates' soft skills, such as communication, timeliness, and trustworthiness. References from previous employers appear to be more effective at conveying a holistic picture of candidates' skills. As a result, microcredentials are not very effective at helping candidates with no work history to win their first project. This limits microcredentials' usefulness in reducing entry barriers to new workers.

Our findings contribute to multiple strands of the empirical literature. Previous research on job testing has focused on settings where job testing is a mandatory part of the recruitment process of a firm, and the tests are tailored to the needs of that firm (Autor and Scarborough, 2008; Hoffman et al., 2018). Our study is the first rigorous evaluation of a voluntary microcredentialing system administered by a labor market intermediary.

Our study also contributes to a growing literature on labor market institutions on remote online labor markets. An influential article by Pallais (2014) reported on a field experiment in which she randomly hired inexperienced workers and provided feedback on their performance. These workers earned considerably more from their subsequent projects than the control group, who received no feedback; she argued that this stemmed from the information that the feedback provided to subsequent employers. Subsequent research has examined how other institutions might increase information and reduce employer uncertainty on online labor markets. Stanton and Thomas (2015) showed that affiliation with an intermediary agency helps inexperienced workers gain projects and earn higher rates. Horton (2017) showed that the algorithmic recommendation of workers to employers can improve the functioning of the market by reducing search frictions. Agrawal et al. (2016) showed that information on past work experience disproportionately benefits workers from less developed countries, who might be statistically discriminated against. Barach and Horton (2021) recently showed that past compensation history is a particularly salient signal of worker quality.

We complement this literature by offering the first thorough analysis of another major online labor market institution, namely skill tests that award digital microcredentials. While it is challenging to compare effect sizes across different empirical designs, data sets, and institutional settings, we generally find that microcredentials' impacts on labor market outcomes are considerably smaller than are the impacts of past work experience, agency affiliation, and algorithmic recommendations.

More broadly, we contribute to the job market signaling literature. A large fraction of empirical literature on the topic (e.g., Tyler et al., 2000; Arcidiacono et al., 2010; Dale and Krueger, 2014) assumes that the underlying ability of the agents determines their level of signaling. We do not make this assumption. Instead, we build on Lang and Manove (2011), who argued that, in addition to differences in ability, agents also take into account the differences in employer uncertainty when making their signaling decisions. According to our model, workers who face greater employer uncertainty obtain a greater marginal return from signaling and are more likely to complete microcredentials.

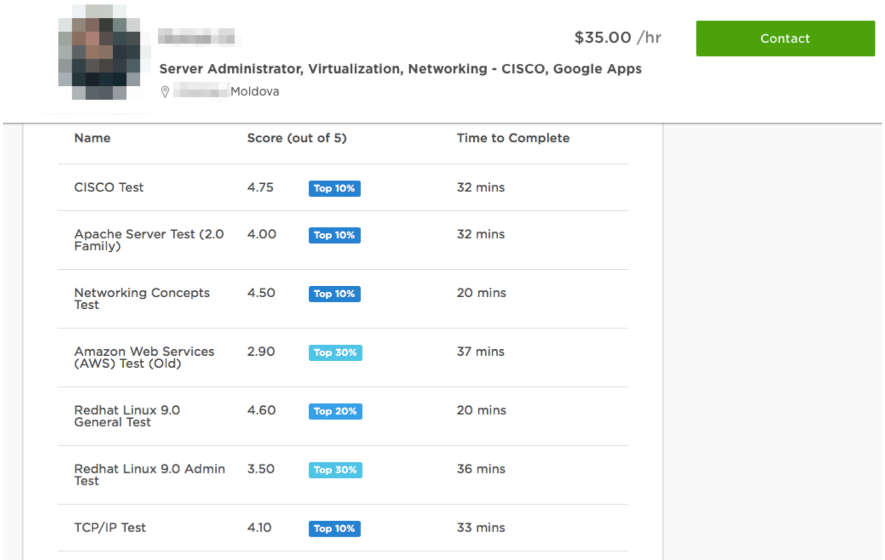
2 Empirical Setting

We collected our data from one of the largest online labor platforms, which did not wish to be identified. Before turning to details, we briefly present a typical workflow of contracting within the platform. Employers looking to hire a worker for a particular task typically start the process by posting a vacancy on the site. The vacancy includes a project category, the skills required, the expected contract duration, preferred worker characteristics, and the contract type (either a fixed sum contract or an hourly pay rate). The project category is chosen from among 89 options, such as mobile development, translation, and graphic design.

After the vacancy is posted, it is visible to registered workers who can apply for the position by submitting private bids. The interview and wage negotiation phases also take place on the platform. Screening the applicants takes a fair amount of effort on the employer's part. The employer must identify a worker whose wage requirement fits the budget and whose credentials are good enough to be trusted to complete the project successfully. Besides facilitating matching, the labor platform offers various other affordances to both workers and employers. For instance, the platform offers tools for payroll management, monitoring, project management, and dispute resolution. It also provides workers with protection against client nonpayment.

The platform also allows workers to take skill microcredential tests, which are computerized tests administered as multiple-choice quizzes and scored automatically. Over 300 different microcredentials are available on skills such as programming languages, graphic design techniques, and office software packages. The tests are highly technical, quizzing test takers about very specific facts within their skill areas.[1] Therefore, the outcomes are likely to depend on their knowledge in the specific skill areas rather than on their general ability. In the empirical part of this article, we will group the tests into six broad, mutually exclusive categories: technology, finance, design, writing and translation, sales and marketing, virtual assistant work, and other tests.

[1] A representative question from the Java programming language skill test is the following: “Assuming the tag library is in place, and the tag handler is correct, which of the following is the correct way to use a custom tag in a JSP page?”



Name	Score (out of 5)		Time to Complete
CISCO Test	4.75	Top 10%	32 mins
Apache Server Test (2.0 Family)	4.00	Top 10%	32 mins
Networking Concepts Test	4.50	Top 10%	20 mins
Amazon Web Services (AWS) Test (Old)	2.90	Top 30%	37 mins
Redhat Linux 9.0 General Test	4.60	Top 20%	20 mins
Redhat Linux 9.0 Admin Test	3.50	Top 30%	36 mins
TCP/IP Test	4.10	Top 10%	33 mins

Figure 1

A screenshot of a worker's profile featuring microcredentials.

Once a microcredential test has been successfully completed, the worker's profile shows that they are certified (see Figure 1). The profile also shows the numerical grade and percentile rank among all test takers. Only the highest passing mark in each test will be visible to potential employers, and if a worker fails a test, a failed mark is not displayed. A failed microcredential test can be retaken after a cooldown period lasting between 30 and 180 days. Workers can also choose to hide the results of microcredential tests they have passed.

Although employers can see skill test results on workers' profiles, the platform does not enable employers to exclude workers from consideration based on their test results. It is also highly uncommon for the free-text field of a vacancy to state that consideration is conditional on a specific microcredential.

3 The Motivating Theoretical Framework

This section introduces a signaling model, which we use to show that employer uncertainty about worker ability creates an incentive for workers to invest in a costly signal such as completing and displaying microcredentials. The model is a slightly modified version of the model presented in Lang and Manove (2011). It provides testable implications for how the level of worker signaling and returns to signaling vary with uncertainty about worker productivity.

The workers differ in their ability. Potential employers observe workers' microcredentials accurately, while workers' true ability is observed with noise. Consequently, workers have an incentive to gain microcredentials to reduce employer uncertainty. Throughout the rest of this section, we assume that a separating equilibrium exists – workers' ability and employer uncertainty uniquely determine how much workers signal.

Following the bulk of the literature in labor economics (e.g., Griliches 1977; Blackburn and Neumark 1993; Harmon and Walker 1995), we define "ability" as everything that is time-invariant, unobservable to the researchers, and positively correlated with both microcredential completion and success in an occupation. It can include specific skills and general intelligence as well as things such as diligence and being serious about online freelancing as a source of income.

Different categories of work require different skills. We denote a category of work as m and ability in that category as a_m. The ability is distributed along a fixed interval $[a_{0,m}, a_{1,m}]$. Workers' productivity, p_m^*, in a project of category m, conditional on their ability, is given by

$$p_m^* = a_m + \varepsilon, \tag{1}$$

where a_m is the worker's ability in category m, and ε is a normally distributed match-specific random variable whose value is only realized after the match between a worker and an employer is formed and has a mean of 0 and variance of σ^2_ε.

A potential employer can observe the number of relevant microcredentials the worker has, s_m, but not their true productivity, p^*_m.[2] Instead, the employer observes a noisy estimate of worker productivity given by

$$p = p^* + u, \tag{2}$$

where u is another normally distributed random error term reflecting employer uncertainty. The error term u has a variance of $\sigma^2_u(s)$, which is common to all employers, with $\frac{\partial \sigma^2{}_u(s)}{\partial s} \leq 0$ and $\frac{\partial^2 \sigma^2{}_u(s)}{\partial s^2} \geq 0$. The terms ε and u are independent of one another, and their distributions are assumed to be common knowledge.

We denote the accuracy of employer inference as $\lambda(s) \epsilon\ [0, 1]$, where

$$\lambda(s) = \frac{\sigma^2{}_\varepsilon}{\sigma^2{}_u(s) + \sigma^2{}_\varepsilon}. \tag{3}$$

For a given value of σ^2_ε, if $\lambda(s)$ is close to zero, then $\sigma^2_u(s)$ must be large, and consequently, the employer's ability to directly observe worker productivity is poor. In this case the employers must give more weight to the microcredential signal. If $\lambda(s) = 1$ then $\sigma^2{}_u(s) = 0$ and the employer observes worker productivity perfectly and does not have to rely on signals.[3]

---

[2] For simplicity, we will omit the subscript m from s_m, a_m, and p^*_m for the rest of this section.

[3] In the special case, where $\frac{\partial \sigma^2{}_u(s)}{\partial s} = 0$, signaling is completely uninformative, and $\frac{\partial \lambda}{\partial s} = 0$. In this case, microcredentials are fully uninformative of earnings, the workers have no incentive to signal, and employers give no weight to workers' microcredentials.

The employers follow the rules of a competitive labor market. The wage, *w,* they pay to a worker is determined by the worker's expected productivity. Their equilibrium inference of the workers' productivity, p^*, depends on the elements they observe, *p* and *s*. Let $\hat{a} = a(s)$ denote employers' equilibrium inference of *a* conditional on *s.* Throughout this article, we assume a unique, continuous, differentiable equilibrium value of s, and *w* which are uniquely determined for every combination of a and λ.

