Conditional Payments Incentivize Online Freelancers to Learn Skills

Gobi Dasu

ABSTRACT

Online freelancers struggle to gain the skills necessary to increase their earning capacity. This study investigates incentivizing freelancers on Upwork to learn new skills with conditional cash transfer (CCT), an incentive scheme traditionally used to promote school attendance and stop child labor. In our case, workers train to become *backend developers* or *marketers* and are paid upon completion of individual training program milestones. To test the program, we perform a month-long field experiment (N=2100), randomizing assignment of compensation level. The ratio between the fraction of invited workers completing any given milestone in the lowest pay condition (forgoing 90% of income) and at earning capacity is on average 1.0 for *marketer* milestones and 0.36 for *backend developer* milestones. This suggests educational CCT incentivizes online workers and bolsters their motivation to learn skills needed for better future opportunities. Moreover, we find *compensation amount* predicts training program *percent completion* for *backend developer* trainees.

ACM Classification Keywords

Crowdsourcing

Author Keywords

online freelancing; conditional cash transfer; skill development

INTRODUCTION

Online freelancing has potential to fill human skill deficits by tapping into a geographically distributed workforce [22,42,49]. The industry can improve social mobility for the world's growing internet-connected population and help developing countries leapfrog by giving their digital workers access to clients in wealthy countries [22,36,52]. While the feasibility of this vision is worth investigating, currently online labor platforms lack industry-wide adoption, quality is poor, and worker skills remain static [28,66]. At the same time, high quality, free online courses suffer from severe dropout [30,47,55,61]. UC Berkeley and Delft University scholars have studied the use of high-skill work opportunities as an incentive for online course completion and skill development. They have found that the skills learned in MOOCs (massive online open courses) can be applied to freelance project completion. Moreover, they suggest workers, particularly those in developing countries, may be motivated to learn skills online when the education is presented as training for work opportunities [22]. Our research introduces educational conditional cash transfer (CCT) to incentivize online workers to persist as they learn individual skills needed for a job.

Given the current state of online learning and earning, we wonder whether it may be worth it for society to actually invest in capable freelancers, paying them to learn individual skills as they train for a job [22]. We seek to answer the following questions: (*Q1*) Will online workers forgo fractions of their immediate income to learn skills, and if so to what extent? (*Q2*) What is the relationship between *compensation amount* and training program *percent completion?* We invoke economics literature of conditional cash transfer. Conditional cash transfer literature tells us that in the domains of education and work, conditional financial incentives more often than not have significant positive impact on behavior change [3,12,13,16,38,64]. Educational CCT can be particularly effective on positive behavior change in traditionally disenfranchised groups (the global poor, minorities, females in many countries), and there is a strong argument that cash incentives do not diminish intrinsic motivation [19–21,43,59,62].

In traditional CCT, payments are made for school attendance or performance, behavior that has long-term benefits [7,13,40]. Our CCT pays online freelancers for completing learning milestones in a contract that converts to an hourly-rate job immediately upon training completion. The benefits are more explicit. So, we may expect CCT to perform at least as well in our case. However, freelancers specifically come to Upwork to earn as much money as they can [66]. While CCT may *convince* families or students to attend school, the attendees may have reasons other than cash to attend – such as learning, social activity, etc. Hence, online freelancing, where CCT can actually scale and sustain [22], presents unexplored challenges and advantages for CCT. Questions of program efficacy, cost, and success are not fully answerable without experimentation.

In this paper we report results from an experiment in which online freelancers are paid to learn skills as a form of training for a promised hourly job. We train *backend developers* and *marketers* across 4 randomly assigned payment levels, the highest of which corresponds to trainees' estimated earning capacity. HCI literature and Economics theory suggest that conditional financial incentives should have an impact on the behavior of the trainees, and specifically that higher payment levels should command greater training program *percent completion*. We proceed to:

- review theoretical and experimental work on online freelancing, online learning, and conditional cash transfer to develop research questions and hypotheses
- detail our experiment to study CCT as an incentive mechanism for online freelancers to learn skills
- present the ratio between the fraction of invited workers completing any given milestone in the lowest pay condition (forgoing 90%) and at earning capacity: 1.0 for *marketers* and 0.36 for *backend developers*, which shows that freelancers will accept pay lower than their estimated earning capacity ($11.43/hr) to learn useful skills to train for a promised high-paying job ($14/hr)
- present that *compensation amount* is a significant predictor of program *percent completion* ($p < 0.01$) in the *backend developer* program but not the *marketer* program, which suggests that the compensation level becomes important when learning skills necessary for a job is a particularly time consuming endeavor

An implication of this study is that conditional cash incentives may enable sustainable and scalable skill development via online freelance markets. However, programs must be sure to fairly compensate workers and uphold promises of future payment.

RELATED WORK

The State of Online Work Platforms

The key barrier to industry adoption of online labor is poor worker quality, so investment in the education and training of teachable freelancers is an important consideration [22,28,66]. Currently, there is a mismatch between the needs of clients on Upwork and the skills that freelancers have [66], necessitating worker skill development. Yet, the skills of online workers often remain static because learning skills consumes time freelancers allocate for earning money [66]. As such, well-trained high-quality freelancers and high-paying clients are scarce and so neither side is trusting or generous [28,34]. In traditional organizations middle management ensure the quality of work produced by junior employees, but lack of such oversight in decentralized online work environments makes quality assurance difficult [65,68]. Independent freelancers' poor communication skills and insufficient expertise lead to lack of trust and micromanagement — a waste of time and money on both ends [28,54]. Online freelancing platforms, in their current state, require autonomous, quality-conscious experts who have proven themselves. We now look to online education to explore the notion of training such workers.

Online Learning meets Online Freelancing

Online courses and tutorials (like those found on edX, Coursera, and Codecademy) let millions of internet users around the world learn from top professors or field experts. They have the potential to change education because of the scalable instruction they facilitate [61]. They can feature progress-tracking [57], gamification [24], peer to peer feedback and discussion [45], and scalable grading [44]. However, these online courses are plagued by notoriously high dropout rates because a university degree is not on the line for the course takers [30,47,55]. Research suggests that incentives may be required to motivate online course takers to complete what they have started. To motivate learners, UC Berkeley and Delft University researchers developed a recommendation engine that connected MOOC takers with relevant freelance jobs. They inserted bonus exercises into MOOC courses which were paid tasks on Upwork related to the topic being learned. Excel work was popular. The results from the study were that the knowledge learned in MOOCs could be applied to Upwork project completion and workers, particularly those in developing countries, may be motivated to learn skills specifically for a paid job or task [22].

Unfortunately, many real world roles require multiple skills and many hours of training. For instance, we found that it takes freelancers on average 74 hours to train as a *backend developer* (9 milestones/skills). One cannot expect a freelancer to forgo income entirely for 70+ hours to learn skills in hopes of a higher paying future job. This problem parallels the one governments face when trying to convince students (and their families) to stay in school and forgo outside income, in the case of minors via child labor. Governments combat this issue with conditional cash transfer, paying students or their families frequent deposits of money under the condition that they attend or perform well in school. *Our* CCT helps freelancers to persist through training by paying them for every milestone they complete (i.e. for every skill they master) in a program that leads to hourly expert work. We predict based on literature that such payment will incentivize freelancers to complete milestones and learn skills needed for better jobs.

Conditional Cash Transfers in Education

Conditional cash transfers are welfare payments made *conditional* on the actions of the receiver. Upholding the *conditional* nature, in that learners are only paid if they meet an objective, has been found to be crucial [18].

Cash Incentives for Education in OECD Countries

Paying students for better performance has been found effective in a number of case studies. Many US schools have been encouraged to provide financial incentives for performance, and guidelines on how to do so have been written [3]. In Texas, incentivized underprivileged high schoolers earned higher AP and SAT scores and there were no undesirable side effects measured. The effect is attributable to a change in social norms, specifically an attitude shift toward viewing exam success as an important and relevant goal [38]. A study on elementary school students in Ohio found financial incentives lead to increased math scores [16], while a study in Maryland, Georgia, and New York found that *paying for A's* leads to improved reading scores. Moreover, the latter study shows that the frequency of payments is key in the success of conditional cash transfer because students are continuously validated [58].

