CESifo WORKING PAPERS

7810
2019

August 2019

# Do Digital Skill Certificates Help New Workers Enter the Market? Evidence from an Online Labour Platform

*Otto Kässi, Vili Lehdonvirta*

CESifo

**Impressum:**

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7810*
*Category 4: Labour Markets*

# Do Digital Skill Certificates Help New Workers Enter the Market? Evidence from an Online Labour Platform

## Abstract

We study the effects of a voluntary skill certification scheme in an online freelancing labour market. We show that obtaining skill certificates increases freelancers' earnings. This effect is not driven by increased freelancer productivity but by decreased employer uncertainty. The increase in freelancer earnings is mostly realised through an increase in the value of the projects won rather than an increase in the number of projects won. Moreover, we find evidence for negative selection to completing skill certificates, which suggests that the freelancers who complete more skill certificates are in a more disadvantaged position in the labour market.

JEL-Codes: J210, J230, J240, J310, I200.

Keywords: signaling, human capital, skill validation, skill certificates, micro-credentials, online freelancing, platforms, gig economy, computer-based assessment.

*Otto Kässi**
*Oxford Internet Institute*
*1 St Giles*
*United Kingdom – OX13JSe Oxford*
*otto.kassi@oii.ox.ac.uk*

*Vili Lehdonvirta*
*Oxford Internet Institute*
*1 St Giles*
*United Kingdom – OX13JSe Oxford*
*vili.lehdonvirta@oii.ox.ac.uk*

*corresponding author

28th May 2019
We are grateful to Fabian Braesemann, Stijn Broecke, Gretta Corporaal, Paolo Falco, Andrew Green, Marguerita Lane, Duncan MacDonald, Rafe Martyn, Andrea Salvatori, and Alex Wood for helpful comments and discussion. Otto Kässi gratefully acknowledges financial support from the OECD Future of Work Fellowship. This research was supported by a grant from the European Research Council (grant number 639652).

# 1 Introduction

Newcomers often face difficulties breaking into labour markets. Even highly educated and skilled workers without work history often struggle to land their first jobs. An important reason for this that it is difficult for prospective employers to be sure of such unproven workers' productivity Pallais (2014). This information friction is particularly relevant in online labour markets, where buyers and sellers of digitally deliverable freelance work are matched remotely through online platforms. Without direct interactions such as face-to-face interviews, worker quality and motivation are especially difficult to ascertain (Autor, 2001; Malone and Laubacher, 1998). These information frictions are further exacerbated by the global nature of online labour markets, as prospective employers are faced with evaluating freelancers from unfamiliar backgrounds. This leads to outcomes where new freelancers face difficulties breaking into the market, whereas older, previously screened freelancers have the upper hand because there is less uncertainty about their productivity. Partly due to these reasons, many online freelancing platforms are extremely unequal in terms of earnings.[1]

Several online labour market platforms have introduced voluntary skill certification schemes to help skilled, but unvetted newcomers break into the market. In these schemes, workers take computer-based skills assessment tests, which grant them digital skill certificates, also known as "digital badges" or "micro-certifications". The purpose of this paper is to empirically and theoretically evaluate whether completing skill certificates helps freelancers. We demonstrate empirically that obtaining skill certificates operates as a type of a signalling device in the spirit of Spence (1973). Skill certificates do not increase freelancers' productivity, but demonstrate their ability, and consequently lead to decreased employer uncertainty and increased freelancer earnings. In a textbook version of a signalling model, the agents' signalling cost depends only on their ability. We argue that in the context of online labour, the net benefit of signalling, and therefore, the freelancers' decision to signal is determined by two parameters: their ability and the uncertainty that prospective employers have about their ability. To formalise this, we present a theoretical model that captures this essential heterogeneity (Heckman et al., 2006).

A recurring challenge in estimating returns to signalling is that the returns to signalling by education are confounded with increases in human capital. For instance, if we observe that education increases wages, it is difficult to tell whether the higher earnings are caused by increased information or by the increase in individuals' productivity (Blackburn and Neumark, 1993; Chevalier et al., 2004). Transaction-level data provided by online freelancing platforms

[1]For example, Wood et al. (2018) reports a 90:10 income inequality ratio of 19 among South-East Asian and Sub-Saharan African online freelancers based on survey data.}

has two appealing features for studying this phenomenon. First, the data contains a rich set of freelancers' background characteristics, which can be used as control variables. Second, the fact that these projects are relatively short and follow each other relatively frequently allows us to use the longitudinal dimension of the data to account for freelancer unobserved heterogeneity.

In an ideal research setting, a researcher would fully control freelancer ability when studying the effect of signalling on earnings. In this paper, we approximate the ideal setting by comparing freelancers' earnings before and after acquiring a skill certificate. This allows us to capture all time-invariant unobservable factors into freelancer fixed effects. In addition, we limit our attention to a 14-day time period around the awarding of the certificate. This way, we can ensure that the return estimates are not contaminated by individual learning or other time-varying human capital effects.

The digital nature of online platforms' skill certification schemes makes them partly non-verifiable, which may reduce their accuracy as signals. For example, a freelancer might ask someone else to take a test in their stead, copy answers from an online 'crib sheet', or hide poor test results from their profile, where permitted by the platform. Despite this, high-ability freelancers can still be expected to find it easier to obtain the skill certificates. Our empirical results can be interpreted as a test of how much information such skill certificates convey. If we find that completing skill certificates does not affect freelancers' labour market outcomes, this implies that the signal from skill certificates is too inaccurate to provide any additional information on freelancer ability. Our empirical results rule out this hypothesis. We find that completing an additional certificate has a positive effect on both the number of projects obtained and the income earned from each project. However, the positive return estimates are only found in models that control unobserved heterogeneity using freelancer and test fixed effects. This finding suggests that the freelancers who are worse off in the labour market, such as discriminated-against groups, tend to complete more skill certificates to offset their disadvantage.

Signalling model has a set of clear-cut empirical predictions which will be used to validate the theory. In particular, freelancers' incentive to signal is smaller if employers' uncertainty about their ability is lower. Our finding that standardised information generated by completing projects on the platform decreases returns to signalling supports this prediction. This implies that signalling ability through skill certificates is to some extent a substitute for other types of standardised information on freelancer quality. In addition, the returns to signalling are found to be decreasing with the number of skill certificates completed. This suggests that the marginal effect of obtaining an additional certificate is smaller for freelancers who have previously earned certificates. On the other hand, we do not find any effects of signalling on project ratings, which suggests that taking skill tests does not increase freelancer performance in projects.

The results of this paper contribute to multiple strands of empirical literature. First, the research links to an emerging literature on how various types of online labour market institutions affect employment outcomes on online platforms. The most closely related papers include Pallais (2014). She discusses a field experiment where she randomly hired inexperienced freelancers and then provided feedback on their performance. She compares her hires' subsequent income to other freelancers who applied for her posted jobs but were not hired. The randomly hired freelancers earned considerably more from their subsequent jobs compared to the control group. She argues that this effect is a result of the information that her feedback provided on the hired freelancers to other potential employers. Extending on Pallais' results, Agrawal et al. (2016) show that standardised information on platform-based work experience benefits all freelancers, but that freelancers from developing countries benefit more. Lehdonvirta et al. (2018) show that platform-generated information on work experience has a greater effect on earnings than employer feedback ratings and skill certificates and that freelancers from lower-income countries benefit from all these signals more. This suggests that employers initially have more difficulty evaluating the quality of freelancers from developing countries compared to developed countries. Relatedly, Stanton and Thomas (2015) show that information from intermediaries helps inexperienced freelancers: a freelancer affiliated with an intermediary agency has a higher job finding probability and wage. Horton (2017) shows that algorithmic recommendations of freelancers to employers can improve the functioning of online freelancing labour markets by reducing search frictions. Effects of skill certificates have not been systematically studied in the literature on online labour platforms. Barach and Horton (2017) show that in addition to work experience, employers use freelancers' past hourly wages as a signal of their quality, and concealing freelancers' past realised hourly wages led to employers having exert more effort in interviewing freelancers.

This paper also contributes to the literature on using standardised tests as a method for revealing information on worker quality in traditional labour markets (Autor and Scarborough, 2008; Hoffman et al., 2018). A recurring theme in this literature is that standardised tests can benefit minorities and other statistically discriminated against groups in the labour market. We provide a detailed analysis of standardised tests' effects on new entry. More broadly, the results link to empirical studies on job market signalling (Tyler et al., 2000; Lang and Manove, 2011; Pinkston, 2003; Arcidiacono et al., 2010; Dale and Krueger, 2014), where we contribute an analysis of the role of employer uncertainty in signalling decisions.

In policy literature, new forms of skill validation – and especially micro-credentials based on computer-based assessments – have been proposed as a potential solution to improving labour market matches in an era of rapidly-changing skill requirements Painter and Bamfield (2015). Our paper presents one of the first empirical analyses of such a scheme.

## 2 Empirical setting

The dataset used in this paper was collected from one of the largest online labour platforms, which did not wish to be identified. Before turning to the details, we briefly present a typical workflow of contracting within the platform. Employers looking to hire a freelancer for a particular task typically start the process by posting an opening on the site. The opening includes the skills required, expected contract duration, preferred freelancer characteristics and the contract type, which can be either a flat rate or an hourly billed contract. A major difference between flat rate and hourly projects is that with the latter, the platform allows employers to use monitoring technologies, namely screenshots taken automatically at semi-regular intervals from freelancers' screens, and records of the rate of keystrokes and mouse clicks. These technologies are not available for flat rate contracts, where the freelancer's work can only be evaluated once the output is delivered.[2]

After the vacancy is posted, it is visible to registered freelancers, who can apply for the position by submitting private bids. The interview and wage negotiation phases also take place on the platform. When posting a project, the employer also chooses a category for the project. Projects are classified into 12 broad categories (software development, graphic design, writing, etc.), which are further broken down to 89 distinct subcategories (such as mobile development, game development, and software testing, which are all subcategories of software development).