To solve for $E\ [p^* \mid p,\ s]$, note that (1) and (2) imply that $p^* = a + \varepsilon$. Moreover, since values of $p - \hat{a}$ and s uniquely determine p, an expectation conditioned on p and s is equivalent to being conditioned on $p - \hat{a}$ and s. Therefore,

$$E\ [w \mid p, s] = E\ [p^* \mid p, s] = E\ [p^* \mid p - \hat{a}, s] = E\ [a \mid p - \hat{a}, s] + E\ [\varepsilon \mid p - \hat{a}, s]. \tag{4}$$

Since, $E\ [a \mid p - \hat{a}, s] = \hat{a}$ and $E\ [\varepsilon \mid p - \hat{a}, s] = \frac{Cov(\varepsilon, u + \varepsilon)}{Var(u + e)}(p - \hat{a}) = \lambda(p - \hat{a})$, Equation (4) is equivalent to:

$$E\ [w \mid p, s] = \lambda p + (1 - \lambda)\hat{a}, \tag{5}$$

which is the equilibrium competitive wage offer of the employer, conditional on *p* and *s.*

It is useful to note that Equation (5) implies that if there are two workers, *L* and *H*, with the same level of *a*, but $\sigma^2_{u,L}(s) > \sigma^2_{u,H}(s)$ worker *H* is at an advantage because the employer can better evaluate their productivity. Therefore, worker *L* will have a larger incentive to invest in signaling. The worker's problem boils down to choosing *s* that solves

$$\max_s\ E[w] - c(a)s, \tag{6}$$

where $c(a)$ ($c(a) > 0$ for all $a \in [a_0, a_1]$) is the effort cost of signaling. It is assumed that $c(a)$ is decreasing and convex in a. In equilibrium, Equation (6) is equivalent to

$$\max_s\ \lambda E[p] + (1 - \lambda)\hat{a} - c(a)s. \tag{7}$$

Its first order condition reads as

$$\lambda_s a - \lambda_s \hat{a} + (1-\lambda)\hat{a}_s = c(a) + c_a a_s, \tag{8}$$

where subscripts denote partial derivatives. In equilibrium, $a = \hat{a}$, so (8) simplifies to

$$(1 - \lambda - c_a)\hat{a}_s = c(a), \tag{9}$$

which implicitly solves s for each combination of λ and a. Finally, solving Equation (8) for a_s and inverting yields

$$s_a = \frac{1-\lambda-c_a}{c(a)}. \tag{10}$$

Equation (10) demonstrates that the equilibrium value of $s(a)$ is strictly increasing with a. Another feature of the equilibrium is that the worker with the lowest level of ability does not invest in signaling, or $s(a_0) = 0$. To see why this the case, note that if $s(a_0) > 0$, a worker with $a > a_0$ could deviate to a smaller s without affecting the employers' equilibrium inference of their ability. The only case where this is impossible is if $s(a_0) = 0$.

Having confirmed that $s(a_0) = 0$, and noting that Equation (10) is continuous and differentiable, we know that Equation (10) uniquely determines $s(a)$ for all combinations of a and λ.

The equilibrium has the following empirical predictions:

1. If there are two workers (L, H) with the same value for a but $\lambda_L < \lambda_H$ then $s(\lambda_L) > s(\lambda_H)$ whenever $a > a_0$. That is, higher employer uncertainty about worker ability results in more signaling by the worker. To see this, note that Equation (10) implies that if $\lambda_L < \lambda_H$, then $s_a(\lambda_L) > s_a(\lambda_H)$. We argue above that $s(a_0; \lambda_L) = s(a_0; \lambda_H)$. By the continuity of s, this is only possible if $s(\lambda_L) > s(\lambda_H)$ when $a > a_0$.

2. If there are two workers (L, H) with same a but $\lambda_L < \lambda_H$, then $\frac{\partial E[w; \lambda_L]}{\partial s} > \frac{\partial E[w; \lambda_H]}{\partial s}$.

   To see why this holds, note that by Equation (9), $\frac{\partial^2 E[w]}{\partial s \partial \lambda} < 0$ for all $a > a_0$.

3. Signaling exhibits decreasing returns to scale, so that $\frac{\partial^2 E[w]}{\partial s^2} < 0$ for all $a > a_0$ (by Equation (9)).

Predictions 1 and 2 are intuitive. Workers for whom productivity uncertainty is higher get a higher marginal return from signaling and signal more when their ability is held constant. Prediction 2 implies that the returns to signaling are lower if employer uncertainty about worker productivity is lower. Prediction 3 implies that the marginal effect of signaling is lower for higher levels of signaling. While signaling models could, in theory, have multiple equilibria, Predictions 1–3 are specific to the equilibrium that defines a unique, continuous in a and λ, differentiable value of s. We later show that Predictions 1-3 hold empirically in our data.

Turning to empirical identification of returns to signaling, Equation (10) demonstrates that the choice of the level of signaling depends on two characteristics that are unobservable to the researcher but affect worker earnings. Workers with a higher ability signal more because their cost of signaling is lower; that is, it takes less effort for them to acquire microcredentials. On the other hand, workers who know that employers have problems evaluating their productivity also signal more. As a result, failing to control for these when regressing the number of microcredentials on earnings will likely lead to a biased estimate of the effect of signaling.

4 Data and Descriptive Statistics

The dataset used in this article was collected with assistance from a major online labor platform, which provided access to their developer API to make the data collection possible, but which was not otherwise involved in any aspect of the study design or sample construction. The data were collected in three steps.

In the first step, we used the platform’s search functionality to collect a sample of workers. The search functionality orders the search results in various ways that are opaque to the user, intending to increase the efficiency of the searches. To obtain a representative sample of the population of workers on the platform, we first used the platform API's search functionality to obtain a near-comprehensive list of workers[4] and then randomly sampled 10% of them. This approach allowed us to collect a reasonably large sample without violating the rate limits set by the API. After removing duplicates, we ended up with a sample of 46,791 workers. Next, we used separate API requests to obtain background information on each worker and the details of their completed projects, which numbered 467,455 projects in total.

To better capture the situation where an employer screens workers with limited information, we applied further selection criteria to filter the sample. First, we excluded projects where the employer explicitly invited a single worker to the project as we assume that, in these cases, the employer had negligible uncertainty about the worker's quality.[5] Second, to filter out projects where all applicants were hired, and no employer screening took place, we excluded projects where more than one worker was hired.[6] Third, projects which had a missing variable for project value (0.02% of observations) were excluded as artifacts. Finally, we

[4] The list is not fully comprehensive because the platform can, at its discretion, exclude workers from search results, including workers who have violated the terms of use and workers who have failed to log in for a long time. Workers can also choose to deactivate their profiles from search results when they stop freelancing, for instance. Our empirical results are robust to any such omissions.

[5] While our data have unique indicators for projects and workers, the data do not include employer id's. Therefore, we are not able to separate re-hires from first-time hires. Nonetheless, we expect a large share of re-hires to be filtered out by excluding projects with only a single applicant.

[6] Some employers try to cut screening costs by following a strategy like one outlined by Lazear (1998). They hire a large number of workers for a test project. The best-performing workers are then hired with a more permanent contract. The motivation for posting and applying to these projects may differ from what we normally associate with employer hiring under uncertainty, hence the exclusion. Agrawal et al. (2016) applied a similar restriction.

excluded the “readiness test for new workers” microcredential from our sample, given that a large majority of workers in our sample had completed it. It is a near-mandatory test used to screen out registrants who are not serious about freelancing. After applying these restrictions to the sample, we ended up with 422,199 project observations. The main summary statistics of the data are presented in Table 1.

Our data form an unbalanced panel. Most of the workers are observed more than once. The worker-specific values of observable variables vary over time as workers win more projects and get rated by their employers. Each of the observations in our data is a combination of worker-level observables and project characteristics. Column 1 of Table 1 presents the descriptive statistics for the project observations.

In our empirical models, we concentrate on the subset of workers who complete microcredentials. Column 2 of Table 1 shows the descriptive statistics for workers at the time of microcredential completion. When a worker has completed more than one microcredential, they are included several times in this filtered data.[7]

We operationalize worker signaling intensity as their number of completed and disclosed microcredentials. Table 1 shows that most projects in our sample have accrued to workers without any microcredentials at the time of project start (median number of microcredentials in Column 1 is zero), whereas the median number of microcredentials of workers at the time of microcredential award is two.

To understand how the workers who take microcredential tests differ from workers who do not take tests, we also present the descriptive statistics of workers who have not completed any

---

[7] If a worker completes a microcredential and completes another one less than 14 days after completing the first one, we exclude the second one and only include observations falling within the first credential’s time window.

microcredentials in Column (3). We find that the test takers tend to be at the start of their freelancing careers. Compared to the full sample, they have fewer completed projects and smaller cumulative earnings. The workers without microcredentials have completed more projects and earned more money than test-takers at the time of microcredential award, but less than the full sample. On the other hand, workers without microcredentials have spent less time on the platform compared to the two other subsamples. We note that the differences are not statistically significant as the three samples are very heterogeneous, as evidenced by the large standard errors for most variables. One variable where statistically significant difference emerges is the share of workers with at least one completed project. About one third of workers who take tests have won at least one project, compared with 45% in the full sample and 19% in the no-test subsample. In summary, we find that microcredentials are relatively common among workers, while most of the workers have only completed a few. Moreover, a large share of projects are won by workers without microcredentials.

Table 1 also reports selected time-invariant background characteristics of workers. Again, there are no discernible differences between the test-takers and the full sample. However, a comparison between workers without microcredentials and the full sample reveals that the latter has a larger share of workers based in the US (15% compared to 8%) and Ukraine (7% compared to 4%), and a smaller proportion of college-educated workers.

We expect workers to disclose their microcredentials strategically: after completing a microcredential, they will compare their ranking with other test takers to decide whether to disclose it. We present the distribution of observed (i.e., disclosed) test score ranks in Figure 2. As is clear from the figure, the probability of workers disclosing their test scores is substantially higher if they have scored above the median.[8] We study how the returns to signaling vary with test scores later in Section 6.4.

Our independent variable, the number of completed and disclosed microcredentials, implies that we are agnostic about any possible heterogeneity in the effect of different microcredentials on labor market success. We assume that all microcredentials act as a signal of the same underlying ability. While this is clearly a simplifying assumption, Table 2 supports our approach. The table reports the distribution of microcredentials and projects by worker specializations, where workers' specialization is defined as the category of project that each worker has worked on the most. We can see from the table that workers tend to specialize. The workers have largely worked on a single category of jobs and have predominantly also completed microcredentials in the same category. This holds for the largest specializations—design, technology, virtual assistance, and writing and translation—to which 87% of the workers belong. We separately study the heterogeneous effects of different microcredentials on worker success in different project categories in Section 6.3.