In the US and abroad, financial incentives have shown particular promise for women, minority ethnic groups, and low-income students [6,39]. College freshman women, more so than men, have benefitted from cash transfer conditional on their educational performance [7]. Similarly in Israel, paying high school students for performance on college entrance exams was effective, particularly for female students [43]. Scholars suggest this is because women have more foresight when making decisions with long-term consequences, have better study habits, and are more prone to take advantage of cash transfer programs [7]. A joint study between Princeton University and the Federal Bank of Chicago found that low-income community college students (who are also parents) took 40% more credits when financially incentivized because cash payments meant they had to work less to support their families [11]. Northwestern University researcher Kirabo Jackson found that payments improved college GPAs significantly, particularly for black and Hispanic students, and moreover found evidence of enduring benefits even after removal of the extrinsic motivator [39]. These effects may be because cash payments have greater value to students belonging to socioeconomically disadvantaged minorities or traditionally disenfranchised groups.

Roland G Fryer Jr. is literature's main skeptic of financially incentivized education and presents cases where payments have had no effect or even a negative effect on students' performance by undermining their intrinsic motivation to learn [3,32]. One of his studies has cautioned that extrinsic motivation can have a negative impact on subjects, who through exertion realize that their abilities are lower than they previously believed [33]. Along these lines, a financial incentive field experiment done on students at the University of Amsterdam found that payment helped capable students, but dampened the intrinsic motivation of low-ability students [46].

Despite the drawbacks of CCT, in the education research community, financial incentives are viewed as a better alternative to punitive measures [56]. With this, we proceed to survey cash transfers in non-OECD countries.

Cash Incentives for Education in non-OECD Countries

Conditional payments for education have had positive results in Mexico [12,13,64], Ecuador [8], Colombia [4,5], the Punjab state in Pakistan [21], Kenya [43], Bangladesh [9], Costa Rica [29], and Nepal [63]. Some of the field experiments report improved school attendance and others improved school performance, but all conclude that individuals and families do respond to financial incentives. The numbers of participants in the Mexico, Colombia, and Punjab studies are in the tens of thousands. The projects in Punjab and Kenya focused on female students, and both report substantial test score improvement [21,43]. Deeper analysis of the Mexico study shows that conditional cash transfer programs are not "silver bullets" and certain techniques such as calibrating transfer frequency for maximum response bolster the program impact [13,18,40]. A study in Vietnam highlights the link between household income and school performance, which explains some of the positive feedback loops that the cash transfer experiments experience [14]. The perceived versus measured benefits of education are an important note to touch upon. Education of the workforce actually has a large impact on per capita economic output of workers [25]. A study within the Dominican Republic confirms this, but also highlights that public perception of benefits does not match up to measured monetary benefits [41]. Indeed citizens may not realize the income growth education will bring them. CCT can therefore seed society with education, the benefit of which takes a generation to realize.

Extrinsic vs. Intrinsic Motivation

Many social scientists argue that *extrinsic motivation* (i.e. paying someone to learn) undermines latent *intrinsic motivation* (i.e. natural curiosity) [2,15,31]. Certain experiments have found that only punitive measures decrease intrinsic motivation, not extrinsic incentives [26]. Other studies and theoretical works have found that task motivation must take into account the personal goals of the trainee being motivated, which our case is freelancers who want to build an online career [48]. Meta-analyses of 100+ experimental studies have challenged the notion that extrinsic motivation undermines intrinsic motivation. In fact we find that extrinsic motivation in some cases is additive and actually supports intrinsic motivation [19,20]. Extrinsic motivation (i.e. paying people to learn) can also help naïve actors who procrastinate immediate cost activities (i.e. learning on one's own time without getting paid for it) and do too soon immediate reward activities (i.e. doing lower paying work) [53]. In fact, in a global context, such behavioral patterns may feed poverty cycles which keep the uneducated poor and the poor uneducated [41].

Intra-enterprise Learning and Scalability

In the workplace, skill-based pay systems have also been effective in encouraging workers to develop work-related skills [27]. In fact, there are some brick and mortar software engineering boot camps that pay students to learn to code [17,23]. Incentivizing participation in organizational learning and knowledge transfer systems has also been effective in industry [1], allowing newcomers and different groups to acclimate. Incentivizing worker performance can also work along the same lines as incentivizing worker skill development. For instance a study on construction firms in the UK and Ireland shows that incentivizing workers to learn the needs of clients leads to higher project success rates [51].

Inspired by the background literature, we make two hypotheses: (*H1*) Online freelancers will forgo immediate income to learn skills. (*H2*) The higher the *compensation amount* is for learning skills (i.e. the closer it is up to their earning capacity), the greater the training program *percent completion* will be for the payment group.

STUDY DESIGN

We performed a month-long field experiment to investigate whether and to what extent online workers will forgo income to learn skills, and to determine the relationship between *compensation amount* and training program *percent completion*. We describe the freelancer recruitment process, the *backend developer* and *marketer* training programs, the various payment levels, our hypotheses, measurements, and analysis methods.

Participants

To test whether online workers would partially forgo income to learn in-demand skills, we invited online freelancers (N=2100) on upwork.com whose profiles had keywords relevant to our *backend developer* and *marketer* jobs. Keywords for the *backend developer* role included terms like *developer* or *python* and keywords for the *marketer* role included terms like *social media* or *viral marketing*. We recruited a subset of those invited that completed our background check and accepted our offer (N=95, 20% female, aged 18-65). The inviting algorithm preferred freelancers who reported being fluent in English and who have billed hours on Upwork before. Online freelancers are not used to payment for learning skills, so descriptions were made clear – freelancers were informed that they are being paid fixed prices to learn skills and after they finish training they would be awarded a $14 per hour job. Specifically, each fixed price payment would be released to a freelancer when they displayed mastery of the skill in online exams.

Method and Conditions

After the freelancers were invited to *backend developer* or *marketer* jobs on Upwork, they were asked to complete a background check before being given an offer to confirm their commitment to the program. They were assigned randomly to one of four payment levels, the highest of which corresponded to the trainees' estimated earning capacity. The background check surveyed freelancers' logic abilities as well as their creativity and divergent thinking abilities. Questions for the latter check included coming up with as many uses of a paperclip as possible and solving riddles.

We had an intent-to-treat setup. As long as freelancers completed the required questions in the background check and submitted it, they were *approved* and made an offer for the specific job they were invited to (*backend developer* or *marketer* at one of four payment levels: *low*, *medium*, *high*, *capacity*). The fixed price milestones that form each role's training program taught freelancers the various skills necessary for the job. The milestones for *backend*

Table 1. Conditional Payments for Completion of *backend developer* Training Program Milestones at each Payment Level

Payment Level	Compensation Amount
low	$1.20/hr
medium	$2.23/hr
high	$5.30/hr
capacity	$11.43/hr

developers covered English communication, mathematics, computer programming, and web frameworks; while the milestones for *marketers* covered English communication, mathematics, social media marketing, and Salesforce. The hourly compensations for the four payment levels are presented in Table 1.

The four payment levels were set based on a pilot survey of 10 *marketers* and 22 *backend developers* (not members of the 2100 invited to the experiment), who reported an average earning capacity of $11.43 per hour (stdev=$10.22). The two roles were not considered separately because we found that the difference in salaries between the two online roles were not significant compared to the pilot data variance. The payment level *low* corresponds to 1 standard deviation (stdev) below the pilot data's reported mean earning capacity ($1.20 per hour), *medium* corresponds to 0.9 stdev below the mean ($2.23 per hour), *high* corresponds to 0.6 stdev below the mean ($5.30 per hour), and *capacity* corresponds to the mean ($11.43 per hour). In the pilot, we also surveyed the freelancers on how long it took them to complete each milestone in the training program and multiplied the average figures by the four hourly payment rates to obtain the necessary compensation for each milestone. Given this process since milestone prices were fixed, actual amounts of time and true hourly wages may have varied slightly between individual freelancers.

