Evaluating Sama's Training and Job Programs in Nairobi, Kenya

David Atkin (MIT) Antoinette Schoar (MIT Sloan)
Kiara Wahnschafft (MIT)

May 12, 2021

Executive Summary

This paper explores the impacts of two related programs provided by Sama in Nairobi, Kenya. The first is a training program, Artificial Intelligence 101 (AI 101), that teaches young workers the skills needed for digital work. The second is the Sama Delivery Center that provides jobs doing digital work to these young workers.

We use a randomized control trial to rigorously evaluate the program over multiple years. Applicants were either provided training, provided training and a referral for a job at Sama, or placed in a control group. The findings reveal:

- Workers receiving both training and a job referral see almost 40 percent higher earnings and 10 percentage points lower unemployment than the control group, an effect that is sustained through the midline and endline surveys.
  - Effects are particularly strong for women, and earnings benefits were particularly dramatic during the Covid-19 pandemic when Sama was able to offer remote work.
  - These workers also experience increases in their current and projected life satisfaction as well as improvements in their savings tendencies.
- Workers receiving only training struggle to find digital work in the first couple of years post training with slightly negative impacts on earnings. Eventually earnings recover, particularly for males, as workers enter other industries.
- There is some evidence that the promise of a job referral induces additional effort during training, particularly in terms of attendance.

Study Motivation

An increasing number of countries, especially those in the developing world, see the Information and Communication Technology (ICT) sector as having the potential to generate a large expansion in well-paying formal-sector jobs and rapid economic growth. In particular, the business process outsourcing (BPO) sub-sector, worth over $130 billion worldwide , provides opportunities for wealth generation and employment creation for countries in the economic periphery. These opportunities are particularly needed in Africa, where population growth has exceeded formal sector job growth. These opportunities seem well-suited for the African population given the significant young computer-literate individuals fluent in several global languages. However, we have a limited understanding of both how to equip these young workers with the skills required for digital work opportunities and of how this type of work impacts incomes and wellbeing.

In this study, MIT researchers David Atkin and Antoinette Schoar partnered with Sama, a mission-driven technology company tackling this challenge. The goal was to measure the effectiveness of the Artificial Intelligence 101 (AI 101) training program and to understand its complementarity with employment opportunities provided at Sama's Delivery Center (SamaDC). The AI 101 training focuses on providing market-aligned occupational skills, driven by real client demands and project work procured through SamaDC. At the same time, a pre-employment training program like AI 101 allows Sama to prepare and hire even more marginalized and vulnerable workers that otherwise would not pass through Sama's hiring process.

The Programs

Artificial Intelligence 101 (AI 101) is a free, classroom-based vocational training program . Trainees attend class 5 hours a day for a ten day period. The curriculum covers basic digital skills, occupational skills relevant to digital work and the business process outsourcing (BPO)

industry, and job search preparation. The curriculum leverages insights gained from Sama's 8+ years of operational experience sourcing and completing digital work projects. The curriculum is largely technical skills based training, but there is a basic module that helps graduates navigate the job application process. Graduates of AI 101 pursue job opportunities at Sama and its vetted network of hiring partners. Since beginning operations in 2015, the training program has graduated 4,134 trainees (count through 4th quarter of 2020), 51% of whom are women.

Sama's Delivery Center (SamaDC) hires some graduates of the training program to complete digital projects, sourced from clients such as Google and Walmart, including image tagging, image annotation, data classification, or dataset creation for machine learning algorithms. The SamaDC is a formal office environment where entry-level employees (agents) can learn both technical digital work skills through client work and soft skills to jumpstart their careers in the formal employment sector. On average, tenure at the DC for agent entry-level positions is 25 months (2020). The monthly salary is, on average, $301 USD including salary and benefits. This wage is generous—approximately 2.5 times the formal minimum wage in Kenya—and workers also receive benefits such as health care, pension, meal subsidies, and transportation for night shift workers. Workers typically work 22 days in a month. Entry level agents typically work one of three shifts at the SamaDC in Nairobi for a span of 9 hours.

Research Question and Research Design

The research study uses a randomized controlled methodology to rigorously answer the following main research questions:

- What are the effects of a digital training program on job seeking behavior, employment and earnings of training participants? Are the skills portable?
- What are the effects of the training program paired with a high likelihood of a job at

SamaDC that requires exactly these skills?

- Does the expectation of receiving a job requiring these skills lead participants to make more of the training program?

Randomization

In order to test the complementarity of the two Sama programs, this evaluation consists of two treatment arms: individuals that are randomly admitted to the AI 101 program but are not eligible to apply for jobs at the SamaDC (training only) and individuals that are admitted to the AI 101 program and are eligible to apply for jobs at the SamaDC (training + job). Although we could not guarantee that participants who successfully completed the training course would be hired by SamaDC, given the close ties between the two programs, participants had strong expectations that they would be hired. Before the start of each training cohort, a sample of eligible AI 101 training applicants (i.e. those who have passed an initial online screening) were randomized into 3 groups as follows:

- Group 1 (Control): Applicants to AI 101 who passed the initial eligibility check but were not selected to attend the training as part of the randomization. They were informed that they were not eligible to apply for 1-2 years.

- Group 2 (Training): Applicants to AI 101 who were informed at the start of the training that they would not receive any job referrals to work for Sama after completing the training. However, they were free to apply for employment at any of Sama's partner work centers or other jobs.

- Group 3 (Training and Job): Applicants to AI 101 who were informed on the first day of training about both Sama partner work centers and employment opportunities within Sama. They were told that jobs were available at Sama, and their candidacy would be forwarded to Sama. Similarly to Group 2, they were told that the AI 101

training program would do its best to prepare graduates for employment and those who applied for jobs would still need to pass the standard candidate interview and vetting process to actually receive an employment offer. This information was given to them after they were accepted into the training program.

Study sample

The primary geography for the research study is Nairobi, Kenya, where the Sama-owned and -operated Delivery Center is based. As the AI 101 training program already existed and was operational, the study participants for this project were pulled from the program's applicant pool. All applicants to the AI 101 program were required to complete an online screening form. The information collected through this form was used to determine if an applicant had basic qualifications for the training program.

The AI 101 training program targets low-income youth in Nairobi with little to no formal experience. The minimum criteria are youth between the ages of 18 and 35 with a high school certificate and a proficiency in English. The Learning & Development department – Sama's workforce development and mobilization arm – works with various community training partners who are located and embedded within low income settlements in Nairobi and thus able to recruit the target demographic for training. This study relied on Sama's standard communication and advertising process for sample recruitment.

The study targeted a sample size of 256 or more participants in each group so as to have sufficient power for a minimal detectable effect of a 25 percent increase in income. Due to a large client contract that necessitated a hiring ramp at SamaDC, for 3 months Sama required that all treated participants were in Treatment Arm 2. Hence we continued enrollment for the study until Treatment 1 group had exceeded the target of 256. In total, the final enrolled study sample was 1115. The final sample size by group is as follows:

- Group 1 (Control): 281

- Group 2 (Training): 286
- Group 3 (Training and Job): 548

Data Collection

At baseline, the enumerators collected data on household demographics, income and expenditure, employment history and respondent job seeking behavior. Data was collected on performance during training exercises to test the hypothesis that participant effort levels may respond to the belief that the skills would be of immediate value at SamaDC. Data on job seeking behavior, employment, income and expenditure of the individual and the individual's household along with well-being and personal motivation were collected in a midline survey that was conducted, on average, 16 months after baseline. This same information was then collected again in an endline survey that was conducted, on average, 2.5 years after baseline. Table 1.1 provides further details on the timing of the survey rounds. All surveys were conducted by experienced operators over the phone.

Baseline Balance, Takeup, and Attrition

Baseline Balance

To better understand the demographics of the sample as well as examine baseline balance, we explore the mean values of key characteristics at baseline. The first column of Table 1.2 reports the mean of each characteristic for the control group, with different rows displaying different characteristics. The second and third columns compare the means of these characteristics in each treatment arm relative to the control group. Our sample is half female, has a mean age of 24, and is generally not married. Consistent with these being applicants to a job training program, sixty percent are currently unemployed with an average hours worked per week of 7.4 across all groups. We find no significant differences across the groups, except

in baseline hours worked, where we see that the training group and the training + job group both worked significantly fewer hours than the control group at baseline. This may bias us to finding muted effects of the treatment arms on employment outcomes in the absence of suitable baseline controls.

