Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda*

Livia Alfonsi Oriana Bandiera Vittorio Bassi
Robin Burgess Imran Rasul Munshi Sulaiman Anna Vitali†

December 2017

Abstract

We design a labor market experiment to compare demand-side and supply-side policies to tackle youth unemployment, a key issue in low-income countries. The experiment tracks 1700 workers and 1500 firms over four years to contrast the effects of offering workers vocational training (VT) to offering firms wage subsidies to train workers on-the-job (FT). Both treatments lead to skill accumulation but whilst VT workers learn sector-specific skills, FT workers learn more firm-specific skills. This is associated with higher employment rates for each type of worker but the effect is 50% larger for VT (21% vs 14%) and their total earnings increase by more (34% vs 20%). Structurally estimating a job ladder model reveals the mechanisms: VT workers receive higher rates of unemployment-to-job offers and higher rates of job-to-job offers. This greater labor market mobility stems from the certifiability and transferability of their skills, and causes the wage profiles of VT workers to diverge away from FT workers. Evidence from the firm-side of the experiment complements these findings: we find that some of the higher returns to VT are driven by workers matching to higher productivity firms. Our evidence shows both firms and workers are constrained in this setting and that subsidies to either side of the labor market would increase workers' employment and earnings. However, VT workers are better off than FT workers as the greater certifiability and transferability of their skills allows them to climb the job ladder more quickly. *JEL Classification: J2, M5.*

*We gratefully acknowledge financial support from the Mastercard Foundation, PEDL, the International Growth Centre and an anonymous donor. We thank Jerome Adda, Orazio Attanasio, Chris Blattman, Richard Blundell, Nick Cerkez, Thomas Dohmen, Dave Donaldson, Armin Falk, Marc Gurgand, John Hardman-Moore, Eliana La Ferrara, Joe Kaboski, Daniel Keniston, Thibaud Lamadon, Thomas Le Barbanchon, Ethan Ligon, Marco Manacorda, David McKenzie, Costas Meghir, Karen Macours, Dilip Mookherjee, Suresh Naidu, Derek Neal, Emily Nix, Jeffrey Nugent, Fabien Postel-Vinay, Barbara Petrongolo, Mark Rosenzweig, Johannes Schmieder, T.Paul Schultz, John Strauss, Miguel Urquiola, Gerard van der Berg, Eric Verhoogen, Alessandra Voena, Chris Woodruff, Fabrizio Zilibotti and numerous seminar participants for comments. All errors are our own.

†Alfonsi: BRAC, livia.alfonsi@berkeley.edu; Bandiera: LSE, o.bandiera@lse.ac.uk; Bassi: USC, vbassi@usc.edu; Burgess: LSE, r.burgess@lse.ac.uk; Rasul: UCL, i.rasul@ucl.ac.uk; Sulaiman: BRAC, Save the Children, munshi.slmn@gmail.com; Vitali: UCL, anna.vitali.16@ucl.ac.uk.

1 Introduction

Youth unemployment and underemployment has become a major challenge in the developing world. A growing mass of unskilled, young workers are failing to find work in manufacturing and service sectors consisting mainly of small-scale firms. This raises two questions about the labor markets in these countries. First, on the supply side, why don't workers acquire the skills that might help them secure these jobs? Second, on the demand side, what is stopping firms from hiring these workers? Answering these questions is important – how development proceeds in the coming decades will be largely determined by whether or not these young workers can be matched to good jobs.

Nowhere is the youth unemployment challenge more keenly felt than in East Africa where the majority of the population is aged below 25, and youth represent 60% of the unemployed. We study interventions to tackle youth unemployment in urban labor markets in Uganda, the country with the second lowest median age in the world where 60% of the population is aged below 20, and where formal sector youth employment rates are below 30% [UN 2017].

To do this we design a two-sided experiment involving both workers and firms which allows us to compare supply and demand side interventions – vocational training and firm-provided training through apprenticeships – commonly used across the world to help workers transition into the labor market.[1] On the supply side, subsidized vocational training might help workers overcome credit market imperfections which prevent them from investing in skills or imperfect knowledge regarding the return to different skills [Jensen 2010]. On the demand side, subsidized apprenticeships might help firms overcome credit market imperfections which prevent them from incurring the costs of hiring and training workers [de Mel *et al.* 2016, Hardy and McCasland 2017] or of learning about the ability and match quality of inexperienced workers [Farber and Gibbons 1996, Altonji and Pierret 2001, Pallais 2014].

As the vocational training and firm training interventions are fielded in the same setting we can directly compare their impacts on worker outcomes. This is our core contribution. The key distinction between them is that formally provided vocational training gives workers certified skills that can be used within any firm in the sector. In contrast, firms have limited incentives to provide skills that apprentices can use elsewhere. Vocational training could therefore have a larger impact on workers' welfare in the long run, as it enhances mobility between jobs. However, the effectiveness of vocational training relies on the existence of job opportunities for the vocationally trained. If these do not exist, only a policy that relaxes firms' hiring constraints will increase employment rates.

[1] Training programs like those evaluated here are popular interventions in both low- and high-income settings to assist workers with transition into employment. The World Bank, for example, has invested $9bn in 93 such programs between 2002 and 2012 [Blattman and Ralston 2015], and expanded training programs were the most common type of labor market policy implemented globally in response to the 2008 financial crisis [McKenzie and Robalino 2010].

Our research design provides evidence on each of these elements. We first use an assessment test to show that vocationally trained (VT) workers learn sector-specific skills whilst firm-trained (FT) workers learn more firm-specific skills. We then combine reduced form analysis with the structural estimation of a job ladder model to show that, in line with the skill difference and certifiability of their skills, VT workers move more easily between jobs and have higher steady state rates of employment and earnings. We then exploit the firm-side of the experiment to show that, in line with firms being constrained, less profitable firms are significantly more likely to take up the offer of using wage subsidies to hire new workers and train them. We use this as motivation to extend the job ladder model to allow for firm side heterogeneity, and then find that in steady state, VT workers also match with significantly more productive firms than FT workers. Overall, the experimental and structural results thus suggest that there are firms that face binding constraints on hiring but also that there are enough job opportunities for skilled workers so that VT is more effective at increasing worker welfare in the long run.

We track 1700 workers and 1500 firms over four years, after randomly allocating them to either control or one of two treatments: (i) vocational training (VT), six months of sector-specific in-class training offered to workers before they enter the labor market; (ii) firm training (FT), incentives to firms (via wage subsidies) to hire and train workers on-the-job for six months, replicating a classic in-firm apprenticeship structure.[2]

Our supply-side subjects are disadvantaged youth entering the labor market. Relative to their counterparts in high income countries these individuals have lower levels of human capital from the formal schooling system (due both to fewer years and lower schooling quality), and fewer options to raise their human capital through colleges, universities or other forms of tertiary education. Moreover, they face stiff labor market competition due to the majority of the population being aged 20 or below and there being high rates of unemployment in this group. On the demand side, we have small and medium size enterprises (SMEs) in both manufacturing and service sectors. These SMEs represent a core segment of firms in Uganda. Unlike the US, the majority of firms in Uganda have fewer than 10 employees, and they employ the majority of workers (see Figure A1). This skewed firm size distribution, which is common across developing countries, implies

[2] Earlier studies have often evaluated a combination of in-class vocational and on-the-job training, e.g. JTPA in the US and the YTS in the UK. In low-income settings, Card *et al.* [2011] and Attanasio *et al.* [2011] both evaluate the impacts of combining three months of vocational training followed by three month apprenticeships, in the Dominican Republic and Columbia respectively. Lalonde [1995] and Heckman *et al.* [1999] survey the earlier literature on job training programs. On-the-job training, internships and wage subsidies are all common policy approaches that have been used to target disadvantaged groups in the labor market [Layard and Nickell 1980, Katz 1998]. The justification for such approaches are typically twofold [Ham and Lalonde 1996, Katz 1998, Bell *et al.* 1999, Blundell 2001]: (i) to reduce employer screening costs; (ii) to provide workers some labor market experience that can have persistent impacts. On the first channel, Autor [2001] and Hardy and McCasland [2017] present evidence on the use of apprenticeships as screening technologies. On the second channel, Pallais [2014] shows via an experiment on an online jobs platform that providing employment to an inexperienced worker helps improve their later employment outcomes, emphasizing that early labor market experiences convey information on the workers skills rather than raising productivity.

that labor mobility tends to occur across as opposed to within firms.

Our analysis has three parts: (i) reduced form evidence on the returns to both forms of training from the worker's perspective; (ii) evidence on the mechanisms driving differences in these returns via the structural estimation of a dynamic job ladder search model; (iii) reduced form and structural evidence from the firm's perspective to assess which types of firms VT and FT workers end up matching to, and how firms incentivized through wage subsidies are affected by the newly hired workers.

The first part of the analysis shows that three years after the intervention: (i) only FT workers report being trained by the firm in their first employment spell; (ii) VT workers score higher than FT workers on a sector-specific skills test, and report having more skills transferable across firms, (iii) by endline, VT and FT workers end up performing different job tasks in firms within the same sector. These results are in line with the predictions of a long literature examining how firm incentives to train workers depends on labor market frictions [Becker 1964, Acemoglu and Pischke 1998, 1999]. VT workers seem to be more mobile because of their certifiable skills, and their skill set is tilted towards sector-specific rather than firm-specific human capital.

Relative to control workers, the employment rates of VT and FT workers significantly increase by 21% and 14% respectively. On the intensive margin, VT and FT workers work significantly more months relative to control and, as above, the impact is larger for VT workers. All this results in increased monthly earnings: VT and FT workers earn 34% and 20% more than the control group. The difference is driven by differences in total labor supply as both VT and FT workers experience similar increases in hourly wages over the control group of 40%. Productivity bounds analysis suggests both sets of workers become significantly more productive [Attanasio *et al.* 2011]. However, we find the majority of workers hired under wage subsidies remain employed at the same firm they were originally assigned to even after the subsidy expires. This suggests firms are constrained to begin with and hints at the key dynamic difference between training routes. Relative to young workers that enter the labor market via vocational training, those that transition through firm-provided apprenticeships do not move up the job ladder as fast.[3]

The second part of the analysis builds on this insight. Under the assumption that by endline (three years post-intervention), workers have reached their steady state wage trajectory, we structurally estimate a job ladder model of worker search. This pins down mechanisms driving the treatment effects on labor market outcomes. The channels investigated are: (i) unemployment-to-job offer arrival rates (UJ); (ii) job-to-job offer arrival rates (JJ). These channels are important because VT workers have certifiable skills from their training in vocational training institutes

[3]In relation to earlier studies that have evaluated a combination of vocational and on-the-job training, Card *et al.* [2011] find no evidence of employment impacts; Attanasio *et al.* [2011] find a 7% increase in employment rates for women and a 20% earnings increase, with impacts being sustained in the long run [Attanasio *et al.* 2017]. Galasso *et al.* [2004], Levinsohn *et al.* [2014] and Groh *et al.* [2016] evaluate wage subsidy interventions. McKenzie [2017] reviews the evidence on training and wage subsidy programs in low-income settings. We later discuss our findings in relation to this literature.

(VTIs). All else equal, this greater ability to signal their skills to potential employers, and the fact that their skills are more sector-specific, might make it easier for them to get back on the job ladder if they become unemployed, and to make job-to-job transitions as they receive outside wage offers.

The structural estimates indeed reveal that: (i) VT workers have higher steady state rates of UJ transitions than FT workers: if they fall off the job ladder into unemployment, they are more likely to get back on it; (ii) FT workers have very similar rates of UJ transition as the control group: their history of labor market attachment seems to count for little if they fall off the job ladder into unemployment; (iii) VT workers are also more likely than FT workers to make JJ transitions: such poaching is in line with them having certifiable skills and having relatively more sector- rather than firm-specific human capital than FT workers. Workers only make JJ transitions if the new job is at least as good as the current one. Hence over time, this dynamic causes the wage profile of VT workers to diverge away from FT workers as they climb up the job ladder.

Indeed, in steady state the earning returns are 34% for VT workers (relative to the control group), and 12% for FT workers. The estimated returns to FT are lower from the structural model than the reduced form evidence because the structural estimates account for the dynamic mechanism of lower labor market mobility. Finally, both training routes significantly reduce steady state unemployment rates for treated workers. Given the transition dynamics above, the supply-side policy of vocational training does so to a far greater extent than the demand-side policy of wage subsidies (15% versus 2.3%).

The third part of our analysis exploits the two-sided experimental design, to measure how training routes differ from the perspective of firms. In particular, we examine for each training route, the kind of firm individuals end up being initially matched to and eventually employed at. This allows us to understand whether the differential returns to VT and FT are partly due to the fact that workers are employed by more productive (and so better paying) firms. We find that firms that were offered and took up the wage subsidy have lower profits per worker than firms in the control group. Informed by this finding, we extend the structural model to allow for firm heterogeneity and to estimate the productivity of firms matched to *in steady state* under each treatment arm. This shows the firms VT workers end up employed at in steady state are far more productive than those firms that FT workers end up being employed at.

We also show that for the majority of workers hired under wage subsidies, retention with the firm lasts longer than the period of the wage subsidy itself, suggesting these firms were constrained to begin with. However, in the long run this leads to full employment displacement of other workers, so that firm size overall is no different relative to the control group of firms. However, their monthly profits increase in the longer run by 11% [$p = .032$], again suggesting firms that take on workers via wage subsidies were constrained. The profit increase corresponds to a magnitude over three times the value of the wage subsidy itself. This reflects the fact that they have more productive workers, but the additional surplus is not so large as to generate additional employment.

We combine program accounting costs with the estimated steady state benefits to derive the internal rate of return from each training route assuming a social discount rate of 5%. If we assume gains last over the working life, the IRR to vocational training is 21%, while those those to firm training are 10%. Both training routes pay for themselves in a decade.

What prevents workers and firms investing in these forms of training given such high rates of return? As described above, a key constraint is credit: the cost of vocational training, or self-financing apprenticeships, are both orders of magnitude higher than young workers average annual earnings at baseline. For firms, the effectiveness of wage subsidies suggests that they are also credit constrained in taking on young workers as in the long term such hires increase firm profits. However, our evidence suggests it is lower profitability/productivity firms that are induced to take-on (and retain) workers hired through this route.

We make four contributions to the literature on training program evaluation. First, we compare the returns to in-class vocational and on-the-job training in the same labor market context. We thus contribute to the literature by providing treatment effect estimates of the productivity and earnings impacts of firm provided training, something that has long been debated [Ham and Lalonde 1996, Katz 1998, Blundell *et al.* 1999], and compare them to the impacts of vocational training in the same setting.

Second, tracking workers and firms for several years allows us to structurally estimate a job ladder model of worker search that identifies the dynamic mechanisms driving the reduced form impacts. Our evidence supports the basic intuitions from economic theory that the mechanisms behind training routes differ. Vocationally trained workers have more sector-specific and certifiable skills, and hence are more mobile in labor markets than FT workers.[4]

Third, the two-sided experimental design allows us to compare the two training routes from the dual perspectives of workers *and* firms. The firm-side of the experiment shows that some of the higher returns to VT are driven by workers matching to higher productivity firms, that the long run net effect of wage subsidies on the number of firm employees is zero, and that there are long run profit impacts generated by firm-trained workers.

Finally, by using the two-sided experimental design to provide a complete economic comparison of a supply-side (vocational training) versus a demand-side intervention (wage subsidies to firms to hire and train workers on-the-job), we provide policy relevant insights on underlying causes of youth unemployment in low-income labor markets. Both worker-side and firm-side constraints are relevant factors driving youth unemployment in this setting. But, from a worker's perspective, tackling the issue by skilling youth using vocational training pre-labor market entry, is far more effective than incentivizing firms through wage subsidies to hire young labor market entrants.

Despite their popularity, the evidence base for training programs is thin. The meta-analyses

[4] Comparing experimental and structural estimates of the returns to training remains rare in the literature. Notable exceptions are Card and Hyslop [2005] and Hoffman and Burks [2017], although the latter compares to quasi-experimental estimates, not those based on random assignment.

of Card *et al.* [2015], Blattman and Ralston [2015] and McKenzie [2017] show relatively weak or short-lived impacts of training programs in low-income settings. We thus close our analysis by highlighting potential explanations for the impacts we document: the experimental design, the selection of workers, and the quality of the vocational training institutes worked with. On each dimension, we make suggestions for future research, and conclude by discussing the wider implications of our findings for the study of youth unemployment in low-income labor markets.[5]

The paper is organized as follows. Section 2 describes the setting, experimental design and data. Section 3 presents treatment effects on worker skills and labor market outcomes. Section 4 develops and structurally estimates the job ladder model to pin down the mechanisms via which training routes differ. Section 5 presents firm side impacts by training route, focusing on firm selection and productivity, employment and profits. Section 6 presents the IRR estimates and discussing external validity. Section 7 concludes. Robustness checks are in the Appendix.

2 Setting, Experimental Design and Data

Our study is a collaboration with the NGO BRAC, who implemented all treatments, and five vocational training institutes (VTIs). The VTI sector is well established in Uganda, with hundreds in operation. We worked with five of the most reputable. Each could offer standard six month training courses in the eight sectors we focus on: welding, motor mechanics, electrical wiring, construction, plumbing, hairdressing, tailoring and catering. These sectors constitute an important source of stable wage employment for young workers in Uganda: around a quarter of employed workers aged 18-25 work in one of these sectors.[6]

2.1 Setting

Workers The experiment is based on an oversubscription design, where we advertized an offer of potentially receiving six months of sector-specific vocational training at one of the five VTIs we collaborated with. As in other training interventions, the eligibility criteria meant that disadvantaged youth were targeted [Attanasio *et al.* 2011, Card *et al.* 2011]. We received 1714 eligible applicants: 44% were women, and applicants were on average aged 20.[7] Table A1 describes ap-

[5] Card *et al.* [2015] discuss over 800 estimates from 200 studies documenting impacts of active labor market programs. They find in contrast to training or wage subsidies, job search assistance (matching) has comparatively large short and long run impacts and are more pronounced for disadvantaged workers. Search costs have been shown to be relevant for workers and firms in labor markets [Abebe *et al.* 2016, Franklin 2016, Hardy and Macasland 2017, Pallais 2014, Bassi and Nansamba 2017].

[6] The VTIs we worked with were: (i) founded decades earlier; (ii) were mostly for-profit organizations; (iii) trained hundreds of workers with an average student-teacher ratio of 10; and (iv) in four VTIs, our worker sample shared classes with regular trainees. We derive the share of employed workers aged 18-25 working in these eight sectors using the Uganda National Household Survey from 2012/13.

[7] The program was advertized throughout Uganda using standard channels, and there was no requirement to participate in other BRAC programs to be eligible. The eligibility criteria were based on: (i) being aged 18-25;

plicant characteristics: the vast majority are out of school and have never received vocational training. Table 1 shows labor market outcomes at baseline for workers: the first row shows that unemployment rates are over 60% for these youth (Columns 2 and 3) with insecure casual work comprising the prevalent form of labor activity. Average monthly earnings are $6, corresponding to less than 10% of the Ugandan average. Hence these are not individuals that could self-finance either vocational training (that costs over $400), or apprenticeships in firms.[8]

The oversubscription design is informative of the impact of marginally expanding vocational training programs. Given Ugandan demographics, there is no shortage of the kind of marginal young labor market entrant that applied to our offer.

Firms To draw a sample of firms for the experiment, we first conducted a firm census in each of 17 urban labor markets. From this census we then selected firms: (i) operating in one of the eight sectors of interest; (ii) having between one and 15 employees (plus a firm owner). The first criteria restricts to manufacturing and service sectors in which we offered sector-specific vocational training, thus limiting skills mismatch in our study. The second restriction excludes micro-entrepreneurs and ensures we focus on SMEs that, as Figure A1 highlights, are central to employment in Uganda. We end up with a sample of 1538 SMEs, that in aggregate employ 4551 workers at baseline.

We also asked SMEs about constraints to expansion: prominent explanations included those related to credit and labor. For example, 65% of firms reported the terms of available finance limiting their growth suggesting they might be credit constrained, and 52% reported the inability to screen workers as a constraint. The offer of wage subsidies might then plausibly help relax demand-side constraints on SMEs related to hiring young labor market entrants in this setting.

Vocational Training Table 2 provides evidence on the supply of, and returns to, vocational training in this setting. It shows: (i) the share of workers employed at baseline in these firms that self-report having ever received vocational training from a VTI; (ii) the coefficient on a dummy for this self-report in an otherwise standard Mincerian wage regression of log wages. The first row pools across all sectors and shows that, as measured at baseline, 31% of all workers in our sample of SMEs have vocational training from some VTI. Vocational training is therefore a common route through which workers acquire skills in Uganda, and SME firm owners are familiar with recruiting

(ii) having completed at least (most) a P7 (S4) level of education (corresponding to 7-11 years); (iii) not being in full-time schooling; (iv) a poverty score, based on family size, assets owned, type of building lived in, village location, fuel used at home, number of household members attending school, monthly wage, and education level of the household head. Applicants were ranked on a 1-5 score on each dimension and a total score computed. A relative threshold score (varying by geography) was used to select eligibles.

[8] Table A2 compares our sample to those aged 18-25 in the Uganda National Household Survey from 2012/3. The program appears well targeted: our sample is worse off in terms of labor market outcomes at baseline, and that remains true when we compare to youth in the UNHS that report being active in the labor market. Irrespective of the precise sample, Table A2 highlights the seriousness of youth unemployment and underemployment in Uganda.

workers with such training in this setting. The Mincerian wage returns to vocational training are over 50%, and these results hold by sector: in each there is demand for these skills and there are potentially high returns to them. Of course, the Mincerian returns are not causal, being upwards biased due to positive selection into employment. Our experimental results shed light on the causal impact of vocational training and quantify the selection bias in these Mincerian returns. This evidence just shows there is demand for vocational training in this setting, and potentially very high returns to vocational training in the sectors SMEs in our study operate in. This is in contrast to high-income settings where many training programs have had low returns or short-lived impacts on workers [Card *et al.* 2015].