Of particular interest to us are the skill tests administered on the platform. The tests are administered as multiple-choice quizzes scored automatically. In all, freelancers can take over 300 distinct skill tests from topics such as the English language, programming languages, graphic design techniques, and office software packages. Once a skill test has been completed, the freelancer gets a "badge" certifying its completion (see Figure 1). The badge also shows the freelancer's numerical grade and percentile rank among all test takers. When inviting freelancers to a project, employers can limit their search to those who have completed a particular skill certificate or have scored over a certain threshold in it. If the freelancer has tried to take a test and failed, a failed mark is not visible to potential employers. The freelancer can retake the test after a cooldown period lasting between 30 and 180 days. Freelancers can also choose to hide the results of tests they have passed.

The tests are highly technical in nature. They are designed to test details on the particulars

[2]For more details on the differences between the two project types, see Lin et al., 2016.

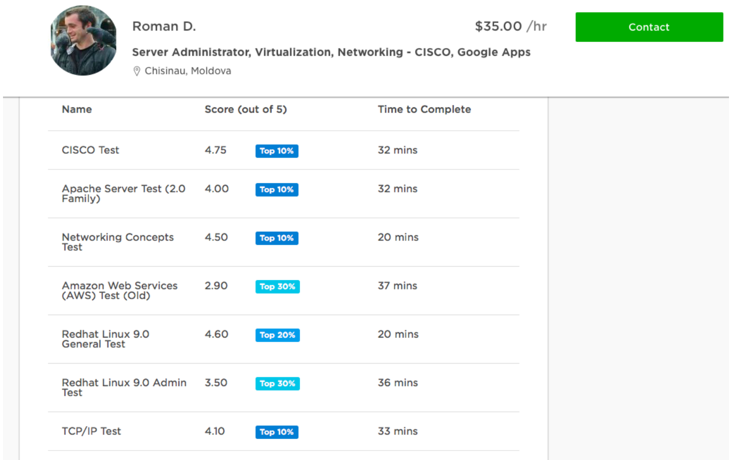



Figure 1: Screenshot of a freelancer's profile featuring skill certificates.

of the topic being tested. Therefore, the outcome of the test is not likely to depend on the general skill level of the freelancer, but rather on their specific knowledge on the topic[3]

A crucial assumption for this paper is that a freelancer does not learn anything from just taking the test. This a reasonable assumption since it is not plausible that a freelancer would pick up a skill such as a programming language or a foreign language – which typically takes months or years to learn – when taking a simple multiple-choice test.

# 3 Motivating theoretical framework

This section introduces a signalling model, which we use to show that employer uncertainty on freelancer ability creates an incentive for freelancers to invest in a costly signal. The model is a slightly modified version of the model presented by Lang and Manove (2011). It provides testable implications on how the level of freelancer signalling varies with uncertainty about their productivity, as well as on how returns to signalling vary with uncertainty about productivity.

Freelancers differ in their ability. Only the freelancers are assumed to know their own ability. Potential employers observe freelancers' certificates accurately, but freelancer ability is observed with noise. Consequently, the freelancers have an incentive to gain certificates to reduce employer uncertainty. Throughout the rest of this section, we assume that there exists a separating equilibrium; freelancers' ability and employer uncertainty determine how much freelancers signal.

Following the bulk of the literature in labour economics (Blackburn and Neumark (1993); Griliches (1977); Harmon and Walker (1995) to name a few), we define "ability" as everything that is time-invariant, unobservable to the researchers and is positively correlated with both the

[3] A typical 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?*"

number of completed skill certificates and success in the labour market. These could include skills and intelligence, but also things such as determination and being serious about online freelancing as a source of income.

Freelancers' ability is distributed along a fixed interval $[a_0, a_1]$. A freelancer's productivity $p^*$ in a given project conditional on their ability is given by

$$p^* = a + \varepsilon, \tag{1}$$

where $a$ is the freelancer's ability. $\varepsilon$ is a normally distributed match specific random variable which is only realised after the match between a freelancer and employer has been formed. It has a mean of 0 and variance of $\sigma_\varepsilon^2$.

A potential employer can observe the number of skill certificates the freelancer has, $s$, but not their true productivity $p^*$. The employer observes a noisy estimate of freelancer's productivity given by

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

where $u$ is another normally distributed random error term. The error term $u$ has a variance of $\sigma_u^2(s)$ which is common to all employers, and continuous with $\dfrac{\partial \sigma_u^2(s)}{\partial s} \leq 0$ and $\dfrac{\partial^2 \sigma_u^2(s)}{\partial s^2} \geq 0$ in signalling. $\varepsilon$ and $u$ are independent of one another, and their distributions are assumed to be common knowledge.

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

$$\lambda(s) = \frac{\sigma_\varepsilon^2}{\sigma_u^2(s) + \sigma_\varepsilon^2}.$$

For a given value of $\sigma_\varepsilon^2$, if $\lambda(s)$ is close to zero, then $\sigma_u^2(s)$ must be large, and, consequently, the employer's ability to observe freelancer productivity directly is poor. In this case the employers have to give more weight to the certificate signal. If $\lambda(s) = 1$ then $\sigma_u^2(s) = 0$ and the employer observes freelancer productivity perfectly and does not have to rely on signals.[4]

The employers follow the rules of a competitive labour market. They pay the freelancers the wage which is determined by their expected productivity. Their equilibrium inference of the freelancers' productivity, $p^*$, depends on the elements they observe, $p$ and $s$. Let $\hat{a} = a(s)$ denote employers' equilibrium inference on $a$ conditional on $s$. Throughout this paper, we assume that there exists a unique, continuous, differentiable abd strictly increasing in $a$ equilibrium which

[4] In a special case, where $\dfrac{\partial \sigma_u^2(s)}{\partial s} = 0$, signalling is completely uninformative, and $\dfrac{\partial \lambda}{\partial s} = 0$. In this case, skill certificates are fully uninformative of earnings, the freelancers have no incentive to signal, and employers give no weight to freelancers' skill certificates.

specifies a unique best response for every ability level and $\lambda$. To solve for $E\left[p^* \mid p, s\right]$, note that, Equations (1) and (2) imply that in equilibrium, $p - \hat{a} = u + \varepsilon$. Therefore,

$$\begin{aligned} E\left[w \mid p, s\right] &= E\left[p^* \mid p, s\right] \\ &= E\left[p^* \mid p - \hat{a}, s\right] \\ &= E\left[a \mid p - \hat{a}, s\right] + E\left[\varepsilon \mid p - \hat{a}, s\right]. \end{aligned} \tag{3}$$

Since, $E\left[a \mid p - \hat{a}, s\right] = \hat{a}$, and $E\left[\varepsilon \mid p - \hat{a}, s\right] = \dfrac{Cov(\varepsilon, u + \varepsilon)}{Var(u - e)}\left(p - \hat{a}\right) = \lambda\left(p - \hat{a}\right)$, equation (3) is equivalent to:

$$E\left[w \mid p, s\right] = \lambda p + \left(1 - \lambda\right)\hat{a}, \tag{4}$$

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

It is useful to note that Equation (4) implies that if there are two freelancers $L$, and $H$ with the same level of $a$, but $\sigma^2_{u,L}(s) > \sigma^2_{u,H}(s)$, freelancer $H$ is at an advantage because the employer can better evaluate their productivity. Therefore, freelancer $L$ will have a larger incentive to invest in signalling.

The freelancers' problem boils down to choosing $s$ to solve for

$$\max_s E\left[w\right] - c(a)s, \tag{5}$$

where $c(a)$ ($c(a) > 0$, for all $a \in [a_0, a_1]$) is the effort cost of getting a certificate. $c(a)$ is assumed to be decreasing and convex in $a$. In equilibrium, equation (5) simplifies to

$$\max_s \lambda E\left[p\right] + (1 - \lambda)\hat{a} - c(a)s. \tag{6}$$

Its first order condition reads as

$$\lambda_s a - \lambda_s \hat{a} + (1 - \lambda)\hat{a}_s = c(a) + c_a a_s, \tag{7}$$

where subscripts denote partial derivatives, ($\hat{a}_s = \dfrac{\partial \hat{a}}{\partial s}$). Therefore, in equilibrium, $a = \hat{a}$, Expression 7) simplifies to

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

which implicitly solves $s$for each combination of $\lambda$ and $a$. Finally solving for $a_s$ and inverting yields

$$s_a = \frac{1 - \lambda - c_a}{c(a)}. \tag{9}$$

Equation (9) demonstrates that the equilibrium value of $s(a)$ is strictly increasing in $a$. We also know that the freelancer with the lowest level of ability does not invest into signalling, or $s(a_0) = 0$. To see why this the case, note that if $s(a_0) > 0$, the freelancer with $a > a_0$, could deviate to smaller $s$ without affecting employers' equilibrium inference on their ability. The only case when this is impossible is if $s(a_0) = 0$. Having confirmed that $s(a_0) = 0$, and noting that Equation (8) is continuous and differentiable, we know that $s(a)$ exists and is uniquely determined for all combinations of $a$ and $\lambda$.

Now, assume that there are two freelancers with the same level of $a$ but $\sigma^2_{u,L} > \sigma^2_{u,H}$ and consequently, $\lambda_L < \lambda_H$. In words, the employers face a more uncertainty when trying to evaluate the expected productivity of freelancer $L$ compared to freelancer $H$. Under this assumption, the theoretical framework laid out generates the following predictions.

1. If there are two freelancers $(L, H)$ with the same value of $a$ but $\lambda_L < \lambda_H$, we have $s(\lambda_L) > s(\lambda_H)$ whenever $a > a_0$. That is, higher employer uncertainty on freelancer ability results in more signalling by the freelancer. To see this, note that equation (8) implies that if $\lambda_L < \lambda_H$, then $s_a(\lambda_L) > s_a(\lambda_H)$. Furthermore, we argue above that $s(a_0; \lambda_L) = s(a_0; \lambda_L)$. By the continuity of $s$, this possible only if $s(\lambda_L) > s(\lambda_H)$.

2. If there are two freelancers with same $a$ but $\lambda_L < \lambda_H$, then $\frac{\partial E\left[w; \lambda_H\right]}{\partial s} > \frac{\partial E\left[w; \lambda_L\right]}{\partial s}$, or returns to signalling are higher if the uncertainty is higher. To see why this holds, note that $\frac{\partial^2 E\left[w\right]}{\partial s \partial \lambda} < 0$ for all $a > a_0$.