Attrition

In Table 1.3, we explore whether there was differential attrition by treatment group across survey rounds. At endline, the training + job group is just as likely to answer our survey as the control group. There is a slightly lower response rate among the training-only group (5.5 percent lower, significant at the 10 percent level). That said, response rates are generally high with rates in the other two groups around 85%. At midline, there does appear to be differential attrition with a 30 percent lower response rate for the training + job group (and so results for that group at midline should be treated with some caution). This decreased attrition is due to the fact that these individuals were harder to survey given their busy work schedules; however, for the latter half of the midline data collection and for the endline data collection, the surveyors made phone calls outside of working hours and thus had more success reaching the respondents.

Empirical Strategy

Our primary approach is an ANCOVA regression specification where we regress an outcome measured in period t, y_{it}, on dummies for whether the individual is in the training-only group ($Training_i$) and training + job treatment groups ($TrainingJob_i$), controls for baseline values of the outcome variable (when available), as well as fixed effects for the year-month of enrollment (δ_{ie}) and the year-month of the survey (δ_{is}):

$$y_{it} = \beta_1 Training_i + \beta_2 TrainingJob_i + \gamma y_{i0} + \delta_{ie} + \delta_{is} + \varepsilon_{it} \quad (1)$$

The coefficients β_1 and β_2 tell us whether we observe different outcomes for individuals assigned to the two treatment groups. This is an intent-to-treat (ITT) specification. We run a similar treatment-on-the-treated specification (TOT) by regressing the outcome on a dummy for whether the individual received AI 101 training and dummy for whether they ever worked at SamaDC, instrumenting with the treatment group dummies above. In this specification we recover the returns to those who take up training or manage to obtain a SamaDC job after training. We report these TOT results in *Appendix Tables 9.1 and 9.2.* As the effects are very similar to the ITT just scaled up by the proportion of the treatment groups taking up the treatment, we focus on the ITTs in the analysis that follows.

In some of our analysis we will explore heterogeneity in the responses among different subgroups by interacting membership of one of these groups with our treatment dummies.

Takeup

We first explore take up, specifically whether the individuals took the AI 101 course and whether they ever obtained a job at SamaDC. Finally, we also explore whether the individuals are currently a SamaDC employee to help us interpret some of the effects we will identify later. In Tables 1.4 and 1.5, we regress these three different outcomes on our treatment group dummy variables along with the controls above (note that baseline values of these outcomes are zero for all individuals). Table 1.4 reports results for the sample surveyed at midline and Table 1.5 reports results for the sample surveyed at endline (with responses correct at the time of the relevant survey).

Column 1 presents the results for whether a participant attended the AI 101 course. Depending on whether we look at the midline or endline sample, between around 83 percent of the training group attended some training, rising to between 91 and 94 percent for the training + job group. (There were a few cases of non compliance with 6 to 8 percent of the control group attending AI 101 due to administrative issues at one of the training centers.) Thus, our intervention was successful in enrolling study participants in the two treatment

arms into the AI 101 training program.

The 8 to 10 percentage point gap in attendance between the training and training + job group provides some initial evidence that the offer of the job increased the incentive to attend training. Column 2 replaces a dummy for whether the participant attended AI 101 with a dummy for whether they completed the course. Consistent with the training-only group being less motivated, the gap in attendance between groups widens to between 12 and 14 percent when examining whether an individual completed training.

As shown in column 3, about half of the training + job group ever worked at SamaDC, with column 4 showing that 37 percent are currently working at SamaDC at the time of the two surveys. These percentages are substantial in comparison to the two other groups, particularly since participants had to complete AI 101 and perform well enough to pass several stages of vetting to obtain a SamaDC job.

In Table 1.6, we look at the amount of time it took for individuals in the training + job group to begin their SamaDC jobs. Within the individuals surveyed at baseline, some were surveyed again at midline and endline, some only at midline, and some only at endline. We find that it took about 3 to 4 months for individuals to receive their jobs, depending on whether we look at those surveyed at midline and baseline, those surveyed at endline and baseline, or those surveyed in all three rounds.

Effects on Income and Employment

We start by analyzing the effects on various outcomes related to income and employment. For clarity of exposition we will discuss the results for Group 3, the Training + Job group, first.

Training + Job Group

Table 2.1 reports the key outcomes at midline and endline. The odd number columns

report the regression above for various outcomes, while the even numbered columns include additional controls for whether the individual had ever been a SamaDC employee at the time of the survey and whether they were currently a SamaDC employee at the time of the survey.

At midline, approximately 16 months after the baseline survey, those in the training + job group significantly increased their monthly earnings, were less likely to be unemployed and worked more hours. By endline, these effects further grew in magnitude, with significant increases in the inverse hyperbolic sine (IHS) of earnings as well as for those other categories. In terms of magnitudes, individuals have earnings that are 37 percent higher than the control group, are 10 percent less likely to be unemployed and are working 22 percent more hours a week.

The odd columns clearly show that these income and employment effects are driven by those who worked at SamaDC (consistent with the lack of positive impacts on the training-only group that we will discuss next), with the largest effects for current employees. However, substantial income effects are present even for those who previously worked at SamaDC but no longer do, suggesting that the experience gained working at SamaDC had returns elsewhere and that this group could be more selective in the search for other opportunities.

One key question in the efficacy of AI 101 and SamaDC job matching is whether these interventions have differential effects on men and women. For this question, we look to Tables 2.3-2.6. Within the training + job group, columns 1 through 3 show a positive effect on earnings and negative effect on unemployment that is noticeably more pronounced for women both at midline and endline. Looking at column 1 for endline in particular, we see that amongst men, the training + job group have earnings that are 26 percent higher than the control group, while amongst women, the training + job group women have earnings that are 60 percent higher than the control group.

Given the overlap of the end of the study period with the COVID-19 pandemic, naturally, the question arises as to whether or not these interventions proved particularly beneficial

in the COVID-19 era. Though there is only a month or so of data from endline surveys collected after Kenya's first country-wide COVID restrictions were implemented on March 15th, we can look for heterogeneity for the 21 percent of the sample interviewed after this date. In Table 3.1, columns 3 and 6 shows that the control group income fell dramatically during this period while the training + job group income did not. Thus, we have reasonably strong evidence that the Sama treatment negated any negative income effects from COVID. This is not surprising given that the SamaDC jobs that the training + job group received continued unabated during this period and work was able to be done from home.

Reassuringly, however, we find that the effect of the intervention is by no means purely driven by COVID times. Indeed, in Table 3.2, we look at key outcomes among the subsample of the training + job group individuals surveyed prior to the beginning of the COVID period, finding that the direction and magnitude of the key coefficients does not change noticeably, barring the hours worked outcome shown in column 12, which does decline slightly from 6.03 hours above the control mean for the full sample to 3.84 hours above the control mean for the pre-COVID subsample.

Finally, the first two orientations at the start of the program suffered from some ambiguous messaging on whether or not a job opportunity at SamaDC was available to the training-only group. For robustness, we further exclude individuals that were in these two orientations. Table 3.3 shows that results are unaffected by these omissions.

Training Group

We find more nuanced results for the training group, the set of individuals that received only digital skills training. We first turn our attention to Table 2.1 where we see that, oddly, the training group experiences a non-monotonic effect in which they initially (at midline, on average 16 months after baseline) have worse outcomes than the control group. These effects are only significant in the case of the IHS earnings; however, monthly earnings, unemployment probability and hours worked all paint the same picture. Looking at Table

2.2, we see that these negative effects at midline dissipate by endline, on average 14 months later. In fact, while not significant, their outcomes appear a little better than the control group (for example, monthly earnings are 15 percent higher as shown in column 3).