Apprenticeships Firm-sponsored training is another central means by which workers accumulate human capital [Acemoglu and Pischke 1998, 1999, Autor 2001]. Apprenticeships are a common labor contract throughout Sub-Saharan Africa, including Uganda [Hardy and McCasland 2017]. Table 3 provides evidence on such contracts from our sample of SMEs at baseline. Panel A shows that half of workers employed in these SMEs at baseline report having received on-the-job training in their current firm, with an average training duration of 10 months. Panel B shows a variety of payment structures for apprentices: the majority are unpaid, while others are paid, and some pay for their training. For those 20% of workers that are paid, apprentices report earning an average monthly wage of $39.[9]

Panel C shows the main opportunity cost to SMEs taking on new hires is the firm owner's time: they are predominantly tasked to train apprentices. Firm owners likely have the skills to do so: they have significantly more years of education than their employees, are significantly more likely to have received vocational training themselves, and have owned their firm for 6.5 years on average. As mentioned above, the majority of SMEs report an inability to screen workers as a constraint to expansion. Hence, if SMEs are credit constrained, it is these kinds of up-front screening costs, or firm owner's opportunity costs of training new hires, that are reduced in our wage subsidy treatment.

2.2 Experiment

Design The left hand side of Figure 1 presents the experiment design from a worker's perspective. The oversubscription design is such that 1714 workers initially applied to the offer of vocational training. Among workers assigned to vocational training, we further randomly assigned them into one of two treatments (the top branch of Figure 1A): the first group completed their six months of

[9] To get a sense of the cost of an apprenticeship we note that: (i) for 52% of all apprentices their main cost is the opportunity cost of labor market opportunities during the apprenticeship as well as fixed costs of work (e.g. travel, tools). For 29% of workers that pay for their apprenticeship, on average the total payment is over $500. Whichever way we might calculate it, the expected cost of an apprenticeship is high and above the annual earnings of our sample of workers at baseline.

training and then transitioned into the labor market. This is the business-as-usual training model, where VTIs are paid to train workers, but not to find them jobs. The second group of trained workers, upon graduation, were matched to firms operating in the same sector as the worker had been trained in, and in the same region.

Workers not offered vocational training were randomly assigned as follows: (i) matched to firms; (ii) matched to firms and those firms offered a wage subsidy in order to hire the worker and train them on-the-job for six months; (iii) held as a control group. Pairwise comparisons across treatment arms are informative of the returns to vocational training (VT), firm-provided training (FT), as well as labor market search frictions.[10]

The right hand side of Figure 1 shows the design from a firm's perspective: firms were randomly assigned to either be matched with vocationally trained workers, matched with untrained workers, matched with untrained workers and given a wage subsidy to hire them, or held as a control group. Comparing firm outcomes between the wage subsidy and control group is informative of the employment displacement and profit impacts of hiring workers through wage subsidies.[11]

Treatments Vocational training provides workers six months of sector-specific training in one of eight sectors. In those treatment arms involving vocational training, BRAC entirely covered training costs, at $470 per trainee.[12] Vocational training lessons were held Monday-Friday, for six hours per day; 30% of the training content was dedicated to theory, 70% to practical work covering sector-specific skills and managerial/business skills.[13]

In the firm training treatment, we offered firms the chance to meet untrained workers and a payment of $50 a month for six months, to hire one such untrained worker. This was designed as an inflexible wage subsidy where $12.5/month was to be retained by the owner, and $38/month

[10] As Figure 1A shows, the control group was purposively larger than the other groups in anticipation of higher attrition rates. At the point of application, workers provided a preference ranking over their top three sectors to be trained in. For those assigned to vocational training, 91% of them were trained in one of their top-3 sectors.

[11] In a companion paper, Bandiera *et al.* [2017] we provide a comprehensive analysis of the firm side impacts of these treatments (and other treatments) in the short and long run, and what light they shed on constraints to expansion that SMEs face. Of relevance for the current analysis is that: (i) firms are balanced on observables across treatments, including on monthly profits, employee numbers, the value of the capital stock, age and owner characteristics; (ii) we find that firms assigned to the wage subsidy treatment are more likely to attrit by the first follow up, and we account for this by weighting observations using inverse probability weights.

[12] The cost per trainee breaks down as the cost to the VTI ($400), plus the worker's out-of-pocket costs during training, such as those related to travel and accommodation ($70). Each VTI received 50% of the payment one week after training began, and the remaining 50% four months later (for trainees still enrolled). Hence VTIs were incentivized to retain trainees, not to find them jobs (as is the norm in Uganda). This simple incentive contract solved drop out problems associated with training programs in low-income settings [Blattman and Ralston 2015]. It also highlights these drop out problems are not driven by worker behavior but by the VTIs themselves. There was no additional stipend paid to trainees during training, and no child care offered either (recall that around 10% of our worker sample have at least one child).

[13] At end of the vocational training, workers were asked about their satisfaction with the training: 75% reported being extremely happy/very happy with the VTI experience; 80% were extremely happy/very happy with the skills gained; 56% reported that six-months of training was enough time for them to learn the skills they had wanted to, and 36% reported skills acquisition as being as expected (60% reported better than expected).

was to be paid to the worker. To assess whether the wage subsidy amount is reasonable in our context, we relate to two anchors: (i) Table 3 showed that during apprenticeships in the sample of SMEs, if workers were paid their mean wage was $39/month; (ii) using the wages of all unskilled workers employed in our SMEs at baseline, our wage subsidy treatment had a subsidy rate (wage subsidy/average wage) of 63% (Figure A2a shows the entire distribution of unskilled wages at baseline among those employed in our SMEs). This is high: for example, de Mel *et al.* [2010, 2016] evaluate a wage subsidy program with a 50% subsidy rate. The wage subsidy lasted six months, conditional on the trainee remaining employed in the firm. Monitoring firms that took up the offer, we found the apprenticeship program to be implemented as intended.[14]

In the matching treatments firms were presented lists of workers that were: (i) willing to work and vocationally trained (T4); (ii) willing to work but untrained (T2, T5). In case (i), the firms knew what sector the workers had been trained in, where they had been trained, but not that training had been paid for by BRAC. Hence, firms might expect workers presented to them to be similar to those able to self-finance such training. In all treatments involving matching, there were a maximum of two workers presented to firms on a list, and the randomly assigned matches took place with firms operating in the same sector as the worker had been trained in (or had expressed an initial desire to be trained in), and both worker and firm were located in the same region in Uganda (Central, North, East and West).

2.3 Data

Timeline and Take-up Figure 2 shows the study timeline: the baseline worker survey took place from June to September 2012 when workers applied to the offer of vocational training. 1714 eligible workers were tracked over follow-up surveys fielded 24, 36 and 48 months after baseline (12, 24 and 36 months after the end of vocational training/firm-training placements).[15] The lower part of Figure 2 shows the timeline of firm surveys over four post-intervention waves.

We use a stratified randomization where strata are region of residence, gender and education. Table 1 shows the labor market characteristics of workers in each treatment, and Table A1 shows demographic and other background characteristics. In both cases, the samples are well balanced across treatments, and normalized differences in observables are small. Attrition is low: 13% of

[14] We monitored the use of wage subsidies via spot checks by BRAC staff to ensure the designated subsidy split was being adhered to. We found that both workers and firms reported the correct subsidy split being made. Figure A2b shows worker and firm reports on the wage subsidy being received by the worker, with a clear spike at $39 /month as intended. Also, we found most apprentices that started working at the matched firm completed the full six months of subsidized on the job training.

[15] We also conducted a tracker survey to those randomized out of vocational training: this was fielded just as vocational trainees were transitioning into the labor market. The purpose of this survey was to construct accurate measures of the opportunity cost of attending the six months vocational training, that is used for the later IRR calculations. The tracker survey had a 19% attrition rate. The work status of respondents were as follows: 15% were currently involved in some work activity, 12% had been involved in a work activity in the last six months (but not on survey date), and 72% had not worked in the last six months.

workers attrit by the 48-month endline.[16] The Appendix describes correlates of attrition in more detail, confirming attrition between baseline and endline is uncorrelated to treatment.

Workers are observationally equivalent at the point of application, when assigned to treatment. Subsequent to treatment assignment there is selective non-compliance by workers, and for the treatments involving worker-firm matches, there is also non-compliance by firms because worker-firm matches only occur if both a worker and the firm express a willingness to meet. Table A3 provides evidence on worker and firm take-up by treatment.

Focusing first on treatments involving vocational training we see that: (i) over 95% of workers that initially apply for vocational training are later found and offered it (Column 1); (ii) around 68% of workers take-up the offer of vocational training (Column 2). We further note that over 95% of them complete training conditional on enrolment.[17]

For workers assigned to the firm-training treatment, 51% of them are actually offered a meeting with a firm (Column 3). Hence in common with earlier studies, firm interest is a key limiting factor preventing worker-firm matches to actually occur [Groh *et al.* 2016]. This might be because of stigma effects where firms perceive workers with attached subsidies as being of low quality [Bell *et al.* 1999]. However, conditional on the worker-firm match, 80% of such meetings actually end up taking place (Column 4), 90% of interviewed workers are offered a job, and two thirds of job offers are accepted (Column 5). Hence overall, 24% of workers initially assigned to the firm training treatment end up being employed at the matched firm.

Firm's lack of demand to meet workers we present to them is more severe in the other treatments involving matching: for workers assigned to either the vocational training plus matching, or pure matching treatments (T4, T5), only 13% and 19% of workers end up being offered even a meeting with a firm (Column 3). This is not altogether surprising given the context: for example, we note from Table 1 that given youth unemployment rates of around 60% firms should have little difficulty in meeting untrained workers, and as Table 2 showed, around one third of employees in SMEs are vocationally trained and so SMEs might have no difficulty meeting trained workers. In short, there is not much evidence for search frictions related to meeting untrained workers or meeting skilled workers in these labor markets.[18]

Given the low worker-firm matching rate in the vocational training plus match treatment (T4), for the remainder of the analysis we combine these workers with those assigned to the vocational training treatment (T3). Moreover, given the low worker-firm matching rate in the pure matching

[16] This attrition rate compares favorably to other studies such as Attanasio *et al.* [2011] (18%), and Card *et al.* [2011] (38%). Indeed, in the meta-analysis of McKenzie [2017], all but one study have attrition rates above 18%.

[17] In the meta-analysis of McKenzie [2017], most studies have training completion rates between 70 and 85%. Among workers that did not take-up the offer: 25% reported not doing so because they had found a job, 8% were in education, 4% were in another form of training, 35% did not take-up for family reasons (they had a child, illness, or family emergency), 15% reported distance as being the main constraint, and 24% reported other reasons.

[18] For treatments T4 and T5 we note from Table A3 that a high percentage of these matches end up leading to job interviews (Column 4), fewer of them convert to job offers (Column 5) and acceptances (Column 6). This suggests worker preferences can also be a source of matches not translating into hires [Groh *et al.* 2016].

treatment for untrained workers (T5), we drop this treatment arm for the bulk of the analysis.

This allows us to focus attention throughout on the comparison between firm-trained (FT) workers (T2) and vocationally trained (VT) workers (T3 and T4).

Two other timing-related points are relevant. First, although workers were randomly assigned to treatment at the point of application, they were only informed about any match that might be offered once vocational trainees had completed their courses. This helps avoid lock-in or threat-effects on worker search [Black *et al.* 2003, Sianesi 2004]. Second, the design ensures vocational trainees and firm-trained workers both come into contact with firms at the same time: this is in line with the underlying motivation for our study, to understand labor market transitions of youth. However, inevitably this means that vocational trainees receive their training before firm-trained workers do. This six month divergence in training times is however unlikely to bias estimates based on the three years of subsequent follow up data.[19]

Estimation As workers are observationally equivalent only at the point of application to vocational training, we mostly present ITT estimates for worker outcomes based on random assignment to treatment at the point of application. We use the following ANCOVA specification for worker i in strata s in survey wave $t = 1, 2, 3$,

$$y_{ist} = \sum\nolimits_j \beta_j T_i + \gamma y_{i0} + \delta \mathbf{x}_{i0} + \lambda_s + \vartheta_t + u_{ist}, \tag{1}$$

where y_{ist} is the labor market outcome of interest, worker i is assigned to treatment T_i (vocational training or firm training), y_{i0} is the outcome at baseline, $\mathbf{x}_{i0}$ are the worker's baseline covariates. λ_s and ϑ_t are strata and survey wave fixed effects respectively. As randomization is at the worker level (i), we use robust standard errors, and we weight ITT estimates using inverse probability weights (IPWs). We later show the robustness of the main results to dropping all covariates except baseline outcomes, randomization strata, and survey wave fixed effects, and to not using IPWs.[20]

The coefficient of interest is β_j: the ITT impact estimate for treatment T_i as *averaged* over the three post-intervention survey waves. We therefore leave all discussion of dynamics to the job ladder model. β_j measures the causal effect of treatment on outcomes under SUTVA. In this setting SUTVA will not hold if treatment displaces control workers because treated workers

[19] We show evidence in support of this by exploiting a small second batch of vocational trainees that received their training between October 2013 and April 2014, so when apprenticeships were implemented. The primary worker outcomes do not differ between the main and second batch of vocational trainees.

[20] The baseline worker characteristics $\mathbf{x}_{i0}$ controlled for are age, a dummy for whether the worker was married, a dummy for whether the worker had any children, a dummy for whether the worker was employed, and a dummy for whether the worker scored at the median or above on a cognitive test administered at baseline. We also control for the vocational training implementation round and month of interview. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the three follow-up survey rounds.

become relatively more attractive to firms. To assess whether this is likely we first need to establish the relevant labor market for these workers. We note that at baseline workers are geographically and sectorally mobile: the majority are willing to travel to other labor markets to find work, and many are willing to consider working in different sectors.[21] Defining a labor market as a sector-region, our firm census shows that on average, there are 156 employed workers and 40 firms in each market. We match an average of 8 workers per market, corresponding to just 5% of all workers or 7% of new hires (those starting employment in the prior three months). Hence we do not expect the control group to be contaminated by treated workers in the same labor market as they are unlikely to be competing for the same exact job.[22]

3 Treatment Effects

3.1 Skills

Our core focus is on comparing supply- and demand-side training interventions for young workers: providing sector-specific vocational training to workers before they enter the labor market (a supply-side policy), versus incentivizing firms via wage subsidies to take on and train workers on-the-job (a demand-side policy). A key distinction between these training routes relates to a fundamental information asymmetry in labor markets: vocationally trained workers have more certifiable skills, because VTIs provide graduates with certificates. Certification makes it more likely that vocationally trained workers can move across firms in the same sector. In turn, this impacts incentives of firms to train such workers in firm-specific skills.[23]

We present three results related to treatment impacts on worker skills. We first test whether workers report having been trained by a firm in their first employment spell. We define two dummies: (i) whether the worker reports having received on-the-job training at her first employer; (ii) whether the worker reported being a 'trainee' in her first employment spell. Columns 1 and 2 of Table 4 show that for both outcomes, workers assigned to firm training are between 14 and 22pp more likely than the control group to report being firm trained.[24] More surprisingly: (i)

[21] At baseline, 33% of workers reported that they had previously attempted to find a job in a different town than the one they come from. Of the ones that had attempted to find a job in another town, 27% had succeeded. Of the ones that did not try to find a job in a different town, 92% said they would like to find a job in a different town than the one of origin. On workers sectoral mobility, at baseline 97% of workers reported being willing to work in more than one sector. Moreover, only 14% of all main job spells of workers in the control group at follow-up are in the same sector as the ideal sector mentioned at baseline.

[22] Crepon *et al.* [2013] provide experimental estimates of the equilibrium impacts of labor market policies in France using a design that randomizes the fraction of treated workers across labor markets, and individual treatment assignment within labor markets.

[23] Evidence of the value of certification in labor markets is provided by Pallais [2014], MacLeod *et al.* [2015] and Bassi and Nansamba [2017].

[24] This is over a baseline of 40% of workers in the control group reporting to have received training in their first employment spell (Column 1), a magnitude that matches up well with the descriptive evidence in Table 3 where 50% of workers employed in the SMEs at baseline reported having been apprentices in the firm.

vocationally trained workers are no more likely than the control group to have been trainees in their first employment spell; (ii) workers assigned to firm training are significantly more likely to view themselves as trainees than vocational trained workers ($p = .000$). This suggests firms are less willing to train workers that have already been vocationally trained in sector-specific skills: we thus find no evidence of firms targeting on-the-job training to workers with higher levels of human capital to begin with, or a complementarity between firm-provided skills and sector-specific skills. The finding is consistent with firms anticipating these workers to be more mobile than others, because their skills are certifiable to other employers.[25]

The second piece of evidence relates to a sector-specific skills test that we developed in conjunction with skills assessors and modulators of written and practical occupational tests in Uganda. This kind of skills test has not been conducted much in the training literature.[26] Each sector-specific test comprises seven questions (five multiple choice, one pairwise-matching questions, and one question requiring tasks to be correctly ordered): Figure A5 shows an example of the skills test for the motor mechanics sector. Workers had 20 minutes to complete the test, and we convert answers into a 0-100 score. If workers answer questions randomly, their expected score is 25. The test was conducted on all workers (including those assigned to the control group) at second and third follow-up. There was no differential attrition by treatment into the test.[27]

Before administering the test, we asked a filtering question to workers on whether they had *any* skills relevant for the sectors in our study. The dependent variable in Column 3 of Table 4 is a dummy equal to one if the worker reported having skills for a sector. The ITT estimates show that VT workers and FT workers all report being significantly more likely to have relevant skills

[25] The investment by firm owners in training workers is well recognized by employees. For example, from the firm-side experiment, we interviewed employees in our sample of SMEs pre-intervention. We asked them about the role of their firm owner in training workers. In the control group of firms, 79% of employees agreed with the statement, "Does the owner put special effort in training and retaining the best workers?", and when asked, "What do you feel makes it better to work at this firm relative to your competitors, if anything?", 43% of employees reported the better training/learning opportunities. This was the most frequent answer given, the next being higher wages, as reported by 22% of employees.

[26] Berniell and de la Mata [2016] present evidence from a 12-month apprenticeship program in Argentina. In comparison to the control group, they find little evidence that the cognitive and non-cognitive skills of apprentices are impacted, but that relative to the control group they are able to certify their skills to a great extent and this drives some of the higher employment probabilities for treated workers. They do not set out to measure the task composition of workers, or develop specific tests to measure sector- or firm-specific skills. Adhvaryu *et al.* [2016] present more detailed evidence on worker knowledge, preferences and task assignment of workers in the context of an evaluation of a soft-skills training program for garment workers in India.

[27] We developed the sector-specific skills tests over a two-day workshop with eight practicing skills assessors and modulators of written and practical occupational tests from the Directorate of Industrial Training (DIT), the Uganda Business and Technical Examinations Board (UBTEB) and the Worker's Practically Acquired Skills (PAS) Skills Testing Boards and Directorate. To ensure the test would not be biased towards merely capturing theoretical/attitudinal skills taught only in VTIs, workshop modulators were instructed to: (i) develop questions to assess psychomotor domain, e.g. trainees ability to perform a set of tasks on a sector-specific product/service; (ii) formulate questions to mimic real-life situations (e.g. "if a customer came to the firm with the following issue, what would you do?"); (iii) avoid using technical terms used in VTI training. We pre-tested the skills assessment tool both with trainees of VTIs, as well as workers employed in SMEs in the eight sectors we study (and neither group was taken from our worker evaluation sample).

than those in the control group. As reported at the foot of the Table, 60% of control group workers report having skills for some sector, and this rises by 11pp for FT workers, and by 27pp for VT workers. This increase is significantly greater for VT than FT workers ($p = .000$), suggesting some share of workers might be left untrained by firms in receipt of a wage subsidy.

Only workers that reported having skills took the skills test: others were assigned a score of 25 assuming they would answer the test at random. Column 4 shows that only VT workers significantly increase their measurable sector-specific skills. Relative to the control group (that on average score just above the random answer baseline), VT workers increase sector-specific skills by 23% (or .3 of a standard deviation in the test score distribution). Strikingly, there is no increase in sector-specific skills among FT workers, and the skills impacts are significantly larger for VT workers than for FT workers ($p = .001$).

One concern might be that ITT impacts among FT workers are hard to detect given the low rate of worker-firm matches that occur in this treatment arm. We therefore estimate a LATE specification using 2SLS for those treatment arms involving worker-firm matches (T2 and T4) relative to the control group, where we instrument whether the offer of a match taking place was made with the original treatment assignment. Under the assumption that assignment to a matching treatment does not impact outcomes for those that do not actually match with a firm, and that there are no individuals that would match with the firm only if they are in the control group, this yields the LATE: the impact of matching with a firm for individuals who match with a firm when assigned to a match treatment, and who do not match otherwise. The LATE specification in Column 5 confirms the earlier result: matched FT workers do not have significantly higher sector-specific skills than the control group. In contrast, VT workers have significantly higher sector-specific skills, and the magnitude of the effect is impressive: their skills test score rises by 57 points over the control group (that score 30/100) and so score near to full-marks on the test.

The third and final piece of evidence relates directly to firm-specific skills, as measured at endline. As Column 6 shows, relative to control workers, VT workers are significantly more likely to report their skills being transferable across firms relative to either the control group or FT workers ($p = .025$), while FT workers are no different to the control group in the skills transferability (indeed, the point estimate is negative). This suggests the labor market mobility of VT and FT workers might differ, a core issue we return to below.