[8] A more complete model for worker microcredentialing could have two steps. In the first step, the worker could observe a noisy estimate of their own a_m and choose whether to take a microcredential test. The worker only takes the test if their expectation of a_m is high enough. In the second step, the worker could observe the result of their microcredential and disclose the result based on the difference between their estimate of a_m and the result of the microcredential test. In this amended version of the model, signaling still strictly increases with a_m so it would not change the comparative static analyses presented in Section 3.

Table 1

Descriptive statistics

	Sample:					
	Full sample		At the time of microcredential award		Workers without any microcredentials	
Panel A: Time varying worker characteristics	(1)		(2)		(3)	
	Mean (standard deviation)	Median	Mean (standard deviation)	Median	Mean (standard deviation)	Median
Number of microcredentials	2.18	0	3.18	2	0	0
	(3.81)		(4.33)		(0)	
Number of completed projects	42.44	19	5.08	0	17.97	2
	(64.79)		(18.03)		(49.58)	
Dollars earned	11193.82	3165.9	2272.35	0	5449.62	195
	(22862.06)		(10244.44)		(15803.23)	
Months active	22.86	17.8	19.7	9.96	13.69	3.7
	(20.29)		(21.75)		(19.15)	
Freelancer rating	3.02	3.43	2.57	3.06	3.31	3.59
	(1.66)		(2.07)		(1.45)	
Average project value	403.59	146.04	673.23	157.98	601.28	168.7
	(1335.32)		(2386.92)		(2459.63)	
Panel B: Time invariant background characteristics		Share		Share		Share
	Male	70%		70%		68%
	College degree or more	74%		72%		58%
Top-5 countries						
	India	27%	India	24%	India	22%
	Bangladesh	12%	Bangladesh	12%	United States	15%
	Philippines	11%	Philippines	9%	Ukraine	7%
	Pakistan	10%	United States	8%	Philippines	6%
	United States	7%	Pakistan	7%	Pakistan	6%
Panel C: Sample sizes						
	Share with at least 1 project won	45%		32%		19%
	Number of projects	442,203		233,791		20,827
	Number of workers	46,791		33,091		13,700

Notes: Column (1) presents the descriptive statistics for the full sample, Column (2) presents the descriptive statistics at the time of microcredential completion for the subsample of workers who take tests, and Column (3) presents the descriptive statistics for workers who have not completed any tests. In Columns (1) and (3) worker characteristics are measured at time of project start and one observation corresponds to a project completed by a worker. In Column (2), time-varying characteristics are measured at time of microcredential completion. In Panel B, worker home countries are self-reported and verified by the platform. Education is reported by workers themselves. Worker gender is inferred by workers' self-reported first name using the Python library SexMachine (https://github.com/ferhatelmas/sexmachine/, accessed 2021-03-17).

Table 2
Project and Test Distribution among Workers

	Design and Creative	Finance	Sales and Marketing	Technology	Virtual Assistant	Writing and Translation	Other
Panel A: Worker Specialization							
	19.5%	1.9%	8.1%	42.9%	13.5%	13.8%	0.3%
Panel B: Project Distribution by Specialization							
Design and Creative	90.4%	0.7%	1.3%	4.6%	1.3%	2.1%	1.8%
Finance	0.3%	73.4%	1.5%	0.5%	1.1%	3.0%	4.6%
Sales and Marketing	1.0%	4.0%	73.7%	1.7%	3.3%	8.7%	0.5%
Technology and IT	5.8%	5.2%	10.2%	90.7%	1.7%	6.0%	3.4%
Virtual Assistant	1.2%	5.0%	4.0%	0.7%	86.4%	7.1%	8.1%
Writing and Translation	1.2%	11.0%	9.3%	1.8%	5.4%	72.9%	2.0%
Other	0.1%	0.6%	0.1%	0.1%	0.7%	0.2%	79.5%
Panel C: Microcredential Distribution by Specialization							
Design and Creative	41.9%	0.9%	3.4%	7.2%	2.9%	1.6%	1.9%
Finance	1.1%	52.1%	3.4%	1.1%	4.2%	4.6%	8.0%
Sales and Marketing	3.3%	3.7%	34.7%	5.3%	7.8%	6.6%	3.2%
Technology and IT	23.9%	19.4%	23.4%	69.5%	34.4%	12.7%	20.5%
Virtual Assistant	2.4%	4.8%	5.8%	1.6%	11.8%	4.8%	7.1%
Writing and Translation	27.4%	19.1%	29.4%	15.4%	38.9%	69.9%	59.4%

Notes: Worker project distribution. In Panel A, workers' specialization is defined as the category of project that each worker has worked on the most. Projects and microcredentials are manually classified into the seven categories.

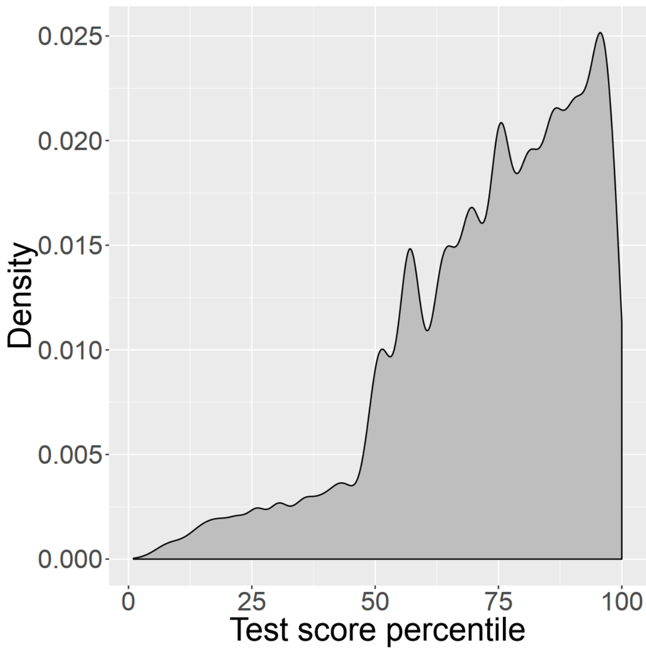


Figure 2 The kernel density of the microcredential test score distribution

5 Empirical Analysis

5.1 The Two Identification Strategies

We now turn to empirically studying how signaling efforts are rewarded in the labor market. As discussed in the theory section, the decision to complete microcredentials is driven by two types of selection biases: worker ability and employer uncertainty about worker ability, which remain unobservable to us. An additional complication related to our setting is that workers can potentially have very long careers, during which they can acquire new skills via education, training, or on-the-job learning. We use two alternative empirical designs to address these concerns: a fixed effects event study design and a conditional difference-in-differences design.

In the fixed effects event study design, we compare a worker's labor market success during a short 14-day window before completing a microcredential test to labor market success during a 14-day window after the test. Under the assumption that the unobservables remain constant over the 28-day period, we can subsume them into event fixed effects (with one fixed effect for each time the worker completes a microcredential).

As with any model that relies on fixed effects, we need to assume strict exogeneity (i.e., the choice to signal should not be determined by the outcome variables), and that the unobservables affecting outcomes are fixed over time.

In practice, our estimation strategy is vulnerable to three plausible scenarios that lead to the violation of these assumptions. First, if the worker faces a transitory shock in their labor market success and decides to signal to overturn the shock, this will create a correlation between signaling and labor market outcomes and bias our results. Second, the strict exogeneity assumption will not hold if a worker decides to put more effort into their freelancing career by simultaneously applying for more projects and completing additional microcredentials. Third, if the workers' outcomes systematically trend upward or downward around the time of microcredential completion, our estimates might be biased due to these secular trends. The strict exogeneity and lack of secular trends assumptions are not directly testable, but we present robustness checks in the form of placebo tests and donut specifications that support the validity of the fixed effects event study design.

An alternative identification strategy is conditional difference-in-differences, which leverages the difference in evolution of outcomes between treatment and control groups. We need to define a suitable control group to apply this strategy. To construct the control group, we find workers who have not completed tests over two or more 14-day blocks and assign a placebo test between two of these blocks. Our difference-in-differences design compares the placebo-treated control events with the actual microcredential completion events described earlier.

For a difference-in-differences design to identify the returns to signaling, it is necessary to make the standard parallel trends assumption. Under this assumption, the treatment and control groups evolve in parallel in the absence of the treatment. In our data, the treatment and control groups are strikingly different both along observable dimensions and the pre-treatment trends of dependent variables. To address this, we resort to a conditional difference-in-differences approach where we first match each treated worker to a set of control workers and then proceed with a standard difference in differences strategy. After matching, we find no statistically discernible differences between the treatment and control groups in terms of pretrends.

The conditional difference-in-differences specification is subject to the same issues as the event study design. First, any unobservable factor that affects both the probability of completing a skill certificate and workers' labor market outcomes will create a spurious correlation between them and confound the estimates. Second, any difference in trends of labor market outcomes before treatment will threaten the parallel trends assumption and consequently confound the return estimates. We evaluate the plausibility of the parallel trends assumption by comparing differences in pre-trends between treated and control groups.

Chabé-Ferret (2017) has shown that matching can introduce additional sources of bias to the difference-in-differences design. Nonetheless, reassuringly for our results, we show that both strategies yield estimates of the effects of signaling on labor market outcomes that are almost identical both qualitatively and statistically.

5.2 Fixed Effect Event Study Results

Our first approach for studying the effect of signaling on workers' labor market outcomes is to compare the outcomes before the completion of a microcredential to the outcomes after the completion of a microcredential in an event study setting. We estimate the following fixed effects regression models:

$$y_{iket} = \alpha_e + X_{ik}\beta + \gamma s_{ik} + v_t + \varepsilon_{iket}, \tag{11}$$

$$Z_{ijet} = \alpha_e + X_{ij}\beta + \gamma s_{ij} + v_t + \varepsilon_{ijet}, \tag{12}$$

Equation (11) concentrates on the intensive margin: the dependent variable is the log-value of a project, k, won by worker i.

In Equation (12), the unit of observation is a 14-day block immediately before or after the microcredential completion, indexed with j. The variable Z_{ij} corresponds to three alternative dependent variables: the number of projects initiated within the 14-day period, an indicator variable which gets the value, 1 whenever the number of won projects is positive and 0 otherwise, and the number of dollars earned in each 14 day pre- or post-test period.[9]

In all specifications, the main parameter of interest is γ, which captures the marginal effect of earning a microcredential – denoted by s_{ik} or s_{ij} – on the platform and captures the effect of signaling on labor market outcomes.[10] All specifications contain controls for time-varying characteristics: the number of previously completed projects, the average reputation rating from previous projects, a dummy variable for no past ratings, and the number of dollars earned on the platform (arsinh transformed), denoted by X_i, and year dummies, v_t.