In Tables 2.3-2.6, we look at gender heterogeneity in these results, finding that the effect at midline is seen amongst both men and women in the training group. However, looking at endline, we find that only the men recover, ultimately with results that parallel those of the training + job group (though with consistently smaller effects). Women, on the other hand, still see decreased earnings and hours worked with increased unemployment probability at the end of the study.

These results suggest that some of the training-only group subjects are still holding out for ICT employment at midline. By the time of the endline, the males at least have either found such a job or have settled for a non ICT job. We show supportive results for the latter explanation, that workers eventually settle for a non ICT job, when we investigate industry of employment next. At the end of this section, we explore and dismiss several alternative explanations for this pattern.

In relation to the pandemic, Table 3.1 shows that having participated in the AI 101 Training Program does not appear to make individuals more resilient to the economic damage caused by the COVID-19 pandemic. Again, the impact of the intervention on the main outcomes does not change when we limit our sample to pre-COVID surveys or exclude the first two orientations in Tables 3.2 and 3.3.

The gender heterogeneity described above is also not driven by the COVID-19 pandemic. Only 20 percent of surveys were conducted after the start of the pandemic. Among those women reporting being unemployed in the training group surveyed post covid, none of them reported having a job in the period just prior to covid (based on answers to question "For how many months have you been unemployed").

Reasons for Unemployment

In Tables 4.1-4.6, we dive deeper into unemployment, looking at the reasons individuals cite for their unemployment and discern differences in these reasons across men and women (the latter is explored later on). At midline, column 1 shows that the training + job group individuals are 8.3 percent less likely than the control group to report being unemployed due to an inability to find work as. This is the case at endline as well, when they are also significantly less likely to report being unemployed due to having lost their job or struggling with a health condition as shown in columns 2 and 4. For the training-only group there are no significant differences between treatment and control, although consistent with the hypothesis above, the training group is more likely to say they cannot find work at midline (an effect that flips at endline).

Effect on Industry of Employment

Training + Job Group

To better understand the nature of the jobs done by subjects in different treatment groups, we delve deeper both into the industry and job type of study participants in Tables 5.1 and 5.2.

In relation to industry, the first column of these tables shows that at both midline and endline, the training + job group individuals shifted from all industries other than ICT towards ICT. Since the strong possibility of a job in the ICT industry is part of the treatment, this result is unsurprising.

In relation to job type, columns 2 and above of these tables show that at midline, individuals in the training + job group worked less in self-employed jobs and more in salaried jobs than their control group counterparts (most notably with 10.4 percent fewer individuals working in casual self employment and 27.3 percent more individuals working in salaried

jobs). At endline, though this effect is slightly less pronounced, it is still significant.

Training Group

The training group is more interesting to evaluate. Recall that earnings were lower than control at midline before rising above control at endline (although the latter effect was driven exclusively by males). The industry and job type patterns shed further light on this unexpected result.

Looking at participation in the ICT industry in Tables 5.1 and 5.2, we see that the training group individuals, after receiving only the AI 101 training, left lower-skilled industries (particularly retail and transportation/storage, as shown in Appendix Tables 9.3 and 9.4) and shifted towards ICT, though the positive ICT shift is significant only at midline. By endline, however, this effect has disappeared. This is consistent with the story that many individuals in the training group are waiting for an ICT job at midline and are therefore unemployed; thus, those with jobs are those in the ICT industry and we see an uptick in ICT at midline. By endline, however, these individuals have likely stopped waiting for ICT jobs and have dispersed across the other industries, thereby decreasing the effect on ICT participation.

Looking at job type in Tables 5.1 and 5.2, we see that at midline and endline, individuals in the training group held fewer self-employed jobs and more salaried jobs than individuals in the control group. The effects were more pronounced at midline, when 7.5 percent fewer training group individuals held casual self employment jobs than control group individuals and 6.1 percent more training group individuals held salaried jobs than control group individuals.

Appendix Tables 9.5-9.8 show the industry heterogeneity by gender for the training group. We see a jump in ICT jobs for men at midline that disappears at endline, while there is not a significant increase in ICT jobs for women either at midline or at endline. This suggests that the men in the training group who were unemployed at midline, holding out for ICT jobs,

eventually found jobs outside of the ICT industry. No such pattern is seen for women. This gender discrepancy is consistent with the fact that women in the training group continued to experience depressed earnings at endline while men had significantly higher earnings than the control.

Happiness and Financial Security

Training + Job Group

In Tables 6.1 and 6.2, we study the effect of the treatments on a set of additional life outcomes. At both midline and endline, we find that the training + job group individuals experienced clear and substantial benefits across a wide range of outcomes beyond earnings and employment. Specifically, we find substantial increases in their life satisfaction, measured on the "ladder of life," wherein the bottom rung (value of 0) represents someone living their worst possible life and the top rung (value of 10) represents someone living their best possible life. The individuals see a substantial increase to their current life satisfaction by half a rung on the ladder of life, or 0.34 control group standard deviations. They also see an increase to their projected life satisfaction, defined as their best guess of life satisfaction five years in the future, by about a third of a rung, or 0.19 control group standard deviations. In addition they became less reliant on money transfers from both the government and relatives, with 7 percent fewer of these individuals receiving monthly transfers than the control group, with an average monthly transfer amount that is 1045 KSH less than that of the control group. In Table 6.2, it's evident that the training + job group individuals also see benefits in money management, as they are 10 percent more likely to have bank accounts, 11 percent more likely to have savings accounts, and have around 7724 more kenyan shillings saved than the control group. Together, these results paint a consistent story of strong gains for the training + job group.

Training Group

Turning to the training group, effects are much more muted. However, two findings are particularly intriguing. First, even though the training group had a dip in earnings at midline, they also had a decreased reliance on money transfers, with 8.2% fewer individuals receiving transfers in the training group than the control group and the training group individuals receiving about 612 KSH less on average in money transfers. One possible explanation for this apparent paradox could be that the training group individuals falsify their financial stability to relatives in order to avoid embarrassment. Alternatively, they could now feel more self reliant given their new skills. Second, though only borderline significant, we do see that the training group at endline has a lower projected life satisfaction, perhaps as a result of the difficulties they experienced finding ICT employment.

Understanding the Earnings Dip at Midline for training-only group

Now, we delve into the IHS earnings dip for the training-only group at midline that we have touched upon at several points above. We hypothesize that some of these workers are still holding out for ICT employment at midline before either finding such a job or settling for a non ICT job later. In support of this hypothesis, we saw that there was significantly higher participation in the ICT industry for men at midline than at endline, with male earnings rising above the control group only after unemployed males at midline had found non-ICT jobs at endline.

We now explore potential causes for these individuals holding out for a better job at midline. At baseline, individuals in the training group were told that if they waited about a year after partaking in the AI 101 training, they would become eligible for employment at SamaDC. Ultimately, however, as the duration of the experiment was extended, producing a backlog of potential employees, these workers remained ineligible. Thus, one explanation for the dip in earnings at midline is because these individuals were still holding out for a Sama job at that time, but by endline, had secured an alternative job.

If this was the case, we would expect that the longer the gap between baseline and midline, the more likely that, by the time of the midline, training group individuals would have given up on the prospect of obtaining a SamaDC job and sought out a different job instead. In this scenario, the earnings would be higher and unemployment would be lower amongst those who had a greater gap between us surveying them at baseline and midline. However, in Table 7.1, we see the opposite: those in the training-only group with gaps larger than 12 months (column 3) or 15 months (column 4) have even lower IHS earnings than those with shorter gaps. We also look at males and females separately in Appendix Tables 9.11 and 9.13, finding that this result holds for both groups.