Figure 2. Random Payment Condition Assignment Probabilities

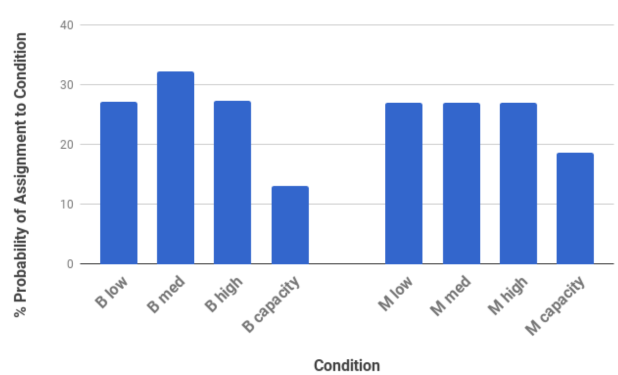


Figure 2 presents the random assignment probabilities across the 4 different conditions for each role. While a given *marketer* or *backend developer* invitee may have had different probabilities of ending up in one of the four payment conditions, amongst *marketers* and amongst *backend developers*, the probability mass functions used to place freelancers in a payment level were the same (i.e. each *marketer* invitee had the same probability of ending up in the *low* payment condition).

All exams were multiple choice, either hosted on our own learning management system (Moodle) or on Upwork itself. The Upwork API was used to verify a freelancer's true completion of Upwork exams. Each exam's description section linked to the online courses or tutorials which the exam covers. Freelancers could study these free courses and tutorials before attempting the exams. The passing threshold on Moodle was 80% (reported to freelancers as 85% to encourage them to shoot higher) but they could be immediately retaken. The passing threshold on the upwork.com exams was 60% but Upwork prevents freelancers from retaking exams for months. *Autorelease milestone* was the software that piped hired freelancers through the *backend developer* and *marketer* training programs, paying them whenever they passed the exams required by a milestone. Once a freelancer completed all the milestones, they were considered fully trained and their contract was converted from fixed price to a $14 per hour job. There was an exit survey for all trainees, regardless of their progress, at the end of the month-long experiment. The exit survey asked freelancers how long each milestone they completed took to do and whether they thought the program compensation was sufficient.

Measures

The dependent variable of interest is training program *percent completion*, which is the percentage of the training program that a freelancer has been successfully incentivized to complete. The binary variant of this variable has precedent as an appropriate choice of dependent variable for financial incentive experiments [37]. Independent variables the background check collected for each trainee were *# hours in work schedule*, *logic score*, *creativity score*, and *gender* [10,67]. We felt that the *logic score* based on IQ test and the *creativity score* based on a divergent thinking assessment are valid control variables because they are known to predict academic achievement and creative potential, both important factors for our training programs [50,60]. Independent variables collected for both trainees and invitees were *compensation amount* (the hourly rate of the assigned payment bracket), *Upwork hours worked*, *Upwork score*, *# Upwork skills*, *Upwork portfolio items count*, *Upwork rate*, and *number of other Upwork tests passed* (these are Upwork tests that we do not require in our training program). Training program *percent completion* (TPC) was determined based on which milestone a *backend developer* or *marketer* trainee was on at the end of the experiment as per Table 3.

Table 3. Training Program Percent Completion after each Completed Milestone

Backend Developer	TPC	Marketer	TPC
English	4.2%	*English*	9.1%
Grammar	8.5%	*Grammar*	18.2%
Arithmetic	12.7%	*Arithmetic*	27.3%
Algebra & Geometry	17.7%	*Algebra & Geometry*	38.0%
Intro CS, Command Line, Source Control, & Internet	38.4%	*Idioms*	47.1%
Idioms	42.7%	*Social Media Marketing*	56.2%
Python/SQL	69.2%	*Salesforce*	100%
HTML/CSS & Javascript	73.4%		
Django	100%		

Logic score was the auto graded result of a multiple choice assessment, and *creativity score* was also an auto graded result. The *creativity score* auto grader ensured that certain keywords were in freelancers' responses to the divergent thinking riddles. For example, the text of one riddle was: "You see all your friends when you're with me. But when you come to me, you look at your own face the most :)". The auto grader ensured the string *facebook* or *fb* or *social media* was in the participants' answers. Finally, *gender* was also surveyed.

Analysis Methods

To determine whether and to what extent online freelancers would forgo immediate income to learn skills, we compared the fraction of invited workers completing any given milestone in each of the different payment conditions. The ratio between the payment conditions would provide insight into how many freelancers would drop out when pay is reduced from estimated earning capacity by 54%, 80%, or even 90%.

To determine how payment scheme relates to training program *percent completion*, we first had to eliminate the sampling bias that is introduced because a specific 95 of the 2100 freelancers invited decided to complete our background check, accept our offer, and participate in our experiment. We employed the Heckman 2-step regression [35], which uses logistic regression to generate a derived variable predicting whether or not an invited freelancer would participate. This propensity was then set as a control variable in the second step of the analysis, a linear regression which used participant data (N=95) to predict training program *percent completion*. Variables available for both invitees and trainees were used in the logistic regression. The variables that were only available to trainees (not all invitees) as well as *compensation amount*, *Upwork portfolio items count*, *Upwork rate*, and *number of other Upwork tests passed* were used in the linear regression. We ensured variables used did not display high correlations, and we did a p-value analysis of the entire Heckman 2-step regression. Finally, we calculated summary statistics like the mean and median *training program progress* in each of the 4 conditions.

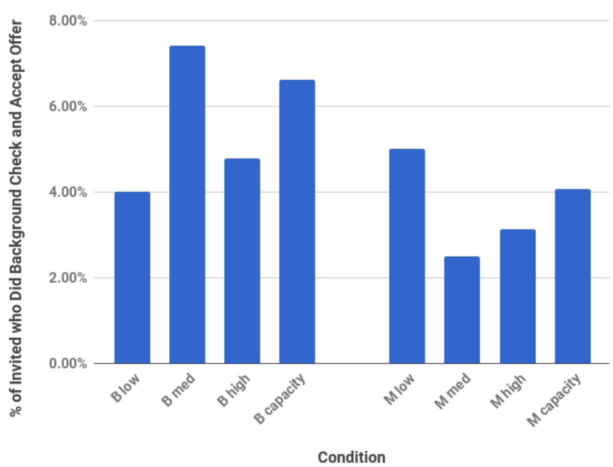
We also considered visual and subjective measures. We generated the shapes of the milestone progress histograms under the various payment conditions and compared subjective impressions of the various payment levels with experimental *percent completion* results.

RESULTS

A total of N=95 participants of the 2100 invited completed the background check survey and accepted our Upwork offer (i.e. participated) across the four pay conditions. The time between when the invites were sent and the data-collection cutoff was a month-long period. Figure 4 shows what percent of invitees in each condition participated (*B* = *backend developer*, *M* = *marketer*, i.e. *B med* indicates *backend developers* in the *med* [medium] pay condition).

Except for in the cases of *B med* and *M low*, it appears that the higher the payment condition, the more the freelancers participated. We therefore did a one-way analysis of variance (ANOVA) across the payment conditions to test for statistical significance between mean *participation rates*. We did one ANOVA for each role, *marketer* and *backend developer*. The differences between the payment conditions in Figure 4 are not statistically significant (one-way ANOVAs for each role yielded $p > 0.1$). Still, we used Heckman 2-step regression when predicting training program *percent completion* to mitigate any sampling bias.