3. Finally, signalling exhibits decreasing returns to scale, so that $\frac{\partial^2 E\left[w\right]}{\partial s^2} < 0$ for all $a > a_0$.

Predictions 1. and 2. are intuitive. Freelancers who are statistically discriminated against, that is, for whom productivity uncertainty is higher, get a higher marginal return from signalling and signal more when their ability is kept constant. Predictions 2. and 3. suggest that signalling exhibits two types of decreasing returns. The marginal effect of signalling is lower for higher levels of signalling. Return to signalling is also lower if employer uncertainty about freelancer productivity is lower.

Equation (9) demonstrates that the choice of the level of signalling depends on two characteristics which are unobservable to the researcher, but which affect freelancer earnings. Freelancers with a higher ability signal more, because their cost of signalling is lower. On the other hand, freelancers who know that employers have problems evaluating their productivity also signal more. As a result, failing to control for these in an OLS regression of the number of skill certificates on earnings likely leads to a biased estimate on the coefficient on $s$.

In the empirical analysis section, we present comparisons between OLS estimates and fixed effect estimates. Since the fixed effects arguably subsume the unobservable ability $a$ and employer uncertainty $\lambda$, the direction of the bias of the OLS estimates can be used to infer which one of the two effects dominates. If OLS estimates are biased downwards, the decision to signal is negatively correlated with earnings, implying the bias due to employer uncertainty dominates the bias due to unobservable freelancer ability. If, instead, the OLS estimates are biased upwards, the unobservables in the earnings equation are net positively correlated with the decision to signal, which suggests that the decision to signal is driven by differences in unobservable ability.

# 4 Data and descriptive statistics

The dataset used in this paper was collected with assistance from the online labour platform, which provided access to their developer API to make the data collection possible, but was not otherwise involved in any aspect of the study design or sample construction. The data was collected in three steps. In the first step, we used the search functionality of the platform to sample freelancers from all job categories. The search functionality of the platform can order the search results in various ways opaque to the user in an attempt to increase the efficiency of the searches, which might lead to a nonrepresentative sample of the underlying population of freelancers. To overcome this, we randomly sampled 10% of the workers returned on each search result page. This approach also allowed us to collect a reasonably sized sample without violating the rate limits set by the API. After removing duplicates, we ended up with a sample of 46,791 freelancers. We used separate API requests to get background information on the freelancers and the details of each project they had completed, which totalled 422,199 projects. The main summary statistics of the data are presented in Table 1.

Our data is a panel. Freelancers are observed more than once; values of freelancer specific observables vary as they win more projects and get rated by employers. Each observation is a combination of freelancer-level observables and project characteristics. We refer to these as freelancer-project observations. The left panel of Table 1 presents the descriptive statistics for all 422,199 freelancer-project observations.

We apply an event study approach where a 14-day pre-test period acts as a control group for each 14-day post-test period. Since learning a new skill such as a new programming language usually takes much longer than 14 days, limiting the investigation to a short time window allows us to assume that freelancer ability remains roughly constant. The right panel thus shows the same descriptive statistics for projects falling within a +/-14-day window around skill certificate completion only. In cases where a freelancer has completed more than one skill test, this filtered sample contains several 14-day pre-test and 14-day post-test observation periods for the freelancer.[5] Overall, the two samples seem fairly similar; their means are within one standard deviation of each other. Nonetheless, the samples are very heterogeneous, as evidenced by the large standard errors on most variables.

To better capture the situation where an employer screens workers in the presence of limited information, we have applied further selection criteria to the filtered sample. We excluded projects where the employer explicitly invited a single freelancer to the job (5% of observations) as we assume that in these cases the employer had neglible uncertainty about the freelancer quality. In addition, to filter out projects where all applicants were hired and consequently no employer screening took place, we excluded projects where more than one freelancer was hired. We also excluded projects without dollars-earned information (0.02% of observations) from the filtered sample as artefacts. Finally, the platform also features ”readiness tests for new freelancers”. These are near-mandatory tests used to screen out freelancers who are not serious about freelancing work. We exclude these tests from our data since virtually every freelancer in our sample has completed them.

Freelancers' average reputation scores tend to be very close to a full five out of five stars. This a common observation in many online markets for both labour and goods (see Nosko and Tadelis 2015; Filippas et al. 2017). Nonetheless, as will become evident from the regression results, the ratings have some – albeit noisy – predictive power on freelancers' expected future earnings.

**The observed test score distribution of skill tests is unlikely to be representative of the underlying distribution of skills in the freelancer population. This is because both taking a skill test and disclosing the result of the test on one's profile is voluntary. The strategic disclosure of test scores implies that we would expect the observed test score distribution to be biased towards good ratings in comparison to the uncensored test score distribution. The test score distribution is plotted in Figure 2. As is clear from the picture, the probability of freelancers disclosing their test score is substantially higher if they have scored above the median. Since the test score distribution is censored from below due to strategic reporting, we operationalise intensity of freelancer signalling using the number of completed skill tests as our independent variable.**[6]

[5] We exclude observations where the freelancer completes another skill test less than 14 days after completing a skill test.

[6] The theoretical model presented in Section 3 does not, strictly speaking, consider the possibility that some freelancers might take skill tests but not disclose the results. This could be modelled by adding an extra step in

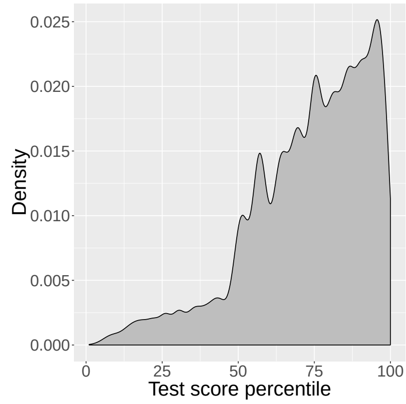



Figure 2: Kernel density of the test score distribution. Note: bandwidth chosen by rule-of-thumb method (Silverman, 1986).

Table 1: Descriptive statistics.

| | Full sample | | Filtered sample | |
|---|---|---|---|---|
| | Mean<br>(st. dev.) | Median | Mean<br>(st. dev.) | Median |
| Number of completed tests | 2.29<br>(3.87) | 0 | 4.69<br>(4.66) | 3 |
| Number of completed projects | 44.45<br>(65.63) | 21 | 19.95<br>(40.95) | 6 |
| Dollars earned | 11724.19<br>(23264.14) | 3604 | 4801.3<br>(13572.34) | 475.56 |
| Months active | 23.95<br>(20.13) | 19.17 | 28.42<br>(22.97) | 23.9 |
| Freelancer rating | 3.16<br>(1.56) | 3.54 | 3.01<br>(1.95) | 3.75 |
| Project value | 474.91<br>(2864.42) | 70 | 316.61<br>(1954.96) | 48 |
| N | 422,199 | | 338,91 | |

Notes: One observation corresponds to one freelancer-project observation. The left panel the descriptive statistics for the full sample; the right panel presents the descriptive statistics for the sample limited to -14, ...., +14 days around the completion of skill tests, and to projects with more than one applicant (see main text for details). Project-freelancer characteristics are measured at the time of project start.

# 5 Empirical analysis

## 5.1 Returns to signalling

We now turn to study empirically how signalling efforts are rewarded in the labour market. As suggested by the equation (9) in the previous section, the decision to complete certificates is driven by two types of selection on unobservable characteristics: freelancer ability and employer

which freelancers observe a noisy estimate of their own draw of $a$, and can choose to take skill tests to reduce their own uncertainty over their draw of $a$. This amendment would not affect the comparative static analyses, so it has been omitted for the sake of clarity.

uncertainty about freelancer ability. In an ideal setting, we would fully control freelancer ability and employer uncertainty when estimating the return to signalling. In the absence of these controls, a fixed effects estimator will subsume time-invariant heterogeneity and enable consistent estimation of returns to signalling. More concretely, we estimate the following fixed effects regression model:

$$y_{ik} = \alpha_i + X_{ik}\beta + \gamma s_{ik} + \nu_t + \varepsilon_{ik}. \tag{10}$$

Here, $y_{ik}$ is the log-value of a project $k$ won by freelancer $i$. On the right-hand side of the equation, $\alpha_i$ are either freelancer or freelancer $\times$ test specific fixed effects. Vector $X_{ik}\beta$ consists of measures of observable time-varying characteristics at the start of the project: number of previously completed projects, average reputation rating from previous projects, and the number of (log) dollars earned on the platform. To account for possible time heterogeneity, the specification includes observation year dummies, $\nu_t$. The main parameter of interest is $\gamma$, which measures the marginal effect of earning a skill certificate on the platform, and captures the effect of signalling on earnings.

We also study the employment margin, that is, how successful the freelancer is in winning projects in the first place. This studied using the specifications,

$$\begin{aligned} NumProjects_{ij} &= \alpha_{ij} + X_{ij}\beta + \gamma s_{ij} + \nu_t + \varepsilon_{ij} && (11) \\ I\left(NumProjects_{ij} > 0\right) &= \alpha_{ij} + X_{ij}\beta + \gamma s_{ij} + \nu_t + \varepsilon_{ij} && (12) \end{aligned}$$

where each 14-day time window is indexed with $j$, and $I(.)$ is an indicator function getting the value 1 whenever $NumProjects_{ij} > 0$. Our fourth specification combines the wage and employment margins. Here, the dependent variable is the (log) number of dollars earned in each 14-day pre- or post-test period:

$$log(earnings_{ij} + 1) = \alpha_i + X_{ij}\beta + \gamma s_{ij} + \nu_t + \varepsilon_{ij}. \tag{13}$$

Model estimates are presented in Table 2. Our preferred models are the ones reported in columns (1), (4), (7) and (10) which control for test $\times$ freelancer fixed effects. We argue that these models tackle both time-invariant unobserved heterogeneity, and time-varying unobserved heterogeneity, as long as it is relatively small in the +/-14 day time window around the time of skill certificate completion. Column (1) of Table 2 presents a specification that looks at the per-project earnings margin. It shows that an additional skill certificate leads to a 9.7% increase in project value. Transformed to dollars, this corresponds to a $9.7\% \times \$287.23 \approx \$30.15$ return to completing a skill certificate. When looking at the number of projects won, column (4) of Table 2 shows that completing a skill certificate leads to a $0.016/0.29 \approx 5.5\%$ increase in the number of projects initiated within the 14-day window. This is a relatively small effect economically speaking; the point estimate implies that freelancers win one new project for approximately 63 completed skill certificates. When the probability of working at least once is used as the dependent variable, we find that the marginal effect of completing a skill certificate is 0.2%. Relative to baseline, this is $0.02/0.18 \approx 11\%$.