Thus, the earnings dip at midline does not appear to be an artefact of our messaging. Instead, our leading hypothesis is that attending the AI 101 training heightened the aspirations of the attendees to attain a high-paying technical job, but did not provide them with the path to actually fulfill those aspirations. By endline those aspirations had faded and they took non-ICT jobs. This recovery at endline is primarily driven by men in our sample. In Tables 2.3-2.6, we saw that unemployment levels of men in the training-only group were 6 percent higher than the control group (though not significantly so) at midline but were 5 percent lower than the control group at endline. IHS earning are significantly higher for training-group men compared to the control, consistent with the training eventually providing these workers with higher incomes. In contrast, unemployment rates remained elevated at both midline and endline for women and, if anything, earnings remain lower at endline. This heterogeneity suggests that the effect on aspirations was more prolonged for women. To explore this further, we look at the reasons for unemployment: because some reasons have a degree of voluntariness, we can see if individuals opt into unemployment while waiting for a high quality job. Looking at Tables 4.3-4.6, we see that for men in the training group, between midline and endline, the coefficients on the "voluntary" reasons for unemployment shift from positive to negative. For women, this flip does not occur. This further bolsters the story that men choose to wait for a higher quality job at midline but eventually give in

to accepting non-ICT jobs.

How the prospect of a job affects effort to learn

Finally, we attempt to test the hypothesis that the effort and motivation of individuals in skill training sessions increases when there is a clear application for the use of those skills. Looking at Table 8.1, we compare the attendance of AI 101 training between the training group and the training + job group. Unlike Tables 1.4 to 1.5, here we use administrative data on attendance and only include the two treatment groups in the regressions. We find significant differences here, with AI 101 attendance 10.2 percentage points higher among training + job group individuals compared to the training-only group and AI 101 completion 13.9 percentage point higher. Further, we look at completion of quizzes (each quiz signifying the completion of a training module) to get a better sense of effort put into the training across groups. The sample size is small here due to the fact that tracking of quiz results began well into the study, in late 2017, by which point over half of the subjects in the training and training + job groups had already completed their training. Though the sample size is small, we do see that the training + job group completes many more quizzes than the training-only group (by 18.1 percentage points), potentially signifying that individuals put more effort into the training when there is the possibility of a job using those skills. Columns 4 through 6 show average quiz scores for the baseline quiz, all quizzes, and the cumulative quiz (the final quiz that covered the material across all modules). Again, sample sizes are quite small, particularly with the cumulative quiz (sample size of only 82 individuals). But, that being said, there are no significant differences between the training and the training+job group in terms of quiz score, but overall, the training + job group does slightly better on the baseline quiz, slightly worse on average across all the module quizzes, and slightly better on the final cumulative quiz.

1 Summary of Sample

Table 1.1 Timeline of Surveys

	Mean	Median	95%
Days between Baseline and Midline	490.21	533.00	646.34
Months between Baseline and Midline	16.34	17.77	21.54
Days between Midline and Endline	408.03	395.71	503.13
Months between Midline and Endline	13.60	13.19	16.77

Table shows mean, median, and 95th percentile, across all study participants, of the amount of time (either in days or months) between the baseline and midline surveys (top two rows) or between the midline and endline surveys (bottom two rows).

Table 1.2 Descriptive Statistics

	Control mean	Training v Control	Training + Job v Control	Fstat	Obs
Earnings	3686.17	-6.23	-44.48	0.00	987
	(6663.38)	(668.60)	(562.17)	(1.00)	
IHS Earnings	3.31	0.01	-0.09	0.04	987
	(4.64)	(0.43)	(0.36)	(0.96)	
Unemployed	0.61	-0.02	0.03	0.67	987
	(0.49)	(0.050)	(0.04)	(0.51)	
Hours Worked	11.66	-4.80	-6.06	10.55	987
	(21.78)	(1.61)	(1.35)	(0.00)	
Life Satisfaction	4.26	-0.17	0.12	1.81	987
	(1.76)	(0.16)	(0.13)	(0.16)	
Projected Life Satisfaction	2.14	-0.08	-0.22	5.59	987
	(0.82)	(0.08)	(0.07)	(0.00)	
Currently Studying	0.13	-0.01	-0.05	2.21	987
	(0.34)	(0.03)	(0.03)	(0.11)	
Has Bank Acount	0.70	0.00	0.02	0.14	987
	(0.46)	(0.04)	(0.04)	(0.87)	
Has Savings	0.45	0.04	-0.02	0.72	987
	(0.50)	(0.05)	(0.04)	(0.49)	
Savings	10304.71	1350.09	-1699.20	0.45	987
	(30233.81)	(3501.58)	(2958.16)	(0.63)	
IHS Savings	8.60	0.06	0.09	0.06	987
	(2.19)	(0.30)	(0.25)	(0.94)	
Household Earnings	7125.69	-817.86	-696.70	0.21	987
	(17921.30)	(1491.74)	(1254.27)	(0.81)	
Male	0.49	0.06	-0.01	1.25	987
	(0.50)	(0.05)	(0.04)	(0.29)	
Age	24.00	-0.56	-0.19	1.47	987
	(3.75)	(0.33)	(0.27)	(0.23)	
Married	0.15	-0.04	-0.06	2.74	987
	(0.36)	(0.03)	(0.03)	(0.06)	
Children	0.34	-0.05	-0.08	1.58	987
	(0.64)	(0.05)	(0.05)	(0.21)	

Table shows baseline balance across the control group and two treatment arms. The first column reports the control group mean for each characteristis. The second and third columns are a regression of the variable in the row header regressed on a constant and two dummies for each of the treatment groups. Fixed effects for the year and month of enrollment are included. The F-Stat is for a test of the covariate imbalance across the three groups.

Table 1.3 Attrition

	(1) Midline Attrition	(2) Endline Attrition
Training	0.007	-0.055*
	(0.033)	(0.030)
Training + Job	-0.291***	0.016
	(0.028)	(0.026)
Control Mean	0.906	0.845
Observations	1218	1204
Groups 2 & 3 p-val	0.000	0.007

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.4 Midline Takeup

	Attended SDB	Completed SDB	Ever SamaDC Employee	Current SamaDC Employee
	(1)	(2)	(3)	(4)
Training	0.771***	0.685***	0.00705	-0.000359
	(0.0274)	(0.0316)	(0.0303)	(0.0293)
Training + Job	0.854***	0.801***	0.502***	0.366***
	(0.0246)	(0.0283)	(0.0271)	(0.0263)
Control Mean	0.06	0.05	0.01	0.01
Observations	921	921	921	921
Groups 2 & 3 p-val	0.002	0.000	0.000	0.000
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.5 Endline Takeup

	Attended SDB	Completed SDB	Ever SamaDC Employee	Current SamaDC Employee
	(1)	(2)	(3)	(4)
Training	0.747***	0.664***	0.00361	-0.00844
	(0.0285)	(0.0337)	(0.0352)	(0.0355)
Training + Job	0.855***	0.805***	0.500***	0.370***
	(0.0243)	(0.0287)	(0.0300)	(0.0302)
Control Mean	0.08	0.06	0.02	0.02
Observations	987	987	987	987
Groups 2 & 3 p-val	0.000	0.000	0.000	0.000
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.6 Timeline of Job Receipt

	Mean	Median	95%
Midline Group: Days between Training End and Job Receipt	113.60	111	216
Midline Group: Months between Training End and Job Receipt	3.79	3.7	7.2
Endline Group: Days between Training End and Job Receipt	86.9	62	202
Endline Group: Months between Training End and Job Receipt	2.90	2.07	6.73
Overlap Group: Days between Training End and Job Receipt	112.06	110.5	226
Overlap Group: Months between Training End and Job Receipt	3.74	3.68	7.53

Table shows mean, median, and 95th percentile of the amount of time (either in days or months) between the last day of SDB training and the day the SamaDC job was received. The top two rows show these statistics for the group 2 individuals surveyed at baseline and midline; rows 3-4 show these statistics for the group 2 individuals surveyed at baseline and endline; and rows 5-6 show these statistics for the group 2 individuals surveyed at baseline, midline, and endline. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2 Key Outcomes