Figure 4. Participation Rates: Percentages of those Invited in each Condition who Completed the Background Check and Accepted our Upwork Offer to enter the Training Program



52 participants studied to become *backend developers* and 43 studied to become *marketers*. The number of freelancers who completed the entire training program (all the milestones) was very low (1 *backend developer* and 6 *marketers*), but freelancers did learn various skills by completing a fraction of the milestones. The average training program *percent completion* for *backend developers* was 21.0% (median=12.7%). The average *training program progress* for *marketers* was 30.3% (median=27.3%).

Table 5. Average Training Program Percent Completion

Condition	Mean	Median
B (low pay)	16.1%	12.7%
B (medium pay)	15.9%	12.7%
B (high pay)	26.5%	17.7%
B (capacity pay)	33.0%	30.2
B overall	21.0%	12.7%
M (low pay)	28.2%	27.3%
M (medium pay)	29.8%	27.3%
M (high pay)	27.5%	18.2%
M (capacity pay)	36.7%	9.1%
M overall	30.3%	27.3%

Table 6. Ratios of the Fraction of Invited Backend Developers Completing each Milestone at each Payment Condition to the Fraction of Invited Backend Developers Completing each Milestone at the Earning Capacity Payment Condition (the *capacity* payment condition)

Milestone	Ratio of *high* to *capacity*	Ratio of *med* to *capacity*	Ratio of *low* to *capacity*
English	0.76	0.99	0.55
Grammar	0.64	0.88	0.65
Arithmetic	0.64	0.88	0.56
Algebra & Geometry	0.56	0.48	0.24
Intro CS, Command Line, Source Control, & Internet	0.48	0.31	0.12
Idioms	0.48	0.31	0.12
Python/SQL	0.72	0.20	0.24
HTML/CSS & Javascript	N/A	N/A	N/A
Django	N/A	N/A	N/A
Average across milestones	0.61	0.58	0.36

Table 5 presents the mean and median program *percent completion* for each condition. We see that higher pay commanded higher *percent completion* amongst *backend developers*.

Tables 6 and 7 compare the fraction of invited freelancers completing each milestone at the earning capacity payment condition (*capacity*) with every other payment condition. Specifically, the tables report ratios between the payment conditions. Ratios for the last two milestones of the *backend developer* role are reported as N/A because no freelancers completed those milestones in some of the conditions yielding denominators of 0 in the fractions, so these ratios were not computable. From the ratio of the *low* to *capacity* payment condition averaged across milestones in the *backend developer* role, we see that over *a third as many* freelancers (*ratio=0.36*) would still learn skills to prepare for a promised $14/hr job when only paid 10% of their estimated earning capacity (forgoing 90% of income). From the ratios of the lower payment conditions to the *capacity* payment condition averaged across milestones in the *marketer* role, we see that *half as many* to *as many* freelancers (*ratios 0.52* to *1.02*) would still learn skills to prepare for a promised $14/hr job when forgoing large fractions of income (50% to 90%).

Next, we examined the relationship between *compensation amount* and training program *percent completion*. Before performing the Heckman 2-step regression we computed a correlation matrix of all dependent and independent variables and confirmed that no pair of variables has high correlation (cutoff was > .60). Specifically, we encoded *compensation amount* as the hourly rate presented in Table 1 so that the linear regression we performed in the second step of the Heckman was informed of the exact ratio between the payment amounts under the four conditions.

Table 7. Ratios of the Fraction of Invited Marketers Completing each Milestone at each Payment Condition to the Fraction of Invited Marketers Completing each Milestone at the Earning Capacity Payment Condition (the *capacity* payment condition)

Milestone	Ratio of *high* to *capacity*	Ratio of *med* to *capacity*	Ratio of *low* to *capacity*
English	0.81	0.69	1.61
Grammar	1.04	0.86	1.90
Arithmetic	1.04	0.86	1.55
Algebra & Geometry	0.52	0.35	0.69
Idioms	0.52	0.35	0.52
Social Media Marketing	0.52	0.17	0.52
Salesforce	0.69	0.35	0.35
Average for milestones	0.73	0.52	1.02

Table 8. Heckman 2-step Regression of Variables to Predict Program *participation* with Logistic Regression and then *percent completion* with Linear Regression (*backend developer*)

Logistic Regression (probit)	**β Coeff. Estimate**	**Std. Err**	**t value**	**Pr(>|t|)**	
(Intercept)	-1.083	0.445	-2.436	0.015	*
Upwork hours worked	0.000	0.000	-0.356	0.722	
Upwork score	-0.166	0.059	-2.822	0.005	**
# Upwork skills	0.043	0.041	1.054	0.292	
Upwork portfolio count	-0.009	0.014	-0.615	0.539	
Upwork rate	-0.017	0.005	-3.121	0.002	**
number of other Upwork tests passed	0.046	0.022	2.082	0.038	*
compensation amount	0.073	0.071	1.027	0.305	
Linear Regression	**β Coeff. Estimate**	**Std. Err**	**t value**	**Pr(>|t|)**	
(Intercept)	-111.402	46.566	-2.392	0.017	*
# hours in work schedule	0.040	0.160	0.250	0.803	
logic score	3.267	2.100	1.556	0.120	
creativity score	7.087	2.881	2.460	0.014	*
gender	15.604	9.402	1.660	0.097	.
Upwork portfolio count	-1.477	0.802	-1.842	0.066	.
Upwork rate	-0.580	0.434	-1.335	0.182	
number of other Upwork tests passed	1.203	1.256	0.957	0.339	
compensation amount	9.451	3.530	2.677	0.008	**

Tables 8 and 9 present the Heckman 2-step regression for the *backend developer* and *marketer* programs respectively. *Compensation amount* was a statistically significant predictor ($p < 0.01$) of program *percent completion* for *backend developers*. Amongst *backend developers*, *Upwork score*, *Upwork rate*, and *number of other Upwork tests passed* were significant predictors of *participation* in the Heckman probit selection model. *Creativity score* was also a statistically significant predictor of program *percent completion* for *backend developers*. In absolute terms as per Table 5, highest paid *backend developers* had an average *training program progress* that is over double that of the

Table 9. Heckman 2-step Regression of Variables to Predict Program *participation* with Logistic Regression and then *percent completion* with Linear Regression (*marketer*)

Logistic Regression	**β Coeff. Estimate**	**Std. Err**	**t value**	**Pr(>|t|)**	
(Intercept)	-2.070	0.973	-2.128	0.034	*
Upwork hours worked	0.000	0.000	-0.893	0.372	
Upwork score	0.054	0.185	0.292	0.770	
# Upwork skills	0.019	0.051	0.373	0.709	
Upwork portfolio count	-0.010	0.015	-0.621	0.535	
Upwork rate	-0.003	0.002	-1.658	0.098	.
number of other Upwork tests passed	0.034	0.018	1.913	0.056	.
compensation amount	-0.055	0.064	-0.847	0.397	
Linear Regression	**β Coeff. Estimate**	**Std. Err**	**t value**	**Pr(>|t|)**	
(Intercept)	-29.545	222.427	-0.133	0.894	
# hours in work schedule	0.100	0.233	0.429	0.668	
logic score	6.146	2.059	2.985	0.003	**
creativity score	-7.906	5.217	-1.515	0.130	
gender	12.119	10.226	1.185	0.236	
Upwork portfolio count	0.070	1.552	0.045	0.964	
Upwork rate	-0.044	0.296	-0.149	0.881	
number of other Upwork tests passed	0.106	3.230	0.033	0.974	

lowest paid *backend developers*. Amongst *marketers*, *Upwork rate* and *number of other upwork tests passed* were significant predictors of *participation* in the Heckman probit selection model. *Compensation amount* was not a statistically significant predictor of program *percent completion* for *marketers*. *Logic score* was the only statistically significant predictor for *marketers*.