Finally, when combining the income and projects won margins, column (5) of Table 2 shows a 2.1% increase in earnings. Transformed into dollars, this corresponds to an average earnings gain of $2.1\% \times \$89.45 \approx \$1.88$ The estimates line up very well. The marginal effect of signalling on both number of projects and earnings estimates is of the same magnitude and quite small, whereas the estimate on the project value margin is larger.

A comparison between OLS and the two fixed effects specifications reveals that the OLS estimates are always smaller. The more unobserved heterogeneity we subsume into the fixed effects, the higher the estimated marginal effect for skill certificates is. This implies that freelancer-specific earnings-related characteristics are negatively correlated with the decision to signal. In other words, there is a negative selection effect for completing skill tests. Freelancers who are in a disadvantaged position in the labour market signal more. In light of the theoretical model presented in the previous section, this suggests that the decision to complete skill certificates is driven by some freelancers having a disadvantaged position in the labour market in terms of employers having more uncertainty about their ability (the differences in $\lambda$), rather than by differences in ability between freelancers (the differences in $a$).

A few points on these specifications are worth making. Since our data does not include information on projects where the freelancer bid for a project but did not win, the results might be confounded by the bidding effort. In particular, if freelancers are more active in applying just after completing a skill certificate compared to just before completion, the estimate for signalling might be biased upwards. If this is the case, then the estimates reported in columns (1), (4), (7) and (10) should be interpreted as an upper limit for the estimate of the true effect of signalling on the probability of employment and earnings. We return to this point later in the robustness analysis section.

We also stress that the theoretical model outlined in the previous section has a set of specific predictions on how the returns to signalling vary. In the following sections, we proceed to show that the returns to signalling tend to be smaller when there is more information available on the freelancers. In addition, we will also demonstrate that, despite paying more, the projects that the freelancers win after a test tend to be similar to projects they won before the test across several observable dimensions.

## 5.2 Signalling as a substitute for other forms of verified information

Establishing that signalling decreases employer uncertainty on freelancers' productivity is complicated by the fact that the information set of the employer, and therefore their uncertainty about freelancer productivity is unobservable to researchers. Nonetheless, Prediction 2, outlined in the theory section, suggests that the marginal effect of signalling is lower for high levels of information. Comparing the marginal effects of completing skill certificates with different levels of platform-provided information allows us to test empirically whether Prediction 2 holds in the data.

We follow Agrawal et al. (2016) and Pallais (2014) and assume that employers have more information on more experienced workers. This assumption is plausible – more experienced workers have earned more feedback from previous employers and have established a work history, which both arguably decrease employer uncertainty.

Our empirical operationalisation of the test is straightforward: we include an interaction term between the number of completed skill certificates and work experience, and test for its significance. As above, we estimate the models using both freelancer $\times$test and freelancer fixed effects. In practice, we estimate variants of the following regression model:

$$y_{ik} = \alpha_i + X_{ik}\beta + \gamma s_{ik} + \delta n_{ik} + \eta\left(s_{ik} \times n_{ik}\right) + \nu_t + \varepsilon_{ik},$$

where, depending on the specification, the dependent variable is either, *log* of project value, number of projects won, dummy variable for freelancer having won at least 1 project, or *log* of 1+earnings. The coefficient on the interaction term between completed skill tests and work experience, $s_{ik} \times n_{ik}$, captures how the marginal benefit of signalling varies with the number of completed projects.

In this, and the following subsection, our main interest is in how the return to signalling varies *within* freelancers' career. Thus, our preferred specification is the one that controls for freelancer fixed effects rather than freelancer $\times$test fixed effects. We report the estimation results from both specifications to facilitate comparison with the previous section and to demonstrate that the results hold across both speicifications.

We present our estimation results in Table 3. In columns (1)-(2) and (4)-(8), where the dependent variables are the log of project value, probability of working, and earnings, we find

Table 2: Returns to signalling.

| | *Dependent variable:* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | log(project value) | | | Num projects | | | Num projects > 0 | | | log(1+dollars) | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Feedback rating | $0.105^{***}$ | $0.033^{*}$ | $0.092^{***}$ | $-0.002$ | $0.019^{**}$ | $0.013^{**}$ | $-0.010$ | $0.009^{***}$ | $0.009^{***}$ | 0.032 | $0.060^{***}$ | $0.064^{***}$ |
| | (0.028) | (0.015) | (0.010) | (0.022) | (0.006) | (0.005) | (0.009) | (0.003) | (0.002) | (0.051) | (0.013) | (0.011) |
| Number of completed projects | $0.054^{***}$ | $0.002^{**}$ | $-0.005^{***}$ | $-0.041^{*}$ | $0.004^{***}$ | $0.008^{***}$ | $0.041^{***}$ | $0.001^{***}$ | $0.002^{***}$ | $0.199^{***}$ | $0.007^{***}$ | $0.010^{***}$ |
| | (0.016) | (0.001) | (0.0005) | (0.017) | (0.001) | (0.001) | (0.009) | (0.0001) | (0.0001) | (0.041) | (0.001) | (0.001) |
| Number of certificates | $0.097^{**}$ | $0.030^{***}$ | $-0.011^{***}$ | $0.016^{***}$ | $0.012^{***}$ | $-0.005^{***}$ | $0.002^{+}$ | $0.006^{***}$ | $-0.003^{***}$ | $0.021^{***}$ | $0.031^{***}$ | $-0.015^{***}$ |
| | (0.035) | (0.006) | (0.003) | (0.003) | (0.002) | (0.001) | (0.001) | (0.001) | (0.0004) | (0.005) | (0.003) | (0.002) |
| 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 | FL × test | FL | No | FL × test | FL | No | FL × test | FL | No | FL × test | FL | 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 | 0.261 | 0.208 | 0.136 |

Notes: In columns (1)-(3) unit of observation is one project. In columns (3)-(12), unit of observation is 14-day pre- or post-test period. In addition to the variables reported, all models include year-dummies, and cumulative dollars earned on the platform (log-transformed). Baseline refers to the mean of the dependent variable; Marginal effect / baseline is the magnitude of the marginal effect of a skill certificate divided by the baseline. Standard errors are clustered on freelancer level. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% and + 10%.

negative and statistically significant effects in both freelancer × test and freelancer fixed effects specifications. In columns (3) and (4), where the dependent variable is the number of projects won, we only find negative and statistically significant estimates in the freelancer fixed effects specification.

The estimates reported in Table 3 suggest that the return to signalling is smaller for more experienced freelancers. The negative interaction terms indicate that the main estimates reported in Table 2 hide considerable heterogeneity in returns. Comparing the estimates between Tables 2 and 3 show that the returns to completing skill certificates are up to 1.5 times larger for workers with zero work history compared to average. Figure 3 illustrates the relationship between completed projects and marginal return to signalling estimated from the freelancer × test specifications.

Figure 3 illustrates the relationship between completed projects and marginal return to signalling estimated from the freelancer × test specifications. To summarise, the estimates reported in Table 5 suggest that return to signalling is smaller for more experienced freelancers. Overall, the results lend support to the hypothesis that signalling is a substitute for experience. Nonetheless, the substitution effect is found to be relatively small, even when statistically significant. For instance, the marginal effect of completing an additional skill certificate is found to be positive even for a median freelancer in the filtered sample (with six completed projects).

One explanation for this could be that employers face uncertainty over both workers' technical ability as well as other worker characteristics, including soft skills, such as communication and trustworthiness, or otherwise hard-to-verify skills such as opportunism. Skill certificates mostly decrease uncertainty over hard skills, while work experience and detailed feedback on completed projects will also reduce uncertainty over other types of skills. If this is the case, skill certificates cannot fully substitute for work experience as a signalling device.

Despite the small magnitude, the results presented in this section suggest that skill certificates are substitutes to other sources of verified information on freelancers' ability. While not conclusive, this evidence is consistent with our theoretical model which implies that the reason why earnings are higher after certificate completion is decreased employer uncertainty on freelancer ability.