Table 2.1 Midline Key Outomces

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-596.2	-674.5	-0.758**	-0.795**	0.0626	0.0673*	-1.361	-1.498
	(1143.2)	(1123.0)	(0.368)	(0.358)	(0.0418)	(0.0405)	(2.172)	(2.107)
Training + Job	3485.1***	-95.36	0.263	-1.022***	-0.0979***	0.0596	5.772***	-2.442
	(1024.0)	(1203.4)	(0.330)	(0.383)	(0.0375)	(0.0434)	(1.948)	(2.248)
T+J * Ever SamaDC Employee		3479.2*		0.462		-0.0691		6.305*
		(1945.8)		(0.620)		(0.0702)		(3.632)
T+J * Current SamaDC Employee		4697.2**		2.749***		-0.321***		13.28***
		(1990.2)		(0.634)		(0.0717)		(3.708)
Control Mean	11716.25	11716.25	8.08	8.08	0.35	0.35	27.76	27.76
Observations	921	921	921	921	921	921	920	920
Groups 2 & 3 p-val	0.000	0.646	0.004	0.570	0.000	0.864	0.001	0.687
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.2 Endline Key Outcomes

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	2033.3	1868.1	0.306	0.253	-0.00799	-0.00321	0.682	0.602
	(1744.2)	(1715.4)	(0.445)	(0.424)	(0.0405)	(0.0391)	(2.280)	(2.194)
Training + Job	4973.2***	-77.39	1.402***	-0.348	-0.104***	0.0306	6.032***	-1.764
	(1492.2)	(1710.6)	(0.381)	(0.423)	(0.0345)	(0.0385)	(1.957)	(2.153)
T+J * Ever SamaDC Employee		8271.3***		0.851		-0.0500		4.185
		(2384.2)		(0.590)		(0.0537)		(3.007)
T+J * Current SamaDC Employee		1512.3		3.189***		-0.276***		14.66***
		(2355.3)		(0.583)		(0.0532)		(2.986)
Control Mean	13440.25	13440.25	6.99	6.99	0.30	0.30	27.32	27.32
Observations	935	935	935	935	987	987	964	964
Groups 2 & 3 p-val	0.081	0.296	0.011	0.192	0.014	0.420	0.014	0.311
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.3 Midline Males Key Outcomes

	Monthly Earnings (KSH)	IHS Earnings	Unemployed	Hours Worked
	(1)	(2)	(3)	(4)
Training	-762.620	-0.574	0.063	-1.924
	(1733.312)	(0.511)	(0.058)	(3.137)
Training + Job	2885.532*	0.167	-0.066	7.448**
	(1612.506)	(0.474)	(0.054)	(2.923)
Control Mean	13746.213	8.600	0.268	31.220
Observations	437	437	437	436
Groups 2 & 3 p-val	0.026	0.124	0.020	0.001
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.4 Endline Males Key Outcomes

	Monthly Earnings (KSH)	IHS Earnings	Unemployed	Hours Worked
	(1)	(2)	(3)	(4)
Training	3753.282	0.974*	-0.052	1.704
	(2919.430)	(0.573)	(0.051)	(3.187)
Training + Job	4398.298*	1.560***	-0.101**	5.974**
	(2623.945)	(0.514)	(0.045)	(2.870)
Control Mean	17218.919	7.657	0.236	31.009
Observations	459	459	487	470
Groups 2 & 3 p-val	0.818	0.283	0.316	0.158
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.5 Midline Females Key Outcomes

	Monthly Earnings (KSH)	IHS Earnings	Unemployed	Hours Worked
	(1)	(2)	(3)	(4)
Training	-648.212	-1.050**	0.086	-1.052
	(1538.855)	(0.534)	(0.060)	(3.048)
Training + Job	3919.533***	0.256	-0.117**	5.131*
	(1370.913)	(0.476)	(0.053)	(2.716)
Control Mean	9997.553	7.642	0.413	24.833
Observations	484	484	484	484
Groups 2 & 3 p-val	0.003	0.013	0.001	0.039
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.6 Endline Females Key Outcomes

	Monthly Earnings (KSH)	IHS Earnings	Unemployed	Hours Worked
	(1)	(2)	(3)	(4)
Training	-100.325	-0.525	0.054	-2.054
	(1875.361)	(0.686)	(0.064)	(3.271)
Training + Job	6145.753***	1.331**	-0.112**	5.422**
	(1530.074)	(0.559)	(0.052)	(2.684)
Control Mean	10213.846	6.414	0.354	23.977
Observations	476	476	500	494
Groups 2 & 3 p-val	0.001	0.006	0.007	0.018
Enrollment FE	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3 Key Outcomes Checks

Table 3.1 Endline Key Outcomes: COVID Check

	Monthly Earnings (KSH)			IHS Earnings			Unemployed			Hours Worked		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Training	2033.3	1868.1	2370.0	0.306	0.253	0.335	-0.00799	-0.00321	-0.0147	0.682	0.602	-0.0383
	(1744.2)	(1715.4)	(1930.0)	(0.445)	(0.424)	(0.493)	(0.0405)	(0.0391)	(0.0449)	(2.280)	(2.194)	(2.531)
Training + Job	4973.2***	-77.39	3842.9**	1.402***	-0.348	1.142***	-0.104***	0.0306	-0.0845**	6.032***	-1.764	3.837*
	(1492.2)	(1710.6)	(1645.5)	(0.381)	(0.423)	(0.420)	(0.0345)	(0.0385)	(0.0384)	(1.957)	(2.153)	(2.185)
T+J * Ever SamaDC Employee		8271.3***			0.851			-0.0500			4.185	
		(2384.2)			(0.590)			(0.0537)			(3.007)	
T+J * Current SamaDC Employee		1512.3			3.189***			-0.276***			14.66***	
		(2355.3)			(0.583)			(0.0532)			(2.986)	
COVID Period			-4851.5			-1.183			0.0892			-9.035*
			(4014.2)			(1.025)			(0.0903)			(5.010)
Training * COVID Period			-388.1			0.147			0.0186			4.542
			(4319.4)			(1.102)			(0.0993)			(5.494)
T+J * COVID Period			5907.0			1.376			-0.0967			10.50**
			(3779.8)			(0.966)			(0.0850)			(4.724)
Control Mean	13440.25	13440.25	13440.25	6.99	6.99	6.99	0.30	0.30	0.30	27.32	27.32	27.32
Observations	935	935	935	935	935	935	987	987	987	964	964	964
Groups 2 & 3 p-val	0.081	0.296	0.434	0.011	0.192	0.093	0.014	0.420	0.110	0.014	0.311	0.112
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.2 Endline Key Outcomes: Pre-COVID

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	2556.6	2452.3	0.429	0.406	-0.0210	-0.0185	0.498	0.602
	(2094.3)	(2062.4)	(0.508)	(0.483)	(0.0465)	(0.0447)	(2.631)	(2.492)
Training + Job	3768.7**	-1399.9	1.125***	-0.674	-0.0841**	0.0543	3.530	-5.454**
	(1789.2)	(2045.1)	(0.434)	(0.478)	(0.0397)	(0.0439)	(2.269)	(2.445)
T+J * Ever SamaDC Employee		8763.7***		0.840		-0.0453		4.256
		(2827.4)		(0.662)		(0.0605)		(3.349)
T+J * Current SamaDC Employee		1405.8		3.436***		-0.303***		18.29***
		(2791.0)		(0.653)		(0.0600)		(3.329)
Control Mean	14121.00	14121.00	7.01	7.01	0.30	0.30	28.04	28.04
Observations	748	748	748	748	781	781	761	761
Groups 2 & 3 p-val	0.559	0.092	0.167	0.043	0.169	0.137	0.239	0.025
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where sample is restricted to individuals surveyed before COVID times. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3 Endline Key Outcomes: Pre-COVID & Post-First 2 Months

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	2702.8	2552.0	0.559	0.522	-0.0320	-0.0287	1.901	1.880
	(2223.5)	(2189.6)	(0.532)	(0.505)	(0.0488)	(0.0469)	(2.720)	(2.563)
Training + Job	3661.6*	-1548.4	1.137**	-0.694	-0.0836*	0.0606	3.659	-5.787**
	(1971.8)	(2224.3)	(0.472)	(0.513)	(0.0432)	(0.0472)	(2.433)	(2.584)
T+J * Ever SamaDC Employee		9524.9***		0.973		-0.0618		5.509
		(3055.4)		(0.705)		(0.0643)		(3.501)
T+J * Current SamaDC Employee		973.8		3.495***		-0.306***		18.49***
		(3022.1)		(0.697)		(0.0639)		(3.485)
Control Mean	13988.83	13988.83	6.85	6.85	0.32	0.32	26.83	26.83
Observations	687	687	687	687	717	717	699	699
Groups 2 & 3 p-val	0.670	0.095	0.283	0.032	0.293	0.086	0.518	0.007
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where sample is restricted to individuals not in either of the two first orientation waves and surveyed before COVID times. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4 Reasons for Unemployment