From Tables 8 and 9, we see for *backend developers* that the coefficients for *compensation amount* and *creativity score* are positive 9.451 and positive 7.087 respectively, suggesting that the greater either of these two variables are for a freelancer, the greater the freelancer's training program *percent completion* is. We see for *marketers* that the coefficient for *logic score* is positive 6.146, suggesting that the greater a freelancer's logic abilities are, the greater the freelancer's training program *percent completion*.

The hypothesis (*H1*) that online freelancers will forgo income to learn lucrative skills is supported by the fact that both *backend developers* and *marketers* made on average positive (non-negligible) *training program progress* in pay conditions *low*, *medium*, and *high*, all of which correspond to compensation that is lower than what the pilot estimated they command. Moreover, the ratios of milestone completion rates at lower pay conditions to the *capacity* condition suggest that large fractions of freelancers would indeed forgo income to learn skills.

The second hypothesis (*H2*) was that the higher the compensation is for learning skills (i.e. the closer it is up to their earning capacity), the greater the *training program progress* will be for the payment group. This hypothesis is only supported in the case of *backend developers*.

The first set of milestone progress histograms (Figure 10) shows that the drop-off is less steep in the higher paying *backend developer* conditions than the lower paying *backend developer* conditions. This same trend is only present in the *marketer* scenario when comparing the highest paying condition with the other three. The next set of histograms (Figure 11) is organized by milestone rather than role type and condition. Each group of four bars corresponds to the four payment conditions, so if higher pay conditions leads to greater completion rates of any given milestone, we would expect bars to be rising. We witness this rising effect in the case of the *backend developer* (*bdv*) bars for all milestones. Except for in the last few milestones, we do not see this rising effect in the case of the *marketer* (*mkr*) bars. These visualizations make clear how pay condition affects the *backend developer* TPC but not *marketer* TPC.

We now report subjective results. Nobody reported that the pay is "too much". 0 out of 4 *backend developers* in the *low* condition, 3/8 in *med*, 5/6 in *high*, 4/5 in *capacity* reported the compensation as "the right amount" as opposed to "too little". Amongst *marketers*, the ratings were 3/5 in *low*, 0/2 in *med*, 5/5 in *high*, and 4/4 in *capacity*.

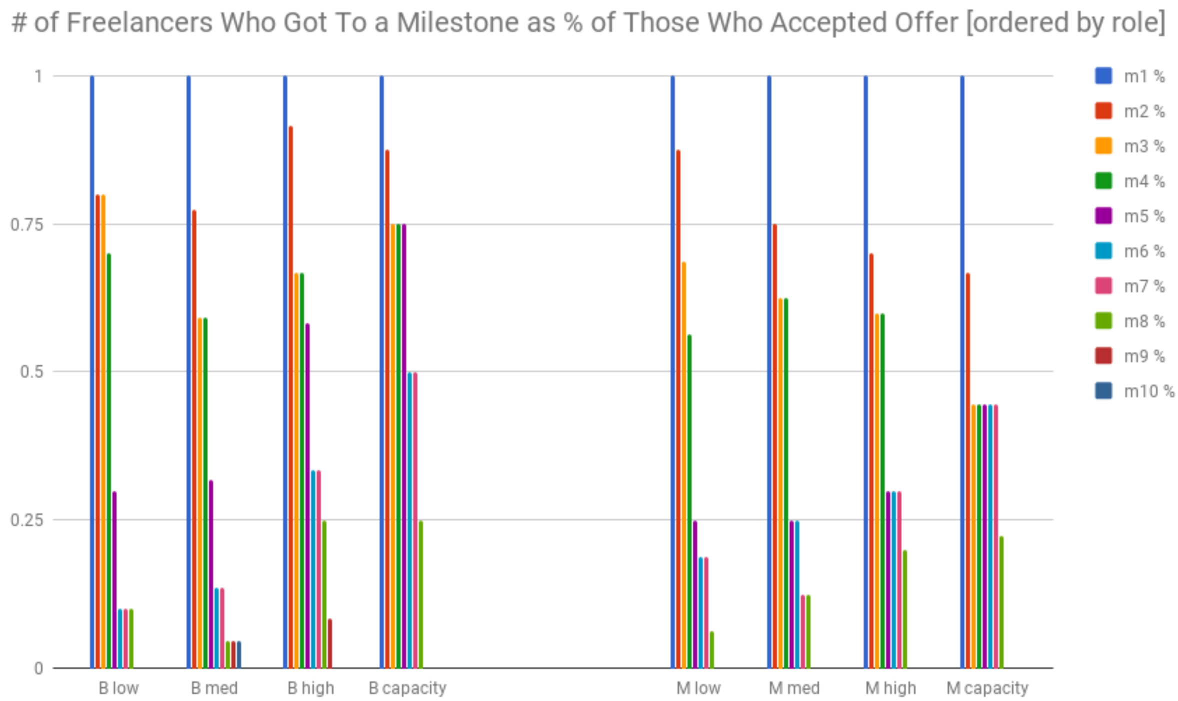


Figure 10. These milestone progress histograms present the percent of participants (y-axis) who entered each role/condition and made it to a particular milestone (x-axis). Drop-off reduces as *backend developers* are paid more.

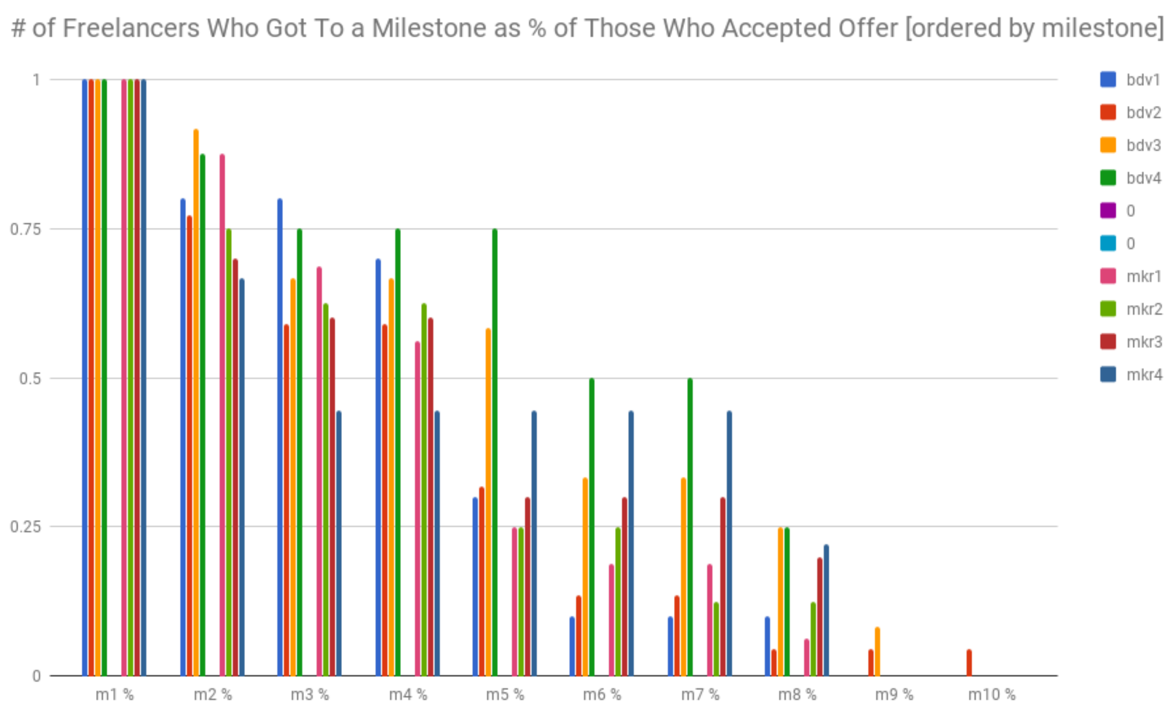


Figure 11. These histograms present the percent of participants (y-axis) entering each role/milestone across conditions (x-axis).