3.2 Tasks

We next explore how these difference in balance between sector- and firm-specific skills translates into the composition of tasks within firms performed by VT workers and FT workers at endline. For each sector, we construct a list of 30 to 40 tasks performed by workers (based on modifying the

O*NET task list).[28] For any given task j in sector k, we construct the share of workers reporting to perform task j, separately for those that have been vocationally trained and trained on-the-job. Figure 3A graphs the difference in these shares for each task j, color coding the Figure by sector. Panels A and B show separately the differential task composition over training types for the manufacturing and service sectors in our evaluation.[29] In each sector we see a stark divergence away from the zero line in the differences in these shares: within a sector, there are some tasks performed relatively more often by VT workers (at the left hand side of the Figures), and other tasks performed relatively more often by FT workers (at the right hand side of the Figures). In short, the two types of worker do not conduct the same tasks within firms, despite these SMEs being in the same sector and relatively homogenous at baseline.

Taking stock of these results on skills and tasks we see that: (i) the balance of evidence suggests VT workers have measurably higher sector-specific skills and indeed report their skills to be transferable across firms; (ii) FT workers report having more firm specific skills; (iii) workers transitioning into the labor market through VT and FT training routes perform very different tasks in firms as measured even years post intervention. This all fits with the broad differences between vocational and firm provided training, as expected, given the greater certifiability of vocational training in labor markets. We next examine how these differences in human capital translate into hard labor market outcomes.

3.3 Employment and Earnings

The relative effectiveness of the two types of training on workers' employment and earnings depend on the constraints faced by workers and, more subtly, by firms. The VT treatment relaxes credit and information constraints for workers and, as shown above, gives them more portable skills. VT should therefore have a larger impact on workers' welfare as it enhances mobility between jobs and hence confers stronger bargaining power within jobs. However, its effectiveness relies on the existence of job opportunities for the vocationally trained. If these do not exist, only a policy that relaxes firms' hiring constraints will increase employment rates.

Table 5 presents ITT estimates for labor market outcomes, starting with extensive margin outcomes on employment. Column 1 shows that, averaged over the three post-intervention survey waves, both forms of worker training positively impact employment probabilities: FT and VT workers are 6pp and 9pp more likely to be employed, corresponding to 14% and 21% impacts over the control group, whose unemployment rate remains over 55%. Hence, the treatment impacts

[28] The Occupational Information Network (O*NET) database contains occupation-specific descriptors for occupations in the US economy. These are designed to reflect the key features of an occupation through a standardized, measurable set of tasks. Further details are here: https://www.onetonline.org/

[29] The data refers to all main job spells reported at endline (so there is one job spell per worker and only employed individuals are included to construct the task composition figures). Workers were asked to report which tasks they performed in each employment spell they had in the year prior to the survey.

on youth unemployment rates are economically significant, and this is so for both supply- and demand-side training policies.[30]

On the total effect margin, Column 2 shows that VT and FT workers significantly increase the months worked in the year by .89 and .52, respectively, corresponding to 19% and 11% increases over the months worked by the control group. Hence the pattern of results is that, through either training route, these young workers increase their overall labor market attachment. This notion is further reinforced if we consider the sectors that workers end up being employed at. Conditional on employment, we find the likelihood of a VT or FT worker being employed in one of the eight sectors we focus on to be around double that for workers in the control group. In contrast, control group workers remain reliant on insecure causal wage labor.

Column 3 then focuses on hourly wage rates in wage employment (set to zero for the unemployed). Given the skewed demographics of Uganda with a large supply of young workers, and a large mass of SMEs in the sectors we study, then absent any search frictions, we might reasonably expect labor markets to be competitive and workers to be paid close to their marginal product. Hence examining wage impacts can be informative of productivity impacts. Building on the earlier results on worker skills in Table 4, we indeed see further evidence of there being wage/productivity impacts of both training routes. For FT workers, hourly wages rise by 38% relative to the control group, and for VT workers they rise by 42%. In line with the earlier skills results, both types of trainee appear to be more skilled than the control group, albeit with a different balance between sector- and firm-specific skills. The fact that hourly wage rates are similar between the two training routes reflects two opposing forces resulting from the fact that FT workers have more firm-specific skills as evidenced in Table 4: (i) their marginal product within the firm can be higher than VT workers whose skills are less firm specific, thus causing wages of FT workers to be higher; (ii) FT workers are less mobile across firms, and so employers hold more monopsony power over FT workers and can thus hold their wages below their marginal product.

Column 4 combines extensive margin and total effect margin effects to derive ITT impacts on total monthly earnings. For FT workers, total earnings rise by 20% over the control group, while VT workers experience a larger rise in total earnings of 34%. This is driven by them having more stable employment and working more months over the year.

This overall ITT impact on earnings combines: (i) an extensive margin employment effect (Table 5, Column 1); (ii) a composition effect, namely those that are employed might be a select group of workers; (iii) a productivity effect, namely the causal change in earnings of those employed.[31]

[30] These increases in employment are all driven by increases in wage employment for FT workers. VT workers are 5.5pp more likely to be in wage employment, a 20% increase over the control group, and 3.6pp more likely to be in self employment, a 23% increase over the control group. We do not see any decrease in casual wage employment for workers in either training route. Rather, their overall attachment to the labor market increases.

[31] On the first two components, we note that descriptively, within each training route, we find no robust differences in time invariant characteristics of workers that are employed or unemployed at endline: for example, they have similar ages, years of schooling, and they score the same on tests we administered at baseline to assess their cognitive abilities or personality traits at baseline. This is suggestive evidence for the differential returns to VT and FT *not*

To isolate the productivity effect component driving the total earnings treatment effect, we follow Angrist *et al.* [2006] and Attanasio *et al.* [2011] in estimating treatment effect bounds. Table A7 gives a detailed breakdown of the bounds, and they are summarized in Column 5 in Table 5, by training route. Both training treatments have productivity bounds strictly above zero, with these bounds being somewhat higher for vocational trainees. These results reinforce the notion that both training routes have substantive impacts on the human capital of (employed) workers.

A contribution we make to the training literature is to quantify the causal productivity impacts on workers of on-the-job training. Much of the earlier evidence has been based on observational data and there has been a long-standing debate over whether there are substantive human capital impacts of such training [Blundell *et al.* 1999], especially once the endogenous selection of workers into training is corrected for [Leuven and Oosterbeek 2008]. Our treatment effect estimates are not subject to this concern. Moreover, our two-sided experimental design allows us to later shed light on the potential selection of firms that VT and FT workers end up being employed at, that can partly drive the treatment effects documented above.

The ITT estimates in Column 4 represent large experimental returns in earnings to both forms of training. This begs the question why don't workers self-invest in either vocational or apprenticeships given such returns? One explanation is that workers are credit constrained: as documented earlier, worker monthly earnings at baseline are \$5, while the vocational training costs over \$400, and taking the evidence from Table 3, if workers were to self-finance training apprenticeships it could cost upwards of \$500. On the firm side, they might also be credit constrained and so unable to pay the up front hiring/screening costs of employing young workers.

An alternative explanation is that workers have incorrect beliefs about the returns to training [Jensen 2010]. We can assess this using information collected from workers at baseline over what they expect their chance to be of finding work, and their earnings conditional on employment, if they received vocational training. This evidence is shown in Table A6. Columns 1 and 2 focus on the extensive margin and show that: (i) at baseline, workers expect their employment probability to be 57%, that is optimistic given baseline employment rates of 40%; (ii) workers expect their likelihood of finding work to rise by 30pp or 53%, if they were to receive vocational training. This is again optimistic given the treatment effect impact on the extensive margin being closer to 14%.

In terms of earnings, Column 3 of Table A6 reports worker beliefs at baseline, over the average monthly earnings given their current skill set (assuming they were employed). These correspond to just under \$60. We then asked workers what they expected their maximum and minimum monthly earnings to be if they received vocational training (and the likelihood they would be able to earn more than the midpoint of the two). Fitting a triangular distribution to their beliefs we can derive an expected earnings from vocational training. This is shown in Column 4: on average, workers report earnings would more than double, so a greater than 100% return. This is higher

being driven by worker unobservables.

than the Mincerian returns shown in Table 2, that are themselves upwards biased. Combining both margins we see that workers expect the returns to vocational training to be close to 300%, that is far higher than the treatment effect returns. In short, workers are overly optimistic with regards to the returns to training, and this is not an explanation for their lack of investment in their own human capital.

3.4 Robustness

The Appendix presents robustness checks on our findings by first combining multiple labor market outcomes into one labor market index. In Table A8: (i) we show treatment impacts on this labor market index by gender, by sector, and by geography; (ii) we exploit the fact that there is a small batch of vocationally trained workers that are trained later in time and so closer chronologically to when apprenticeships were being implemented, to show that there is little impact on the exact timing of entry into the labor market on this labor market index as measured three years post-intervention. For these checks, we also present Lee bounds treatment effect estimates and address multiple hypothesis testing concerns by showing the core results to be robust to comparing critical values to unadjusted and adjusted Romano-Wolf p-values. Finally, the last two Columns show the robustness of the main results to dropping all covariates except baseline outcomes, randomization strata, and survey wave fixed effects, and to additionally not using IPWs.

3.5 Retention

To bridge between the treatment effect estimates and the structural estimation of the job ladder model, we discuss the retention of firm trained workers after the six month wage subsidy expired. More precisely, in each survey wave we asked workers assigned to the firm training treatment arm whether they were employed at the same firm that they were originally matched to. We find high rates of retention for firm trained workers as measured up to 36 months post-placement: on average over this period, 11.4% of all workers assigned to the firm training treatment arm were at the same firm as they were initially matched to. Recall from Table A3 that in this treatment arm: (i) 51% were actually offered a meeting with a firm; (ii) 80% of such meetings actually end up taking place; (iii) 90% of interviewed workers were offered a job; (iv) two thirds of job offers were accepted. Hence the highest feasible percentage of workers that could have been in retained is $.51 \times .80 \times .90 \times .66 = 24\%$. Our results thus suggest slightly higher retention rates for firm-trained workers in contrast to earlier wage subsidy studies. Averaged over the post-intervention period, a third of these workers are still employed, up to 30 months after the wage subsidy has expired.[32]

[32] For apprentices retained for the duration of the wage subsidy, those with above (below) the median employment spell duration with the firm have an average earning of \$47.6/month (\$40.9/month), so that earnings of retained apprentices do rise slightly over time from the wage subsidy of \$39/month (the top 1% earnings are excluded from the analysis).

Figure 4 plots the survival function for employment in the firm matched to for workers in the wage subsidy treatment. Among those actually hired by the firm, the share of FT workers who were employed at the matched firm for at least 6 months is 57%. Hence for the majority of hired workers, retention with the firm lasts longer than the period of the wage subsidy itself, suggesting these firms were constrained to begin with. Yet, it does not last so much longer: the average duration of employment at the matched firm, conditional on being higher than six months is 8.9 months, and by endline, almost none of these workers remain in the firm they were originally matched to. Hence FT workers do transition out of the firm they were originally trained by.

There are two key issues that follow, and these shape the remainder of the analysis. First is whether they transition to other firms, and if so whether this occurs at a lower rate than job-to-job transitions of VT workers, given that the skills of VT workers are more certifiable and transferable to other firms in the same sector. The job ladder model we develop and estimate in the next Section structurally estimates these transition rates for workers in each training route. The second key issue is whether they transition into unemployment or onto worse firms because the SMEs causally induced to meet workers *because* of the offer of a wage subsidy are negatively selected relative to SMEs in general. Worse worker-firm matches for young workers transitioning into the labor market might have long lasting impacts. Again the job ladder model below sheds light on this, as do the firm side treatment impacts investigated in Section 5 that fully exploit the two-sided experimental design.

4 Job Ladder Model

We now examine whether and how the dynamic impacts of the supply- and demand-side training interventions differ. As motivation, Figure 5 shows the outcome of estimating a specification that allows treatment effects to vary by survey wave. For outcomes related to employment and earnings (Panels A and B), as well as an overall index of labor market outcomes (Panel C), there are dynamic treatment effects. In particular, it is worth contrasting the steady improvement in outcomes among vocational trainees, with the more static outcomes of youth entering the labor market through firms offered wage subsidies.[33]

The job ladder model of worker search we now develop and estimate helps pinpoint the exact mechanisms driving these dynamics. The labor market features the model estimates are steady state unemployment-to-job transition rates (UJ) and job-to-job transition rates (JJ). If VT workers have more certifiable skills than FT workers, and their balance of skills is more tilted towards

[33]The Labor Market Index reported in Panel C of Figure 5 is computed using the following variables: any paid work in the last month (dummy), any wage employment in the last month (dummy), any casual work in the last month (dummy), hours worked in wage employment last week, hourly wage rate, total earnings in the last month. Hourly earnings and total earnings are set to zero for workers with no earnings. The index is constructed by converting each component into a z-score, averaging these and taking the z-score of the average. z-scores are computed using means and standard deviations from the control group at baseline.

sector- rather than firm-specific skills, then they should make such transitions more frequently, all else equal. As a result they will more quickly move up the job ladder and their wage profiles will diverge away from FT workers, as suggested in Figure 5.

4.1 Set-up

We develop a standard job ladder model of worker search, in the line of Burdett and Mortensen [1998]. Workers are assumed risk neutral. Workers are homogeneous except along two dimensions: their treatment (training) status is denoted T (where $T = VT$, FT or C), and their employment status in any period (where a worker can be unemployed (u) or employed (e)). We assume workers have reached their steady state labor market trajectories/transition rates by November 2015, so more than two years since the end of vocational training. Firms post a wage w and make take-it-or-leave-it offers, where the offer is a commitment to pay wage w to the worker until the worker is laid off or quits. Employed and unemployed workers sample wages from the same distribution $f(w|T)$, with CDF $F(w|T)$.[34]

Unemployed workers choose search intensity c each period, and their value function is:

$$V^n(T) = -\varphi(c) + \beta \left[\lambda_0(c,T) \max\left\{ \int V(w,T)dF(w|T), V^n(T) \right\} + (1 - \lambda_0(c,T))V^n(T) \right]. \quad (2)$$

We assume unemployed workers earn zero every period, so the first term is the cost of search effort, $-\varphi(c)$. β is the discount rate. In the next period the worker receives a job offer with probability $\lambda_0(c,T)$, that depends positively on her search effort c and training status T. The worker takes up this job offer if the expected value of the job is higher than the value of remaining unemployed. With probability $(1 - \lambda_0(c,T))$ no job offer arrives and the worker remains unemployed.

Employed workers choose search intensity c and whether to accept or reject wage offers (their reservation wage), and their value function if employed at wage w is:

$$V(w,T) = w - \varphi(c) + \beta \left[\delta V^n(T) + \lambda_1(c,T) \max\left\{ \int V(w,T)dF(w|T), V(w,T) \right\} + (1 - \delta - \lambda_1(c,T))V(w,T) \right] \quad (3)$$

where such workers are assumed to be able to search on-the-job at the same cost $-\varphi(c)$ as unemployed workers. In the next period, the worker's employment terminates with probability δ: this job destruction rate captures both the quality of jobs and the expected duration of the employment relation. With probability $\lambda_1(c,T)$ the worker receives an outside job offer that is also increasing in c and T. She takes up this opportunity if the expected value of the job offer exceeds the current job value: this is the notion of a job ladder. With probability $(1 - \delta - \lambda_1(c,T))$ the worker's job

[34] There is an established literature on job ladder search models, the defining characteristic of which is always that workers agree on the ranking of available jobs, hence the notion of a job ladder [Bontemps *et al.* 2000, Moscarini and Postel-Vinay 2017].

neither destructs, nor does she receive an outside offer and so she remains in her current job.[35]

The model makes precise that training affects behavior through two mechanisms: (i) the probabilities of receiving job offers: ($\lambda_0(c,T)$, $\lambda_1(c,T)$); (ii) the distribution of offered wages ($F(w|T)$). Through these mechanisms training impacts the worker's endogenous choices over search effort (c) and whether to accept or reject wage offers (reservation wage).

4.1.1 Supportive Evidence

We provide three pieces of evidence supporting the basic structure of this simplified job ladder model in our context, relating to: (i) wage growth; (ii) expectations over job offers; (iii) search behavior. First, the model predicts wage growth occurs between, not within, job spells because a worker's wage only increases when making JJ or UJ transitions. Within a spell, the wage is fixed at w. To examine this prediction we decompose workers' wage growth into that occurring within and between job spells. When doing so over a reference period of a year, we find the average wage growth of job movers is at least twice as high as that of job stayers, irrespective of the exact reference period used.[36]

We next present reduced form ITT estimates for outcomes related to the mechanisms in the structural model. The results are in Table 6. In Column 1 we examine how each treatment impacts workers stated belief that they will find a job in the next six months (on a 0-10 scale), relating to $\lambda_i(c,T)$. We see that FT and VT workers are both significantly more optimistic than the control group. Moreover, VT workers are significantly more optimistic than FT workers in this regard ($p = .000$). Columns 2 to 4 show how this translates into expected wage offers. FT workers do not expect their offered wages to be any higher than the control group. In sharp contrast, VT workers expect a rightward shift of the distribution of $F(w|T = VT)$ so their minimum and maximum wage offers are significantly higher. Using information on their report of how likely their wage is to be above the midpoint of the two and fitting a triangular distribution, Column 4 shows VT workers' expected offer wages to rise by around 27% relative to the control group. Finally, the last two Columns examine worker search effort, relating to c in the model. In line with offered

[35] There is no wage bargaining in this set-up. Our empirical setting is not well suited to such a version of the model: unionization rates are less than 1% in Uganda, and the demographic structure ensures there is no shortage of potential labor hires available. Both factors dampen workers bargaining power [Rud and Trapeznikova 2016]. The model also assumes a stationary environment so that there is no accumulation of human capital or assets over time, or directed search/memory. Such extensions lie outside the scope of the current paper.

[36] To decompose worker's wage growth, we first exploit the fact that for each job spell we have information on the wage in the first month and the last month of the spell. We then choose some reference date and linearly interpolate wages from the first and last month of the spell ongoing at the reference date. We then calculate the wage growth between two reference dates (e.g. between April 2015 and April 2016) for: (i) workers employed in the same job throughout the reference period (job stayers); (ii) workers who change job at least once in the reference period (job movers). To avoid sensitivity to outliers, the top 1% of wages are excluded. Self-employed workers and workers with at least one unemployment spell in the reference period are excluded. We then take the ratio of the average wage growth of the job movers to job stayers. Using the reference period of April 2015 to April 2016 this ratio is 2.06 (the ratio of medians is 2.31).

wages, VT workers are significantly more likely to report actively searching for work, and also to switch towards using more formal search channels. Strikingly, for five out of six outcomes in Table 6 relating to mechanisms in the model, the response of VT workers is significantly higher than that of FT workers.

However, because these reduced form ITT impacts average over unemployed and employed workers, and workers with different employment histories, it is hard to precisely interpret the findings. While all the evidence is suggestive of the main assumptions and mechanisms of the job search model being at play in our setting, we need to structurally estimate the model to properly quantify the mechanisms for each treatment.

4.2 Steady State

To close the model we derive steady state conditions. We make the simplifying assumption that $c = \phi(c) = 0$. The implication is that the UJ and JJ transition probabilities, $\{\lambda_0(T),\ \lambda_1(T)\}$, combine worker and firm search effort, leading to job offer arrivals. As unemployed workers are homogeneous, firms have no incentive to make a wage offer to an unemployed worker that she would refuse, so we assume unemployed workers always accept any offer received in equilibrium. In steady state the following condition must then hold for unemployed workers, where N is the total size of the labor force, and u is the unemployment rate:

$$Nu\lambda_0 = N(1-u)\delta, \tag{4}$$

The left hand side is the outflow from unemployment into work, and the right hand side is the inflow into unemployment as jobs destruct. Hence the steady state unemployment rate is:

$$u = \frac{\delta}{\delta + \lambda_0}. \tag{5}$$

For employed workers with wages $\leq w$, the following flow condition defines the steady state:

$$[\delta + \lambda_1(1 - F(w))]N(1-u)G(w) = \lambda_0 F(w) N u, \tag{6}$$

where the left hand side is the outflow of workers employed at wage w (layoffs plus quits due to better wage offers), and the right hand side is the inflow into employment from unemployment. The cross sectional CDF of accepted job values among the employed, $G(.)$, differs from the offer sampling CDF $F(w)$. $G(w)$ is observed in the data, while $F(w)$ is not. However, given (5), we can derive the following steady state relationship between $F(w)$ and $G(w)$:

$$F(w) = \frac{(\delta + \lambda_1)\, G(w)}{\delta + \lambda_1 G(w)}, \tag{7}$$

$$\frac{F(w) - G(w)}{(1 - F(w))G(w)} = \frac{\lambda_1}{\delta} = \kappa_1. \tag{8}$$

We see that $G(w)$ FOSD $F(w)$ unless there are no JJ transitions ($\lambda_1 = 0$), i.e. because on-the-job search leads to outside offers, there exists a wedge between offered and accepted wages. Finally, κ_1 measures the intensity of interfirm competition for workers (labor market tightness): it corresponds to the number of outside offers a worker receives before being laid off.

4.3 Estimation, Data and Descriptives

In the Appendix we detail the construction of the likelihood function. We structurally estimate the model using maximum likelihood following the two-step procedure in Bontemps *et al.* [2000]. For all workers in each treatment arm T and the control group, the model estimates a separate set of parameters: δ, λ_0 and λ_1 (and their asymptotic standard errors). We can then straightforwardly derive steady state unemployment rates and interfirm competition (u, κ_1). To be clear, as the model is estimated for each training route and control group separately, we do not model the coexistence of differentially treated workers in the same labor market. In the next Section we examine within-firm employment displacement effects more directly.