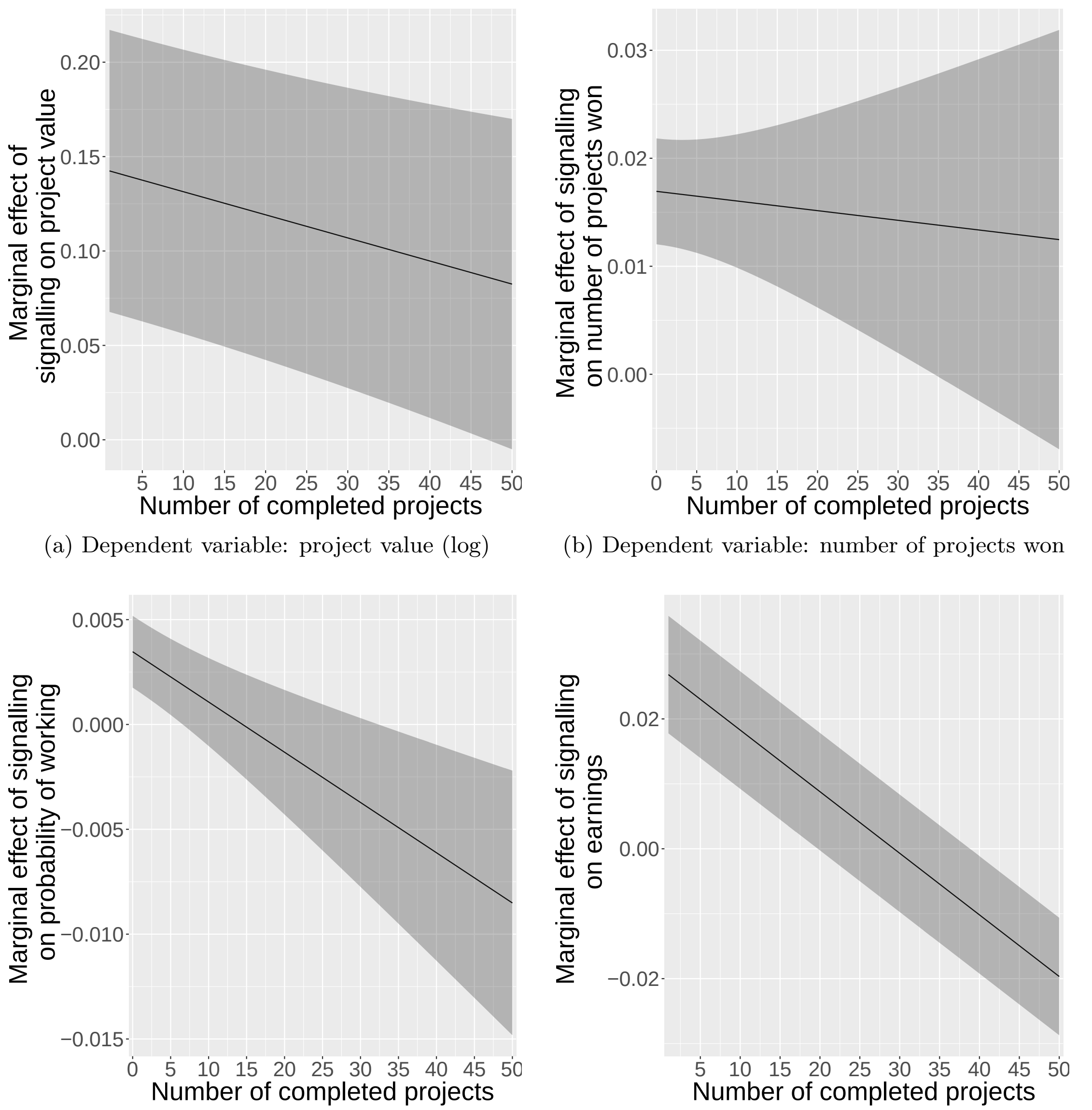
### 5.3 Decreasing returns to signalling

We now turn to study how returns to signalling vary with the level of signalling. As suggested by Prediction 3, one would expect the gains from signalling to be lower for higher levels of signalling.

Table 3: Returns to signalling by different levels of experience

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | log(project value) | | Num projects | | Num projects > 0 | | log(1+dollars) | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Num certificates | $0.144^{***}$ | $0.046^{***}$ | $0.017^{***}$ | $0.015^{***}$ | $0.003^{***}$ | $0.007^{***}$ | $0.028^{***}$ | $0.036^{***}$ |
| | (0.038) | (0.008) | (0.003) | (0.002) | (0.001) | (0.001) | (0.005) | (0.004) |
| Num certificates × Num projects / 100 | $-0.122^{**}$ | $-0.030^{**}$ | $-0.009$ | $-0.019^{+}$ | $-0.024^{***}$ | $-0.008^{***}$ | $-0.095^{**}$ | $-0.036^{***}$ |
| | (0.047) | (0.010) | (0.019) | (0.011) | (0.006) | (0.002) | (0.033) | (0.008) |
| Fixed effects | FL × test | FL | FL × test | FL | FL × test | FL | FL × test | FL |
| 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 | 0.249 | 0.273 | 0.214 | 0.261 | 0.209 |

In addition to the variables reported, all models include year-dummies, average rating for completed projects, cumulative dollars earned on the platform (log-transformed) and cumulative number of completed projects on the platform measured in the time of project start. Standard errors are clustered on freelancer level. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% and + 10%.




(a) Dependent variable: project value (log)

(b) Dependent variable: number of projects won

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

(d) Dependent variable: log(1+earnings)

Figure 3: Marginal effect of signalling for different levels of experience. Notes: the estimates are from regression models that control for freelancer fixed effects. The gray band corresponds to a 95% confidence interval calculated as $+/-1.96 \times s.e.$

We implement the test for decreasing returns to signalling in the form of the regression model

$$y_{ik} = \alpha_i + X_i\beta + \gamma_1 s + \gamma_2 s^2 + \nu_t + \varepsilon_{ik}. \tag{14}$$

Introducing the quadratic term for signalling into the regression allows us to test for the possible nonlinearity in return to signalling. As above, while our preferred specifications are the ones controlling for freelancer fixed effects, we estimate separate models using freelancer and freelancer $\times$test fixed effects.

Table 4 reports the estimation results. As evidenced by the consistently negative estimates for $\gamma_2$, the returns to signalling are found to be decreasing in $s$. The monetary returns for signalling are fairly high (up to 16%) for the first few skill certificates completed, but quickly go down after that. Comparison between Tables 4 and 2 demonstrates that the average returns reported in Table 2 conceal considerable heterogeneity.. The effect of the first completed skill certificate is up to 1.5 times as high as the average effect. The evidence on the decreasing returns to signalling 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 tiny effect sizes of signalling altogether. The results are nevertheless broadly consistent with Prediction 3.

Table 4: Nonlinear return to signalling

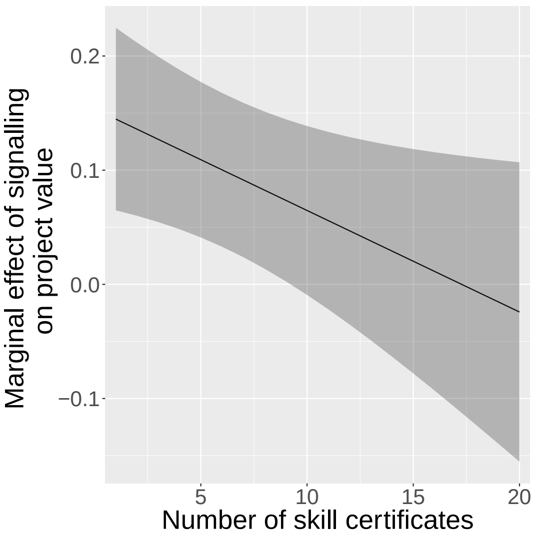
| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | log(project value) | | Num projects | | Num projects $>$ 0 | | log(1+dollars) | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Num certificates | $0.154^{***}$ | $0.057^{***}$ | $0.023^{***}$ | $0.024^{***}$ | $0.004^{*}$ | $0.012^{***}$ | $0.033^{***}$ | $0.059^{***}$ |
| | (0.043) | (0.011) | (0.002) | (0.003) | (0.001) | (0.001) | (0.007) | (0.006) |
| Num certificates$^2$ / 10 | $-0.044^{*}$ | $-0.010^{*}$ | $-0.004^{**}$ | $-0.004^{**}$ | $-0.001$ | $-0.002^{***}$ | $-0.007^{+}$ | $-0.010^{***}$ |
| | (0.021) | (0.004) | (0.001) | (0.001) | (0.001) | (0.001) | (0.004) | (0.003) |
| Fixed effects | FL $\times$ test | FL | FL $\times$ test | FL | FL $\times$ test | FL | FL $\times$ test | FL |
| 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 | 0.261 | 0.208 |

In addition to the variables reported, all models include year-dummies, average rating for completed projects, cumulative dollars earned on the platform (log-transformed) and cumulative number of completed projects on the platform measured in the time of project start. Standard errors are clustered on freelancer level. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% and + 10%.

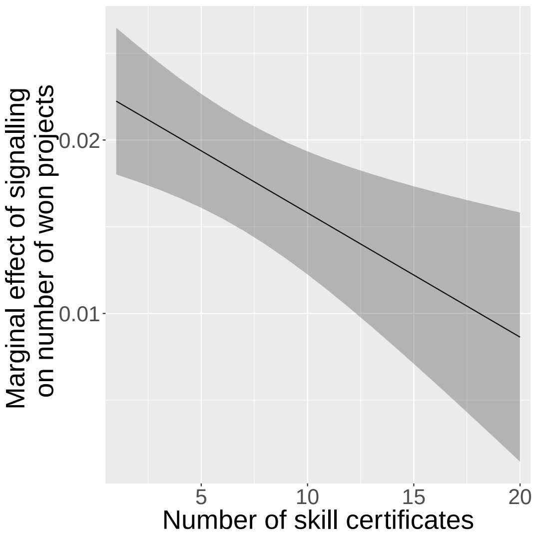
### 5.4 Does return to signalling vary by test score?

If completing skill tests reveals information about freelancers' ability to employers, we would expect the returns to signalling to be higher for higher test scores. The strategic choice of freelancers to publish their test scores might bias our estimates, but it is still instructive to examine how the return to signalling varies across the observed distribution of test scores.

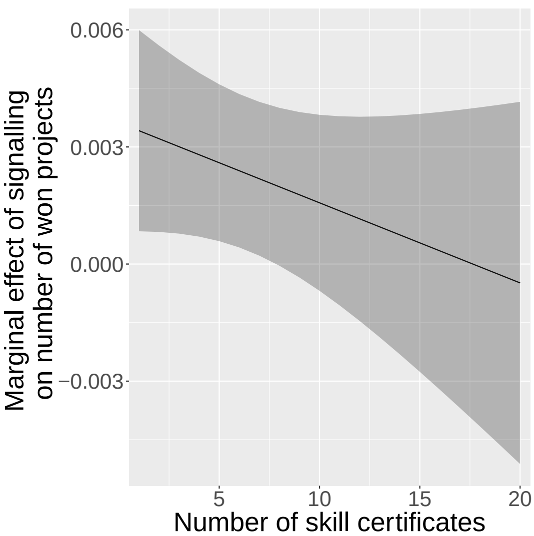
We operationalise the test for the effect of scores by the following regression specification:



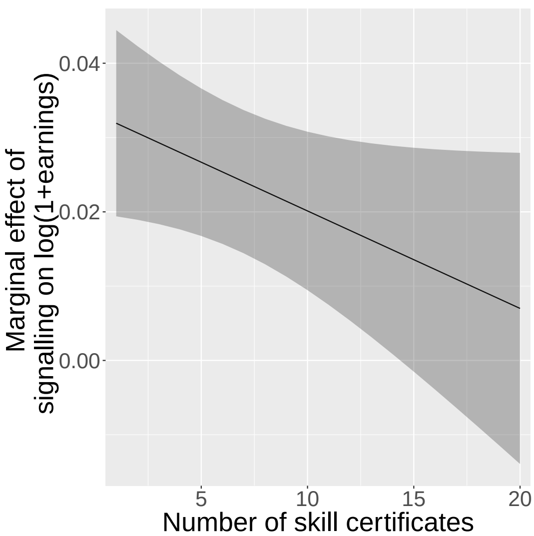

(a) Dependent variable: project value (log)




(b) Dependent variable: number of won projects




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




(d) Dependent variable: log(1+earnings)

Figure 4: Marginal effect of signalling for different levels of experience. Notes: the estimates are from regression models that control for freelancer fixed effects. The grey band corresponds to 95% confidence interval calculated as $+/-1.96 \times s.e.$

$$y_{ik} = \alpha_i + X_{ik}\beta + \gamma s_{ik} + \delta p_{ik} + \theta\left(s_{ik} \times p_{ik}\right) + \nu_t + \varepsilon_{ik},$$

where $p_{ik}$ is the average percentile of completed tests for each freelancer. To further ease interpretation, the average test score percentile of each freelancer is standardised by subtracting the average percentile among all observed test scores and dividing by standard deviation. The resulting coefficients have the following interpretation: $\gamma$ corresponds to the marginal effect of completing a skill certificate at the mean percentile of the observed test score distribution. $\theta$, on the other hand, is the estimate for the change in the return signalling for one standard deviation increase in observed test score.