DISCUSSION

This study shows that financial incentives, akin to literature's governmental conditional cash transfer, can play a positive role in online skill development. The hypothesis (*H1*) that online freelancers will forgo income to learn lucrative skills is supported by the fact that both *backend developers* and *marketers* made nontrivial *training program progress* in payment conditions lower than their estimated earning capacity. The reason why may be two-fold. The main reason may indeed be because the freelancers viewed the opportunity to learn skills as a form of extra compensation because they see it as a positive influence on their online career. One marketer in the *high* payment condition said: "You get to learn and earn money both at the same time. The bad part is you don't know how long you'll be working." On the other hand, a *backend developer* in the *capacity* condition said: "[It's] preparing me [for] … some big projects." A *marketer* in the *capacity* condition said the program "need[s] a lot of self-drive factors (which not everyone has much)", which highlights that self-motivation is important for completing the training. On the other hand, the quote *"*This is really good for the trainee to increase patience and learn something.*"* from a marketer in the *med* condition underscores quite well that cash incentives are able to tackle this hurdle – a notion supported empirically both by literature and our study. People want to engage in good habits, but the fact that good habits require patience and hard work make that difficult. Money, even in sums less than peoples' earning capacities, can act as a subsidy or impetus to help people get over the hurdles that stop them from doing what they already intrinsically want to do. The other reason why freelancers made nontrivial *training program progress* even under sub-earning-capacity pay conditions may be logistical. While freelancers in our pilot study reported earning \$11.43/hr on average, they also reported working only ~16 hours online and ~14 hours offline on average, which suggests that finding enough work may be an issue for Upwork freelancers. This means that these freelancers may forgo income by doing a lower paying job because the lower paying job is immediately available, even though their earning capacity is higher.

The hypothesis (*H2*) that the higher the compensation is for learning skills the greater the *training program progress* will be for the payment group is only supported in the case of *backend developers*. In fact even before the *backend developer* training program began, this notion was present as the number of intent-to-treat participants who passed the background check and accepted our offer rises with *payment level* (see Figure 5). It makes sense that the significant variables for predicting *training program progress* of *backend developers* are *payment level* and *creativity score* because backend development takes freelancers significantly longer to learn than marketing before achieving the prize \$14 per hour job. One *backend developer* in the *low* condition said "Some concepts are hard to grasp and it takes quite a long time to understand them thus making progress very slow." We see that freelancers need time to learn full-on programming languages like Python as per this testimonial and expect to be compensated for that time. Moreover, computer science concepts require creativity and out-of-the-box thinking. It is strange that *logic score* is not also a significant predictor for *backend developer* TPC though this may be because creativity (particularly our measure of it which emphasizes riddle-solving) is a better indicator. It seems that *logic score* and not *payment level* is a significant predictor of *marketer* TPC because smart marketers would be able to whiz through the objective social media marketing and Salesforce exams even at the lowest pay level. These exams require much less dedicated time for learning skills than the programming exams require for the *backend developer* role. It may be easier for *marketers*, particularly quick smart ones, to bear low pay for learning skills because the prize \$14 per hour job is much closer in sight than for *backend developers* who have to get past the hurdles of learning Python and Django. It seems that big hurdles are what increased cash incentives help mitigate, a notion consistent with literature. In fact, even in the sub-portions of the *marketer* training program that require relatively more studying like the Salesforce exam (last histogram of Figure 11), we witness rising bars of greater milestone completion rate sat higher pay levels. This study suggests that large cash incentives will be particularly effective when society needs to produce more highly-specialized highly-skilled workers.

We do not find evidence that support literature's claim that cash incentives benefit women more than average. The mean TPC for female *backend developers* and *marketers* are 10.6% and 22.9% respectively, as opposed to the overall averages of 21.0% and 30.3%. An important limitation to consider here is insufficient data (8 female *backend developers*, 11 female *marketers*).

Participant testimonials suggest that the extrinsic motivator of increased financial incentives actually enhance intrinsic motivation to learn, a notion supported by meta-analyses presented in the literature review. In the low paying conditions, both *backend developers* and *marketers* were concerned about the time commitment of learning. A *marketer* in the *low* condition said: "Good part is that you can renew your knowledge from the past and practice some things that maybe you forgot … bad part is that it takes a lot of time." On the other hand, a *marketer* in the *med* condition was more blunt saying simply "there were too many questions." Some even questioned the necessity of prerequisites, saying "It should be related to technical skills only like I am back end developer, like c#, .net, SQL, jQuery, database etc. but not like arithmetic, grammar, algebra." In the higher pay conditions, freelancers seemed in general less worried about finances and hence more enthusiastic to learn skills. For instance, a *backend developer* in the *high* condition said "Excellent idea. Like the idea of giving people the incentive to learn and improve current skills.” Finally, a *marketer* in the *med* condition said, *“This is really good for the trainee to increase patience and learn something.”*, which suggests that cash incentives help freelancers, who always wanted to learn skills, buckle down and develop the patience to actually do it.

This study was designed to train freelancers from all over the world for real jobs. We looked at two different roles *backend developer* and *marketer*, varied payment with respect to freelancers’ estimated earning capacity from a pilot, and based milestone pay on reported milestone completion times from the pilot. Because of the real-world nature of the work and training, it could take over $1000 to train a single freelancer (*backend developer* in the *high* pay condition) so making this experiment use more payment levels, more roles, and more participants is an expensive feat. More payment levels may be particularly important for testing whether the *marketer* role has sensitivity at the lower end of the spectrum due to the ease of its milestones compared to the *backend developer* role. The pilot itself could have been larger to better estimate the earning capacity of freelancers and the milestone completion times, upon which the entire payment scheme is based. This study has good global representation, but more gender diversity is also important for achieving more generalizable results. Still, this study is important because it suggests a method for society to expedite and scale complex skill-development.

CONCLUSIONS AND FUTURE WORK

The nexus of online freelancing, online education, and conditional cash transfer has potential to meet global demand for high-skill labor and also promote global economic development through education and employment. It is particularly interesting due to recent rapid growth in internet penetration. This study tests whether cash incentives, priced at fractions of estimated participant earning capacity, can convince freelancers to learn lucrative skills and finds that they can by bolstering intrinsic motivation. However, this study only finds evidence for a strong relationship between specific compensation level and learning progress in highly complex training programs that take significant time and effort to complete. Future work would be to create a system to test the self-sustainability and scalability of online educational conditional cash transfer. Testing whether human capital investments in education make sense would require training freelancers like is done in this paper, creating teams out of them, and matching those with real projects from paying clients. The system would have to deal with issues of attrition of trained freelancers and the experiential gap between a “trained” freelancer and an “expert”. Developing such a system would be an impressive feat for global human and economic development.

ACKNOWLEDGMENTS

We thank Michael Bernstein for advice on the study design and feedback on the paper.