In each survey wave, we asked workers to provide their monthly labor market history since the last survey wave. To estimate the model we use this information to convert our panel data into a job spells format data set: for each worker i we construct a complete monthly history of their employment status $e_i \in \{0, 1\}$ from August 2014, when the matching interventions took place, to November 2016, our endline survey date. We assume workers have reached their steady state trajectories by November 2015. Consistent with the model, we set one wage per employment spell, w_i, and then estimate transition probabilities (τ_{UJ}, τ_{JJ}) using a maximum of two spells since the steady state has been reached. Hence the model is estimated off the last two survey waves, the dynamic treatment effects in which are shown in Figure 5.[37]

Table A9 provides descriptive evidence on steady state employment spells, pooling the data across treatments. We see that the steady state unemployment rate is 52%: this is lower than among our workers at baseline, reflecting the fact that young workers do transition into employment. Initial unemployment spells (those being experienced in November 2015) are more likely to be right censored than initial employment spells. The average duration of employment spell is 15 months, and the duration of unemployment spells is 21 months: these are long periods out of

[37] August 2014 coincided with the start of the recall period for job spells reported in the second follow-up worker survey. Hence we use this starting point to balance the trade-off between sufficiently far from the intervention period, and using a wide enough span of worker employment history data to precisely estimate the model. Figure A3 illustrates some of the possible cases how the worker spells data is constructed for worker i, where the spell duration is d_i and transition indicators between spells are τ_{JU_i}, τ_{JJ_i}, τ_{UJ_i}. In the top panel, we consider a scenario in which worker i is unemployed in November 2015, the unemployment spell is not left censored, and the worker transitions into employment (so τ_{UJ_i} is recorded) after d_i months of unemployment. The bottom panel considers a scenario in which worker i is employed in November 2015, the employment spell is left censored, and the worker transitions into a new state after d_i months of employment.

the labor force, and initial spells of unemployment when first transitioning into the labor market might have lasting impacts on later labor market outcomes. Hence any impact either training route has on this margin can have substantial impacts on the lifetime welfare of young workers.

4.4 Results

Table 7 presents the model estimates. Panel A shows the parameter estimates (δ, λ_0, λ_1, κ_1, u) for the control group and each treatment arm. On job destruction rates δ in steady state: (i) vocational trainees end up being employed in jobs with lower job destruction rates than do FT workers; (ii) both groups of trained workers have job destruction rates lower than the control group of workers, hinting that trained workers have better jobs or greater employment spell durations.[38]

On transition rates, we see that for UJ transitions (λ_0): VT workers have transition rates that are 24% higher than FT workers, and also higher than the control group. Remarkably, the rate of transitions out of unemployment is almost identical for firm trained workers and those in the control group ($\widehat{\lambda}_0 = .019$ for both groups): the additional labor market attachment shown earlier for firm-trained workers counts for little if they fall off the job ladder into unemployment.

For JJ transitions (λ_1), VT workers transition rates are 14% higher than FT workers. Indeed, they again have the highest transition rate, while FT workers have the lowest JJ transition rates of any group of workers. We thus see the stark difference in labor market mobility of vocationally trained workers relative to those that transition into the labor market through firm provided training. In short, vocational trainees are far more mobile: when unemployed they get back onto the job ladder more quickly. When employed, they are more likely to receive outside offers and make job-to-job transitions. Of course they only accept such job offers if the value of the offered job is greater than their current one. As a result of these key dynamics, vocational trainees pull away from FT workers in terms of their overall labor market performance: this is all consistent with the earlier evidence from Figure 5, except the structural model precisely quantifies the mechanisms driving this difference. These dynamics are all in line with VT workers having more certifiable skills than FT workers, and their balance of skills being more tilted towards sector- rather than firm-specific skills.[39]

This all feeds through into the inter-firm competition for VT workers: $\kappa_1(VT) > \kappa_1(FT)$ so that in steady state they receive 1.79 job offers per employment spell, while FT workers receive 1.45 job offers per employment spell. This is driven by vocational trainees receiving more outside job offers when employed ($\lambda_1(VT) > \lambda_1(FT)$), and also because their job destruction rates are lower for firms in which vocationally trained workers end up at ($\delta(VT) < \delta(FT)$).

[38] The estimated destruction rates match closely other literature estimates. In particular, Rud and Trapeznikova [2016] estimate a different structural model of job search suing UNHS 2010-11 data for Uganda and find a very similar implied annual job destruction rate of 32% as we find for the control group.

[39] We note that for all groups of worker, the monthly rate of JJ transitions is higher than observed for US workers, that is usually below 3% across the business cycle [Moscarini and Postel-Vinay 2017].

On steady state unemployment rates, relative to the counterfactual of no intervention, both demand- and supply-side policies effectively reduce youth unemployment for treated workers. The impact on FT workers is to lower their unemployment rates in steady state by 1.37pp (corresponding to a 2.3% reduction); (ii) for VT workers the reduction is 8.68pp (14.7%). Both impacts are of economic significance, assuming there are no displacement effects.

Panel B shows offered and accepted wage CDFs, $F(w|.)$ and $G(w|.)$. When unemployed, the offered wage distribution for VT workers, FT workers and the control group are similar. This is in line with the assumption that unemployed workers take any job. The steady state gap between treatment arms opens up in relation to the distribution of *accepted* wages, $G(w|.)$. $G(w|VT)$ has a slightly higher monthly mean wage than $G(w|FT)$ and $G(w|C)$. This is as expected: vocational trainees have moved further up the job ladder, and so are only willing to accept jobs with higher wage offers.

The implied annual earnings impact of having received firm provided training in steady state relative to the control group is \$37 (corresponding to a 12% increase), the implied annual earnings impact of vocational training is \$108 (34%). Contrasting the experimental and structural returns to training, we saw earlier the treatment effect on earnings from vocational training were 34%. The steady state returns estimated are almost the same. In contrast, the treatment effect on earnings from firm-training was 20%, and the structural model estimates are just over half that at 12%. This contrast arises because the steady state calculations account for the lower UJ and JJ transition rates of FT workers. In steady state they do not move up the job ladder as fast as VT workers, and the control group of workers slowly catches up with them. This dynamic is masked by the reduced form impacts that measure average effects in the post-intervention survey waves.[40]

These earnings impacts are larger than the percentage impacts on steady state unemployment rates. This reinforces the fact that each training route not only reduces unemployment risk, but also leads to higher wages when employed, consistent with the reduced form evidence on the human capital impacts and productivity bounds results of each training route. Comparing earnings impacts across training routes, in line with the dynamics above and with the reduced form evidence in Table 5, the earnings gap between them occurs because VT workers spend less time unemployed, and more time employed (as job destruction rates are lower).

[40] An alternative hypothesis for these dynamics is that the training routes differ in how they enable workers to learn-how-to-learn, rather than enhancing their productive capacity *per se* [Neal 2017]. Dynamic impacts are then driven by intertemporal complementarity in worker's capacity to learn. Although it is difficult to find skills that impact learning capacity but not productivity, we partially explore this hypothesis by estimating whether workers cognitive abilities, and other preference parameters, change over time and differentially by training route. We do not find any evidence of such mechanisms in our setting.

5 Treatment Effects on Firms

We now exploit the two-sided experimental design to shed light on the types of firm that workers are initially matched to, and end up being employed at in steady state. This provides insight on whether the differential returns to workers across VT and FT training routes are partly due to workers matching to differentially productive (and so better paying) firms.

The right hand side of Figure 1 summarizes the experimental design from firms' perspective. Recall that we drew a sample of 1538 SMEs, operating in one of the eight sectors of interest, and having between one and 15 employees (plus a firm owner). Firms are assigned to either be held in a control group or to be matched with: (i) an untrained worker and offered a wage subsidy for six months to hire the worker; (ii) a vocationally trained worker; (iii) an untrained worker.

5.1 Selection

5.1.1 Initial Worker-Firm Matches

To understand whether firm-side selection drives returns to training routes, we first consider the characteristics of firms that initially express an interest to meet a worker they are matched to. As before, given the rate of worker-firm matches that take place across treatments, we focus attention on comparing firms in the wage subsidy treatment to those matched to untrained workers in T5. Table 8 shows the firm characteristics that predict firm's expressing an interest to meet at least one worker they are matched to. Column 1 shows that firms incentivized with wage subsidies are significantly more likely to express an interest in meeting the worker relative to firms in the pure matching treatment. Column 2 shows this to be robust to controlling for firm characteristics. However, Column 3 shows that firms interested in the wage subsidy treatment have significantly lower profits per worker than firms interested in the other matching treatment. Hence, those firms interested in meeting workers when given a wage subsidy incentive appear to be negatively selected relative to firms offered to meet with similar workers but absent any financial inducement. The lower returns to FT workers relative to VT workers might in part reflect this *initial* match with less profitable firms just as they transition into the labor market.

5.1.2 Heterogeneous Firms in the Job Ladder Model

We can extend the job ladder model to estimate the productivity of firms that workers in each training route end up being employed at *in steady state*. To add firms to the job ladder model we first assume they only use labor inputs with a CRTS technology. For a given treatment/training route T, workers are homogenous and assumed to supply one unit of labor per period, but firms are heterogenous in their productivity p, distributed with CDF $P(.)$. A firm is just a collection of jobs with equal wage w. Let M denote the total number of firms, and recall that $\kappa_1 = \frac{\lambda_1}{\delta}$ measures inter-firm competition for workers. The average firm size for those offering wage w is given by:

$$n(w) = \frac{(1-u)N}{M}\frac{g(w)}{\gamma(w)} = \frac{(1-u)N}{M}\frac{1+\kappa_1}{[1+\kappa_1(1-F(w))]^2}\frac{f(w)}{\gamma(w)}, \tag{9}$$

where $g(w) = G'(w)$, obtained by differentiating (7), and $\gamma(w)$ denotes the density of job offers at wage w in the population of firms, and $\frac{f(w)}{\gamma(w)}$ is the sampling weight of type-w firms. With random search, $\frac{f(w)}{\gamma(w)} = 1$ and the size of firms offering wage w is, $n(w) = \frac{(1-u)N}{M}\frac{1+\kappa_1}{[1+\kappa_1(1-F(w))]^2}$. Profits for firms of size $n(w)$ are,

$$\pi(p,w) = (p-w)n(w) \propto \frac{(p-w)(1+\kappa_1)}{[1+\kappa_1(1-F(w))]^2}, \tag{10}$$

The wage offer function solves the firm's maximization problem,[41]

$$w(p) = \arg\max_w \frac{(p-w)(1+\kappa_1)}{[1+\kappa_1(1-F(w))]^2}, \tag{11}$$

The wage offer function can be inverted to give:

$$p(w) = w + \frac{1+\kappa_1 G(w)}{2\kappa_1 g(w)}. \tag{12}$$

This is used to retrieve the underlying $P(.)$ needed to rationalize the observed distribution of wages. Hence in this extended version of the job ladder model, the firm productivity distribution is a non-trivial function of the CDF of accepted wages $G(w)$ (and hence of offered wages, $F(w)$).

The result is in Panel C of Table 7. Focusing on the average productivity of firms workers are employed at in steady state, we see that relative to firms that the control group of workers end up employed at: (i) FT workers end up in firms with 30% *lower* productivity; (ii) VT workers end up in firms with 55% *higher* productivity (albeit with a high dispersion of productivity).

Pulling together the treatment effects and above evidence on firms, suggests the differential returns to training route are reflective of at least three channels: (i) the fact that FT workers have less certifiable skills than VT workers; (ii) the differential forms of human capital impacted to workers through each training route, as evidenced in Section 3.1 and 3.2; (iii) the productivity of firms that VT and FT workers end up initially matched with, and in steady state. As discussed in Section 3.3, the descriptive evidence suggests that in each training route, there are no robust differences in time invariant characteristics of workers that are employed or unemployed at endline. Hence the differential returns to VT and FT are *not* much driven by worker unobservables.[42]

[41] We consider the Nash equilibrium of the wage posting game, played by a large number of firms, assuming firms maximize their steady state profit flow subject to the workers' reservation wage policy. We focus on pure strategy equilibria, where all type-p firms post a unique wage $w(p)$. It can be shown that $w(p)$ is increasing in p, so that more productive firms pay higher wages.

[42] To fully disentangle the relative importance of these three channels would require a substantively more complex

5.2 Employment Displacement

The job ladder model implies large reductions in steady state youth unemployment rates from both supply- and demand-side training policies for beneficiary youth. The firm-side experiment sheds light on whether within-firm employment displacement effects exist over the longer term. We focus on the comparison between firms assigned to the wage subsidy offer and the control group of firms. For the other treatments involving matching, workers are retained within firms too infrequently to say anything on employment displacement impacts on firms (this remains true even if we restrict ourselves to the short run as measured in the first follow-up firm survey). We estimate the effect of being offered to meet a worker, where we instrument the offer with treatment assignment, and average the impact over the four post-intervention survey waves of firm side data (that run until July 2017 as shown in the lower panel of the timeline in Figure 2), that is long after any wage subsidy has expired. We thus estimate the following specification for firm f in randomization strata s:

$$y_{fst} = \sum\nolimits_j \beta_j Offer_f + \gamma y_{f0} + \delta \mathbf{x}_{f0} + \lambda_s + \vartheta_t + u_{fst}, \tag{13}$$

where y_{fst} is the firm outcome of interest in post-intervention survey wave t ($t = 1...4$), $Offer_f$ is a dummy equal to one if the firm is offered to meet with a worker, that is instrumented by the treatment firm f is assigned to, T_f. y_{f0} is the same firm outcome at baseline, $\mathbf{x}_{f0}$ are the firm's baseline covariates and λ_s and ϑ_t are strata and survey wave fixed effects respectively. We cluster standard errors by sector-branch, and we weight observations using inverse probability weights. The coefficient of interest is β_j: the LATE impact of being offered a worker as averaged over the four post-intervention survey waves.[43]

The results are in Table 9. Column 1 shows that in wage subsidy treatment, matched workers are significantly more likely to be retained within the firm four years post-baseline, well after wage subsidies have expired. This suggests firms were labor constrained to begin with. However, Column 2 reveals that there is no impact on overall employment in firms that were offered wage subsidies, although given take-up rates, the impact is imprecisely estimated. The result tentatively suggests that over the long run, there is no change in overall firm size. The wage subsidy treatment therefore changes the allocation of jobs to workers by allowing some to queue jump, but there is full employment displacement of other workers not in our evaluation sample. We discuss this issue further below when commenting on the external validity of our findings.[44]

model. For example, the job ladder model has no wage bargaining so does not allow for differential monopsony power of firms across treatments. Differences in monopsony power could show up in the $F(.)$ and $G(.)$ functions faced by FT workers, that translates into productivity estimates in the extension.

[43] $\mathbf{x}_{f0}$ controls include owner's gender and years of education, and firm size. The strata are BRAC branch and sector fixed effects. The instruments for the IPW estimates are a dummy for whether the owner reported an intention to relocate, and the number of firms economically or socially connected to the firm owner at baseline.

[44] Our two-sided experimental design adds to a nascent literature examining impacts of wage subsidy programmes on firms (above and beyond the impacts on workers). De Mel *et al.* [2016] conduct a field experiment in Sri Lanka

5.3 Profits

The two-sided experiment allows us to measure whether the (past) presence of firm-trained workers impacts average firm profits over the four post intervention survey years. The result is shown in Column 3 of Table 9. Firms initially offered to match with an untrained worker and offered a wage subsidy to hire them, have 11.4% higher monthly profits on average over the four-year period than control firms [p = .032], corresponding to .114 x 191 = \$21.8 monthly increase in firm profits over the four years post baseline. This cannot be attributed to the wage subsidy alone (that corresponds to a $(50 \times 6)/48$ =\$6.3 monthly increase in firm profits over this time frame. Column 4 shows no capital stock adjustments take place in these firms, and so given no change in net employment, this all suggests the additional profitability arises from the higher productivity of firm trained workers earlier hired under wage subsidies, than would otherwise have been hired.

Along with the fact that the majority of workers hired under wage subsidies are retained longer than six months (as shown in Figure 4), this profit impact is the second piece of evidence suggesting firms are constrained in making new hires. Hence demand-side policies do tackle binding constraints in this setting.

6 IRR and External Validity

6.1 IRR

The supply- and demand-side training interventions we evaluate are costly big-push style policies. Hence, it is important to establish whether the returns are sufficiently high to warrant a social planner implementing either policy. Table 10 presents IRR calculations for each treatment arm, where our benchmark case assumes a social discount rate of 5%.

Panel A shows the cost breakdown of each treatment. Total costs comprise: (i) training costs: the cost per individual of vocational training was \$470, while the wage subsidy amounted to \$302 per trainee (\$50.2/month for six months); (ii) program overhead costs: these vary by treatment depending on whether worker-firm matches needed to be organized, the firm monitored etc.; (iii) the opportunity cost to workers of attending the vocational training: these turn out to be relatively small (comprising less than 10% of the total cost) because levels of youth unemployment and underemployment are so high.[45]

that provides wage subsidies to SMEs. They find firms increased employment during the subsidy period, but there was no lasting impact on employment, profitability, or sales. McKenzie *et al.* [2016] also find positive (short run) employment impacts of a wage subsidy and matching intervention with firms in Yemen. Hardy and McCasland [2017] evaluate an apprenticeship program with small firms in Ghana, and find firms retain this extra labor for at least six months, and earn higher profits in doing so.

[45] The tracker survey administered to workers randomized out of VT and interviewed just as vocational trainees were completing their courses confirms this is so: workers that are randomized out of vocational training find few employment opportunities during the six months when other workers are being trained.

Panel B shows the NPV of the lifetime earning gains, as derived from the job ladder model. We first assume the remaining expected productive life of beneficiaries is 38 years (the average life expectancy in Uganda in 2012 minus the average age of workers at baseline). The lifetime gains to FT workers are around half those accruing to VT workers. The benefit-cost ratio is above one for both interventions: it is 1.69 for FT workers and 3.56 for VT workers. The ratio falls to one for FT if the social discount rate doubles to 10%. Finally, the IRR to each treatment arm is 9.8% for FT, and 21% for VT.[46]

Panel C shows the sensitivity of these IRR estimates to alternative assumptions on: (i) the remaining productive life of beneficiaries; (ii) varying the foregone earnings from attending vocational training. As productive life shortens, the IRR to FT drops off quickly, but not for VT. This is as expected given the different wage profiles to the interventions. however it remains the case that the FT intervention still pays for itself in 10 years (over that time frame the IRR = 0). On foregone earnings, only under very extreme assumptions does the IRR for VT ever fall below 10%.

These calculations are based on the cost structure of the NGO BRAC that we collaborated with. This is an established NGO in Uganda. Hence their program overhead costs represent the *marginal* costs to them of extending their activities to the kind of training program evaluated. To get a more accurate sense of the social return of starting such programs from scratch, Panel D shows what the *total* cost per individual would have to be in order for the IRR to equal the social discount rate, 5%. We see total costs per beneficiary would have to increase from $368 to $624 for the wage subsidy intervention to break even, and total costs would have to rise threefold for vocational training to break even. The final row performs the same calculation assuming a 10% social discount rate. In this case the wage subsidy intervention would just break even and the costs for vocational training would still need to more than double.

6.2 External Validity

In meta-analyses of training interventions in low-income settings, Blattman and Ralston [2015] and McKenzie [2017] document most interventions have a very low IRR. Figure 6 compares our treatment impacts relative to other experimental studies discussed in McKenzie [2017], on employment and earnings outcomes. Our effect sizes are large relative to earlier studies, although the ranking across treatment types is in line with earlier work. We speculate over four reasons why our returns are high relative to other studies, each of which opens up avenues for future work.

First, there are design issues: our experiment has a precise sectoral focus limited to eight sectors. All workers receive vocational training in one of these sectors, and all sampled firms operate in

[46] Two further points are of note. First, these IRR figures match up closely with the short and long run IRRs for the combined in-class vocational and on-the-job training intervention evaluated in Colombia by Attanasio *et al.* [2011, 2017]. Second, they likely underestimate the utility gains from each intervention as we measure benefits only through earnings, and take no account of reduced earnings risk, or how such human capital investments can reduce worker vulnerability to macroeconomic and other shocks over the longer term.

one of these sectors (and have at least one employee plus a owner at baseline). This reduces the possibility of worker-firm mismatch. Moreover, our treatments are intensive. Specifically, we separate out in-class vocational training from a wage subsidy program, both treatments last six months, and the wage subsidy treatment had a subsidy rate higher than some other studies.

Second, only 13% of workers attrit over our four year evaluation. This attrition rate compares favorably to other studies such as Attanasio *et al.* [2011] (18%), and Card *et al.* [2011] (38%). Indeed, in the meta-analysis of McKenzie [2017], all but one study has attrition rates above 18%. As Figure 6 shows, other studies have similar or larger point estimates, but more imprecise treatment effects, that might in part arise from attrition. Moreover, our payment structures to VTIs ensured that the vast majority of workers completed training conditional on starting it, mitigating drop-out problems that earlier studies have faced.

Third, given our oversubscription design, workers that select into our sample are those willing to undergo six months of vocational training. On the one hand, such individuals might be negatively selected in that they are willing to bear the opportunity cost of lost labor market opportunities during the vocational training period. On the other hand, such individuals might be positively selected in that they are more patient than unemployed youth in general (as our intervention ensures any benefits are back-loaded). A natural contrast for future work would be to recruit workers using the offer of being hired through a potential wage subsidy scheme, that front loads benefits to workers relative to non-participation.