Table 4.1 Midline Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.043	-0.013	0.017	0.000	0.000	0.004	0.006
	(0.039)	(0.009)	(0.014)	(0.007)	(.)	(0.007)	(0.019)
Training + Job	-0.083**	-0.009	0.001	0.002	0.000	0.007	-0.016
	(0.034)	(0.008)	(0.013)	(0.006)	(.)	(0.006)	(0.017)
Control Mean	0.78	0.04	0.05	0.01	0.00	0.00	0.14
Observations	921	921	921	921	921	921	921
Groups 2 & 3 p-val	0.001	0.670	0.245	0.730	.	0.711	0.229
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.2 Endline Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	-0.022	-0.009	0.004	-0.008	-0.000	0.021	0.004
	(0.036)	(0.008)	(0.013)	(0.007)	(.)	(0.017)	(0.011)
Training + Job	-0.074**	-0.017***	-0.009	-0.013**	-0.000	0.009	0.002
	(0.030)	(0.006)	(0.011)	(0.006)	(.)	(0.015)	(0.009)
Control Mean	0.72	0.05	0.07	0.04	0.00	0.08	0.04
Observations	987	987	987	987	987	987	987
Groups 2 & 3 p-val	0.123	0.321	0.306	0.500	.	0.462	0.791
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for column of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.3 Midline Males Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.031	-0.002	0.000	0.018**	0.000	0.005	0.011
	(0.056)	(0.011)	(.)	(0.009)	(.)	(0.011)	(0.026)
Training + Job	-0.062	-0.014	0.000	0.006	0.000	0.008	-0.012
	(0.051)	(0.010)	(.)	(0.008)	(.)	(0.010)	(0.024)
Control Mean	0.78	0.04	0.05	0.01	0.00	0.00	0.14
Observations	437	437	437	437	437	437	437
Groups 2 & 3 p-val	0.074	0.247	.	0.169	.	0.766	0.347
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to males. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.4 Endline Males Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	-0.060	-0.006	0.000	-0.020*	0.000	0.014	0.013
	(0.044)	(0.009)	(.)	(0.011)	(.)	(0.022)	(0.017)
Training + Job	-0.075*	-0.010	0.000	-0.015	0.000	0.011	-0.009
	(0.039)	(0.008)	(.)	(0.009)	(.)	(0.019)	(0.015)
Control Mean	0.72	0.05	0.07	0.04	0.00	0.08	0.04
Observations	487	487	487	487	487	487	487
Groups 2 & 3 p-val	0.731	0.636	.	0.662	.	0.866	0.185
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to males. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.5 Midline Females Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.071	-0.023*	0.043	-0.011	0.000	0.003	-0.001
	(0.055)	(0.014)	(0.027)	(0.010)	(.)	(0.008)	(0.028)
Training + Job	-0.086*	-0.011	-0.001	0.005	0.000	0.002	-0.023
	(0.049)	(0.013)	(0.024)	(0.009)	(.)	(0.007)	(0.025)
Control Mean	0.78	0.04	0.05	0.01	0.00	0.00	0.14
Observations	484	484	484	484	484	484	484
Groups 2 & 3 p-val	0.004	0.354	0.099	0.107	.	0.941	0.415
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to females. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.6 Endline Females Reasons for Unemployment

	Cannot Find Work	Lost Job, Not Looking Yet	Caretaking	Health Condition	Not Interested In Working	In School	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.019	-0.012	0.015	0.003	-0.000	0.032	-0.003
	(0.057)	(0.013)	(0.027)	(0.009)	(.)	(0.028)	(0.013)
Training + Job	-0.085*	-0.024**	-0.022	-0.010	-0.000	0.005	0.014
	(0.047)	(0.011)	(0.022)	(0.007)	(.)	(0.023)	(0.011)
Control Mean	0.72	0.05	0.07	0.04	0.00	0.08	0.04
Observations	500	500	500	500	500	500	500
Groups 2 & 3 p-val	0.059	0.317	0.156	0.124	.	0.305	0.178
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to females. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Industries & Job Types

Table 5.1 Midline Job Industry & Type

	ICT Industry	Salaried	SE (Serious)	SE (Casual)	SE (Business)
	(1)	(2)	(3)	(4)	(5)
Training	0.097***	0.061	-0.001	-0.075***	-0.035**
	(0.037)	(0.043)	(0.022)	(0.028)	(0.016)
Training + Job	0.423***	0.273***	-0.035*	-0.104***	-0.031**
	(0.033)	(0.039)	(0.020)	(0.025)	(0.014)
Control Mean	0.072	0.336	0.076	0.184	0.054
Observations	921	921	921	921	921
Groups 2 & 3 p-val	0.000	0.000	0.115	0.286	0.793
Enrollment FE	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}

Table 5.2 Endline Job Industry & Type

	ICT Industry	Salaried	SE (Serious)	SE (Casual)	SE (Business)
	(1)	(2)	(3)	(4)	(5)
Training	0.050	0.074	-0.026	-0.014	-0.027
	(0.041)	(0.046)	(0.028)	(0.028)	(0.023)
Training + Job	0.381***	0.318***	-0.093***	-0.065***	-0.061***
	(0.035)	(0.039)	(0.024)	(0.024)	(0.020)
Control Mean	0.107	0.316	0.134	0.146	0.103
Observations	987	987	987	987	987
Groups 2 & 3 p-val	0.000	0.000	0.013	0.062	0.126
Enrollment FE	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}

6 Sub-Outcomes

Table 6.1 Midline Sub-Outcomes

	Life Satisfaction	Projected Life Satisfaction	Currently Studying	Receives Money Transfers	Monthly Money Transfers
	(1)	(2)	(3)	(4)	(5)
Training	0.192	0.144	-0.043	-0.082**	-612.028*
	(0.142)	(0.121)	(0.032)	(0.037)	(327.537)
Training + Job	0.686***	0.436***	-0.029	-0.176***	-1200.746***
	(0.127)	(0.108)	(0.029)	(0.034)	(292.797)
Control Mean	4.451	8.089	0.170	0.318	1743.633
Observations	919	901	921	921	911
Groups 2 & 3 p-val	0.000	0.012	0.646	0.009	0.062
Enrollment FE	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	No	No

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.2 Endline Sub-Outcomes

	Life Satisfaction	Projected Life Satisfaction	Currently Studying	Receives Money Transfers	Monthly Money Transfers	Has Bank Acount	Has Savings	Savings	IHS Savings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Training	0.065	-0.237*	0.002	-0.011	-727.779*	0.033	-0.065	776.610	-1.715**
	(0.139)	(0.123)	(0.032)	(0.034)	(436.787)	(0.028)	(0.047)	(4922.390)	(0.837)
Training + Job	0.519***	0.290***	0.035	-0.070**	-1045.272***	0.101***	0.109***	7724.176*	-0.022
	(0.118)	(0.105)	(0.027)	(0.029)	(372.058)	(0.024)	(0.040)	(4254.173)	(0.722)
Control Mean	4.664	8.064	0.119	0.186	1595.635	0.836	0.494	13131.019	4.059
Observations	984	975	987	987	981	979	977	342	342
Groups 2 & 3 p-val	0.001	0.000	0.288	0.075	0.451	0.011	0.000	0.146	0.037
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7 Key Outcomes Delay