REFERENCES

1. Per J. Agrell, Peter Bogetoft, and Jorgen Tind. 2002. Incentive plans for productive efficiency, innovation and learning. *International Journal of Production Economics* 78, 1: 1–11. https://doi.org/10.1016/S0925-5273(00)00097-9

2. By Alfie, Kohn Boston, and Houghton Mifflin. 1995. Punished by Rewards : The Trouble with Gold Stars , Incentive Plans , A ’ s , Praise , and Other Bribes. *Sociology The Journal Of The British Sociological Association*: 92–96. https://doi.org/10.1177/0022487197048002009

3. Bradley M Allan and Roland G Fryer. 2011. *The Power and Pitfalls of Education Incentives*. Retrieved from http://www.brookings-tsinghua.cn/~/media/Files/rc/papers/2011/09_incentives_fryer_allen/092011_incentives_fryer_allen_appendix.pdf

4. Joshua Angrist, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. 2002. Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review* 92, 5: 1535–1558. https://doi.org/10.1257/000282802762024629

5. Joshua Angrist, Eric Bettinger, and Michael Kremer. 2006. Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. *American Economic Review 96*, 847–862. https://doi.org/10.1257/aer.96.3.847

6. Joshua Angrist, Daniel Lang, and Philip Oreopoulos. 2006. Lead Them to Water and Pay Them to Drink: An Experiment with Services and Incentives for College Achievement. *The National Bureau of Economic Research*. Retrieved September 17, 2017 from http://www.nber.org/papers/w12790

7. Joshua Angrist, Daniel Lang, and Philip Oreopoulos. 2009. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics* 1, 1: 136–163. https://doi.org/10.1257/app.1.1.136

8. Maria Caridad Araujo and Norbert Schady. 2006. Cash transfers, conditions, school enrollment, and child work : evidence from a randomized experiment in Ecuador. *Annals of Human Biology* 36, 6: 717. https://doi.org/10.3109/03014460903058992

9. Mary Arends-Kuenning and Sajeda Amin. 2004. School Incentive Programs and Children's Activities: The Case of Bangladesh. *Comparative Education Review* 48, 138: 295–317. https://doi.org/10.1086/421376

10. Autumn group. IQTest.com. Retrieved December 5, 2017 from http://www.iqtest.com/test.html

11. Lisa Barrow, Lashawn Richburg-Hayes, Rouse Cecilia Elena, and Thomas Brock. 2014. Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-Income Adults. *Journal of Labor Economics 32*, 563–599. https://doi.org/10.1086/675229

12. Jere R. Behrman, Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. 2015. Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools. *Journal of Political Economy* 123, 2: 325–364. https://doi.org/10.1086/675910

13. Jere R. Behrman, Pilali Segupta, and Petra Todd. 2001. Progressing Through Progresa: An Impact Assessment of a School Subsidy Experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.286295

14. Jere R Behrman and James C Knowles. 1999. Household Income and Child Schooling in Vietnam. *The World Bank Economic Review* 13, 2: 211–256. https://doi.org/10.1093/wber/13.2.211

15. Roland Benabou and Jean Tirole. 2003. Intrinsic and extrinsic motivation. *Review of Economic Studies* 70, 3: 489–520. https://doi.org/10.1111/1467-937X.00253

16. Eric P. Bettinger. 2012. Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores. *Review of Economics and Statistics* 94, 3: 686–698. https://doi.org/10.1162/REST_a_00217

17. Allie Bidwell. 2015. Andela: An African Company Paying People to Learn Computer Science | STEM Solutions | US News. *US News*. Retrieved September 17, 2017 from https://www.usnews.com/news/stem-solutions/articles/2015/05/14/andela-an-african-company-paying-people-to-learn-computer-science

18. Alan De Brauw and John Hoddinott. 2011. Must conditional cash transfer programs be conditioned to be effective? The impact of conditioning transfers on school enrollment in Mexico. *Journal of Development Economics* 96, 2: 359–370. https://doi.org/10.1016/j.jdeveco.2010.08.014

19. J. Cameron and W. D. Pierce. 1994. Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research* 64, 3: 363–423. https://doi.org/10.3102/00346543064003363

20. J Cameron, K M Banko, and W D Pierce. 2001. Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior analyst / {MABA}* 24, 1: 1–44. https://doi.org/10.1016/j.neuron.2009.04.009

21. Nazmul Chaudhury and Dilip Parajuli. 2006. *Conditional cash transfers and female schooling: The impact of the female school stipend program on public school enrollments in Punjab, Pakistan*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract%7B_%7Did=953570

22. G Chen, D Davis, M Krause, E Aivaloglou, C Hauff, and G J Houben. 2016. Can Learners be Earners? Investigating a Design to Enable MOOC Learners to Apply their Skills and Earn Money in an Online Market Place. *IEEE Transactions on Learning Technologies* PP, 99: 1. https://doi.org/10.1109/TLT.2016.2614302

23. Kim Clark. 2017. Coding Bootcamp Revature Offers Free Tuition, Guaranteed Job | Money. *Time Magazine*. Retrieved September 17, 2017 from http://time.com/money/4711742/coding-bootcamp-apprentice-revature/

24. GabrielCulbertson R Culbertson, Erik L Andersen, Walker M White, Daniel Zhang, and Malte F Jung. 2016. Crystallize: An Immersive, Collaborative Game for Second Language Learning. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, 635–646. https://doi.org/10.1145/2818048.2820020

25. Joshua Daron Achemoglu; Angrist. 2000. How large are human capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomic Annual* 15, 2000: 9–59.

26. Edward L. Deci. 1972. The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance* 8, 2: 217–229. https://doi.org/10.1016/0030-5073(72)90047-5

27. Erich C. Dierdorff and Eric A. Surface. 2008. If You Pay for Skills, Will They Learn? Skill Change and Maintenance Under a Skill-Based Pay System. *Journal of Management* 34, 4: 721–743. https://doi.org/10.1177/0149206307312507

28. Alpana Dubey, Kumar Abhinav, Sakshi Taneja, Gurdeep Virdi, Anurag Dwarakanath, Alex Kass, and Mani Suma Kuriakose. 2016. Dynamics of Software Development Crowdsourcing. In *2016*

IEEE 11th International Conference on Global Software Engineering (ICGSE), 49–58. https://doi.org/10.1109/ICGSE.2016.13

29. Suzanne Duryea, Andrew Morrison, Suzanne Duryea, Andrew Morrison, Suzanne Duryea, Andrew Morrison, Suzanne Duryea, and Andrew Morrison. 2016. The Effect of Conditional Transfers on School Performance and Child Labor: Evidence from an Ex-Post Impact Evaluation in Costa Rica. Retrieved September 17, 2017 from https://publications.iadb.org/handle/11319/3293

30. Brent Evans and Rachel Baker. 2016. MOOCs and Persistence: Definitions and Predictors.: EBSCOhost. In *New Directions in Institutional Research*. 69–86. https://doi.org/10.1002/ir.20155

31. Bruno S. Frey and Margit Osterloh. 2002. Motivation — A Dual-Edged Factor of Production. 3–26. https://doi.org/10.1007/978-3-662-10132-2_1

32. Roland G. Fryer. 2011. Financial incentives and student achievement: Evidence from randomized trials. *Quarterly Journal of Economics* 126, 4: 1755–1798. https://doi.org/10.1093/qje/qjr045

33. Roland G. Fryer Jr and Richard T. Holden. 2012. *Multitasking, learning, and incentives: A cautionary tale*. https://doi.org/10.1007/s13398-014-0173-7.2

34. S S Gaikwad, D Morina, A Ginzberg, C Mullings, S Goyal, D Gamage, C Diemert, M Burton, S Zhou, M Whiting, K Ziulkoski, A Ballav, A Gilbee, S S Niranga, V Sehgal, J Lin, L Kristianto, J Regino, N Chhibber, D Majeti, S Sharma, K Mananova, D Dhakal, W Dai, V Purynova, S Sandeep, V Chandrakanthan, T Sarma, S Matin, A Nassar, R Nistala, A Stolzoff, K Milland, V Mathur, R Vaish, and M S Bernstein. 2016. Boomerang : Rebounding the Consequences of Reputation Feedback on Crowdsourcing Platforms. *Proc. UIST '16*: 625–637. https://doi.org/10.1145/2984511.2984542

35. James J. Heckman. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47, 1: 153. https://doi.org/10.2307/1912352

36. John J. Horton. 2010. Online labor markets. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 515–522. https://doi.org/10.1007/978-3-642-17572-5_45

37. Kazushi Ikeda and Michael S. Bernstein. 2016. Pay It Backward. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 4111–4121. https://doi.org/10.1145/2858036.2858327

38. C. Kirabo Jackson. 2010. A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program. *Journal of Human Resources* 45, 3: 591–639. https://doi.org/10.1353/jhr.2010.0019