Moreover, the potential positive selection of unemployed youth into our evaluation has important implications for how we think about training interventions even when they lead to full employment displacement of other workers. Given youth unemployment rates of 60%, there still might be an improvement in the allocation of talent in the economy if we think of the large pool of unemployed workers as heterogeneous, and that those attracted to the sample through the offer of vocational training as being positively selected relative to the average unemployed youth in Uganda. It is exactly these kinds of motivated unemployed youth that the economic gains to matching to jobs might be highest for.

Fourth, we worked with a limited set of VTIs in Uganda, pre-selected to be of high quality based on their reputation. There is no shortage of VTIs in Uganda, and as in other low-income contexts, there are concerns over a long tail of low-quality training providers existing in equilibrium. Hence, although our treatments essentially relax credit constraints for workers, it is not obvious the results would be replicated through an unconditional cash transfer: this would rely on workers having knowledge over training providers. Rather a conditional cash transfer (conditioned on having to attend one of these VTIs) is likely to have higher returns, all else equal. This might explain why similar programs providing vouchers to workers redeemable at any training provider within the VTI market have met with more limited success [Galasso *et al.* 2004, Groh *et al.* 2016].

7 Conclusion

The development path of low income countries in the coming decades will largely depend on whether or not young workers can be matched to good jobs. High levels of youth unemployment are a symptom of the mismatch between supply and demand for labor in these countries – a growing mass of young, mainly unskilled workers are failing to find work in manufacturing and service sectors consisting mainly of small-scale firms. If workers cannot acquire the skills to access these jobs and if firm-level constraints on hiring these workers cannot be relaxed then industrialization will not proceed and living standards will stagnate.

This paper contributes to the classic literature on the value of human capital [Becker 1964, Schultz, 1981, 1993] by looking at whether and how different forms of worker training can ease the transition into manufacturing and service sector jobs. Transitions into the labor market mark a key stage in the life cycle, and a body of evidence documents how initial experiences and first job opportunities during this transition have persistent impacts on later employment trajectories and welfare [Becker 1994, Pissarides 1994]. This paper provides experimental and structural evidence on this transition from a novel two-sided experiment in the context of urban labor markets in a low-income country: Uganda.

Training of young workers whether through vocational training institutes or apprenticeships has a particular salience in low income economies for three main reasons: (i) very young populations imply that transitioning new workers into the labor market is the dominant challenge, (ii) the quality and duration of schooling is low and therefore young people are ill-equipped to access jobs in the manufacturing and service sectors of the economy and (iii) there are limited opportunities to use colleges, universities or other forms of tertiary education as a means of transitioning young people into good jobs.

What the paper reveals is that both types of training when provided over an extended period can have highly positive effects on employment and earnings within a disadvantaged set of young people transitioning into the labor market in Uganda. This is in sharp contrast with workers who receive neither type of training and who remain largely unemployed or employed in highly itinerant casual work. The labor market outcomes of these *status quo* workers is likely symptomatic of the fate of young, unskilled workers across the developing world.

What is even more revealing is that the effects for vocational training are almost twice as large as those for firm-trained workers. This result speaks directly to value of general versus specific training [Becker 1964, Acemoglu and Pischke 1998, 1999]. We find that VT workers have sector specific skills that are more transferable across firms whereas FT workers have skill that are more firm specific and less transferable. This combined with the fact that vocational training is certifiable implies that it has a much larger effect on workers welfare as it enhances their mobility between jobs. Structural estimation of a job ladder model of worker search reveals this difference in mobility as being the main mechanism for the divergence in wage profiles between VT and

FT workers. VT workers receive higher rates of job offers when unemployed, and higher rates of job-to-job offers. Consistent with this using the firm side of the experiment we find that in steady state, VT workers end up matching with significantly more productive firms than those transitioning into the labor market through firm-provided training.

As most firms in these economies are small, certification in effect enables workers to move across firms thus raising both their wages and skill levels over time. In strict contrast, both the lack of certification and the specificity of skills acquired by firm-trained apprentices may lead to a situation where workers secure jobs but remain trapped in firms with limited ability to move, to extract rents from the firm, or to get back onto the job ladder if they fall into unemployment. All this leads firm-trained workers to have flatter wages profiles as well as more limited acquisition of skills over the their working lives.

The paper thus points to the value for workers of acquiring vocational skills that are valuable to a particular sector as opposed to a particular firm. This in turn opens a rich set of research possibilities for analyzing how vocational education might be best organized in these countries, and on how the process of certification would fit into this.[47]

A Appendix

A.1 Attrition

Table A4 presents evidence on the correlates of worker attrition. Attrition is generally low, with only 13% of workers attriting by the 48-month endline. Focusing on attrition between baseline and endline, Column 1 shows that: (i) attrition is uncorrelated to treatment assignment; (ii) worker characteristics do not predict attrition in general but workers that score higher on a cognitive ability test at baseline are more likely to attrit. Column 2 shows there to be little evidence of heterogeneous attrition across treatments by baseline cognitive scores at baseline. Any bias that might arise from selective attrition on unobservables cannot be signed *a priori*. Tracked workers would be negatively selected if attriters are more likely to find employment themselves, or they would be positively selected if attriters are least motivated to find work and remain attached to the labor market. To account for attrition, we weight our ITT estimates using inverse probability weights (IPWs). We also show the robustness of the main treatment impacts when using conditional Lee bounds [Lee 2009].

On the IPWs, we proceed as follows. At each survey wave t we define a dummy s_{it} such that we observe (y_{it}, x_{it}) for observations for which $s_{it} = 1$. We then first estimate a probit of s_{it} on z_{it} for each post-intervention survey wave separately, where z_{it} includes: (i) $\mathbf{x}_{i0}$: the vector of baseline

[47] For example, if as economies develop, workers acquire credible means by which to certify their skills (both vocational and those acquired on-the-job) and to signal their labor market histories, then this might explain why in many high-income settings, training programs more commonly provide a combination of vocational training and apprenticeships, such as JTPA in the US and the YTS in the UK.

covariates used as controls throughout in (1); (ii) strata and implementation round dummies; (iii) $\mathbf{z}_{i0}$, baseline measures excluded from regression analysis: dummies for orphan, anyone in household has a phone, willing to work in multiple sectors, and; (iv) the survey team the respondent was assigned to in each survey round ($Team_{it}$). The underlying assumption is that conditional on z_{it}, y_{it} is independent of s_{it}. $\hat{p}_{it}$ are fitted probabilities from this regression using survey wave t, and so at a second stage, we weight our OLS ITT estimates with weights $1/\hat{p}_{i1}$, $1/\hat{p}_{i2}$, $1/\hat{p}_{i3}$.

A.2 Robustness Checks

To conduct robustness checks we first combine multiple labor market outcomes (beyond those shown in Table 5) into one index. The index is computed using the following variables for which a baseline value is available: any paid work in the last month, any wage employment in the last month, any casual work in the last month, hours worked in wage employment last week, the hourly wage rate, and total earnings in the last month. Hourly earnings and total earnings are set to zero for unemployed workers. The index is constructed by converting each component into a z-score, averaging these and taking the z-score of the average. z-scores are computed using means and standard deviations from the control group at baseline. Column 1 of Table A8 shows the ITT estimates on this labor market index: all treatment effects remain statistically significant.

In addition to the ITT estimates, we also report: (i) conditional Lee bounds on the treatment effects (where we use the convention that the bound is underlined if it is statistically different from zero); (ii) unadjusted and adjusted Romano-Wolf p-values to account for multiple hypothesis testing. For each ITT estimate we see that: (i) the Lee bounds are nearly always significantly different from zero (the exceptions are the lower bounds on firm training); (ii) the effect is significant against the Romano-Wolf p-values.[48]

Columns 2 and 3 split the labor market index by gender. Women have been found to benefit more from some training interventions [Friedlander 1997, Attanasio *et al.* 2011], although this finding is far from universal [McKenzie 2017]. We generally find larger impacts on men. Columns 4 and 5 split treatment effects by sector: we generally find larger productivity impacts in manufacturing. Given the correlation between gender and sector (manufacturing sectors tend to be male dominated), it is hard to definitively separate out whether the impacts are driven by gender or sector. Fourth, we consider impacts in labor markets outside of Kampala, where 81% of workers reside: the result in Column 6 largely replicates the main findings.

[48] We bound the treatment effect estimates using the trimming procedure proposed by Lee [2009]. The procedure trims observations from above (below) in the group with lower attrition, to equalize the number of observations in treatment and control groups. It then re-estimates the program impact in the trimmed sample to deliver the lower (upper) bounds for the true treatment effect. The bounding procedure relies on the assumptions that treatment is assigned randomly and that treatment affects attrition in only one direction so there are no heterogeneous effects of the treatment on attrition/selection, in line with the evidence in Table A4. As Lee [2009] discusses, using covariates to trim the samples yields tighter bounds. The covariates we use are the strata dummies. To adjust the individual p-values, we implement the Romano-Wolf [2016] step-down procedure based on re-sampling bootstrap methods, and using the publicly provided code (http://www.econ.uzh.ch/en/faculty/wolf/publications.html9).

Finally, we examine the sensitivity of the treatment effects to the timing of labor market entry. To do so, we exploit the fact that we have two batches of vocationally trained workers: the majority of trainees from the first round of applicants started training in January 2013, as shown in Figure 2. For logistical reasons, a second round of randomized-in applicants received vocational training between October 2013 and April 2014 (and so receive their training at the same time as when the apprenticeships are being implemented). In Column 7 we allow the impacts of vocational training to differ by the first and second batch of trainees: we see no evidence that workers in the second batch have different outcomes as measured by the labor market index.

Throughout Columns 2 to 7, in most cases the Lee bounds remain significantly different from zero, and the treatment effect estimates remain significant against Romano-Wolf p-values.

Finally, the final two Columns show the robustness of the main results to dropping all covariates except baseline outcomes, randomization strata, and survey wave fixed effects, and to additionally not using IPWs.

A.3 Likelihood

We first assume all random events $(\lambda_0, \lambda_1, \delta)$ are realizations of Poisson processes, so that the residual durations are exponentially distributed. As unemployed workers are assumed to be made offers by firms that they would accept, the unemployment spell hazard is λ_0. The hazard rate of job spells with wage w is $\delta + \lambda_1(1 - F(w))$. Hence for a given e_i and initial wages w_{i1} the individual likelihood contributions are as follows. The likelihood contribution for $e_i = 1$ is:

$$\begin{aligned} L(\mathbf{x}_i|e_i &= 1) = g(w_{i1}) \times (\delta + \lambda_1(1 - F(w)))^{(1-c_i)} e^{-(\delta+\lambda_1(1-F(w_{i1})))d_i} \\ &\times \left(\frac{\delta}{\delta + \lambda_1(1 - F(w_{i1}))}\right)^{\tau_{JU_i}} \left(\frac{\lambda_1(1 - F(w_{i1}))}{\delta + \lambda_1(1 - F(w_{i1}))}\right)^{\tau_{JJ_i}} \end{aligned} \tag{14}$$

where $g(.)$ is the density of $G(.)$, and c_i is an indicator for censoring, and the spell duration is d_i. The transition indicators between spells are τ_{JU_i}, τ_{JJ_i}. The likelihood contribution for $e_i = 0$ is:

$$L(\mathbf{x}_i|e_i = 0) = \lambda_0^{(1-c_i)} e^{-\lambda_0 d_i}.f(w_{i0})^{(1-c_i)}, \tag{15}$$

where $f(.)$ is the density of $F(.)$. Given steady state, $p(e_i = 0) = u = \frac{\delta}{\delta+\lambda_0}$, and $p(e_i = 1) = 1 - u = \frac{\lambda_0}{\delta+\lambda_0}$. Hence the generic likelihood contribution of observation $\mathbf{x}_i$ is:

$$L(\mathbf{x}_i) = \left(\frac{\lambda_0}{\delta + \lambda_0} L(\mathbf{x}_i|e_i = 1)\right)^{e_i} . \left(\frac{\delta}{\delta + \lambda_0} L(\mathbf{x}_i|e_i = 0)\right)^{1-e_i} . \tag{16}$$

The model is then estimated following the two-step procedure in Bontemps *et al.* [2000] for each treatment arm and the control group. In Step 1, $G(.)$ is non-parametrically estimated from the CDF of observed wages among those employed. In Step 2, we substitute $G(.)$ into $L(\mathbf{x}_i)$ using the relationship between $G(.)$ and $F(.)$ in (7). $\lambda_0, \lambda_1, \delta$ are then estimated, and their asymptotic standard errors calculated using maximum likelihood.

References

[1] ABEBE.G, S.CARIA, M.FAFCHAMPS, P.FALCO, S.FRANKLIN AND S.QUINN (2016) Curse of Anonymity or Tyranny of Distance? The Impacts of Job-Search Support in Urban Ethiopia, NBER WP22409.

[2] ACEMOGLU.D AND J.S.PISCHKE (1998) "Why Do Firms Train? Theory and Evidence," *Quarterly Journal of Economics* 113: 79-119.

[3] ACEMOGLU.D AND J.S.PISCHKE (1999) "The Structure of Wages and Investment in General Training," *Journal of Political Economy* 107: 539-72.

[4] ADHVARYU.A, N.KALA AND A.NYSHADHAM (2016) The Skills to Pay the Bills: Returns to On-the-job Soft Skills Training, mimeo Michigan.

[5] ALTONJI.J.G AND C.R.PIERRET (2001) "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics* 116: 313-50.

[6] ANGRIST.J, E.BETTINGER AND M.KREMER (2006) "Long-term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review* 96: 847-62.

[7] ATTANASIO.O, A.KUGLER AND C.MEGHIR (2011) "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics* 3: 188-220.

[8] ATTANASIO.O, A.GUARIN, C.MEDINA AND C.MEGHIR (2017) "Vocational Training for Disadvantaged Youth in Colombia: A Long Term Follow up," *American Economic Journal: Applied Economics* 9: 131-143.

[9] AUTOR.D.H (2001) "Why Do Temporary Help Firms Provide Free General Skills Training," *Quarterly Journal of Economics* 116: 1409-48.

[10] BANDIERA.O, V.BASSI, R.BURGESS, I.RASUL, M.SULAIMAN AND A.VITALI (2017) The Missing Middle? Experiment Evidence on Constraints to SME Expansion in Uganda, mimeo UCL.

[11] BASSI.V AND A.NANSAMBA (2017) Information Frictions in the Labor Market: Evidence from a Field Experiment in Uganda, mimeo UCL.

[12] BEAM.E (2016) "Do Job Fairs Matter? Experimental Evidence on the Impact of Job Fair Attendance," *Journal of Development Economics* 120: 32-40.

[13] BECKER.G.S (1964) *Human Capital*, Chicago: University of Chicago Press.

[14] BECKER.G.S (1994) "Human Capital Revisited," in *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Chicago: University of Chicago Press.

[15] BELL.B, R.BLUNDELL, AND J.VAN REENEN (1999) "Getting the Unemployed Back to Work: The Role of Targeted Wage Subsidies," *International Tax and Public Finance* 6: 339-60.

[16] BELL.D.N AND D.G.BLANCHFLOWER (2011) "Youth Unemployment in Europe and the United States," *Nordic Economic Policy Review* 1: 11-37.

[17] BERNIELL.L AND D.DE LA MATA (2016) Starting on the Right Track: Experimental Evidence from a Large-scale Apprenticeship Program, mimeo CAF.

[18] BLACK.D.A, J.A.SMITH, M.C.BERGER AND B.J.NOEL (2003) "Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System," *American Economic Review* 93: 1313-27.

[19] BLATTMAN.C AND L.RALSTON (2015) Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs, mimeo Chicago.

[20] BLUNDELL.R (2001) "Welfare Reform for Low Income Workers," *Oxford Economic Papers* 53: 189-214.

[21] BLUNDELL.R, L.DEARDEN, C.MEGHIR AND B.SIANESI (1999) "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy," *Fiscal Studies* 20: 1-23.

[22] BONTEMPS.C, J.M.ROBIN AND G.J.VAN DEN BERG (2000) "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review* 41: 305-58.

[23] BURDETT.K AND D.T.MORTENSEN (1998) "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* 39: 257-73.

[24] CARD.D, AND D.R.HYSLOP (2005) "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers," *Econometrica* 73: 1723-70.

[25] CARD.D, P.IBARRAN, F.REGALIA, D.ROSAS-SHADY AND Y.SOARES (2011) "The Labor Market Impacts of Youth Training in the Dominican Republic," *Journal of Labor Economics* 29: 267-300.

[26] CARD.D, J.KLUVE AND A.WEBER (2015) What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations, NBER WP21431.

[27] CONTI.G (2005) "Training, Productivity and Wages in Italy," *Labour Economics* 12: 557-76.

[28] CREPON.B, E.DUFLO, M.GURGAND, R.RATHELOT AND P.ZAMORA (2013) "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics* 128: 531-80.

[29] DASGUPTA.U, L.GANGADHARAN, P.MAITRA, S.MANI AND S.SUBRAMANIAN (2012) Choosing to Be Trained: Evidence from a Field Experiment, mimeo Wagner College.

[30] DE MEL.S, D.MCKENZIE AND C.WOODRUFF (2010) "Wage Subsidies for Microenterprises," *American Economic Review Papers and Proceedings* 100: 614-8.

[31] DE MEL.S, D.MCKENZIE AND C.WOODRUFF (2016) Labor Drops: Experimental Evidence on the Return to Additional Labor in Microenterprises, World Bank Policy Research Working Paper WPS7924.

[32] FARBER.H.S AND R.GIBBONS (1996) "Learning and Wage Dynamics," *Quarterly Journal of Economics* 111: 1007-47.

[33] FRANKLIN.S (2016) "Location, Search Costs and Youth Unemployment: A Randomized Trial of Transport Subsidies in Ethiopia," *Economic Journal*, forthcoming.

[34] FRIEDLANDER.D, D.H.GREENBERG AND P.K.ROBINS (1997) "Evaluating Government Training Programs for the Economically Disadvantaged," *Journal of Economic Literature* 35: 1809-55.

[35] GALASSO.E, M.RAVALLION AND A.SALVIA (2004) "Assisting the Transition from Work-fare to Work: A Randomized Experiment," *Industrial and Labor Relations Review* 58: 128-42.

[36] GOLDSTONE.J.A (2002) "Population and Security: How Demographic Change can lead to Violent Conflict," *Journal of International Affairs* 56: 3-21.

[37] GROH.M, N.KRISHNAN, D.MCKENZIE AND T.VISHWANATH (2016) "Do Wage Subsidies Provide a Stepping Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan," *Review of Economics and Statistics* 98: 488-502.

[38] GROH.M, D.MCKENZIE, N.SHAMMOUT AND T.VISHWANATH (2015) "Testing the Importance of Search Frictions and Matching through a Randomized Experiment in Jordan," *IZA Journal of Labor Economics* 4:7.

[39] HAM.J.C AND R.J.LALONDE (1996) "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training," *Econometrica* 64: 175-205.

[40] HARDY.M AND J.MCCASLAND (2017) Are Small Firms Labor Constrained? Experimental Evidence from Ghana, mimeo UBC.

[41] HECKMAN.J.J, R.J.LALONDE AND J.A.SMITH (1999) "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics* 3: 1865-2097.

[42] HOFFMAN.M AND S.V.BURKS (2017) Worker Overconfidence: Field Evidence and Implications for Employee Turnover and Returns from Training, NBER WP23240.

[43] HSIEH.C AND B.OLKEN (2014) "The "Missing" Missing Middle," *Journal of Economic Perspectives* 28: 89-108.

[44] JENSEN.R (2010) "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics* 125: 515-48.

[45] KATZ.L.F (1998) "Wage Subsidies for the Disadvantaged," in R.Freeman and P.Gottschalk (eds.) *Generating Jobs,* Russell Sage Foundation, New York.

[46] KAUFMANN.K.M (2014) "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns," *Quantitative Economics* 5: 583-630.

[47] LALONDE.R.J (1995) "The Promise of Public Sector-sponsored Training Programs," *Journal of Economic Perspectives* 9: 149-68.

[48] LAYARD.P.R.G. AND S.J.NICKELL (1980) "The Case for Subsidising Extra Jobs," *Economic Journal* 90: 51-73.

[49] LEE.D.S (2009) "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies* 76: 1071-102.

[50] LEUVEN.E AND H.OOSTERBEEK (2008) "An Alternative Approach to Estimate the Wage Returns to Private-sector Training," *Journal of Applied Econometrics* 23: 423-34.

[51] LEVINSOHN.J., N.RANKIN, G.ROBERTS AND V.SCHOER (2014) Wage Subsidies and Youth Employment in South Africa: Evidence from a Randomised Control Trial, mimeo Yale.

[52] MACLEOD.W.B., E.RIEHL, J.E.SAAVEDRA AND M.URQUIOLA (2015) "The Big Sort: College Reputation and Labor Market Outcomes," *American Economic Journal: Applied Economics* 9: 223-61.

[53] MAITRA.P AND S.MANI (2016) "Learning and Earning: Evidence from a Randomized Trial in India," *Labour Economics*, forthcoming.

[54] MCKENZIE.D (2017) How Effective Are Active Labor Market Policies in Developing Countries?, mimeo World Bank.

[55] MCKENZIE.D AND D.ROBALINO (2010) Jobs and the Crisis: What Has Been Done and Where to go From Here?, Viewpoint Note 325, World Bank Finance and Private Sector Development Vice Presidency.

[56] MCKENZIE.D, N.ASSAF AND A.P.CUSOLITO (2016) "The Demand for, and Impact of, Youth Internships: Evidence from a Randomized Experiment in Yemen," *IZA Journal of Labor and Development* 5: 1-15.