The fixed effects, α_e, refer to event fixed effects, where each microcredential completion event of each worker corresponds to a separate fixed effect intercept. In terms of empirical identification, these parameters subsume all unobserved heterogeneity that remains constant

---

[9] We study total earnings in levels rather than in logs to retain the substantial number of zeros in the data.

[10] Note that the workers can complete more than one microcredential at the same date. In this case, the variable s_{ik} and s_{ij} are incremented by the number of microcredentials the worker completes on the date.

within the event window into fixed effects. For comparison, we also report the results from a worker fixed effects specification (where one fixed effect parameter corresponds to the full career of each worker) and a cross-sectional OLS specification. These specifications allow us to evaluate the direction and magnitude of the bias caused by unobservables. We also return to worker fixed effects in the next section when we study how returns to signaling vary within a worker's career.

Estimation results are presented in Table 3. When comparing cross-sectional OLS and fixed effects estimates, we find that the cross-sectional estimates are consistently smaller than the estimates from the worker fixed effects specifications, which in turn are almost always smaller than the estimates from the event fixed effects specifications. This implies that worker-specific, earnings-related unobservable characteristics are negatively correlated with the decision to signal. In other words, there is a negative selection effect related to completing microcredentials – workers who are in a disadvantaged position in the labor market signal more. Considering the theoretical model presented in the previous section, this suggests that the decision to complete microcredentials is driven by employers having more uncertainty about some workers' ability (differences in λ) rather than by differences in ability between workers (differences in a).

Our preferred event study fixed effects models are the ones reported in Columns 1, 4, 7, and 10 of Table 3. These models tackle both time-invariant unobserved heterogeneity and time-varying unobserved heterogeneity if it varies smoothly in the +/– 14-day time window around the time of skill microcredential completion.

Column 1 of Table 3 presents a specification that looks at the per-project earnings margin. It shows that an additional microcredential leads to a 9.7% increase in project value. Transformed into dollars, this corresponds to a $9.7\% \times \$310.84 \approx \$30.15$ return on completing a microcredential. Column 4 looks at the number of projects won, showing that completing a microcredential leads to a 0.016/0.29 ≈ 5.5% increase in the number of projects initiated within the next 14 days. This is a relatively small effect economically speaking; the point estimate implies that workers win one new project for approximately 63 completed microcredentials. When the probability of working at least once is used as the dependent variable, we find that the marginal effect of completing a microcredential is 0.2 percentage points. Relative to baseline, this is an improvement of 0.02/0.18 ≈ 11%.

Finally, Column 10 of Table 3 shows the combined effects on the income and employment margins. The column shows a marginal increase of $8 in earnings from microcredential completion. Transformed into percentages, this corresponds to an average earnings gain of 8.9% in two weeks. Again, the estimates line up very well; the marginal effect of signaling on both the number of projects and the probability of working is of the same magnitude and economically small, whereas the effects on project value and earnings are larger.

Since our data do not include information on cases where a worker bid for a project but did not win it, the results might be confounded by the bidding effort of workers. If workers are more active in bidding for projects just after completing microcredentials, then our estimate of the signaling effect might be biased upwards. If this is the case, the event fixed effect estimates should be interpreted as upper limits for the true effects of signaling on employment and earnings. We demonstrate in Appendix 2 that there are no differences in the observable characteristics of projects won before microcredential completion vs. after its completion. This indicates that workers earn more from completing observationally similar projects after microcredential completion. To the extent that varying effort is reflected in the types of projects that workers win, this finding supports our assumption that varying worker effort is not driving our results.

5.3 Conditional Difference-in-Differences Results

An alternative to the within-worker event study described above is a between-workers difference-in-differences study that compares the labor market outcomes of workers who have completed microcredentials with the outcomes of workers who are observationally similar but have not completed them. Our difference-in-differences specifications are the following:

$$y_{ik} = \alpha_0 + X_{ik}\beta + \eta \times treated_i + \gamma s_{ik} + \delta(treated_i) \times s_{ik} + v_t + \varepsilon_{ik}. \quad (13)$$

$$Z_{ij} = \alpha_0 + X_{ij}\beta + \eta \times treated_i + \gamma s_{ij} + \delta(treated_i \times s_{ij}) + v_t + \varepsilon_{ij}. \quad (14)$$

Here, the term "treated" is a dummy variable with the value 1 in the treatment group and 0 in the control group. As before, our parameter of interest is δ, which captures the marginal effect of signaling for the workers whose number of microcredentials increases compared with the control group whose number of microcredentials remains constant. Specifications (13) and (14) include control variables for the number of previously completed projects, the average reputation rating from previous projects, a dummy variable for no past ratings, and the number of (arsinh) dollars earned on the platform and year dummies. The term α_0 corresponds to an OLS intercept. Both models are estimated using estimated matching weights from the matching step, which we discuss next.

In the difference-in-differences research design, a challenge is that selection into treatment is endogenous. To overcome the selection issue, we first match worker microcredential completion events with "placebo events" where no microcredential completion took place. This conditional difference-in-differences design is unbiased to the extent that matching captures time-varying unobservable characteristics and to the extent that time-invariant observables are subsumed into the treated and control indicator variables. We can evaluate these assumptions by examining post-match covariate balances and differences in pre-treatment trends between treatment and control groups.

Since the model for the log of project value is not based on a balanced panel while the other three models are, we implement the matching slightly differently for the two estimation samples.

When estimating Equation (13), where the dependent variable is the log of project value, we limit our donor pool of potential control events to workers who have completed at least one project within 28 days but have not completed any microcredentials during that period. We then assign a placebo microcredential in the middle of the 28-day block.[11] We perform the matching using worker-level covariates: the number of previously completed projects, the average reputation rating from previous projects, a dummy variable for no past ratings, and the number of (arsinh) dollars earned on the platform and year dummies. We use exact matching for integer-valued and discrete variables (the number of completed microcredentials, the number of completed projects, year dummies, the no-past-rating dummy). For continuous-valued variables (cumulative earnings, past ratings), we use Coarsened Exact Matching (CEM) introduced by Iacus et al. (2012).[12]

For Equation (14), where the dependent variable is either the number of completed projects, probability of working, or earnings, we include three additional pretrend periods in the estimation sample so that we compare averages over four two-week pretest blocks with one post-test block. In this case, we failed to reach a satisfactory pre-treatment balance by only matching on worker-level covariates. Instead, we perform the matching on both worker-level covariates and pre-treatment outcomes for all four periods. Because we match on pre-treatment

---

[11] Without the restriction of at least one project within the 28-day block, the control group would be dominated by workers whose covariates match the treated workers extremely well but who have no attachment to the labor market at the time of test completion.

[12] We implement matching using the MatchIt R library of Stuart, et al. (2011).

outcomes, we do not need to filter the donor pool for inactive workers like in the case of the log project value model.[13]

A significant advantage of using CEM over methods based on a propensity score is that it does not rely on specifying a model for the probability of selection into treatment. Instead, the method relies on splitting the estimation data into "histogram" cells defined by covariates, and matching treated and control units within the cell. Since our sample sizes are large, we can match the continuous variables within relatively small cells. We adjust the standard errors to account for the matching by clustering the errors by match group (Abadie and Spiess 2021).

Intuitively, the CEM algorithm matches each microcredential completion event by a worker with a set of placebo events with covariate (and outcome pre-trend) values close to each other. The goal of matching is to prune observations from the set of control events so that the remaining data have better balance between the treated and control group in terms of matching variables. We use many-to-many matching to use our data as efficiently as possible. This leads to different numbers of control units being matched to each treated unit within a cell. To account for this, the CEM algorithm provides matching weights for each observation, with the pruned observations having a weight of zero.

After matching, we proceed with a standard difference-in-differences study using estimated CEM weights as sample weights. We report the balance measures between treatment and control groups in Appendix 3. None of the differences in the matched sample are statistically significantly different from one another at conventional risk levels.

---

[13] Monte Carlo evidence discussed by Chabe-Ferret (2017) suggests that matching on several pre-treatment outcomes can eliminate the bias in standard difference-in-differences when the parallel trends assumption does not hold. Similar approaches have also been applied in job-training program evaluation literature (e.g., Heckman et al. 1997; Dehejia and Wahba 2002, and Lechner & Wunsch 2018)

Our estimates are reported in Table 4. In Column 1 row "Number of microcredentials × treated", we show that the estimate for the marginal effect of completing an extra microcredential is 8.1%. We also note that the negative coefficient on the treated dummy in Table 4 supports our argument that workers who complete skill certificates, on average earn less than their comparable peers who do not complete skill certificates.

We report the effects of signaling on the number of projects and probability of working in Columns 2 and 3. We find statistically significant but economically marginal effects analogously to FE event study estimates. Completing an additional microcredential results in .009 additional projects and an increase of .006 in the probability of working. Finally, Column 4 shows that the marginal effect of a microcredential completion on earnings is \$3.8[14].

We consistently find positive and statistically significant estimates for the marginal effect of signaling. Analogously to our event study estimates, we find economically small effects on the employment margin but larger estimates for the effect of signaling on average project value and earnings.

The difference-in-differences analyses rely on the assumption that, in the absence of treatment, the two groups' time trends would have moved in parallel. While this assumption is not directly testable, we can test if trends statistically differ between treatment and control groups prior to the event. In the bottom row of Table 4, we report the differences in linear time trends between treatment and control groups in the pre-treatment period estimated from a linear

---

[14] The positive coefficients on the Treated dummies in Columns 2-4 of Table 4 runs counter to the argument that skill certificate completing workers would be negatively selected from the total pool of workers. This is driven by the fact that we match on pre-treatment observables when estimating the models presented in Columns 2-4.

event study model. Reassuringly, the pre-treatment time trends are always statistically indistinguishable from zero.[15]

Figure 3 presents a comparison between the conditional difference-in-differences and event study results. The estimates from the two models are very close to one another in all cases. For models examining the probability of working and the number of won projects, the two estimates' 95% confidence intervals do not overlap, but the two estimates are qualitatively similar even in those cases.

---

[15] Note, that equations (13) and (14) do not correspond to a standard 2-by-2 difference-in-differences design with after and treated dummy variables. Instead, s_{ik} and s_{ij} are counters that are incremented by the number of microcredentials the treated workers complete on the treatment date. The pre-trend differences reported in the bottom of Table 4 are based on a standard difference-in-differences model with after and treated dummy variables and their interaction.