We present the regression results in Table 5. These results are somewhat inconclusive. On the project value margin, the returns are found to be largely invariant to changes in test scores. On other margins, the effects are found to be positive, albeit small. The findings imply that only the highest scoring freelancers win more projects by signalling, and even for them, the effect is very small. A possible interpretation for the differences found between the project value and employment margins is that platform recommendation algorithms, described in Horton (2017), increase the probability of getting hired but not the the value of projects won.

According to the estimate in column (5), the effect of signalling on the probability of working is negative and statistically significant – albeit economically miniscule – for a freelancer revealing their test score. When interpreting the estimates in Table 5, it is notable that the distribution of test scores is censored by freelancers' choices of not revealing them.

Table 5: Returns to signalling by different test scores

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | log(project value) | | Num projects | | Num projects $> 0$ | | log(1+dollars) | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Num certificates | $0.100^{**}$ | $0.028^{***}$ | $0.009^{***}$ | $0.011^{***}$ | $-0.004^{**}$ | $0.006^{***}$ | $-0.007$ | $0.028^{***}$ |
| | (0.037) | (0.006) | (0.002) | (0.002) | (0.002) | (0.001) | (0.007) | (0.004) |
| Num certificates × Test score | $-0.049$ | $-0.012^{+}$ | $0.004^{+}$ | 0.001 | $0.002^{*}$ | $-0.0002$ | $0.013^{*}$ | 0.004 |
| | (0.034) | (0.006) | (0.002) | (0.002) | (0.001) | (0.001) | (0.006) | (0.003) |
| Fixed effects | FL × test | FL | FL × test | FL | FL × test | FL | FL × test | FL |
| Observations | 29,758 | 29,758 | 149,117 | 149,117 | 149,117 | 149,117 | 149,117 | 149,117 |
| Adjusted $R^2$ | 0.456 | 0.362 | 0.322 | 0.250 | 0.261 | 0.221 | 0.246 | 0.216 |

In addition to the variables reported, all models include year-dummies, average rating for completed projects, cumulative dollars earned on the platform (log-transformed) and cumulative number of completed projects on the platform measured in the time of project start. Standard errors are clustered on freelancer level. Test scores are standardised by subtracting the population mean and dividing by population standard deviation. Only observations with with a non-missing test score are included. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% and + 10%.

### 5.5 Does skill certification increase productivity?

Thus far the discussion has hinged on the assumption that signalling does not increase freelancer productivity. Is this assumption justified? The online labour platform offers a particularly intuitive measure for freelancer productivity in the form of feedback ratings given by the employers to the workers. We use this measure to operationalise freelancer success in a project. Table 6 presents the results of a regression analogous to regression Model (10), but with the feedback rating on the left-hand side of the regression. Columns (1) and (2) present the results from regression models where the feedback rating ranging between 1 and 5 is the dependent variable. To account for the upward skewed distribution of ratings, Columns (3) and (4) present an alternative specification in which the dependent variable gets a value of 1 if the feedback rating given to a freelancer is above 4.5. In both cases, the effect of signalling on ratings is statistically indistinguishable from zero. We interpret these results as supporting the assumption that signalling does not increase freelancer productivity.

Table 6: Effect of signalling on ratings.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Feedback rating | | Feedback rating > 4.5 | |
| | (1) | (2) | (3) | (4) |
| Number of certificates | −0.0188 | −0.0027 | −0.0073 | −0.0016 |
| | (0.0112) | (0.0017) | (0.0079) | (0.0012) |
| Fixed effects | FL × test | FL | FL × test | FL |
| Observations | 26,193 | 26,193 | 26,193 | 26,193 |
| Adjusted $R^2$ | 0.2620 | 0.1522 | 0.2036 | 0.1072 |

Notes: In addition to the variables reported, all models include year-dummies, average rating for completed projects, cumulative dollars earned on the platform (log-transformed) and cumulative number of completed projects on the platform measured in the time of project start. Only projects with a non-missing rating are included in the regression models. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% . 10%.

## 6 Robustness analyses

### 6.1 Long- and short-term effects of completing skill certificates

By design, the results we have so far presented concentrate on the short-term effects of signalling. Nonetheless, potential longer-term effects might also be relevant. For instance, signalling can lead to higher earnings, which, in turn, might result in an increased probability of being hired, since freelancers' total earnings on the platform are visible to employers.

To study this, we examine how the return estimate changes when we extend the time window from +/- 14 days. Basing the estimation on time windows longer than +/- 14 days increases the possibility that time-varying unobservables such as ability are affecting the estimates. Nonetheless, estimation results from extended time windows can indicate potential longer-term effects of signalling. This exercise also acts as a robustness check, because it shows that a cherry-picked time window does not drive the results.

The effects of varying time window widths are plotted in Figure 5. In Figures 5b and 5c where the dependent variables are the number of projects won and the probability of working, respectively, the treatment effect estimate increases mechanically as the time window width increases. To account for this, in Figure 5b and 5c the parameter estimate is divided by the mean of the dependent variable. Comparing the parameter estimate at 14 days to parameter estimates at longer and shorter time windows shows that the estimates are fairly close to one another at different time window lengths. This suggests that focusing on the +/-14-day time window does not conceal any longer-term effects. Consequently, we argue that the return estimates presented in Table 2 are also reasonable estimates for longer-term effects of completing skill certificates.

### 6.2 Donut estimates

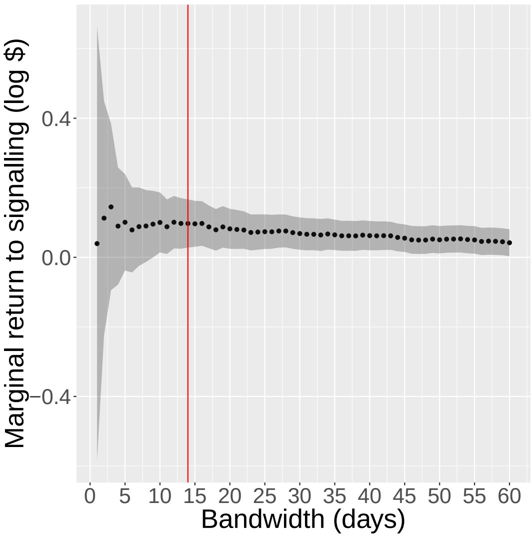
Another potential worry for the validity of our results is that our estimates are biased by varying effort around the time of skill test completion. In particular, a transitory dip in earnings just before skill test completion could bias our estimate for return to skill certificates upwards. A similar upward bias could emerge if workers strategically apply to more or better-paid jobs after the completion of the skill test.

To the extent that this increase in effort is transitory, these effects would be picked up in the time profile of the returns to signalling discussed in the previous section. Another way to assess the possible bias due to transitory changes in effort, recommended in Hausman and Rapson (2018), is to estimate a so-called "donut specification" where we drop the observations between -7 and 7 days of skill test completion. Comparing the donut estimates to the main specification can help us detect any short term effects.

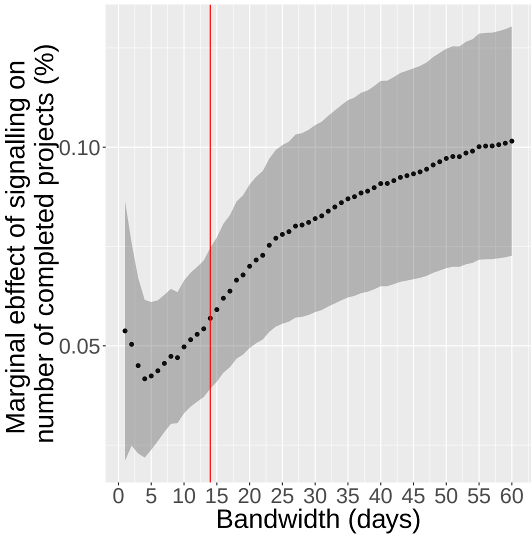
These estimates are reported in Table 7. According to p-values for the test for differences between donut and main estimates, the two are statistically indistinguishable from one another at conventional significance levels. Accordingly, we argue that there are no short-term changes in freelancer effort that would bias our estimates.[7]

Furthermore, it is worth noting that even the donut fixed-effect estimates reported in Table 7 are larger in absolute value than the OLS estimates reported in Table 2, which supports our initial claim that the OLS estimates are biased downwards because of a time-invariant permanent selection effect.