[57] MOSCARINI.G AND F.POSTEL-VINAY (2017) "The Cyclical Job Ladder," forthcoming *Annual Reviews.*

[58] PALLAIS.A (2014) "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review* 104: 3565-99.

[59] PISSARIDES.C.A (1994) "Search Unemployment with On-the-job Search," *Review of Economic Studies* 61: 457-75.

[60] ROMANO.J.P AND M.WOLF (2016) "Efficient Computation of Adjusted P-values for Resampling-based Stepdown Multiple Testing," *Statistics and Probability Letters* 113: 38-40.

[61] RUD.J.P. AND I.TRAPEZNIKOVA (2016) Wage Dispersion, Job Creation and Development: Evidence from Sub-Saharan Africa, mimeo RHUL.

[62] SCHULTZ.T.W (1981) *Investing in People*, University of California Press.

[63] SCHULTZ.T.W (1993) *The Economics of Being Poor*, Cambridge, Mass.: Blackwell.

[64] SIANESI.B (2004) "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s," *Review of Economics and Statistics* 86: 133-55.

[65] UN DESA POPULATION DIVISION (2017) *World Population Prospects: The 2017 Revision Population Database.*

[66] WORLD BANK (2009) *Africa Development Indicators 2008/9: Youth and Employment in Africa*, Washington DC: The World Bank.

Table 1: Baseline Balance on Worker Labor Market Outcomes

Means, robust standard errors from OLS regressions in parentheses
P-value on t-test of equality of means with control group in brackets
P-value on F-tests in braces

	Number of workers	Currently working	Has worked in the last month	Has done any wage employment in the last month	Any self employment in the last month	Has done any casual work in the last month	Total earnings in the last month [USD]	F-test of joint significance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Workers	1714	.360	.383	.130	.046	.257	5.93	
		(.045)	(.044)	(.023)	(.013)	(.508)	(1.11)	
T1: Control	451	.381	.401	.120	.038	.296	5.11	
		(.049)	(.048)	(.025)	(.015)	(.047)	(1.27)	
T2: Firm Trained	283	.369	.387	.103	.064*	.266	6.44	{.939}
		(.035)	(.035)	(.023)	(.017)	(.032)	(1.35)	
		[.979]	[.985]	[.520]	[.096]	[.592]	[.239]	
T3: Vocationally Trained	390	.358	.389	.149	.034	.253	7.29*	{.754}
		(.032)	(.032)	(.023)	(.013)	(.029)	(1.26)	
		[.763]	[.990]	[.188]	[.802]	[.265]	[.063]	
T4: Vocationally Trained + Matched	307	.320	.360	.149	.050	.205*	5.25	{.882}
		(.033)	(.034)	(.026)	(.015)	(.030)	(1.20)	
		[.316]	[.747]	[.229]	[.266]	[.070]	[.758]	
T5: Untrained, Matched	283	.364	.367	.127	.057	.251	5.58	{.998}
		(.033)	(.034)	(.025)	(.016)	(.031)	(1.25)	
		[.707]	[.386]	[.821]	[.207]	[.210]	[.713]	
F-test of joint significance		{.882}	{.908}	{.301}	{.214}	{.433}	{.379}	

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. All data is from the baseline survey to workers. Column 1 reports the number of workers assigned to each treatment. Columns 2 to 7 report the mean value of each worker characteristic, derived from an OLS regression of the characteristic of interest on a series of dummy variables for each treatment group. All regressions include strata dummies and a dummy for the implementation round. The excluded (comparison) group in these regressions is the Control group. Robust standard errors are reported throughout. Column 8 reports the p-values from F-Tests of joint significance of all the regressors from an OLS regression where the dependent variable is a dummy variable taking value 0 if the worker is assigned to the Control group, and it takes value 1 for workers assigned to treatment group j (with j going from 2 to 5) and the independent variables are the variables in Columns 2 to 6. Robust standard errors are also calculated in these regressions. The p-values reported in the last row are from the F-test of joint significance of the treatment dummies in each Column regression where the sample includes all workers. In Column 6 casual work includes any work conducted in the following tasks where workers are hired on a daily basis: loading and unloading trucks, transporting goods on bicycles, fetching water, land fencing and slashing the compound. Casual work also include any type of agricultural labor such as farming, animal rearing, fishing and agricultural day labor. In Column 7 workers who report doing no work in the past month (or only did unpaid work in the last month) have a value of zero for total earnings. The top 1% of earnings values are excluded. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Table 2: The Mincerian Returns to Vocational Training, by Sector

	Worker is skilled: self-reported VTI attendance			
	Share of firms in sector	% workers skilled in sector	Coefficient and SE from worker wage regressions [USD]	Coefficient and SE from worker log(wage) regressions [USD]
	(1)	(2)	(3)	(4)
All Sectors		**31.0%**	**26.2*****	**.515*****
			(3.15)	(.045)
Manufacturing				
Welding	14.57%	24.9%	34.5***	.381***
			(6.40)	(.084)
Motor-mechanics	9.80%	23.5%	16.1*	.294*
			(9.41)	(.153)
Electrical wiring	6.37%	41.9%	27.3***	.486**
			(7.60)	(.189)
Construction	4.38%	28.8%	11.5	.289*
			(9.39)	(.170)
Plumbing	3.08%	49.1%	60.9***	.719**
			(19.0)	(.281)
Services				
Hairdressing	39.64%	29.2%	22.9***	.444***
			(5.97)	(.069)
Tailoring	14.96%	41.6%	15.9	.898***
			(9.76)	(.182)
Catering	7.20%	40.2%	26.8**	.330***
			(11.6)	(.109)

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the Census of firms, which includes 2309 firms and 6306 workers. A worker is defined as skilled if he/she was reported as having attended formal vocational training at any point in the past. Coefficients and standard errors in Columns 3 and 4 are from a regression of workers' total earnings in the last month (or the logarithm of workers' total earnings in the last month) on a dummy for being a skilled worker (as defined above). Control variables in these regressions include: employee's age and age squared, gender, tenure and tenure squared, firm size, BRAC branch dummies and firm sector dummies. Robust standard errors are reported. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD. The top 1% wages and capital stock values are excluded.

Table 3: Characteristics of Apprenticeships

A. Availability	
Worker received on-the-job training at the current firm	.498
Duration of on-the-job training [months]	10
B. Payments	
In the first month of training, the worker:	
Was paid	.198
Was unpaid	.515
Was paying the firm owner	.288
Earnings (conditional on > 0) [US$] (median)	39.2 (40.1)
Amount worker was paying to owner (conditional on > 0) [US$] (median)	51.9 (33.3)
C. Trainers	
Who was mainly involved in training the worker:	
Firm owner only	.457
Other employees only	.091
Firm owner as well as other employees	.452

Notes: The data is from the first firm follow-up, and the sample is restricted to those workers employed in Control firms. The sample includes 955 workers employed in 332 firms. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD. The top 1% monetary values are excluded.

Table 4: Skills

OLS regression coefficients, IPW estimates, robust standard errors in parentheses

	Firm-Provided Training		Sector-Specific Skills Test			Firm-Specific Skills
	Received OTJ-T at First Employer	Position in First Job is "Trainee"	Report Some Skills	ITT	LATE: Offered Worker-Firm Match	Skills Transferability
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Trained	.144***	.215***	.110***	2.13	4.49	-.051
	(.052)	(.041)	(.032)	(1.53)	(3.00)	(.109)
Vocationally Trained	-.029	-.019	.269***	6.96***	56.8***	.150**
	(.042)	(.025)	(.023)	(1.20)	(15.5)	(.068)
Mean (SD) Outcome in Control Group	.404	.092	.596	30.1 (22.9)	30.1 (22.9)	-
Control for Baseline Value	No	No	No	No	No	No
P-values on tests of equality:						
Firm Trained = Vocationally Trained	[.000]	[.000]	[.000]	[.001]	[.000]	[.025]
N. of observations	792	794	1,818	1,818	1,302	650

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the baseline, second and third worker follow-up survey. We report OLS regressions, where we use inverse probability weighting and robust standard errors are reported in parentheses. All regressions include strata dummies, survey wave dummies, a dummy for the implementation round and dummies for the month of interview. In Columns 1 and 2 we use information on the first employment spell reported by a worker in the post-intervention period (so the sample only includes workers that had at least one job in the post-intervention period). In Column 1 the dependent variable is a dummy=1 if the worker reported having received on the job training at her first employer. In Column 2 the dependent variable is a dummy=1 if the worker reported being a "Trainee" when asked about her position at her first employer. In Column 3 we report a linear probability model on whether the respondent reports having any sector specific skills or not at second and third follow-up. In Columns 4 and 5 the dependent variable is the skills test score, from the test administered to workers in the second worker follow-up. In Column 6 the dependent variable is based on a question on the perceived transferability of the skills learned at the current firm. This question is asked only to individuals who are working and is only available at third follow-up. The variable is standardized using the mean and standard deviation in Control. A higher value of the variable corresponds to more transferable skills. For the regressions in Columns 4 and 5 workers that reported not having any sector specific skills are assigned a test score equal to what they would have got had they answered the test at random. Workers that refused to take the skills test are excluded from the regressions in Columns 3-5. Column 4 reports OLS estimates, while in Column 5 we report 2SLS regressions, where we instrument treatment take-up with the original treatment assignment. The sample in Column 5 excludes individuals in T3 who received vocational training only. Take-up in is defined as the worker being offered the treatment, meaning the worker could be traced and the firm he/she was matched with expressed an interest in meeting him/her. We also control for the following baseline characteristics of workers: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, a dummy for whether the worker was employed at baseline, and a dummy for whether the worker scored at the median or above on the cognitive test administered at baseline. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the three follow-up survey rounds. At the foot of each Column we report p-values on the null that the impact of the vocational training is equal to the impact of firm training.

Table 5: Employment, Earnings and Productivity Bounds

OLS regression coefficients, IPW estimates, robust standard errors in parentheses

	Has done any work in the last month	Number of months worked in the last year	Hourly wage rate [USD]	Total earnings in the last month [USD]	Productivity Bounds [USD]
	(1)	(2)	(3)	(4)	(5)
Firm Trained	.063**	.518**	.028**	5.80**	[3.75, 18.2]
	(.025)	(.259)	(.012)	(2.53)	
Vocationally Trained	.090***	.879***	.031***	9.75***	[7.30, 27.7]
	(.020)	(.207)	(.009)	(2.01)	
Mean Outcome in Control Group	.438	4.52	.074	28.7	-
Control for Baseline Value	Yes	No	Yes	Yes	-
P-values on tests of equality:					
Firm Trained = Vocationally Trained	[.255]	[.134]	[.799]	[.111]	
N. of observations	3,256	3,256	3,099	3,111	-

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the baseline and first three follow-up worker surveys. We report OLS regressions, where we use inverse probability weighting and robust standard errors are reported in parentheses. All regressions control for the value of the outcome at baseline (except in Column 2), as well as strata dummies, survey wave dummies, a dummy for the implementation round and dummies for the month of interview. We also control for the following baseline characteristics of workers: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, a dummy for whether the worker was employed at baseline, and a dummy for whether the worker scored at the median or above on the cognitive test administered at baseline. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the two follow-up survey rounds. In Column 3 the hourly wage rate is computed as total earnings from any wage employment in the last month divided by total hours worked in wage employment in the last month (where total hours worked in the last month are computed using information on weekly hours and assuming these do not vary over the month). Individuals with no hours worked (and no wage employment income) in the last month are assigned a value of zero. In Column 4 the dependent variable is total earnings from any wage or self-employment in the last month. Individuals reporting no wage employment earnings and no self-employment earnings are assigned a value of zero. The top 1% of earnings values are excluded. In Column 5 we show the upper and lower bound of the productivity effect reported in Table A7 for the full sample of workers. We assume that the non-program earnings of those individuals who would have found employment regardless of the Treatment (the "always employed") are at least as high as the non-program earnings of the individuals that were induced by the Treatment to switch from unemployment to employment (the "compliers"), so that the lower bound of the productivity effect corresponds to the earnings effect (see Table A7 for more details). At the foot of each Column we report p-values on the null that the impact of the vocational training is equal to the impact of firm training. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Table 6: Worker Beliefs and Job Search

OLS regression coefficients, IPW estimates, robust standard errors in parentheses

	Job Offer Probability	Offered Wages			Search Intensity and Method	
	Expected probability of finding a job in the next 6 months (0 to 10 scale)	Minimum expected monthly earnings [USD]	Maximum expected monthly earnings [USD]	Average expected monthly earnings (triangular distribution) [USD]	Has actively looked for a job in the last year	Main channel through which looked for a job is formal [yes=1]
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Trained	.593***	.422	-.256	.528	.001	.014*
	(.137)	(2.24)	(4.11)	(3.34)	(.025)	(.008)
Vocationally Trained	1.87***	13.5***	23.2***	19.4***	.093***	.017**
	(.108)	(1.82)	(3.36)	(2.75)	(.020)	(.007)
Mean Outcome in Control Group	2.81	49.9	93.6	72.6	.509	.017
Control for Baseline Value	Yes	Yes	Yes	Yes	No	No
P-values on tests of equality:						
Firm Trained = Vocationally Trained	[.000]	[.000]	[.000]	[.000]	[.000]	[.661]
N. of observations	3,136	2,247	2,246	1,905	3,255	3,254

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the baseline and first three follow-up worker surveys. We report OLS regressions, where we use inverse probability weighting and robust standard errors are reported in parentheses. All regressions control for the value of the outcome at baseline (except in Columns 5 to 8), as well as strata dummies, survey wave dummies, a dummy for the implementation round and dummies for the month of interview. We also control for the following baseline characteristics of workers: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, a dummy for whether the worker was employed at baseline, and a dummy for whether the worker scored at the median or above on the cognitive test administered at baseline. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the three follow-up survey rounds. In Column 4 we assume a triangular distribution to calculate the average expected monthly earnings. At the foot of each Column we report p-values on the null that the impact of the vocational training is equal to the impact of firm training. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Table 7: Estimates of the Job Ladder Search Model

Two-step estimation procedure in Bontemps, Robin and van den Berg [2000]

		Steady State: November 2015 (Data from Second and Third Follow Up)		
Panel A: Parameter Estimates		**Control**	**Firm Trained**	**Vocationally Trained**
		(1)	**(2)**	**(3)**
Job destruction rate (monthly):	δ	.0272	.0259	.0239
		(.0030)	(.0037)	(.0021)
Arrival rate of job offers if UNEMPLOYED (monthly):	λ_0	.0189	.0191	.0237
		(.0019)	(.0024)	(.0019)
Arrival rate of job offers if EMPLOYED (monthly):	λ_1	.0388	.0376	.0428
		(.0096)	(.0117)	(.0080)
Interfirm competition for workers	κ_1	1.426	1.452	1.791
Unemployment Rate	u	**.5892**	**.5755**	**.5024**
% Impact:			**2.3%**	**14.7%**
Panel B: Function and Income Estimates				
Average (sd) monthly OFFERED wage [USD]	***F(.)***	44.8	47.0	46.3
		(37.4)	(43.6)	(41.9)
Average (sd) monthly ACCEPTED wage [USD]	***G(.)***	63.7	68.9	70.6
		(45.5)	(54.5)	(54.4)
Treatment Effect Impact on Annual Income [USD]			**37.0**	**107.6**
% Impact:			**11.8%**	**34.3%**
Panel C: Firm Productivity Distribution				
Average (sd) firm productivity	***P(.)***	174.8	122.9	271.8
		(610.7)	(452.9)	(1117.4)
% Impact:			**-30%**	**55%**

Notes: The Vocationally Trained group combines both T3 and T4. The dataset is a cross-section of workers, and for each worker it contains information on: spell type (employment, unemployment), spell duration (in months), earnings in employment spells (in USD), dates of transitions between spells and type of transition: (i) job to unemployment, (ii) unemployment to job, or (iii) job to job. Wages are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD. The dataset contains at most two spells (and one transition) per individual. The data comes from the second and third follow-up survey of workers, and the initial spell is identified as the (employment or unemployment) spell that was ongoing in November 2015. Spells are right censored at the date of interview. Spells are left censored at 1 August 2014. Casual and agricultural occupations are coded as unemployment. Self-employment is coded as employment (but self-employment spells are assigned a separate spell). The estimation protocol follows the two-step procedure in Bontemps, Robin and van den Berg [2000]: in the first step the G function is estimated non-parametrically from the data (so this is just the empirical CDF of observed wages for those workers that are employed in their first spell), and is then substituted into the likelihood function. In the second step, maximum likelihood is then conducted using information from both the first and second spells for each individual to recover the parameter estimates. In Panel B, the average yearly income is defined as the average monthly accepted wage (in levels) multiplied by the unemployment rate and multiplied by 12. The impact of the program on average income is defined as the difference in average income between the respective treatment groups and the control group. In Panel C, the productivity distribution is the one implied by the job ladder model. This is the (exogenous) productivity distribution needed to rationalize the observed distribution of wages in the model. A higher value of productivity means that the firm is more productive.

Table 8: Firm-Side Selection in Initial Worker-Firm Matches

OLS regression coefficients, standard errors clustered by sector-branch in parenthesis

	Firm was Interested in Meeting a Matched Worker		
	(1)	**(2)**	**(3)**
Matched to Untrained Worker, Wage Subsidy Offer	.308***	.304***	.720***
	(.047)	(.046)	(.202)
Firm size		.021	.033
		(.018)	(.020)
Log profits per worker		-.067	.004
		(.049)	(.056)
Log capital per worker		.074**	.056
		(.037)	(.045)
Matched to Untrained Worker, Wage Subsidy Offer x Firm Size			-.034
			(.027)
Matched to Untrained Worker, Wage Subsidy Offer x Log profits per worker			-.259**
			(.112)
Matched to Untrained Worker, Wage Subsidy Offer x Log capital per worker			.056
			(.087)
Mean outcome in T5: Untrained, Matched group		.085	
Number of observations (firms)		602	

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The Table reports results from OLS regressions. Firms in the Control group are not included in the analysis. All independent variables are measured at baseline. All regressions further include the following baseline controls: dummy for whether owner is female, years of education of owner, sector and branch dummies. The excluded category in the treatment dummies is the Match group. The top 1% monetary values are excluded. Manufacturing sectors are: motor-mechanics, plumbing, construction, electrical wiring and welding. Service sectors are: hairdressing, catering and tailoring.

Table 9: Firm Side Outcomes, LATE Estimates

2SLS regression coefficients, and standard errors clustered by sector-branch in parenthesis

	At Least One Worker Hired and Retained	Number of Employees	Log (Average Monthly Profits)	Log (Net Investment)
	(1)	(2)	(3)	(4)
Matched to Untrained Worker, Wage Subsidy Offer	.088***	-.055	.114**	.032
	(.017)	(.186)	(.053)	(.112)
Mean of outcome in T1 Control group	.000	2.34	191	60.8
Controls for baseline value of outcome	No	Yes	Yes	No
Number of firms	657	657	591	651

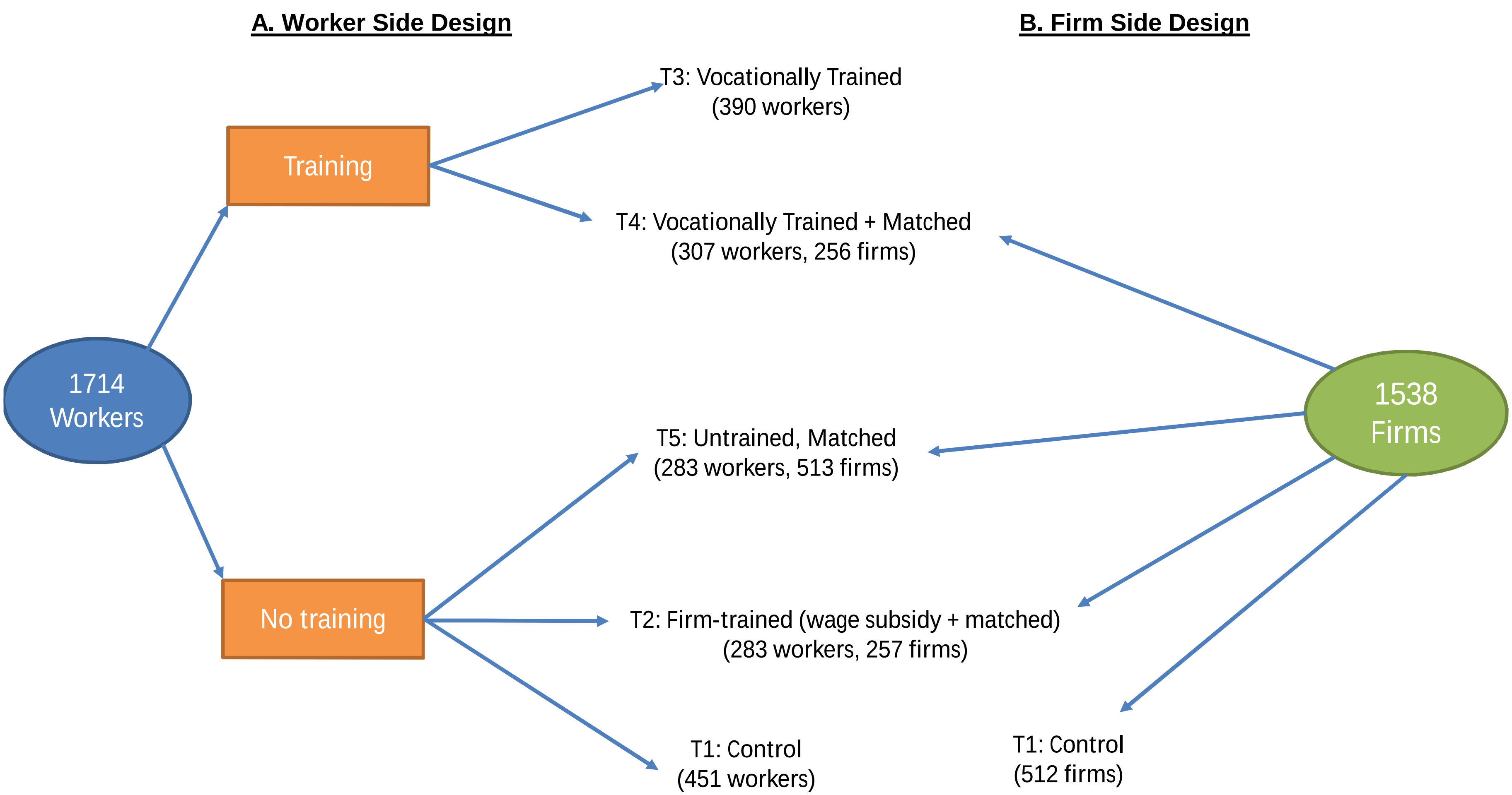
Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. We report results from regressions run on a panel dataset including four rounds of follow-up data, and controlling for the value of the outcome at baseline when available. We show 2SLS regressions coefficients, where we instrument treatment take-up with the original treatment assignment, and standard errors in parentheses (clustered by sector-BRAC branch). Observations are weighted by IPW. The excluded category in the treatment dummies is the Control group. Take-up is equal to one if the firm was successfully contacted by the implementation team and offered to meet a matched worker. All regressions include baseline controls, branch and trade fixed-effects, survey wave fixed effects and dummies for month of interview. Baseline controls include owner's sex, owner's years of education and firm size at baseline. Standard errors are adjusted for heteroskedasticity and clustered at the branch-trade level. The instruments for the IPW estimates are a dummy for whether the owner reported at baseline an intention to relocate in the future, and the number of network firms reported at baseline. All monetary amounts are deflated and expressed in terms of the price level in January 2013 using the monthly Producer Price Index for the manufacturing sector (local market), published by the Uganda Bureau of Statistics. The monetary amounts are then converted in January 2013 USD (1USD=2385UGX). Monthly profits are truncated at 99th percentile. Net investments are truncated at the bottom percentile and at the 99th percentile. The dependent variable in Column 3 is the log of one plus average monthly profits. The dependent variable in Column 4 is the log of one plus net investment.