Table 3
Returns to Signaling

	Dependent variable:											
	Project value (log)			Number of projects			Number of projects > 0			Earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Number of microcredentials	0.097**	0.030***	-0.011***	0.016***	0.012***	-0.005***	0.002+	0.006***	-0.003***	7.953***	5.522***	-0.724
	(0.024)	(0.005)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.0004)	(1.443)	(1.102)	(0.917)
Feedback rating	0.105***	0.33*	0.092***	-0.002	0.019**	0.013**	-0.010	0.009***	0.009***	15.567	26.477**	22.725***
	(0.019)	(0.012)	(0.010)	(0.022)	(0.006)	(0.005)	(0.009)	(0.002)	(0.002)	(25.744)	(8.637)	(4.353)
Number of completed projects	0.054***	0.002**	-0.005***	-0.041*	0.004***	0.008***	0.041***	0.001***	0.002***	16.140	0.810*	-0.357
	(0.011)	(0.001)	(0.0005)	(0.012)	(0.001)	(0.001)	(0.006)	(0.0001)	(0.0001)	(7.715)	(0.353)	(0.219)
Baseline	$ 310.84	$ 310.84	$ 310.84	0.29	0.29	0.29	0.18	0.18	0.18	$ 89.45	$ 89.45	$ 89.45
Fixed effects	Event	Worker	No	Event	Worker	No	Event	Worker	No	Event	Worker	No
Observations	32,975	32,975	32,975	178,478	178,478	178,478	178,478	178,478	178,478	178,478	178,478	178,478
Adjusted R^2	0.463	0.369	0.101	0.363	0.248	0.125	0.273	0.214	0.134	0.056	0.018	0.011

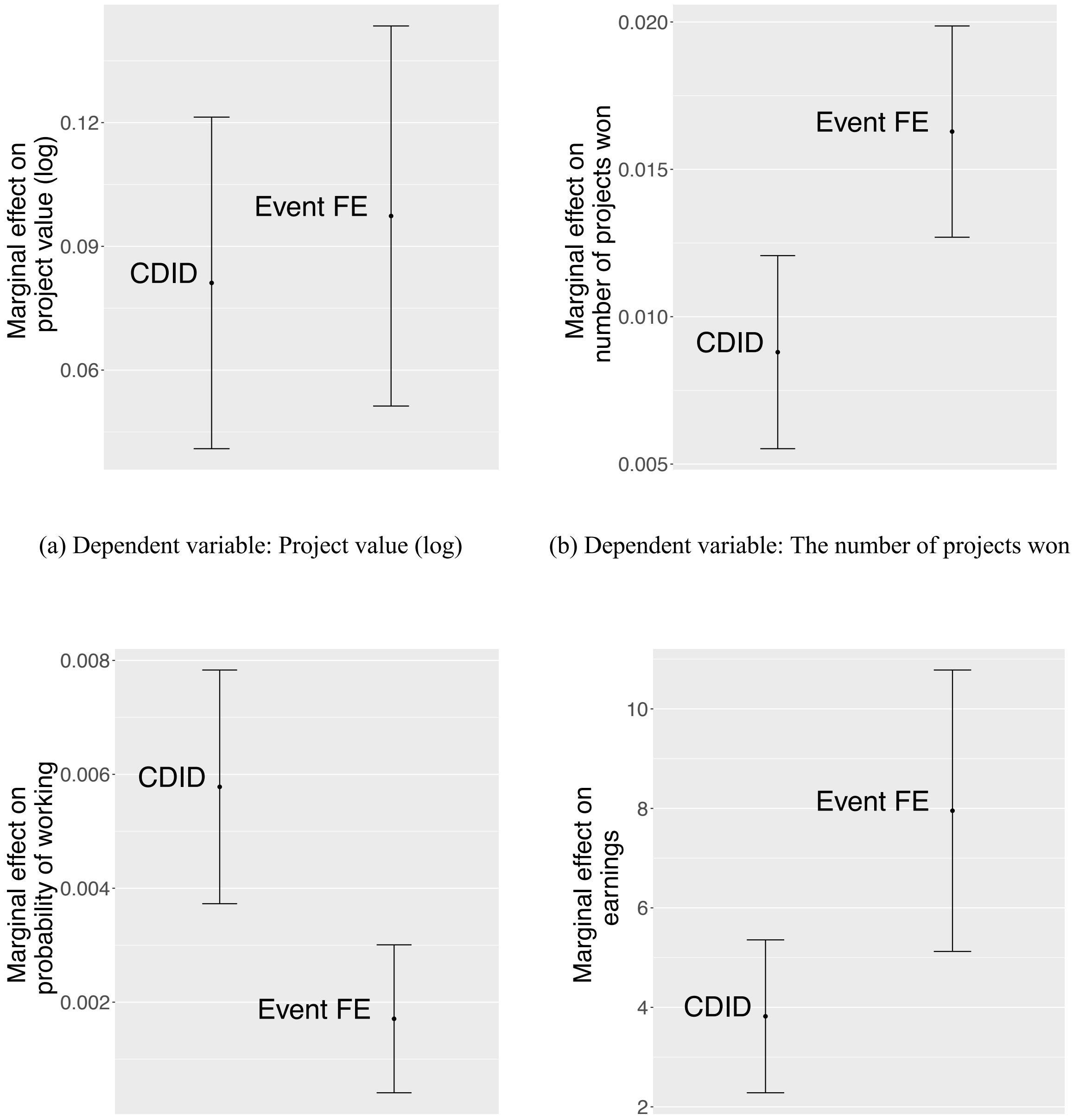
Notes: In Columns 1–3, the unit of observation is one project. In Columns 4-12, the unit of observation is a 14 day pre- or post-test period. In addition to the variables reported, all models include year dummies and cumulative (arsinh-transformed) dollars earned on the platform. “Baseline” refers to the mean of the dependent variable. Standard errors are clustered on the worker level. The significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%.

Table 4

Returns to Signaling: Conditional Difference in Differences

	Dependent variable:			
	Project value (log)	Number of projects	Number of projects > 0	Earnings
	(1)	(2)	(3)	(4)
Treated	-0.499***	0.021***	0.015***	5.081*
	(0.083)	(0.005)	(0.003)	(2.582)
Number of microcredentials × treated	0.081***	0.009***	0.006***	3.820***
	(0.021)	(0.002)	(0.001)	(0.784)
Number of microcredentials	-0.042**	-0.001	-0.001	-0.220
	(0.014)	(0.001)	(0.001)	(0.148)
Fixed effects	No	No	No	No
Observations	10,805	1,501,761	1,501,761	1,501,761
Adjusted R^2	0.106	0.035	0.029	0.003
Difference in pretrends	-0.018	0.000002	0.000002	-0.035
	(0.015)	(0.000002)	(0.000002)	(0.023)

Notes: In Column 1, the unit of observation is one project. In Columns 2–4, the unit of observation is a 14-day pre- or post-test period. In addition to the variables reported, all models include year dummies, the average rating for completed projects, a dummy variable with value 1 if a worker does not have any ratings from past projects, cumulative (arsinh-transformed) dollars earned on the platform, and the cumulative number of completed projects on the platform, measured at the time of project start. Estimation is based on CEM matching without replacement. Variables used for matching in Column 1 are: the number of completed tests, the number of completed projects, year dummies, a no-past-rating dummy, and cumulative earnings and past ratings. In Columns 2–4, matching is done using the same covariates as in Column 1 and the values of the dependent variable in periods t = -1, …, -4. "Treated" is an indicator variable that equals 1 for each worker who belongs to the treatment group and 0 otherwise. "Difference in pretrends" refers to an interaction term between a linear time trend and a treated group dummy variable for the pre-test time trend (see the text for details). Standard errors are clustered on the match group level. Significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%



(a) Dependent variable: Project value (log)

(b) Dependent variable: The number of projects won

(c) Dependent variable: Number of projects won > 0

(d) Dependent variable: Earnings

Figure 3
Comparison between the fixed effects event (Event FE) study and the conditional difference-in-differences estimates (CDID).

6 Further Results

Having shown that signaling has a positive effect on labor market outcomes, in this section, we examine how the returns to signaling vary across various dimensions as predicted by our theoretical model. We show that the returns to signaling tend to be smaller when there is more information available on the workers and that signaling tends only to be beneficial if the test is in the same category as the project. In addition, we demonstrate that generally positive returns from signaling accrue to workers who have scored at the top of the rank distribution of test takers.

6.1 Signaling as a Substitute for Other Forms of Verified Information

Establishing that signaling decreases employer uncertainty about workers' productivity is complicated by the fact that the information set of the employer, and therefore their uncertainty about worker productivity, is unobservable to us. Nonetheless, Prediction 2 (put forward in the theory section) suggests that the marginal effect of signaling is lower when employers have higher levels of information. We can test this empirically by examining how the marginal effect of completing a microcredential varies as the level of the other information on a worker provided by the platform to prospective employers varies. The more projects a worker completes on the platform, the more publicly visible their work history and employer feedback are, which arguably decreases employer uncertainty. We, therefore, follow Agrawal et al. (2016) and Pallais (2014) in assuming that previous work experience on the platform can be used as a proxy measure of employers' level of information about a worker.

To test Prediction 2, we included the interaction term between work experience and the number of microcredentials completed to the event study models (Equations (11)-(12)). Since our main interest is in examining how the returns to skill certification vary *within* workers' careers, our preferred specification is the one that controls for worker fixed effects instead of event fixed effects. We report the estimation results from both worker and event fixed effects specifications to facilitate comparison with the previous section and to demonstrate that the results hold across both specifications.

We present our estimation results in Table 5. In Columns 1–2 and 5–6, where the dependent variables are the log of project value and the probability of working, we find negative and statistically significant effects in both worker-event and worker fixed effects specifications. In Columns 3 and 4, where the dependent variable is the number of projects won, we only find a negative and marginally statistically significant estimate for the worker fixed effects specification. When using earnings as the dependent variable (Columns 7–8), neither of the specifications are statistically significant for the interaction term.[16]

The estimation results reported in Table 5 and Figure 4 suggest that the returns to completing microcredentials are smaller for more experienced workers. The negative interaction terms indicate that the main estimates reported in Table 3 hide considerable heterogeneity in returns. Comparing the estimates between Table 3 and Table 5 shows that the returns to completing microcredentials are up to 1.5 times greater for workers with a work history of 1 completed project compared with the average in the test-taker sample.

While we find that the marginal effect of signaling is smaller for more experienced workers, Figure 4 demonstrates that a vast majority of workers who complete microcredentials get at least a marginally positive benefit from them. For example, panel (c) in Figure 4 shows that the regression results imply that some workers might have a lower probability of working after

---

[16] This result appears to be partly driven by the extremely right-skewed distribution of the $Earnings_{ij}$ variable, leading to large standard errors. If we winsorize the top 0.1% of the distribution of $Earnings_{ij}$, the parameter estimate for the regression coefficient on the interaction term in the worker fixed effect specification is -3.545 with a standard error of 1.587 (p-value = 0.012).

completing a microcredential. These workers are a minority, though. 99% of the microcredential completion events are by workers with less than 15 projects.