39. C. Kirabo Jackson. 2010. The Effects of an Incentive-Based High-School Intervention on College Outcomes. *The National Bureau of Economic Research*. Retrieved September 17, 2017 from http://www.nber.org/papers/w15722

40. Alain de Janvry and Elisabeth Sadoulet. 2004. Conditional Cash Transfer Programs: Are they really magic bullets? *ARE Update*: 1–5. Retrieved from http://pocketknowledge.tc.columbia.edu/home.php/viewfile/download/55924/CCT_programs_Are_they_really_magic_bullets.pdf

41. Robert T. Jensen. 2010. The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics* 125, 2: 515–548. https://doi.org/10.1257/app.1.1.219

42. a Kittur, J Nickerson, and M Bernstein. 2013. The Future of Crowd Work. *Proc. CSCW '13*: 1–17. https://doi.org/10.1145/2441776.2441923

43. Michael Kremer, Edward Miguel, and Rebecca Thornton. 2005. Incentives to learn. *Education Next 5*, 57–64. https://doi.org/10.1162/rest.91.3.437

44. Chinmay E. Kulkarni, Richard Socher, Michael S. Bernstein, and Scott R. Klemmer. 2014. Scaling short-answer grading by combining peer assessment with algorithmic scoring. *Proceedings of the first ACM conference on Learning @ scale conference - L@S '14*: 99–108. https://doi.org/10.1145/2556325.2566238

45. Chinmay Kulkarni, Koh Pang Wei, Huy Le, Daniel Chia, Kathryn Papadopoulos, Justin Cheng, Daphne Koller, and Scott R. Klemmer. 2013. Peer and self assessment in massive online classes. *ACM Transactions on Computer-Human Interaction* 20, 6: 1–31. https://doi.org/10.1145/2505057

46. Edwin Leuven, Hessel Oosterbeek, and Bas van der Klaauw. 2010. The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment. *Journal of the European Economic Association* 8, December 2010: 1243–1265. https://doi.org/10.2139/ssrn.383160

47. Tharindu Rekha Liyanagunawardena, Patrick Parslow, and Shirley Ann Williams. 2014. Dropout: MOOC Participants' Perspective. In *Proceedings of the European MOOC Stakeholder Summit 2014*, 95–100. Retrieved from http://centaur.reading.ac.uk/36002/

48. Edwin A. Locke. 1968. Toward a theory of task

motivation and incentives. *Organizational Behavior and Human Performance* 3, 2: 157–189. https://doi.org/10.1016/0030-5073(68)90004-4

49. Ilaria Maselli. 2012. The evolving supply and demand of skills in the labour market. Retrieved September 17, 2017 from http://econpapers.repec.org/paper/epscepswp/7911.htm

50. Susan Dickerson Mayes, Susan L. Calhoun, Edward O. Bixler, and Dennis N. Zimmerman. 2009. IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences* 19, 2: 238–241. https://doi.org/10.1016/j.lindif.2008.09.001

51. Xianhai Meng and Brendan Gallagher. 2012. The impact of incentive mechanisms on project performance. *International Journal of Project Management* 30, 3: 352–362. https://doi.org/10.1016/j.ijproman.2011.08.006

52. Miniwatts Marketing Group. 2017. World Internet Users Statistics and 2017 World Population Stats. Retrieved September 17, 2017 from http://www.internetworldstats.com/stats.htm

53. Ted O'Donoghue and Matthew Rabin. 1999. Doing it now or later. *American Economic Review* 89, 1: 103–124. https://doi.org/10.1257/aer.89.1.103

54. G M Olson and J S Olson. 2000. Distance matters. *Human-Computer Interaction* 15, 2: 139–178. https://doi.org/10.1207/S15327051HCI1523_4

55. D.F.O Onah, J Sinclair, and Boyatt. 2014. DROPOUT RATES OF MASSIVE OPEN ONLINE COURSES : BEHAVIOURAL PATTERNS MOOC Dropout and Completion : Existing Evaluations. *Proceedings of the 6th International Conference on Education and New Learning Technologies (EDULEARN14)*: 1–10. https://doi.org/10.13140/RG.2.1.2402.0009

56. Lawrence O. Picus. 1991. *USING INCENTIVES TO STIMULATE IMPROVED SCHOOL PERFORMANCE : AN ASSESSMENT OF ALTERNATIVE*.

57. Chris Piech, Jonathan Spencer, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas Guibas, Jascha Sohl-dickstein, and Khan Academy. 2015. Deep Knowledge Tracing. *arXiv*: 1–13.

58. Margaret Raymond. 2008. Paying for A's: An Early Exploration of Student Reward and Incentive Programs in Charter Schools. Retrieved September 17, 2017 from https://credo.stanford.edu/downloads/paying_for_a.pdf

59. William J Reid, Cynthia A Bailey-Dempsey, Elizabeth Cain, and Toni V Cook. 1994. Cash incentives versus case management: Can money replace services in preventing school failure? 4977. *Social Work Research 18*, 227–236. https://doi.org/10.1093/SWR/18.4.227

60. Mark A. Runco and Selcuk Acar. 2012. Divergent Thinking as an Indicator of Creative Potential. *Creativity Research Journal 24*, 66–75. https://doi.org/10.1080/10400419.2012.652929

61. Daniel M Russell, Scott Klemmer, Armando Fox, Celine Latulipe, Mitchell Duneier, and Elizabeth Losh. 2013. Will massive online open courses (moocs) change education? In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2395–2398. https://doi.org/10.1145/2468356.2468783

62. Beng Huat See, Carole J Torgerson, Stephen Gorard, Hannah Ainsworth, Graham Low, Kath Wright, and Carole Torgerson Beng Huat See Stephen Gorard Hannah Ainsworth Graham Low and Kath Wright. 2011. Factors that promote high post-16 participation of some minority ethnic groups in England: a systematic review of the UK-based literature. *Research in Post-Compulsory Education* 16: 85–100. https://doi.org/10.1080/13596748.2011.549736

63. Dhiraj Sharma. 2010. *The Impact of Financial Incentives on Academic Achievement and Household Behavior: Evidence from a Randomized Trial in Nepal*. https://doi.org/10.2139/ssrn.1681186

64. Robert E. Slavin. 2010. Can financial incentives enhance educational outcomes? Evidence from international experiments. *Educational Research Review 5*, 68–80. https://doi.org/10.1016/j.edurev.2009.12.001

65. Stanford HCI. 2015. Daemo: a Self-Governed Crowdsourcing Marketplace. In *UIST*, 2–3. https://doi.org/10.1145/2815585.2815739

66. Ryo Suzuki, Niloufar Salehi, Michelle S. Lam, Juan C. Marroquin, and Michael S. Bernstein. 2016. Atelier. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2645–2656. https://doi.org/10.1145/2858036.2858121

67. Todd Anderson. Test Your Creativity: 5 Classic Creative Challenges - 99U - 99U. Retrieved December 5, 2017 from http://99u.com/articles/7160/test-your-creativity-5-classic-creative-challenges

68. Mark E. Whiting, Dilrukshi Gamage, Snehalkumar

S. Gaikwad, Aaron Gilbee, Shirish Goyal, Alipta Ballav, Dinesh Majeti, Nalin Chhibber, Angela Richmond-Fuller, Freddie Vargus, Tejas Seshadri Sarma, Varshine Chandrakanthan, Teogenes Moura, Mohamed Hashim Salih, Gabriel Bayomi Tinoco Kalejaiye, Adam Ginzberg, Catherine A. Mullings, Yoni Dayan, Kristy Milland, Henrique Orefice, Jeff Regino, Sayna Parsi, Kunz Mainali, Vibhor Sehgal, Sekandar Matin, Akshansh Sinha, Rajan Vaish, and Michael S. Bernstein. 2016. Crowd Guilds: Worker-led Reputation and Feedback on Crowdsourcing Platforms. https://doi.org/10.1145/2998181.2998234