Table 10: Internal Rate of Return

	Firm Trained	Vocationally Trained
	(1)	(2)
Social discount rate = 5%		
Remaining expected productive life of beneficiaries	38 years	38 years
Panel A. External parameters		
Total cost per individual at year 0 [USD]:	**368**	**510**
(i) Training costs (for 6 months)	302	470
(ii) Program overheads costs	31	4
(iii) Foregone earnings (for 6 months) - average at baseline	36	36
Panel B. Estimated total earnings benefits		
1 NPV change in total earnings year 1 and beyond-forever (from structural model)	624	1815
2 Benefits/cost ratio	**1.69**	**3.56**
Social discount rate = 10%	0.98	2.06
3 Internal Rate of Return (IRR)	**0.098**	**0.211**
Panel C. Sensitivity		
Sensitivity to different expected remaining productive life of beneficiaries		
Remaining expected productive life = 20 years	**0.078**	**0.206**
Remaining expected productive life = 10 years	**0.001**	**0.166**
Sensitivity to different earnings		
Foregone earnings = 90th percentile at baseline (120USD)	**0.077**	**0.181**
Foregone earnings = double 90th percentile at baseline (241USD)	**0.057**	**0.150**
Foregone earnings = max earnings at baseline (794USD)	**0.012**	**0.080**
Foregone earnings = double max earnings at baseline (1588USD)	**NA**	**0.041**
Panel D. Programme Costs for IRR to equate social discount rate		
5 Total cost per individual at year 0 [USD]	**624**	**1814**
Sensitivity to different discount rates/time horizons		
Social discount rate = 10%	360	1047

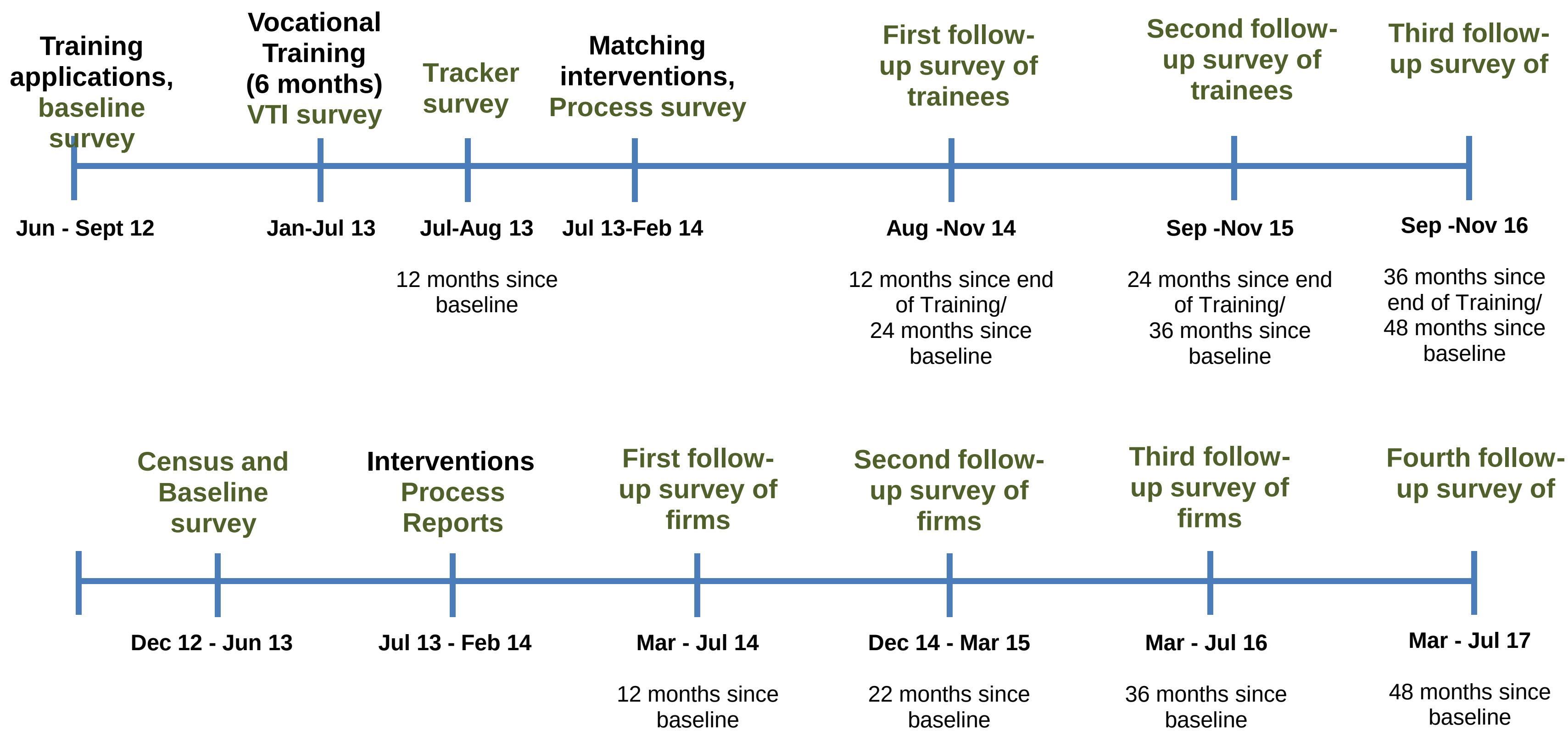
Notes: The Vocationally Trained group combines both T3 and T4. Forgone earnings are calculated as the average monthly earnings at baseline (6 USD) multiplied by six (as the duration of both types of training was six months). The remaining expected productive life of beneficiaries is calculated as average life expectancy in Uganda in 2012 (58 years) minus average age in the sample at baseline (20 years). The computation of the IRR uses as input for the benefit the treatment effect impact on annual income from the structural model. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Figure 1: Experimental Design



Note: Numbers in parentheses refer to the number of applicants originally assigned to each treatment, and the number of firms assigned to each treatment.

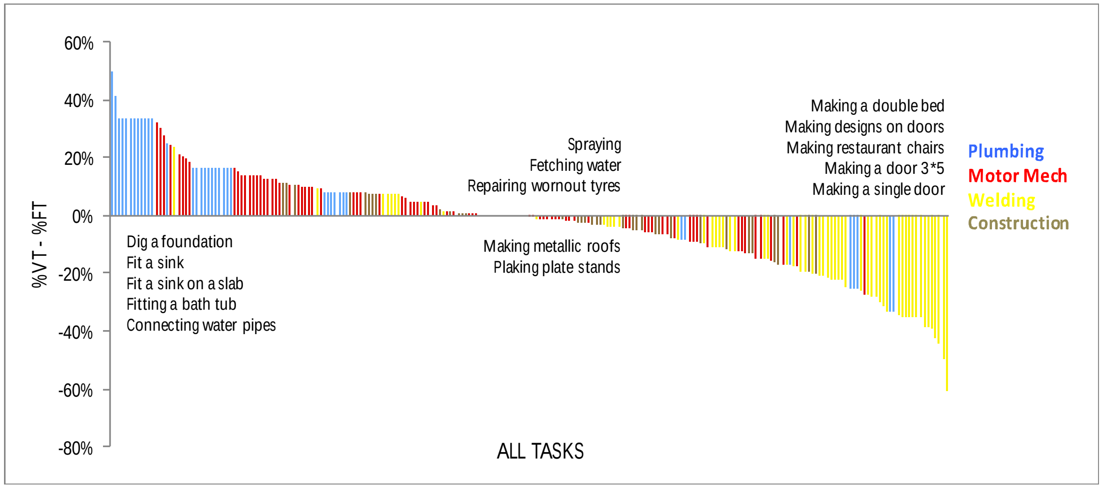
Figure 2: Timeline



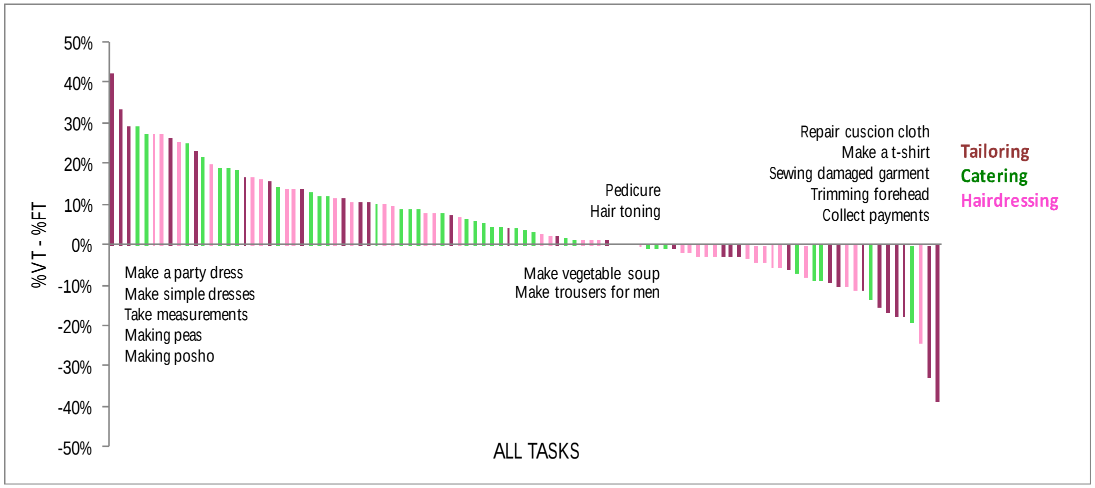
Notes: The timeline highlights the dates relevant for the main batch of worker applications and baseline surveys. A second smaller round of applications and baseline surveys were conducted in May and June 2013. The majority of trainees from the first round of applicants started training in January 2013, as shown in the timeline. For logistical reasons, a smaller group received training between April and October 2013. The trainees from the second round of applications received vocational training between October 2013 and April 2014. VTI surveys were collected towards the end of the training period while trainees were still enrolled at the VTIs. Workers from the second round of applicants were not included in the Tracker Survey. The remaining interventions (the matching treatments and firm training placements) and all follow-up surveys were conducted at the same time for workers from the first and second round of applicants. On the firms' timeline, the firm level interventions include: Matching, Vocational Training + Matching, and Firm Training. There were two rounds of Matching and Vocational Training + Matching interventions, in line with the two batches of trainees from the vocational training institutes. The first round of the Vocational training + Matching interventions took place in August-September 2013. The second round took place in December 2013-February 2014. The Firm Training intervention took place in September-November 2013.

Figure 3: Tasks Performed by Vocationally Trained and Firm Trained Workers

Panel A: Manufacturing



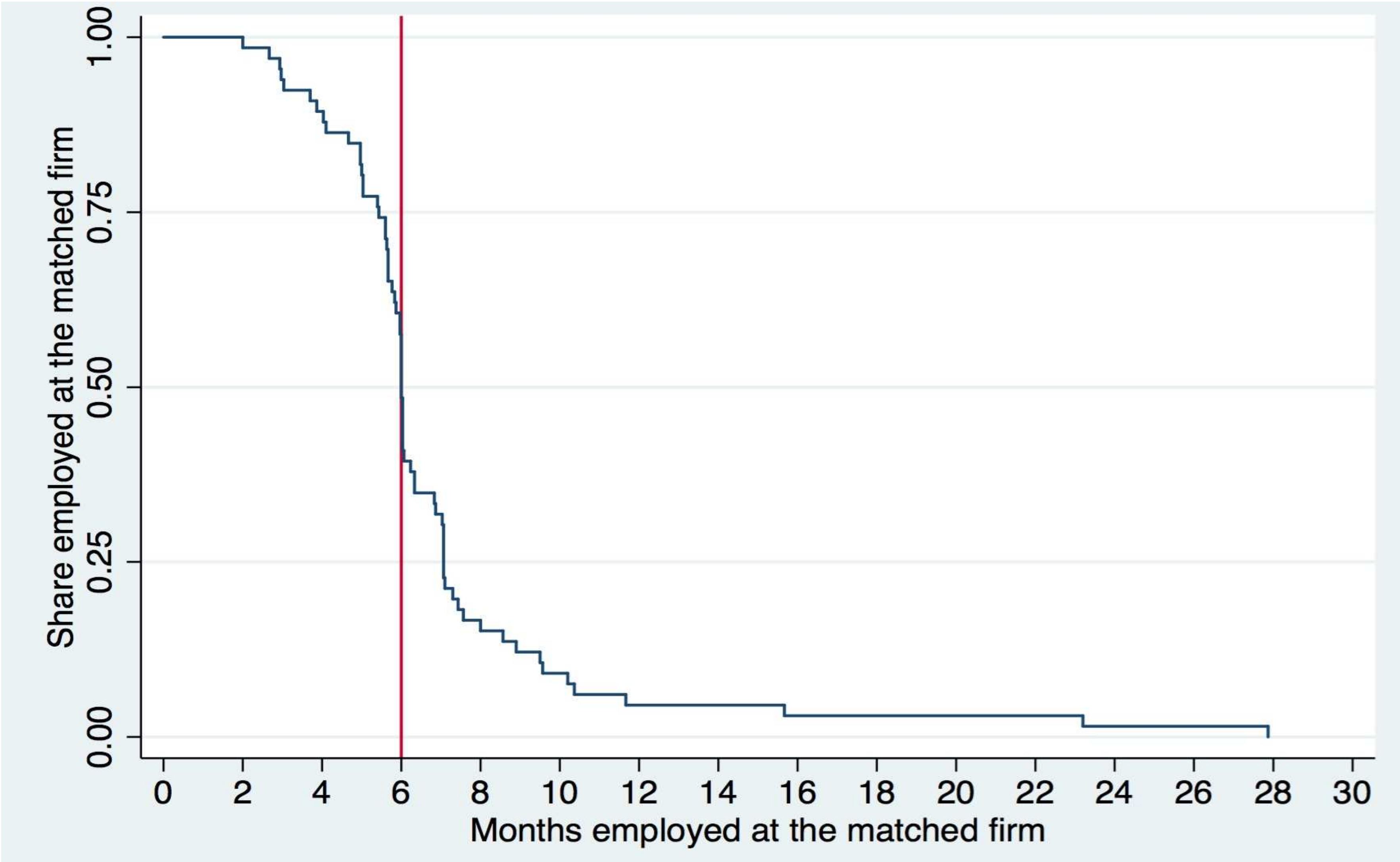
Panel B: Services



Notes: The Figures plot the difference in the percentage of workers performing each given task while employed, between workers who were assigned to receive Vocational Training (T3 and T4) , and workers who were assigned to receive Firm Training (T2). The data refers to all main job spells reported at third follow-up (so there is one job spell per worker and only employed individuals are included in the sample), where workers were asked to report which tasks they performed in each employment spell they had in the year prior to the survey (the tasks had to be indicated from an initial list compiled together with BRAC). Manufacturing sectors are: motor-mechanics, plumbing, construction, electrical wiring and welding. Service sectors are: hairdressing, catering and tailoring.

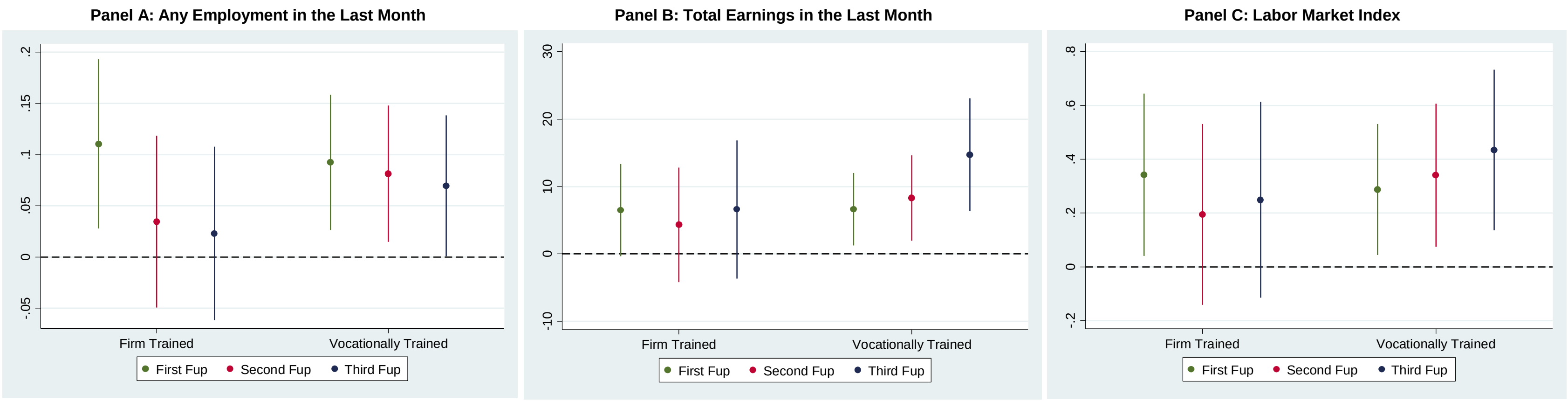
Figure 4: Survival Function for Workers in the Firm Training Treatment

Matched Workers Who Started Employment at a Firm Offered the Wage Subsidy



Notes: The Figure plots the survival function for workers in the firm-training treatment who started a job at the matched firm. The Figure is based on 67 workers (information on employment duration at the matched firm is missing for two workers).

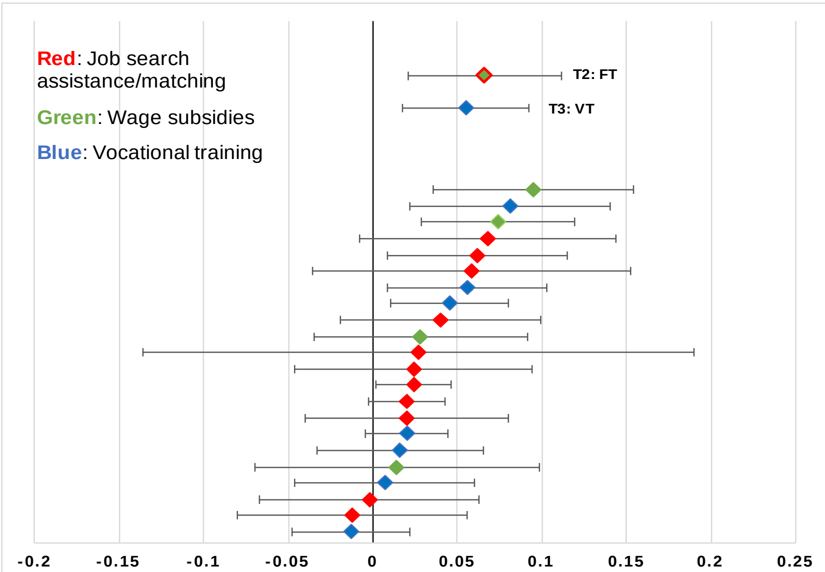
Figure 5: Treatment Effects on Labor Market Outcomes, by Survey Wave



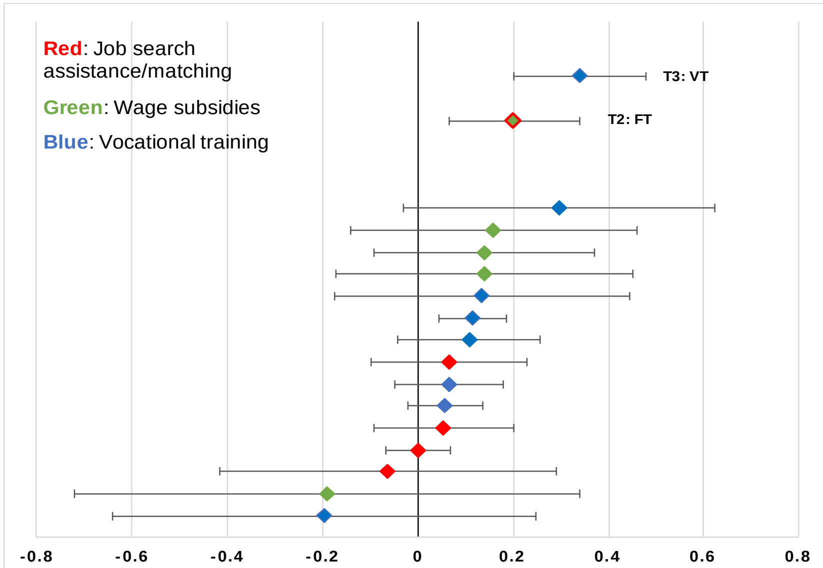
Notes: The data used is from the baseline and first three follow-up worker surveys. We report OLS regression coefficients and 95% confidence intervals, where we use inverse probability weighting and robust standard errors. Impacts on each outcome are estimated from the same panel regression using all three follow-up waves, where we interact the treatment indicators with dummies for the different follow-up waves. The Labor Market Index is computed using the following variables: any paid work in the last month (dummy), any wage employment in the last month (dummy), any casual work in the last month (dummy), hours worked in wage employment last week, hourly wage rate, total earnings in the last month. Hourly earnings and total earnings are set to zero for workers with no earnings. The index is constructed by converting each component into a z-score, averaging these and taking the z-score of the average. z-scores are computed using means and standard deviations from the control group at baseline. All regressions control for the value of the outcome at baseline, as well as strata dummies, a dummy for the implementation round and dummies for the month of interview. We also control for the following baseline characteristics of workers: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, a dummy for whether the worker was employed at baseline, and a dummy for whether the worker scored at the median or above on the cognitive test administered at baseline. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the two follow-up survey rounds.