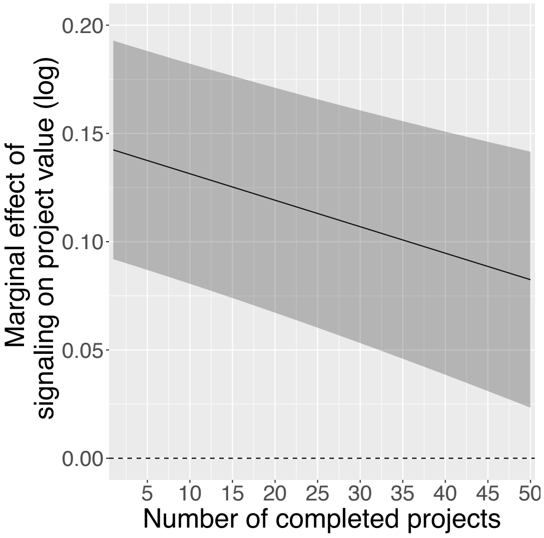
Overall, the results support the hypothesis that signaling is a substitute for other sources of information. Nonetheless, as evidenced by Figure 4, the substitution effect is relatively small, even when statistically significant. For instance, the marginal effect of completing an additional microcredential on log project value is still positive for a worker with 45 completed projects. One explanation for this could be that employers face uncertainty over both workers' technical ability and other dimensions of ability, including soft skills such as communication and trustworthiness. Microcredentials mostly decrease uncertainty over hard skills, while work experience and detailed feedback on completed projects will also reduce uncertainty over such other dimensions of ability. If this is the case, then microcredentials cannot fully substitute for work experience as a source of information.

Despite the small magnitude, the results presented in this section suggest that microcredentials are substitutes for other sources of verified information on workers' ability. While not conclusive, this evidence is consistent with our theoretical model, which implies that the reason why earnings are higher after microcredential completion is decreased employer uncertainty about worker ability.

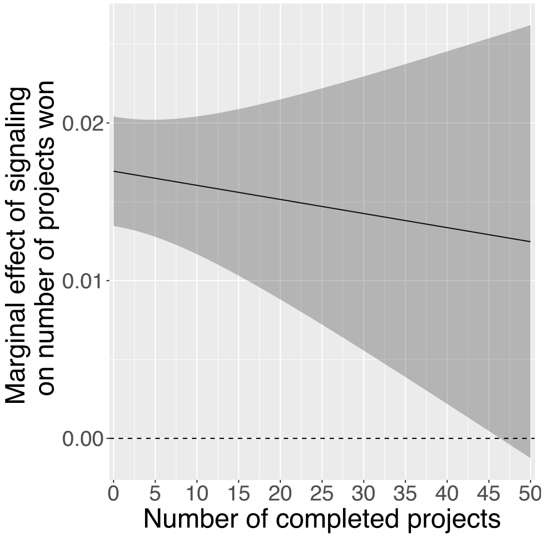
Table 5
Returns to Signaling by Level of Experience

	Dependent variable:							
	Project value (log)		Number of projects		Number of projects > 0		Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of microcredentials	0.144***	0.046***	0.017***	0.015***	0.003***	0.007***	8.384***	5.799***
	(0.020)	(0.008)	(0.002)	(0.002)	(0.001)	(0.001)	(1.426)	(1.226)
Number of microcredentials × Number of projects / 100	-0.122**	-0.030**	-0.009	-0.019+	-0.024***	-0.008***	-5.873	-1.949
	(0.047)	(0.010)	(0.0014)	(0.10)	(0.004)	(0.002)	(7.001)	(2.174)
Fixed effects	Event	Worker	Event	Worker	Event	Worker	Event	Worker
Observations	32,975	32,975	178,478	178,478	178,478	178,478	178,478	178,478
Adjusted R^2	0.463	0.370	0.363	0.273	0.214	0.214	0.056	0.018

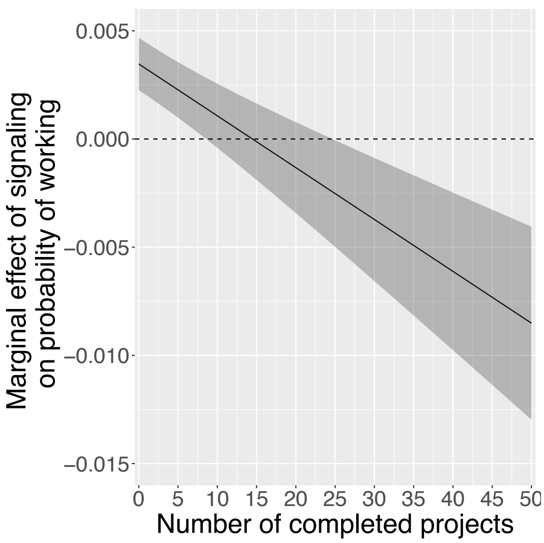
Note: In addition to the variables reported, all models include year dummies, an average rating for completed projects, a dummy variable with value 1 if a worker does not have ratings from past projects, cumulative (arsinh-transformed) dollars earned on the platform, and the cumulative number of completed projects on the platform, measured at the time of project start. Standard errors are clustered on the worker level. The significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%.



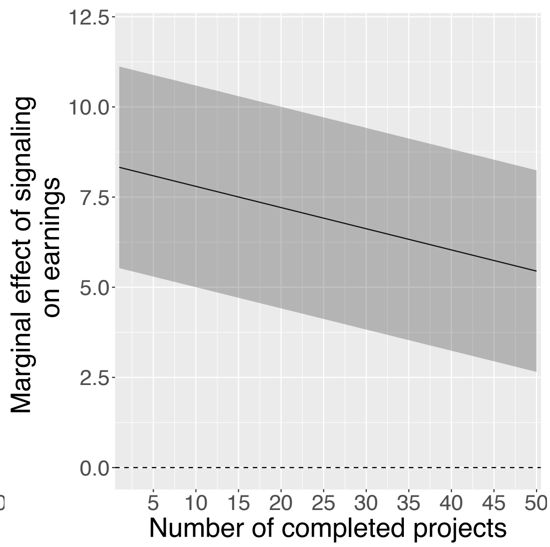
(a) Dependent variable: Project value (log)



(b) Dependent variable: The number of projects won



(c) Dependent variable: The number of projects won > 0



(d) Dependent variable: Earnings

Figure 4
The Marginal Effect of Signaling for Different Levels of Experience

Notes: The estimates are from regression models that control for worker fixed effects. The gray bands correspond to 95% confidence intervals, calculated as $+/-1.96 \times$ standard error.

6.2 Decreasing Returns to Signaling

As suggested by Prediction 3, one would expect the gains from signaling to be lower for higher levels of signaling. We implement the test for decreasing returns to signaling by introducing a quadratic term s^2 into Equations (11)–(12). As above, while our preferred specifications are the ones controlling for worker fixed effects, we estimate separate models using worker and worker-event fixed effects.

Table 6 and Figure 5 report the estimation results. As evidenced by the consistently negative estimates for the quadratic term, the returns to signaling are found to be decreasing in s. The monetary returns for signaling are high for the first few microcredentials completed (up to 16%) but quickly go down after that. According to Figure 5, the effect of signaling on project value is indistinguishable from zero after the 12th completed microcredential.

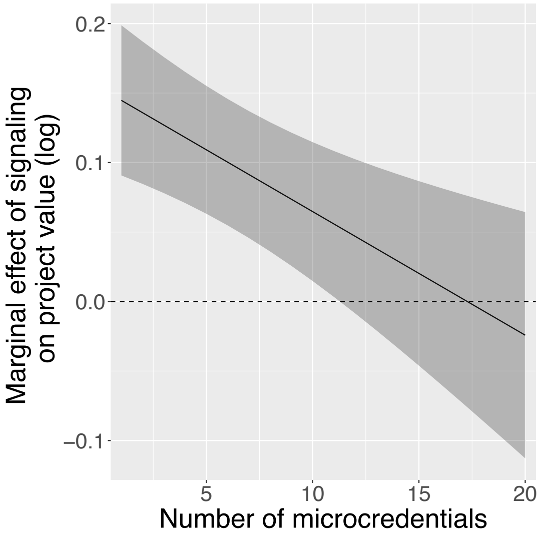
A comparison between Table 6 and Table 3 shows that the average returns reported in Table 3 conceal considerable differences. The effect of the first completed microcredential is up to 1.5 times greater than the average effect. Evidence of decreasing returns to signaling is stronger in the project value and number of projects margins. For the probability-of-working margin and the earnings margin, the decreasing returns are less pronounced, which might be attributable to the tiny effect sizes of signaling on those margins to begin with. The results are nevertheless consistent with Prediction 3.

As a specific caveat, we cannot rule out that the decreasing returns capture the fact that workers who specialize in certain types of projects are likely to prioritize microcredentials matching their specialization. As a concrete example, earning a microcredential in Spanish translation might be more valuable to a translator specializing in Spanish than earning a microcredential in some other language. We turn to skills signaling across different dimensions of skills and the project-microcredential fit in the next section.