Table 7.1 Midline Key Outcomes: Waiting for a Sama Job

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-896.1	405.2	-0.428	-0.245	0.130*	0.0673	-3.038	-1.270
	(2072.7)	(1825.6)	(0.668)	(0.588)	(0.0757)	(0.0668)	(3.923)	(3.465)
Training + Job	3495.0	3162.3*	0.573	0.571	-0.0746	-0.0922	8.277**	5.622*
	(2217.9)	(1722.4)	(0.714)	(0.554)	(0.0810)	(0.0630)	(4.198)	(3.283)
Over 12mo. BM gap	2215.1		0.779		-0.109		9.235*	
	(2860.0)		(0.921)		(0.105)		(5.392)	
Training * 12mo gap	367.1		-0.487		-0.0926		2.218	
	(2487.2)		(0.801)		(0.0909)		(4.691)	
T+J * 12mo gap	50.90		-0.400		-0.0366		-2.812	
	(2499.6)		(0.805)		(0.0913)		(4.712)	
Over 15mo. BM gap		-3370.1		-0.973		0.176		-3.671
		(3127.5)		(1.007)		(0.115)		(5.927)
Training * 15mo gap		-1498.6		-0.823		-0.0128		-0.0516
		(2346.8)		(0.755)		(0.0859)		(4.440)
T+J * 15mo gap		495.6		-0.475		-0.0159		0.339
		(2154.3)		(0.694)		(0.0788)		(4.081)
Control Mean	11716.25	11716.25	8.08	8.08	0.35	0.35	27.76	27.76
Observations	921	921	921	921	921	921	920	920
Groups 2 & 3 p-val	0.054	0.161	0.173	0.197	0.014	0.027	0.008	0.064
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Training Effort

Table 8.1 SDB Training Effort

	Attended SDB	Completed SDB	% Quizzes Completed	Baseline Quiz Score	Quiz Score Average	Cumulative Quiz Score
	(1)	(2)	(3)	(4)	(5)	(6)
Training + Job	0.102***	0.139***	18.135***	3.495	-3.075	2.123
	(0.030)	(0.037)	(6.614)	(3.764)	(2.431)	(3.688)
Training Mean	1.000	1.000	64.462	65.476	41.821	89.722
Observations	734	734	262	234	242	98
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Fixed effects for the year and month of enrollment are included where indicated. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

9 Appendix

Table 9.1 Midline Key Outcomes

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-596.2		-0.758**		0.0626		-1.361	
	(1143.2)		(0.368)		(0.0418)		(2.172)	
Training + Job	3485.1***		0.263		-0.0979***		5.772***	
	(1024.0)		(0.330)		(0.0375)		(1.948)	
Completed SDB		-940.2		-1.131**		0.0952		-2.040
		(1663.8)		(0.536)		(0.0604)		(3.146)
Ever Sama Employee		8445.6***		2.330***		-0.347***		14.89***
		(2435.7)		(0.787)		(0.0888)		(4.554)
Control Mean	11716.25	11716.25	8.08	8.08	0.35	0.35	27.76	27.76
Observations	921	920	921	920	921	920	920	919
p-val	0.000	0.013	0.004	0.005	0.000	0.001	0.001	0.017
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Odd columns in the table show results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Even columns in the table show results of an IV regression of variable in column header on dummies for whether the individual received SDB training and whether the individual ever worked at SamaDC, instrumenting with the treatment group dummies. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value for the odd columns tests the difference between the training group and the training + job group for the variable in the corresponding column. The p-value for the even columns tests the difference between the group that completed the SDB training and the group that was ever a sama employee for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.2 Endline Key Outcomes

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	2033.3		0.306		-0.00799		0.682	
	(1744.2)		(0.445)		(0.0405)		(2.280)	
Training + Job	4973.2***		1.402***		-0.104***		6.032***	
	(1492.2)		(0.381)		(0.0345)		(1.957)	
Completed SDB		3061.8		0.460		-0.0108		1.200
		(2611.3)		(0.655)		(0.0601)		(3.423)
Ever Sama Employee		4932.2		2.028**		-0.192**		10.51**
		(3773.2)		(0.947)		(0.0888)		(4.945)
Control Mean	13440.25	13440.25	6.99	6.99	0.30	0.30	27.32	27.32
Observations	935	934	935	934	987	986	964	963
p-val	0.081	0.754	0.011	0.294	0.014	0.193	0.014	0.232
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Odd columns in the table show results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Even columns in the table show results of an IV regression of variable in column header on dummies for whether the individual received SDB training and whether the individual ever worked at SamaDC, instrumenting with the treatment group dummies. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value for the odd columns tests the difference between the training group and the training + job group for the variable in the corresponding column. The p-value for the even columns tests the difference between the group that completed the SDB training and the group that was ever a sama employee for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.3 Midline Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.097***	-0.002	0.027	-0.128***	-0.021	0.009	0.002	-0.010	-0.029	0.055
	(0.037)	(0.010)	(0.018)	(0.031)	(0.018)	(0.014)	(0.012)	(0.022)	(0.021)	(0.042)
Training + Job	0.423***	-0.001	-0.025	-0.164***	-0.043***	0.003	0.012	-0.062***	-0.043**	-0.099***
	(0.033)	(0.009)	(0.016)	(0.028)	(0.016)	(0.013)	(0.011)	(0.020)	(0.019)	(0.038)
Control Mean	0.072	0.011	0.047	0.249	0.065	0.018	0.014	0.090	0.087	0.347
Observations	921	921	921	921	921	921	921	921	921	921
Groups 2 & 3 p-val	0.000	0.904	0.002	0.223	0.218	0.660	0.385	0.013	0.500	0.000
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.4 Endline Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.050	-0.013	-0.003	-0.086**	0.002	0.007	0.010	0.002	0.019	-0.001
	(0.041)	(0.009)	(0.018)	(0.034)	(0.024)	(0.016)	(0.012)	(0.023)	(0.023)	(0.041)
Training + Job	0.381***	-0.018**	-0.044***	-0.083***	-0.062***	-0.005	-0.014	-0.021	-0.037*	-0.101***
	(0.035)	(0.007)	(0.015)	(0.029)	(0.021)	(0.014)	(0.010)	(0.019)	(0.019)	(0.035)
Control Mean	0.107	0.020	0.063	0.217	0.103	0.032	0.020	0.071	0.071	0.296
Observations	987	987	987	987	987	987	987	987	987	987
Groups 2 & 3 p-val	0.000	0.562	0.018	0.923	0.006	0.425	0.038	0.305	0.010	0.011
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.5 Midline Males Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.122**	0.016	0.035	-0.102**	-0.050**	0.005	0.012	-0.021	-0.058*	0.042
	(0.056)	(0.017)	(0.031)	(0.045)	(0.024)	(0.023)	(0.019)	(0.032)	(0.031)	(0.059)
Training + Job	0.462***	0.008	-0.048	-0.167***	-0.060***	0.012	0.012	-0.077**	-0.064**	-0.077
	(0.052)	(0.016)	(0.029)	(0.042)	(0.022)	(0.022)	(0.017)	(0.030)	(0.029)	(0.055)
Control Mean	0.07	0.01	0.05	0.25	0.06	0.02	0.01	0.09	0.09	0.35
Observations	437	437	437	437	437	437	437	437	437	437
Groups 2 & 3 p-val	0.000	0.618	0.006	0.127	0.664	0.774	0.990	0.069	0.844	0.033
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to males. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.6 Endline Males Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.033	-0.001	-0.007	-0.080	-0.023	0.005	0.007	0.059*	0.054	-0.051
	(0.059)	(0.012)	(0.028)	(0.051)	(0.031)	(0.019)	(0.017)	(0.032)	(0.034)	(0.052)
Training + Job	0.373***	-0.001	-0.057**	-0.044	-0.091***	-0.026	-0.037**	0.001	-0.010	-0.100**
	(0.053)	(0.011)	(0.025)	(0.045)	(0.027)	(0.017)	(0.015)	(0.028)	(0.030)	(0.046)
Control Mean	0.11	0.02	0.06	0.22	0.10	0.03	0.02	0.07	0.07	0.30
Observations	487	487	487	487	487	487	487	487	487	487
Groups 2 & 3 p-val	0.000	0.967	0.065	0.470	0.021	0.097	0.010	0.059	0.052	0.331
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to males. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.7 Midline Females Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.065	-0.021*	0.018	-0.152***	0.004	0.013	-0.011	0.006	-0.008	0.085
	(0.049)	(0.012)	(0.018)	(0.044)	(0.027)	(0.018)	(0.016)	(0.030)	(0.030)	(0.061)
Training + Job	0.376***	-0.002	-0.007	-0.161***	-0.032	-0.001	0.010	-0.037	-0.034	-0.112**
	(0.044)	(0.010)	(0.016)	(0.040)	(0.024)	(0.016)	(0.015)	(0.027)	(0.027)	(0.055)
Control Mean	0.07	0.01	0.05	0.25	0.06	0.02	0.01	0.09	0.09	0.35
Observations	484	484	484	484	484	484	484	484	484	484
Groups 2 & 3 p-val	0.000	0.102	0.163	0.825	0.168	0.417	0.200	0.149	0.372	0.001
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to females. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.8 Endline Females Primary IGA Industry