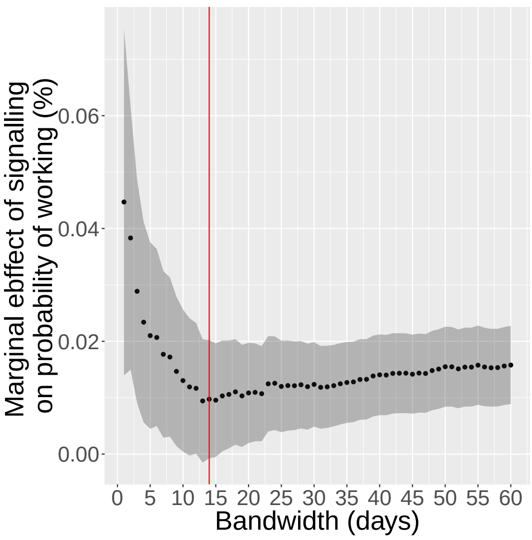
[7] As in the previous section, to transform the level estimates to marginal effects, we have divided the regression coefficient estimate by the mean of the dependent variable.



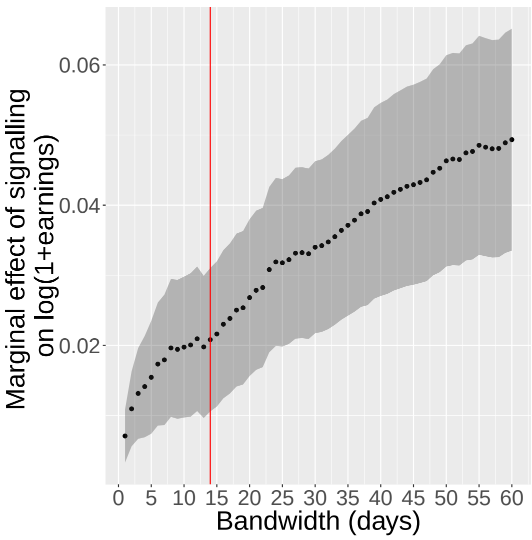

(a) Dependent variable: project value (log)




(b) Dependent variable: number of completed projects




(c) Dependent variable: Num projects $> 0$




(d) Dependent variable: log(1+earnings)

Figure 5: The sensitivity of results to varying the time window length. Note: in all graphs, the grey band corresponds to 95% confidence interval calculated by $+/-1.96 \times s.e.$

Table 7: Returns to signalling: donut estimates

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | log(project value) | | Num projects | | Num projects > 0 | | log(1+dollars) | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Marginal effect of signalling | 0.1** | 0.11 | 0.06*** | 0.04*** | 0.01+ | 0.01+ | 0.02*** | 0.02*** |
| | (0.04) | (0.07) | (0.009) | (0.009) | (0.005) | (0.008) | (0.01) | (0) |
| P-value for diff | 0.84 | | 0.13 | | 0.63 | | 0.94 | |
| Fixed effects | FL × test | FL × test | FL × test | FL × test | FL × test | FL × test | FL × test | FL × test |
| Time window | [-14,...,14] | Donut | [-14,...,14] | Donut | [-14,...,14] | Donut | [-14,...,14] | Donut |
| Observations | 32,975 | 16,434 | 178,478 | 178,478 | 178,478 | 178,478 | 178,478 | 178,478 |
| Adjusted $R^2$ | 0.463 | 0.457 | 0.363 | 0.239 | 0.273 | 0.198 | 0.261 | 0.177 |

Notes: In columns (1)-(2) unit of observation is one project. In columns (3)-(8), unit of observation is a 14 or 7 day pre- or post-test period. "Donut" refers to a model estimated using data from days [-14,.., -7], [7,..., 14] around completion of skill test. In columns (1), (2), (7) and (8) marginal effect corresponds to the point estimate from a linear regression model; in columns (3)-(6) marginal effect corresponds to the point estimate from a linear regression model. Standard errors are calculated using delta-method. P-value for diff is the p-value of a Z-test for for the difference betweeen the regression coefficients estimated from the standard sample and the donut sample. In addition to the variables reported, all models include year-dummies, and cumulative dollars earned on the platform (log-transformed). Standard errors are clustered on freelancer level. Significance levels in all specifications: *** 0.1%, ** 1%,* 5% and +%.

### 6.3 Balance tests for project types

Another way to show that there is no observable change in freelancers' effort at around the time of skill certificate completion is to study if there are some systematic differences between projects won just before the completion of the skill test compared to those won just after the test. We compare the projects won across various observable dimensions by re-estimating the specification in Equation 10 using various observable characteristics of the projects as the dependent variable.

First, we study if the freelancers are more likely to win a fixed-price project after completing a skill test compared to before. Under hourly contracts, employers are obligated to pay freelancers for their time regardless of the quality of their work. In contrast, under fixed-price contracts, freelancers are not monitored, but the employers can withhold their payment if they are not satisfied with the quality of the freelancers' output. If completing skill tests is associated with freelancers applying for more ambitious projects, then we would expect to see more freelancers winning fixed price projects, which are riskier but often pay more, after they have completed a skill test.

Second, we study if there is any difference in how competitive the projects won pre-test and post-test are. Competitiveness is understood as the number of freelancers bidding for the project. If completing skill tests is associated with greater freelancer effort, then freelancers might apply to more lucrative and therefore more competitive projects after completing a test.

Third, we study the preferred freelancer tier defined by the employer. When creating a project, the employer can define what type of freelancer they are looking for. This variable gets three values: "I am willing to hire an inexperienced freelancer for cheap", "I am looking for a balance between value and experience" and "I am willing to pay more for an experienced freelancer". If freelancers' effort increases after completing a skill test, we would expect them to win more projects targeted at higher tiers after completing a skill test.

Finally, we study the proposed contract length defined by the employer. These range from small project with under 10 hours to a full-time contract. If completing skill tests is associated with a higher freelancer effort, freelancers might apply and win longer-term contracts after completing a skill certificate.

The regression results are reported in Table 8. We find no effect from skill test completion on any of the project characteristics. We, therefore, conclude that increased earnings after skill test completion are driven by employers' preference for freelancers with validated skills, rather than by freelancers applying for different types of projects after skill test completion compared to before.

Table 8: Comparison of types of won projects pre-test vs post-test.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Hourly project | log(N of applicants) | Lowest freelancer tier | On-call contract |
| | (1) | (2) | (3) | (4) |
| Number of certificates | 0.003 | −0.006 | −0.011 | −0.006 |
| | (0.011) | (0.040) | (0.045) | (0.013) |
| Fixed effects | FL × test | FL × test | FL × test | FL × test |
| Observations | 32,975 | 23,380 | 9,004 | 24,805 |
| Adjusted $R^2$ | 0.324 | 0.355 | 0.211 | 0.142 |

Notes: In columns (1)-(3) unit of observation is one project. In columns (3)-(12), unit of observation is 14-day pre- or post-test period. In addition to the variables reported, all models include year-dummies, and cumulative dollars earned on the platform (log-transformed).Standard errors are clustered on freelancer level. Only observations with a non-missing value of the dependent variable are included in the regression models. Significance levels in all specifications: *** 0.1%, ** 1%, * 5% and + 10%.

### 6.4 Falsification tests

An additional validity concern in our empirical strategy is that despite limiting our attention to short time windows around the time of skill certificate award, an increase in the number of completed skill certificates might be correlated with other time-varying unobservable characteristics, which might affect the estimated returns to completing skill certificates. We show that this not the case in Table 9 by re-estimating the models in Table 2 while substracting 30 days from the date of the award of the skill certificate. The results of this exercise are either statistically indistinguishable from zero, or, in the case of models where the dependent variable is the probability of winning a project or earnings, negative but economically insignificant. The results from this falsification test further increase confidence in the main estimation results.

Table 9: Falsification tests for returns to signalling.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | (log) project value | Num projects | log(1+dollars) | log(1 + earnings) |
| | (1) | (2) | (3) | (4) |
| Num certificates | 0.029 | −0.001 | −0.005*** | −0.020*** |
| | (0.043) | (0.001) | (0.001) | (0.004) |
| Freelancer fixed effects | FL × test | FL × test | FL × test | FL × test |
| Observations | 16,678 | 178,478 | 178,478 | 178,478 |
| Adjusted $R^2$ | 0.482 | 0.439 | 0.369 | 0.334 |

Notes: This table presents falsification tests where the certification award date is moved 30 days into the past. In column (1) the unit of observation is one project. In columns (2)-(3)-(4), unit of observation is 14-day pre- or post-test period. In addition to the variables reported, all models include year-dummies, average rating for completed projects, cumulative dollars earned on the platform (log-transformed) and cumulative number of completed projects on the platform measured in the time of project start. Standard errors are clustered on freelancer level. Significance levels in all specifications: *** 0.1%, ** 1%, and * 5%.

# 7 Discussion

## 7.1 Why do skill certificates help?

The main result of this paper is that signalling in the form of taking computer-based tests that award digital skill certificates increases worker earnings in an online labour market. Since the skill tests are administered remotely, they are probably easier to cheat in than tests administered at a test site, which would make them less reliable indicators of worker ability. But the positive marginal effects of completing these skill certificates indicate that employers nevertheless trust that the skill certificates do convey some real information about worker ability. The positive effect of the skill certificates can in theory take place through two channels: increasing the number of projects won, and increasing the value of projects won. According to our estimates, the effect mostly operates on the project value margin; increases in the number of projects won are very small.

New workers entering online labour markets do not have validated information on their skills. Therefore, their prospective employers do not have much data to evaluate them on, and they rarely get hired. Skill certificates appear to do relatively little to overcome this hurdle. Instead, they appear to be more useful for workers who have already passed the initial hurdle of winning at least one project. These workers can increase the value of their subsequent projects by completing skill certificates. For new entrants who have yet to win a single project, skill certificates are less useful, as they increase the chances of winning a project only marginally.

We also found that the returns to signalling vary across levels of platform-verified information; in other words, signalling is a partial substitute for other forms of verified information that increases the value of projects won. Freelancers with only a few completed projects in their work history can earn a return of almost 15% for completing skill certificates, while freelancers with long work histories obtain no benefit from it.

Completing skill certificates is found to have decreasing returns. The first few certificates tend to bring returns of around 15%, but the returns decrease as freelancers accumulate more skill certificates. Consequently, more experienced freelancers get a higher return from working and applying for jobs than from signalling. Realising this, most freelancers tend to complete only a few skill certificates.

We quite confidently rule out the alternative explanation that unobservable increases in freelancer productivity would be driving the results. It is likely that the freelancers' skills would remain approximately constant over the short time periods we concentrate on. In addition, the effect of signalling on the feedback scores awarded after completed projects, and direct evaluations of freelancer performance is statistically indistinguishable from zero.