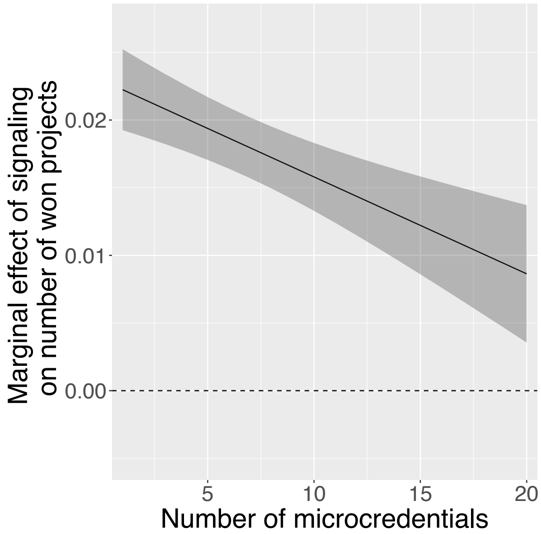
Table 6
Nonlinear returns to signaling

	Dependent variable:							
	Project value (log)		Number of projects		Number of projects > 0		Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of microcredentials	0.154***	0.057***	0.023***	0.024***	0.004***	0.012***	10.740***	8.328***
	(0.029)	(0.010)	(0.002)	(0.002)	(0.001)	(0.001)	(1.695)	(1.339)
Number of microcredentials^2 / 10	-0.044*	-0.010*	-0.004**	-0.004**	-0.001	-0.002***	-1.487**	-1.106**
	(0.014)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.382)	(0.310)
Fixed effects	Event	Worker	Event	Worker	Event	Worker	Event	Worker
Observations	32,975	32,975	178,478	178,478	178,478	178,478	178,478	178,478
Adjusted R^2	0.463	0.370	0.363	0.247	0.272	0.213	0.056	0.018

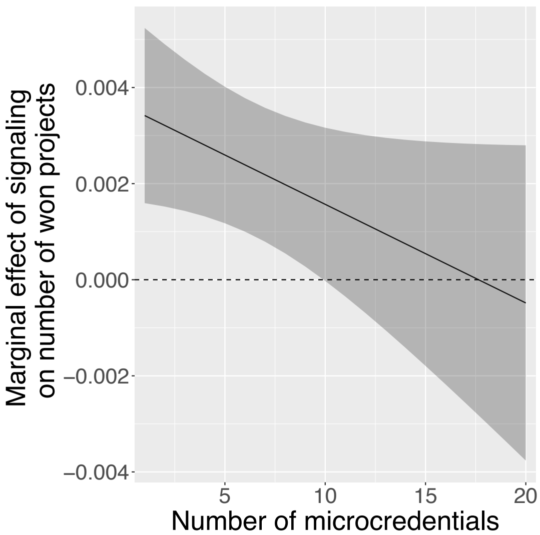
Note: In addition to the variables reported, all models include year dummies, an average rating for completed projects, a dummy variable with the value 1 if a worker does not have ratings from past projects, cumulative (arsinh-transformed) dollars earned on the platform, and the cumulative number of completed projects on the platform, measured at the time of project start. Standard errors are clustered on the worker level. The significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%.



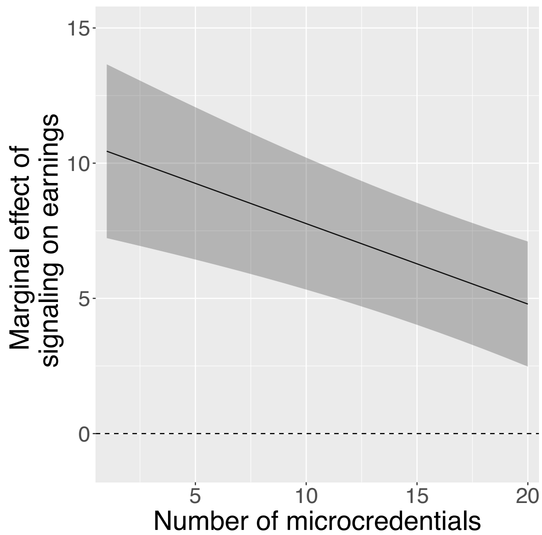
(a) Dependent variable: Project value (log)



(b) Dependent variable: The number of projects won



(c) Dependent variable: The number of projects won > 0



(d) Dependent variable: Earnings

Figure 5

Marginal effect of signaling for different levels of experience. Notes: the estimates are from regression models that control for worker fixed effects. The gray bands correspond to 95% confidence intervals, calculated as +/- 1.96 × standard error.

6.3 Multidimensional Skills and Signaling Across Project Types

Thus far, we have treated ability as a one-dimensional variable and assumed that all microcredentials act as signals for the same ability. This is obviously a simplifying assumption. While the data show that workers tend to specialize in a single category of projects and complete microcredentials within their specialization (Table 2), some workers complete microcredentials and projects across multiple categories. This feature of the data allows us to study how completing a microcredential for a skill of the type m affects the labor market outcomes in projects of the type n $(n \neq m)$.

We implement this test by first filtering out all microcredential completion events other than type m and then rerunning the fixed effect event study regressions from Equations (11)–(12):

$$y_{iket} = \alpha_e + X_{ik}\beta + \gamma s_{\substack{m \\ ik}} + \delta\left(I(project\ type = m) \times s_{\substack{m \\ ik}}\right) + v_t + \varepsilon_{iket}, \tag{15}$$

$$Z_{ijet} = \alpha_e + X_{ij}\beta + \gamma s_{\substack{m \\ ij}} + \delta\left(I(project\ type = m) \times s_{\substack{m \\ ij}}\right) + v_t + \varepsilon_{ijet}. \tag{16}$$

As in Section 5.2, we use event fixed effects (denoted by α_e). The parameter γ identifies the returns to completing microcredentials for projects of the type n $(n \neq m)$, while δ captures the effect for projects of the type m. The regression results are reported in Table 7. We find that the estimate for δ is either imprecisely estimated or highly positive, while the estimates for γ are usually (but not always) statistically indistinguishable from zero. This suggests that, in general, microcredentials of category m increase labor market success in projects of category m, but the effect on projects in other categories is statistically indistinguishable from zero.

Table 7
Return to signaling by project and test types

	Dependent variable:			
	Project value (log)	Number of projects	Number of projects > 0	Earnings
	(1)	(2)	(3)	(4)
Panel A: Only including design microcredentials				
Number of microcredentials	-0.024	0.005	0.003	9.864**
	(0.071)	(0.004)	(0.002)	(3.698)
Number of microcredentials × design project	0.057	0.052***	0.027***	6.208***
	(0.044)	(0.007)	(0.003)	(1.364)
Observations	3,726	41,440	41,440	41,440
Panel B: Only including finance microcredentials				
Number of microcredentials	0.277	-0.006	-0.005*	2.322
	(0.196)	(0.004)	(0.003)	(2.669)
Number of microcredentials × finance project	-0.056	0.026***	0.018***	10.450***
	(0.050)	(0.004)	(0.003)	(2.215)
Observations	743	12,892	12,892	12,892
Panel C: Only including sales and marketing microcredentials				
Number of microcredentials	0.141	0. 012*	0.004	23.136+
	(0.136)	(0.004)	(0.003)	(13.439)
Number of microcredentials × sales and marketing project	0.125	0.035***	0.026***	-3.566
	(0.210)	(0.004)	(0.003)	(15.688)
Observations	18,869	28,884	28,884	28,884
Panel D: Only including technology microcredentials				
Number of microcredentials	0.190****	0.0002	-0.002**	3.271+
	(0.052)	(0.001)	(0.001)	(1.957)
Number of microcredentials × technology project	-0.049+	0.020**	0.012***	13.376***
	(0.025)	(0.002)	(0.001)	(2.204)
Observations	10,776	165,396	165,396	165,396
Panel E: Only including virtual assistant microcredentials				
Number of microcredentials	0.180	0.040***	0.025***	11.860+
	(0.272)	(0.007)	(0.004)	(6.639)
Number of microcredentials × virtual assistant project	-0.046	0.012	0.009+	21.582*
	(0.260)	(0.008)	(0.005)	(8.856)
Observations	988	18,584	18,584	18,584
Panel F: Only including writing and translation microcredentials				
Number of microcredentials	0.042	0.006*	0.003*	6.435***
	(0.054)	(0.003)	(0.001)	(1.677)
Number of microcredentials × writing and translation project	0.019	0.040***	0.022***	7.473***
	(0.025)	(0.004)	(0.002)	(1.178)
Observations	8,394	132,272	132,272	132,272

Notes: Estimates are from the fixed effects event study specification. In Column 1 the unit of observation is one project. In Columns 2–4, the unit of observation is a 14 day pre- or post-test period. In addition to the variables reported, all models include year dummies, an average rating for completed projects, a dummy variable with the value 1 if the freelancer does not have ratings from past projects, cumulative (arsinh-transformed) dollars earned on the platform, the cumulative number of completed projects on the platform, measured at the time of project start, and event fixed effects. Standard errors are clustered on the worker level. The significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%.

6.4 Do Returns to Signaling Vary by Test Score?

Worker's ranking within the test score distribution is highly salient in their profile. Therefore, we expect the returns to signaling to be higher for higher test scores. The strategic choice of workers to publish their test scores might bias our estimates, but it is still instructive to examine how the returns to signaling vary across the observed distribution of test scores. We operationalize the test for the effect of scores by using the following regression specification:

$$y_{iket} = \alpha_T + X_{ik}\beta + \gamma s_{ik} + \theta(s_{ik} \times p_{iT}) + v_t + \varepsilon_{iket}, \tag{17}$$

$$Z_{ijet} = \alpha_T + X_{ij}\beta + \gamma s_{ij} + \theta(s_{ij} \times p_{iT}) + v_t + \varepsilon_{ijet}. \tag{18}$$

Here, we repeat the event study specification from Equations (11)-(12) with the exception that α_T are fixed effects for tests (one fixed effect for each test topic), and v_t are year fixed effects. The term p_{iT} is the percentile rank of worker $i's$ test score in the distribution of disclosed test scores related to test T. Intuitively, the term θ captures how the returns to signaling vary between workers who have completed the same test in the same year but have scored differently. The score percentiles are standardized by subtracting the average test score percentile of disclosed tests ($\overline{p_T}$) and dividing it by their corresponding standard deviation to ease interpretation.

We present the regression results in Table 8. We find that the coefficient on the interaction term is consistently positive, which implies that scoring better leads to a larger increase in labor market success. The estimates for γ (returns to signaling for a worker at the mean of the disclosed score distribution) are either indiscernible from zero or negative. This suggests that the workers who rank highest are the only ones benefiting from signaling.

Table 8
Returns to signaling by test score

	Dependent variable:			
	Log (project value)	Number of projects	Number of projects > 0	Earnings
	(1)	(2)	(3)	(4)
Number of microcredentials	-0.012^{***}	0.0004	0.0004	0.491
	(0.003)	(0.002)	(0.002)	(1.758)
Number of microcredentials × standardized test score rank	0.013^{***}	0.001^{+}	0.001^{+}	5.450^{***}
	(0.002)	(0.001)	(0.001)	(1.346)
Fixed effects	Test	Test	Test	Test
Observations	99,379	255,674	255,674	255,674
Adjusted R^2	0.145	0.222	0.222	0.02

Notes: Estimates are from event study specification with test fixed effects. In addition to the variables reported, all models include year dummies, an average rating for completed projects, a dummy variable with the value 1 if a worker does not have ratings from past projects, cumulative (arsinh-transformed) dollars earned on the platform, and the cumulative number of completed projects on the platform, measured at the time of project start. The significance levels in all specifications are: *** = 0.1%, ** = 1%, * = 5%, and + = 10%.