	(1) ICT	(2) 1-2	(3) 3-6	(4) 7-8	(5) 9-21	(6) 13-14	(7) 15	(8) 16-19	(9) 20	(10) Unemployed
Training	0.049	-0.026**	-0.008	-0.095**	0.042	0.013	0.009	-0.053	-0.023	0.070
	(0.058)	(0.013)	(0.023)	(0.048)	(0.039)	(0.027)	(0.018)	(0.034)	(0.030)	(0.064)
Training + Job	0.366***	-0.035***	-0.029	-0.121***	-0.030	0.016	0.010	-0.031	-0.040	-0.112**
	(0.047)	(0.011)	(0.019)	(0.039)	(0.032)	(0.022)	(0.015)	(0.028)	(0.025)	(0.052)
Control Mean	0.11	0.02	0.06	0.22	0.10	0.03	0.02	0.07	0.07	0.30
Observations	500	500	500	500	500	500	500	500	500	500
Groups 2 & 3 p-val	0.000	0.472	0.348	0.581	0.055	0.904	0.968	0.504	0.553	0.003
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to females. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Industry Categorization: ICT = {Business Process Outsourcing (BPO) & Online Freelancing & Information Technology}; 1,2 = {Agriculture, forestry, or fishing & Mining}; 3-6={Manufacturing & Electricity, gas, steam, or air conditioning supply & Water supply, sewerage, waste management and remediation activities & Construction}; 7,8 = {Wholesale or retail trade & Transportation or storage}; 9,21 = {Accommodation and food service & Domestic services}; 13,14 = {Financial and insurance services & Real Estate}; 15 = {Professional, scientific, and technical activities}; 16-19 = {Administrative and support services & Public administration, defense & Education}; 20 = {Arts, entertainment, and recreation}

Table 9.9 Midline Primary IGA Type

	Salaried	SE (Serious)	SE (Casual)	SE (Business)	Internship	Transfers	Unemployed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.061	-0.001	-0.075***	-0.035**	-0.005	0.000	0.055
	(0.043)	(0.022)	(0.028)	(0.016)	(0.004)	(.)	(0.042)
Training + Job	0.273***	-0.035*	-0.104***	-0.031**	-0.001	0.000	-0.102***
	(0.039)	(0.020)	(0.025)	(0.014)	(0.004)	(.)	(0.037)
Control Mean	0.336	0.076	0.184	0.054	0.004	0.000	0.347
Observations	921	921	921	921	921	921	921
Groups 2 & 3 p-val	0.000	0.115	0.286	0.793	0.356	.	0.000
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.10 Endline Primary IGA Type

	Salaried	SE (Serious)	SE (Casual)	SE (Business)	Internship	Transfers	Unemployed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Training	0.074	-0.026	-0.014	-0.027	-0.014	0.000	-0.010
	(0.046)	(0.028)	(0.028)	(0.023)	(0.009)	(.)	(0.040)
Training + Job	0.318***	-0.093***	-0.065***	-0.061***	-0.008	0.000	-0.104***
	(0.039)	(0.024)	(0.024)	(0.020)	(0.007)	(.)	(0.034)
Control Mean	0.316	0.134	0.146	0.103	0.016	0.000	0.296
Observations	987	987	987	987	987	987	987
Groups 2 & 3 p-val	0.000	0.013	0.062	0.126	0.497	.	0.016
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.11 Midline Males Key Outcomes: Waiting for a Sama Job

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-1861.6	-376.1	-0.166	0.0927	0.129	0.0993	-0.930	-1.313
	(3170.7)	(2811.9)	(0.932)	(0.824)	(0.107)	(0.0945)	(5.704)	(5.043)
Training + Job	4256.1	4333.7*	0.682	0.891	-0.0233	-0.0453	11.55**	9.132*
	(3203.5)	(2579.6)	(0.941)	(0.757)	(0.108)	(0.0865)	(5.757)	(4.648)
Over 12mo. BM gap	1607.4		0.570		-0.00919		11.21	
	(4503.0)		(1.323)		(0.152)		(8.026)	
Training * 12mo gap	1559.4		-0.582		-0.0938		-1.530	
	(3789.5)		(1.113)		(0.128)		(6.772)	
T+J * 12mo gap	-1703.7		-0.695		-0.0612		-5.145	
	(3719.8)		(1.092)		(0.125)		(6.635)	
Over 15mo. BM gap		-902.7		-1.020		0.289*		-8.994
		(4758.4)		(1.393)		(0.160)		(8.502)
Training * 15mo gap		-692.3		-1.056		-0.0759		-0.657
		(3599.9)		(1.054)		(0.121)		(6.427)
T+J * 15mo gap		-2226.3		-1.104		-0.0587		-1.999
		(3337.4)		(0.978)		(0.112)		(5.956)
Control Mean	13746.21	13746.21	8.60	8.60	0.27	0.27	31.22	31.22
Observations	437	437	437	437	437	437	436	436
Groups 2 & 3 p-val	0.056	0.102	0.365	0.342	0.156	0.133	0.028	0.042
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to males. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9.12 Midline Females Key Outcomes: Waiting for a Sama Job

	Monthly Earnings (KSH)		IHS Earnings		Unemployed		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Training	-403.8	807.2	-0.991	-0.722	0.182*	0.0652	-6.182	-1.595
	(2755.1)	(2408.0)	(0.956)	(0.838)	(0.107)	(0.0941)	(5.451)	(4.803)
Training + Job	949.2	592.3	-0.409	-0.333	-0.0327	-0.0796	2.215	1.653
	(3269.0)	(2468.8)	(1.135)	(0.859)	(0.127)	(0.0966)	(6.471)	(4.934)
Over 12mo. BM gap	1987.4		0.593		-0.166		6.563	
	(3727.4)		(1.294)		(0.145)		(7.370)	
Training * 12mo gap	-511.1		-0.124		-0.129		7.076	
	(3335.2)		(1.158)		(0.129)		(6.587)	
T+J * 12mo gap	3493.0		0.782		-0.109		3.970	
	(3598.5)		(1.249)		(0.140)		(7.107)	
Over 15mo. BM gap		-7310.7*		-1.795		0.149		-3.435
		(4291.7)		(1.494)		(0.168)		(8.561)
Training * 15mo gap		-2185.2		-0.499		0.0319		1.098
		(3133.8)		(1.090)		(0.123)		(6.239)
T+J * 15mo gap		4534.7		0.801		-0.0513		4.958
		(2989.1)		(1.040)		(0.117)		(5.959)
Control Mean	9997.55	9997.55	7.64	7.64	0.41	0.41	24.83	24.83
Observations	484	484	484	484	484	484	484	484
Groups 2 & 3 p-val	0.689	0.938	0.620	0.687	0.103	0.183	0.208	0.556
Enrollment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table shows results of regression of variable in column header regressed on a constant and two dummies for each of the two treatment groups, where the sample is restricted to females. Controls for baseline values of the outcome variable are included where indicated. Fixed effects for the year and month of enrollment and the year and month of the survey are included where indicated. The p-value tests the difference between the training group and the training + job group for the variable in the corresponding column. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.