We also tested for and rejected an alternative hypothesis of varying freelancer effort. That is, we do not find that freelancers would systematically win more demanding projects after the skill test completion. Rather, our results indicate that after completing a skill test freelancers earn more from projects that appear similar to the projects they did before completing the test. Using data on hires only, we cannot fully rule out that freelancers would apply to more projects after the completion of a skill test. Nonetheless, if there was a temporary increase in the freelancer's application activity after completing a skill test, we would have picked that up either when comparing 14 day time windows to shorter or longer estimates, or when re-estimating our model using a donut estimate.

If freelancers' effort would permanently increase after skill certificate completion, our estimates for the return to signalling would be biased upwards. In this case, the estimates reported in this paper could be interpreted as an upper limit of the true effect. While this scenario is possible, it is inconsistent with the observations that signalling does not seem to have an effect on the types of projects won, and that the decision to signal is negatively correlated with freelancer-level unobservables.

Given these findings, we argue that the most plausible explanation for why completing skill certificates increases worker earnings is that it decreases employers' uncertainty over worker ability. Why is uncertainty costly? The theoretical model presented in this paper argues that employers prefer workers with validated skills, and therefore pay them more for the same jobs. Some employers on online labour platforms first hire workers into small "digital internship" projects before hiring them into larger projects Corporaal and Lehdonvirta (2017). The purpose of these test-piece projects is to screen workers. It might be the case that platform-administered skill certificates allow the employers to forgo some of this or other types of screening, and this reduced employer screening cost is compensated to the freelancers in the form of higher pay.

Our research design does not allow us to make direct inferences on the general equilibrium effects of skill certificate. We cannot rule out the possibility that skill tests simply cause employers to substitute non-certified freelancers with certified freelancers. Nonetheless, this seems fairly unlikely since the effects of signalling are mostly found on the project value and earnings margins, and not on the number of won projects margin.

### 7.2 Where skill certification helps and where it does not?

Why does signalling via digital skill certificates increase earnings more than it increases the number of projects won? A likely explanation is that employers have uncertainty on both freelancers' hard and soft skills. Skill certificates, in principle, only certify hard skills, while soft skills and general cooperativeness need to be signalled by other means, such as feedback from previous projects. This explanation is consistent with recent literature emphasising the important role of soft skills in the labour market (for example Deming (2017); Almlund et al. (2011); Heckman and Kautz (2012)).

It is also instructive to contrast these estimates with other effect sizes presented in the literature. In particular, when studying a "digital apprenticeship" model, Stanton and Thomas (2015) find that becoming affiliated with an intermediary leads to a roughly 10% increase in project value. In contrast to our results, they also find that becoming affiliated with an intermediary also leads to substantially higher job-finding probabilities. Our estimates for increases in earnings for new workers roughly coincide with Stanton & Thomas's estimates, but their estimates for job finding probabilities are considerably larger. This suggests that online intermediaries and skill certificates signal different dimensions of freelancer quality. In particular, it seems likely that skill certificates provide reasonably reliable information on freelancers' hard skills, whereas intermediaries can provide information on soft skills and cooperativeness in addition to hard skills. More generally, given the high returns to experience for newcomers shown in Pallais (2014), it is likely that a "digital internship" model giving employers the possibility to "sample" several freelancers for a low cost to learn their skills would be more efficient in reducing employer uncertainty than freelancer signalling by skill certificates.

Recommendation algorithms described in Horton (2017) are more likely to recommend freelancers with completed skill certificates to employers. Nonetheless, the observation that the effect of signalling on freelancer success is more pronounced in the project value margin is inconsistent with the idea that an algorithmic recommendation system, which mostly affects the probability of winning a project Horton (2017), is driving the results. Additionally, the effects are constant between years. When interacting year-dummies with the return estimate, the interaction terms remain insignificant across specifications. Therefore, the positive effects of signalling were already present before the algorithmic recommendation systems described by Horton were rolled out.

## 8 Conclusions

Newcomers face considerable difficulties when entering online labour markets. Before a new worker is hired and screening information about them consequently published, there is considerable employer uncertainty about their quality, which prevents them from getting hired in the first place. Skill certificates are an online labour market institution that was designed to allow new workers to break out from this vicious cycle, by making it possible for them to demonstrate their skills to prospective employers at their own expense. However, the findings presented in this paper show that skill certificates, at least as they are implemented on the platform under study, are not very useful for this purpose. They have a statistically and practically significant positive impact on freelancer earnings conditional on winning a project, but their impact on the likelihood of winning a project is limited.

At the same time, experienced workers who have already accumulated a significant work history on the platform do not benefit from skill certificates. This is because the platform-verified work history and employer feedback scores are a substitute for skill certificates in reducing employer uncertainty. Combined, these effects leave a fairly narrow range of workers who are likely to benefit significantly from obtaining skill certificates: early-career freelancers, who have won their first few projects and broken into the market, but who still lack a more extensive work history.

This result has two clear implications to platform operators. The tests should be more challenging to improve the informativeness of skill certificate signals. This would allow high-ability newcomers to separate themselves from low-ability ones, improving matches and making the market more efficient. In its current form, the skill certification scheme can substitute for other types of verified information only to a limited extent.

To the extent that evidence from an online labour market setting can be generalised to conventional labour markets, the findings from this paper suggest that private digital skill certification schemes can decrease information asymmetries. This suggests that they could indeed help improve labour market matches in situations where public qualification schemes are too slow to keep up with rapidly changing skill demands Painter and Bamfield (2015). Such "badges" or "micro-credentials" could moreover help skilled members of statistically discriminated against groups such as immigrants and other minorities to improve their position in labour markets. However, in this study, the average returns to skill certificates remained fairly small even in a fairly low-friction online labour market environment, where the effort cost of taking the tests was low. This underscores the proposition that to facilitate separation between more and less skilled workers, the skill tests should be made "cheap" to take but difficult to do well in. This would facilitate the informativeness of skill certificates and would presumably increase labour market efficiency by reducing skill uncertainty. However, due to the imperfect substitutability between skill certificates and work experience, skill certification schemes are still likely to be only a partial solution, with other institutions also continuing to be needed.

## References


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.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011). Personality Psychology and Economics. *Handbook of the Economics of Education 4*, 1–181.

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*, 1409–1448.

Barach, M. and J. J. Horton (2017). How do employers use compensation history?: Evidence from a field experiment.

Blackburn, M. L. and D. Neumark (1993). Omitted-ability bias and the increase in the return to schooling. *Journal of Labor Economics 11*(3), 521–544.

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. Technical report, Oxford Internet Institute: Oxford.

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), 323–358.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*.

Filippas, A., J. J. Horton, and J. M. Golden (2017). Reputation in the Long-Run. Technical report, CESifo.

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. *The American Economic Review 85*(5), 1278–1286.

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. and T. Kautz (2012). Hard evidence on soft skills. *Labour Economics 19*(4), 451–464.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Hoffman, M., L. B. Kahn, and D. Li (2018). Discretion in hiring. *The 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).

Lang, K. and M. Manove (2011). Education and Labor Market Discrimination. *The American Economic Review 101*(4), 1467–1496.

Lehdonvirta, V., O. Kässi, I. Hjorth, H. Barnard, and M. Graham (2018). The global platform economy: A new offshoring institution enabling emerging-economy microproviders. *Journal of Management 0*(0), 0149206318786781.

Lin, M., Y. Liu, and S. Viswanathan (2016). Effectiveness of Reputation in Contracting for Customized Production: Evidence from Online Labor Markets. *Management Science* (January 2017), mnsc.2016.2594.

Malone, T. W. and R. J. Laubacher (1998). The dawn of the e-lance economy. *Harvard business review 76*(5).

Nosko, C. and S. Tadelis (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Working Paper 20830, National Bureau of Economic Research.

Painter, A. and L. Bamfield (2015). The new digital learning age: how we can enable social mobility through technology. Technical report, The Royal Society of Arts, London.

Pallais, A. (2014). Inefficient Hiring in Entry-Level Labor Markets. *American Economic Review 104*(11), 3565–3599.

Pinkston, J. C. (2003). Screening discrimination and the determinants of wages. *Labour Economics 10*(6), 643–658.

Silverman, B. (1986). *Density estimation for statistics and data analysis.* London, United Kingdom: Chapman and Hall.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Stanton, C. T. and C. Thomas (2015, sep). Landing the First Job: The Value of Intermediaries in Online Hiring. *The Review of Economic Studies 83*(2), 810–854.

Tyler, J. H., R. J. Murnane, and B. Willett, John (2000). Estimating the Signaling Value of the GDE. *Quarterly Journal of Economics 115*(2), 431–468.

Wood, A. J., M. Graham, V. Lehdonvirta, and I. Hjorth (2018). Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy. *Work, Employment and Society*, 095001701878561.

# Appendix A1 Variable definitions for variables used in regression models

Table 10: Definition of variables used in regression models

| | |
|---|---|
| Project characteristics | |
| Variable name | Description |
| Hourly project | A dummy variable getting a value 1 If the project is billed on a hourly basis and 0 if it is billed with a flat price. |
| Project value | Dollars paid to the freelancer after succesful completion of project |
| Hourly rate | Hourly rate of a freelancer hired in a project (only hourly projects) |
| Star rating given to worker | Rating given to the freelancer by the employer after project completion |
| Competitiveness of project | Number of applicants to a project |
| Freelancer tier | Desired freelancer tier set by the employer for each project. Ranges from 1 (looking for cheap freelancers) to 3 (willing to hire an expert for a higher cost). |
| Expected hours | Expected hours needed for project set by the employer. Ranges from under 10 hours to full-time. |
| Freelancer characteristics | |
| Variable name | Description |
| Months active | Difference in full months between the start of the first freelancer project and date of data collection |
| Number of completed projects | Number of completed projects at the time of project start. |
| Dollars earned | Dollars earned at the time of project start ($log(1+)$-transformed) |
| Feedback rating | Average star rating of past projects. The past projects are weighed using the same algorithm that is being used on the online labour platform. |
| Test score | Average test score percentile of completed skill tests. Standardised by subtracting the population mean and dividing by population standard deviation. |