6.5 Robustness analyses

In this subsection, we present two additional robustness tests that support our fixed effect event study results. A potential worry for the validity of these results is that they are biased by varying effort around the time of skill test completion. A transitory dip in earnings a short while before microcredential completion could bias upwards the estimate for the return to microcredentials. A similar upward bias could emerge if workers strategically apply to more or better-paid jobs after the completion of the microcredential. We follow Hausman and Rapson (2018) in estimating a so-called donut specification, where we drop the observations that fall within –7 and 7 days of microcredential completion. Comparing the donut estimates to the main specification can help us detect any short-term effects (if the short-term effects last under 7 days). The results of this test are reported in Appendix 4. The donut and main estimates are statistically indistinguishable from one another at conventional significance levels.

An additional validity concern for our empirical strategy that any secular trends in the labor market might bias our return estimates. We provide evidence against this in Appendix 5, where we have re-estimated the event fixed effect models while adding 365 days to the microcredential completion date. The results of this exercise are either statistically indistinguishable from zero or negative but economically insignificant.

7 Discussion and conclusions

The main result of this paper is that taking computer-administered tests that award digital microcredentials increases worker earnings in an online labor market. Gaining an additional microcredential resulted in an average earnings gain of 8.9% over the following two weeks. We argue that the most plausible explanation for this is that microcredentials decrease employers' uncertainty over worker ability. Our theoretical model posits that employers prefer workers with validated skills and therefore pay them more for the same jobs.

We quite confidently rule out the alternative explanation that unobservable increases in worker productivity would drive the results. It is likely that the workers' skills remain approximately constant over the short time periods we concentrate on. We also find evidence against an alternative hypothesis of varying worker effort. That is, we do not find that workers systematically win more demanding projects after microcredential completion. Instead, our results indicate that after completing a microcredential, workers earn more money from projects that appear otherwise similar to the projects they worked on before completing the test.

An alternative for employers is to perform skill validation themselves, but this is costly. According to Corporaal and Lehdonvirta (2017), some employers on online labor platforms first hire workers into what could be termed "micro-internships": small test projects that are used to screen workers before hiring them into larger projects. Platform-administered microcredentials allow the employers to forgo some of their own skills validation, and this reduced employer screening cost is compensated for in the form of higher pay for credentialed workers. Since the skill tests are administered remotely, they could be easier to cheat in than tests administered at a test site; however, the positive marginal effects of completing the tests indicate that employers nevertheless trust that they convey some information about worker ability.

Employer uncertainty makes it difficult for candidates new to the labor market to gain work, as they lack references from previous employers (Pallais, 2014). Microcredentials are highlighted in labor policy literature as a possible intervention to address this problem (Painter and Bamfield, 2015; OECD, 2019; European Commission, 2020; Cedefop, 2021). However, we find that microcredentials' positive effect on worker earnings is mainly driven by an increase in project value, which increases 9.7% following microcredential completion. The number of projects initiated within the next 14 days increases by only 5.5%. Given the extremely low baselines, the point estimate implies that workers win one new project for approximately 63 microcredentials completed. Thus, microcredentialing does not appear to be very useful to new workers struggling to land their first job on the market. Experienced workers who have already accumulated a significant work history also fail to benefit from microcredentials because verified work history and feedback from previous employers are substitutes for credentials in reducing employer uncertainty. Taken together, these effects leave a narrow range of workers who benefit significantly from microcredentials: early-career workers who have successfully broken into the market but still lack a more extensive work history. These workers were able to obtain a marginal income increase of almost 15% by completing microcredentials.

An alternative explanation for our findings could be that the online marketplace platform itself might favor workers who obtain microcredentials. Horton (2017) describes worker recommendation algorithms that recommend workers to employers. The algorithms could conceivably be designed to recommend workers who complete microcredentials. However, according to Horton, the systems primarily affect the probability of winning a project; this is inconsistent with our finding that microcredentials mainly affected the project value margin.

These findings imply that microcredentialing system designers should ensure that their skill tests are sufficiently challenging to make credential completion informative to employers. The tests should be “cheap” to take but difficult to ace. However, even then, microcredentials may only be able to reduce employer uncertainty over candidates’ hard skills, but not soft skills and general trustworthiness, for which references from previous employers are more effective. Therefore, microcredentialing may not be an effective intervention to reduce entry barriers to new workers, such as unemployed youth.

Microcredentialing could be a somewhat more effective intervention to help skilled members of groups who face statistical discrimination in a labor market, such as immigrants (Oreopoulos, 2011) and minorities (Lang and Manove, 2011), to move to jobs better matching their skills. Microcredentialing could also help improve labor market matches in situations where public qualification schemes are too slow to keep up with rapidly changing skill demands (Painter and Bamfield, 2015). However, due to the imperfect substitutability between credentials and work experience, credentialing schemes are still likely to be only a partial solution, with other institutions also continuing to be needed.

Our research design does not allow us to make direct inferences about the general equilibrium effects of microcredentials. We cannot rule out the possibility that microcredentials simply cause employers to substitute non-certified workers with certified workers.

Nonetheless, this seems unlikely since the effects of signaling are mostly found on the project value and earnings margins, and not on the number of projects won margin.

References

Abadie, A., and J. Spiess (2021). "Robust post-matching inference." *Journal of the American Statistical Association*, forthcoming.

Agrawal, A., N. Lacetera, and E. Lyons (2016). "Does standardized information in online markets disproportionately benefit job applicants from less developed countries?" *Journal of International Economics* 103, 1–12.

Altonji, J. G., & Pierret, C. R. (2001). "Employer learning and statistical discrimination." *The Quarterly Journal of Economics*, *116*(1), 313-350.

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

Autor, D. H. (2001). "Wiring the Labor Market." *Journal of Economic Perspectives* 15(1):25–40.

Autor, D. H., and D. Scarborough (2008). "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments." *Quarterly Journal of Economics* 116, 1409–48.

Barach, M. A., & Horton, J. J. (2021). How do employers use compensation history? Evidence from a field experiment. *Journal of Labor Economics*, *39*(1), 193-218.

Blackburn, M. L., and D. Neumark (1993). "Omitted-ability bias and the increase in the return to schooling." *Journal of Labor Economics* 11(3):521–44.

Cedefop (2018). "*European inventory on validation of non-formal and informal learning 2018 update: Final Synthesis Report.*"
https://www.cedefop.europa.eu/files/8610_en_executive_summary.pdf

Cedefop (2020). “Developing and matching skills in the online platform economy”. https://www.cedefop.europa.eu/files/3085_en.pdf

Cedefop (2021). “Microcredentials: a labour market megatrend.” *Skillset and Match*: *Cedefop’s Magazine Promoting Learning for Work,* Issue 22, May 2021. https://www.cedefop.europa.eu/files/9156_en.pdf

Chabé-Ferret, S. (2017). “Should we combine difference in differences with conditioning on pre-treatment outcomes?” https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2017/wp_tse_824.pdf

Chevalier, A., C. Harmon, I. Walker, and Y. Zhu (2004). “Does education raise productivity, or just reflect it?” *Economic Journal* 114(499).

Corporaal, G. F., and V. Lehdonvirta (2017). “*Platform Sourcing: How Fortune 500 Firms are Adopting Online Freelancing Platforms.”* Oxford: Oxford Internet Institute.

Dale, S. B., and A. B. Krueger (2014). “Estimating the effects of college characteristics over the career using administrative earnings data.” *Journal of Human Resources* 49(2):323–58.

Dehejia, R. H., & S. Wahba (2002). “Propensity score-matching methods for nonexperimental causal studies.” *Review of Economics and Statistics*, 84(1):151–61.

European Commission (2020). “*Questions and Answers: European Skills Agenda for Sustainable Competitiveness, Social Fairness and Resilience”* (https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1197, accessed 2020-07-29).

Griliches, Z. (1977). “Estimating the returns to schooling: Some econometric problems.” *Econometrica* 45(1):1–22.

Harmon, C., and I. Walker (1995). “Estimates of the economic return to schooling for the United Kingdom.” *American Economic Review* 85(5):1278–86.

Hausman, C., and D. S. Rapson (2018). “Regression discontinuity in time: Considerations for empirical applications.” *Annual Review of Resource Economics* 10():533–552.

Heckman, J. J., H. Ichimura, & P. E. Todd (1997). “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme.” *Review of Economic Studies* 64(4):605–54.

Hoffman, M., L. B. Kahn, and D. Li (2018). “Discretion in hiring.” *Quarterly Journal of Economics* 133(2):765–800.

Horton, J. J. (2017). “The effects of algorithmic labor market recommendations: Evidence from a field experiment.” *Journal of Labor Economics* 35(2).

Iacus, S. M., King, G., & Porro, G. (2012). “Causal inference without balance checking: Coarsened exact matching.” *Political analysis*, *20*(1), 1-24.

Kato, S., V. Galán-Muros, and T. Weko. 2020. “*The Emergence of Alternative Credentials”*. 216. OECD Education Working Paper. Paris. doi:10.1787/b741f39e-en.

Lang, K., and M. Manove (2011). “Education and labor market discrimination.” *American Economic Review* 101(4):1467–96.

Lazear, E. P. (1998). “Hiring risky workers.” In *Internal Labour Markets, Incentives and Employment,* pp. 143–58. Springer.

Lechner, M., & Wunsch, C. (2013). Sensitivity of Matching-based Program Evaluations to the Availability of Control Variables. *Labour Economics*, *21*, 111-121.

Ozimek, A. (2021). “When Work Goes Remote.” https://ssrn.com/abstract=3777324

Painter, A., and L. Bamfield (2015). “*The New Digital Learning Age: How We Can Enable Social Mobility through Technology."* London: The Royal Society of Arts.

Pallais, A. (2014). “Inefficient hiring in entry-level labor markets*.” American Economic Review* 104(11):3565–99.

Spence, M. (1973). “Job market signaling.” *Quarterly Journal of Economics* 87(3):355–74.

Stanton, C. T., and C. Thomas (2015). “Landing the first job: The value of intermediaries in online hiring.” *Review of Economic Studies* 83(2):810–54.

Stephany, F., Dunn, M., Sawyer, S., & Lehdonvirta, V. (2020). “Distancing Bonus or Downscaling Loss? The Changing Livelihood of Us Online Workers in Times of COVID-19.” *Tijdschrift voor Economische en Sociale Geografie*, *111*(3), 561-573.

Stuart, E. A., King, G., Imai, K., & Ho, D. (2011). “MatchIt: nonparametric preprocessing for parametric causal inference.” *Journal of statistical software*, 42(8).

Terviö, M. (2009). “Superstars and mediocrities: Market failure in the discovery of talent.” The Review of Economic Studies, 76(2), 829-850.

Tyler, J. H., R. J. Murnane, and J. B. Willett (2000). “Estimating the signaling value of the GDE.” *Quarterly Journal of Economics* 115(2):431–68.