Figure 6: Comparison of Treatment Impacts to Meta-analysis by McKenzie [2017]

Panel A: Employment impacts



Panel B: Earnings impacts



Notes: The Figures compare the treatment impacts from this study to the treatment impacts reported in the meta-analysis by McKenzie [2017]. The green estimates correspond to wage subsidy programs, the blue estimates to vocational training programs, and the red estimates to job search and matching assistance programs. Panel A reports treatment impacts (ITT) on the probability of paid employment, together with 95% confidence intervals. The estimates from our study are taken from Column 2 of Table 4, where we use as outcome variable "Any wage employment in the last month". Alongside our estimates, Panel A further reports 22 estimates of treatment impacts taken from Table 1, 3 and 4 of McKenzie [2017]. These correspond to all the available program estimates for this outcome reported in McKenzie [2017], a part from the estimate from Galasso et al. [2004], which is omitted as no standard error is provided, and the estimate from Groh et al. [2016] with time frame 6 months, as that is estimated while the wage subsidy was still ongoing (while our estimates for T2: FT and all the other estimates for wage subsidy programs reported in the Figure refer to the period after the wage subsidy ended). Panel B reports treatment impacts (ITT) on earnings, in terms of percentage increase relative to the earnings level of the Control group, together with 95% confidence intervals. The estimates from our study are taken from Column 4 of Table 5, where we use as outcome variable "Total earnings in the last month". Alongside our estimates, Panel B further reports 15 estimates of treatment impacts taken from Table 1, 3 and 4 of McKenzie [2017]. These correspond to all the available program estimates for this outcome reported in McKenzie [2017], apart from the estimate from Groh et al. [2016] with time frame six months, as that is estimated while the wage subsidy was still ongoing (while our estimates for T2: FT and all the other estimates for wage subsidy programs reported in the Figure refer to the period after the wage subsidy ended), and the estimate from Maitra and Mani [2016], which is excluded as that is very large relative to all the other estimates: Maitra and Mani [2016] estimate a treatment impact on earnings of .957, with confidence interval [.056 ; 1.86]. However, this corresponds to only a $2.40 monthly increase in earnings in absolute terms, and so the large treatment impact is due to the women in their sample having extremely low earnings to begin with.

Table A1: Baseline Balance on Worker Characteristics

Means, robust standard errors from OLS regressions in parentheses
P-value on t-test of equality of means with control group in brackets
P-value on F-tests in braces

	Number of workers	Age [Years]	Married	Has child(ren)	Currently in school	Ever attended vocational training	F-test of joint significance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All Workers	1714	20.0	.040	.118	.016	.036	
		(.198)	(.016)	(.024)	(.008)	(.018)	
T1: Control	451	20.1	.027	.102	.011	.042	
		(.211)	(.016)	(.025)	(.009)	(.020)	
T2: Firm Trained	283	20.1	.040	.121	.018	.038	{.999}
		(.139)	(.014)	(.024)	(.009)	(.015)	
		[.970]	[.271]	[.260]	[.576]	[.897]	
T3: Vocationally Trained	390	20.0	.056*	.127	.018	.032	{.882}
		(.134)	(.014)	(.022)	(.008)	(.013)	
		[.781]	[.056]	[.339]	[.553]	[.461]	
T4: Vocationally Trained + Matched	307	20.0	.030	.123*	.029	.038	{.845}
		(.146)	(.012)	(.023)	(.011)	(.015)	
		[.975]	[.128]	[.075]	[.248]	[.792]	
T5: Untrained, Matched	283	20.0	.047*	.122	.007	.027	{.875}
		(.148)	(.015)	(.024)	(.007)	(.014)	
		[.429]	[.084]	[.201]	[.468]	[.359]	
F-test of joint significance		{.933}	{.243}	{.449}	{.445}	{.752}	

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. All data is from the baseline survey to workers. Column 1 reports the number of workers assigned to each treatment. Columns 2 to 6 report the mean value of each worker characteristic, derived from an OLS regression of the characteristic of interest on a series of dummy variables for each treatment group. All regressions include strata dummies and a dummy for the implementation round. The excluded (comparison) group in these regressions is the Control group. Robust standard errors are reported throughout. Column 7 reports the p-values from F-Tests of joint significance of all the regressors from an OLS regression where the dependent variable is a dummy variable taking value 0 if the worker is assigned to the Control group, and it takes value 1 for workers assigned to treatment group j (with j going from 2 to 5) and the independent variables are the variables in Columns 2 to 6. Robust standard errors are also calculated in these regressions. The p-values reported in the last row are from the F-test of joint significance of the treatment dummies in each Column regression where the sample includes all workers.

Table A2: External Validity

Means, standard deviations in parentheses

	Number of individuals	Age [Years]	Gender [Male=1]	Married	Currently in school	Ever attended vocational training	Has worked in the last week [Yes=1]	Has had any wage employment in the last week	Has done any casual work in the last week	Total earnings from wage employment in the last month [USD]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Baseline, aged 18-25	1,608	20.1	.567	.038	.014	.037	.362	.142	.156	2.60
		(1.86)	(.496)	(.190)	(.116)	(.189)	(.481)	(.350)	(.363)	(9.74)
Uganda National Household Survey 2012/13:										
B. All, aged 18-25	4,696	21.1	.465	.395	.309	.062	.681	.241	.514	9.13
		(2.32)	(.499)	(.489)	(.462)	(.241)	(.466)	(.428)	(.500)	(28.2)
C. Labor Market Active, aged 18-25	3,456	21.4	.475	.448	.207	.064	.902	.320	.682	12.2
		(2.33)	(.499)	(.497)	(.405)	(.245)	(.297)	(.467)	(.467)	(32.0)

Notes: We present characteristics of individuals from three samples: (i) those individuals in our baseline sample aged 18-25; (ii) individuals aged 18-25 and interviewed in the Uganda National Household Survey 2012/13 (UNHS) conducted by the Ugandan Bureau of Statistics; (iii) individuals aged 18-25 and interviewed in the UNHS who self-report being active in the labor market (either because they are employed or actively seeking employment). The UNHS was fielded between June 2012 and June 2013. Our baseline survey was fielded between June and September 2012. In the UNHS respondents are considered to have attended vocational training if the highest grade completed is post-primary specialized training/diploma/certificate or post-secondary specialized training/diploma/certificate. In the baseline survey questions on employment status did not refer to work activities performed in the last week, but to work activities performed at the time of the survey. Therefore, for the baseline survey the variable "Has worked in the last week" corresponds to the worker being "Currently employed or involved in a work activity". Similarly, Columns 8-10 for the baseline survey are based on the most recent activity performed by the individual, conditional on him/her saying to be currently employed or involved in a work activity. For UNHS, the outcomes in Columns 8-10 are based on the main activity performed in the week before the survey. In Column 9 casual work includes occupations that are casual in nature, as well as agricultural occupations. In Column 10 workers who report doing no wage employment in the past month (or only did unpaid work in the last month) have a value of zero for total earnings.

Table A3: Take-Up of Treatments

	Vocational Training		Matching and Firm Training			
Sample of Workers:	**All Workers**	**Offered Training**	**All Workers**	**Invited to interview**	**Met at least one Firm**	**Worker received a Job Offer**
Outcome:	**% Workers Offered Training**	**% Workers Trained**	**% Workers Invited to Interview**	**% Workers That Met at Least One Firm**	**% Workers Who Received a Job Offer**	**% Workers Hired**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
T3: Vocationally Trained	**97.9**	73.8	-	-	-	-
T4: Vocationally Trained + Matched	**95.4**	63.1	**12.7**	74.4	58.6	23.5
T2: Firm Trained	-	-	**50.5**	80.4	90.4	66.4
T5: Untrained, Matched	-	-	**19.1**	85.2	34.8	18.8

Notes: The data used is from the tracker survey and process reports. The tracker survey was collected in July-August 2013, at the end of the main round of vocational training. Process reports were collected during the implementation of the firm-level interventions (September 2013-February 2014). In Columns 1 and 3 the sample includes all workers assigned to the respective treatment groups. In Column 1 only workers that were traced and successfully informed about the treatment offer are considered as having been offered treatment. In Columns 2 the sample includes those workers who could be traced and were offered the treatment by BRAC staff, and the percentage of workers who took up training includes the workers who completed the 6 months vocational training. For Matching and Firm Training (Column 3) the treatment offer is defined as firms having invited the worker for an interview (so those workers matched to firms that were not interested in the program are not included, as they were not offered treatment). In Column 4 the sample includes workers who were invited for an interview, in Column 5 it includes those workers who met with at least one firm, in Column 6 the sample includes workers who received an offer to start at the firm. In Column 6 the percentage of workers who took up treatment is calculated as the percentage of workers who accepted the offer received by the firm, and so started work/training at the firm.

Table A4: Attrition

OLS regression coefficients, robust standard errors in parentheses

	Worker attrited by endline	
	With covariates	Heterogeneous
	(1)	(2)
T2: Firm Trained	-.000	.002
	(.026)	(.035)
T3: Vocationally Trained	-.018	.022
	(.024)	(.034)
T4: Vocationally Trained + Matched	-.011	-.012
	(.027)	(.036)
T5: Untrained, Matched	.013	.014
	(.027)	(.035)
High Score on Cognitive Test at Baseline [Yes=1]	.045**	.061*
	(.018)	(.032)
T2: Firm Trained X High Cognitive Score		-.005
		(.051)
T3: Vocationally Trained X High Cognitive Score		-.071
		(.047)
T4: Vocationally Trained + Matched X High Cognitive Score		.001
		(.051)
T5: Untrained, Matched X High Cognitive Score		-.002
		(.053)
Mean of outcome in T1 Control group	.134	.134
Strata and Implementation round dummies	Yes	Yes
Other baseline characteristics	Yes	Yes
Test of joint significance of baseline characteristics		
F-statistic	2.35	1.57
P-value	.071	.196
Test of joint significance of Treatment X High Score interactions		
F-statistic		.79
P-value		.529
Number of observations (workers)	1,561	1,561

Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. Data is from baseline, first, second and third follow-up of applicants to the vocational scholarships. Standard errors are adjusted for heteroskedasticity in all regressions. Other baseline characteristics include: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, and a dummy for whether the worker was employed at baseline. The variable High Score on Cognitive Test at Baseline is a dummy=1 if the applicant scored at the median or above on the cognitive test administered with the baseline survey.

Table A5: Sector Skills Test for Motor Mechanics

<table>
<tr><td colspan="3">1. MOTOR-MECHANICS</td></tr>
<tr><td colspan="2"></td><td></td></tr>
<tr><td>1</td><td>multiple-choice
What are you advised to do when servicing the engine by changing oil?</td><td>A. Top up lubricating oil
B. Replace oil filter
C. Over hand engine
D. Over hand cylinder head
Correct Answer: B</td></tr>
<tr><td>2</td><td>multiple-choice
What immediate remedy can you give to a vehicle with a problem of excessive tyre wear in the center more than other parts?</td><td>A. Increase tyre pressure
B. Reduce tyre pressure
C. Inflate pressure
D. Remove the vehicle tire
Correct Answer: B</td></tr>
<tr><td>3</td><td>multiple-choice
If a customer reports to you that his/her vehicle charging system works at lower rate, how can you help him?</td><td>A. Replacing the charging system
B. Adjusting the alternator tension
C. Replacing alternator housing
D. Renewing wire insulator
Correct Answer: B</td></tr>
<tr><td>4</td><td>multiple-choice
Which of the following set of systems or component call for mechanical adjustment during general vehicle service?</td><td>A. Tyres, cooling system, master cylinder
B. Break shoes, alternator, and valve clearance
C. Distributor, radiator, propeller shaft
D. Tank, crank shaft, Turbo charger
Correct Answer: B</td></tr>
<tr><td>5</td><td>multiple-choice
What solution would you give a customer with a vehicle engine producing blue smoke?</td><td>A. Top up lubricant
B. Time the engine
C. Replace piston rings
D. Remove carbon deposits
Correct Answer: C</td></tr>
<tr><td>6</td><td>matching
What should you do to stop the following vehicle troubles?</td><td><table>
<tr><td>1</td><td>Battery over charging</td><td>A</td><td>Leaking fuel tank</td></tr>
<tr><td>2</td><td>Engine over heating</td><td>B</td><td>Renew regulator</td></tr>
<tr><td>3</td><td>Lubricant leakage</td><td>C</td><td>Reduce oil to the correct level</td></tr>
<tr><td>4</td><td>Smoke in exhaust</td><td>D</td><td>Renew piston rings</td></tr>
<tr><td>5</td><td>Engine fails to start</td><td>E</td><td>Charge the battery</td></tr>
</table>
Correct Answer: 1B, 2A, 3C, 4D, 5E</td></tr>
<tr><td>7</td><td>order
When changing engine oil, in which order should you perform the following steps?</td><td>A. Drain oil through drain plug
B. Remove oil filter cup
C. Run engine to check leaks
D. Fill new oil through filler cup to level
E. Remove oil filter
F. Warm up the engine
Correct Answer: B, E, A, D, F, C</td></tr>
</table>

Table A6: Worker Expectations

Means, standard deviations in parenthesis

All amounts in 2012 USD

	Expected probability of finding a job in the next 12 months		Average expected monthly earnings (triangular distribution)	
	With Current Skill Set	**If Received VT**	**With Current Skill Set**	**If Received VT**
	(1)	**(2)**	**(3)**	**(4)**
All Workers (Baseline Interview)	.567	.867	57.8	118.3
	(.288)	(.144)	(46.9)	(71.5)
N. of observations	1,611	1,589	1,243	1,411

Notes: Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the baseline and first three follow-up worker surveys. Columns 1 to 4 report the mean and standard deviation (in parentheses) of the average expected probability of finding a job and the average monthly earnings (assuming a triangular distribution of expected earnings) with the current skill set (columns 1 and 3), or if the worker were to receive vocational training (columns 2 and 4). This is based on all workers interviewed at baseline (across all treatments). All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD. The top 1% values of each variable are excluded from the analysis.

Table A7: Productivity Bounds

	Treatment impact on earnings	Treatment impact on paid employment	Earnings in the control group		Earnings effect	Composition effect		Productivity effect	
			10th percentile	90th percentile		Lower Bound	Upper Bound	Lower Bound	Upper Bound
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Firm Trained	5.80	0.06	19.35	139.29	3.752	-14.451	14.451	-10.699	18.203
Vocationally Trained	9.75	0.09	19.35	139.29	7.297	-20.445	20.445	-13.148	27.742
Mean of outcome in T1 Control Group	28.7	0.44							

Notes: We follow Angrist et al. [2006] and Attanasio et al. [2011] for the computation of the bounds on productivity. Specifically, we decompose the overall treatment effect on earnings into: (i) an extensive margin effect on employment; (ii) a composition effect, corresponding to the impact of the treatment on the selection of workers into employment; and (iii) a productivity effect, namely the causal impact on earnings for those employed. Column 1 reports the treatment effects on total earnings in the last month, and the estimates are taken from Column 4 of Table 6. Column 2 reports the extensive margin impact on paid employment, and the estimates are taken from Column 1 of Table 5. Following Attanasio et al. [2011], we define the overall treatment impact on earnings reported in Column 1 as: $E(LS|T = 1) - E(LS|T = 0)$, where T is the treatment variable (either standing for VT or FT), S stands for total earnings in the last month, and L is one for individuals in paid employment and zero for individuals not in paid employment. Also, we define the impact on paid employment reported in Column 2 as: $Pr(L = 1|T = 1) - Pr(L = 1|T = 0)$. The earnings effect is the change in the earnings of those employed, defined as: $E(S|L = 1,T = 1) - E(S|L = 1,T = 0)$. Following Attanasio et al. [2011], we can rewrite the earnings effect as: $E(S|L = 1,T = 1) - E(S|L = 1,T = 0) = \{[E(LS|T = 1) - E(LS|T = 0)] - [Pr(L = 1|T = 1) - Pr(L = 1|T = 0)] * E(S|T = 1)/Pr(L = 1|T = 1)\} / Pr(L = 1|T = 0)$. We use the treatment effects and control group means reported in Columns 1 and 2 as inputs in this formula to calculate the estimated earnings effects reported in Column 5. Following again Attanasio et al. [2011], the bounds for the composition effect are calculated as: (i) upper bound: $\{(E(S(p(0.90)) - E(S(p(0.10))) * [Pr(L = 1|T = 1) - Pr(L = 1|T = 0)] / Pr(L = 1|T = 1)\}$ and: (ii) lower bound $\{(E(S(p(0.10)) - E(S(p(0.90))) * [Pr(L = 1|T = 1) - Pr(L = 1|T = 0)] / Pr(L = 1|T = 1)\}$, where E(S(p(90))) and (S(p(10))) are, respectively, the mean total earnings for those in the 90th and 10th quantile in the control group (these quantiles are reported in Columns 3 and 4). Columns 6 and 7 then report bounds on the composition effect calculated using this formula. The upper bound of the productivity effect is the difference between the earnings effect and the lower bound of the composition effect. The lower bound of the productivity effect is the difference between the earnings effect and the upper bound of the composition effect. These are reported in Columns 8 and 9. Attanasio et al. [2011] further show that if we assume that the non-program earnings of those individuals who would have found employment regardless of the Treatment (the "always employed") are at least as high as the non-program earnings of the individuals that were induced by the Treatment to switch from unemployment to employment (the "compliers"), then the earnings effect corresponds to the lower bound of the productivity effect (see Attanasio et al. [2011] for more details). In Columns 1 and 2, earnings for individuals out of paid employment are equal to zero. In Columns 3 and 4, earnings for individuals out of paid employment are excluded. All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Table A8: Robustness Checks

Dependent Variable: Labor Market Index

OLS regression coefficients, IPW estimates in Columns 1 to 7, robust standard errors in parentheses

Lee [2009] Bounds in brackets

Bootstrap pvalues in braces: unadjusted p-values (left) and Romano and Wolf [2016] adjusted p-values (right)

	(1) All	(2) Women	(3) Men	(4) Services	(5) Manufacturing	(6) Non-Kampala	(7) Batches	(8) No Covariates	(9) No IPW, No Covariates
Firm Trained	.264***	.226	.311**	.153	.366***	.351***	.265***	.269***	.280***
	(.098)	(.138)	(.137)	(.144)	(.136)	(.111)	(.098)	(.099)	(.098)
	[.140 ; .295]	[.267 ; .306]	[.055 ; .289]	[.087 ; .254]	[.165 ; .327]	[.251 ; .341]			
	{.008 ; .008}	{.110 ; .110}	{.025 ; .025}	{.291 ; .291}	{.008 ; .008}	{.001 ; .001}	{.007 ; .014}	{.007 ; .007}	{.006 ; .006}
Vocationally Trained	.353***	.313***	.382***	.289**	.398***	.380***	.372***	.354***	.364***
	(.080)	(.110)	(.115)	(.117)	(.111)	(.088)	(.087)	(.080)	(.080)
	[.309 ; .425]	[.276 ; 429]	[.332 ; .424]	[.188 ; .416]	[.360 ; .412]	[.287 ; .462]			
	{.000 ; .000}	{.004 ; .007}	{.001 ; .002}	{.015 ; .025}	{.000 ; .001}	{.000 ; .000}	{.000 ; .000}	{.000 ; .000}	{.000 ; .000}
Vocationally Trained x Second Batch of Trainees							-.108		
							(.183)		
							{1.00 ; 1.00}		
Mean Outcome in Control Group	.836	.541	1.06	.542	1.04	.780	.836	.836	.836
Control for Baseline Value	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P-values on tests of equality:									
Firm Trained = Vocationally Trained	[.339]	[.506]	[.592]	[.319]	[.802]	[.782]	[.281]	[.365]	[.369]
N. of observations	2,830	1,249	1,581	1,161	1,658	2,244	2,830	2,830	2,830

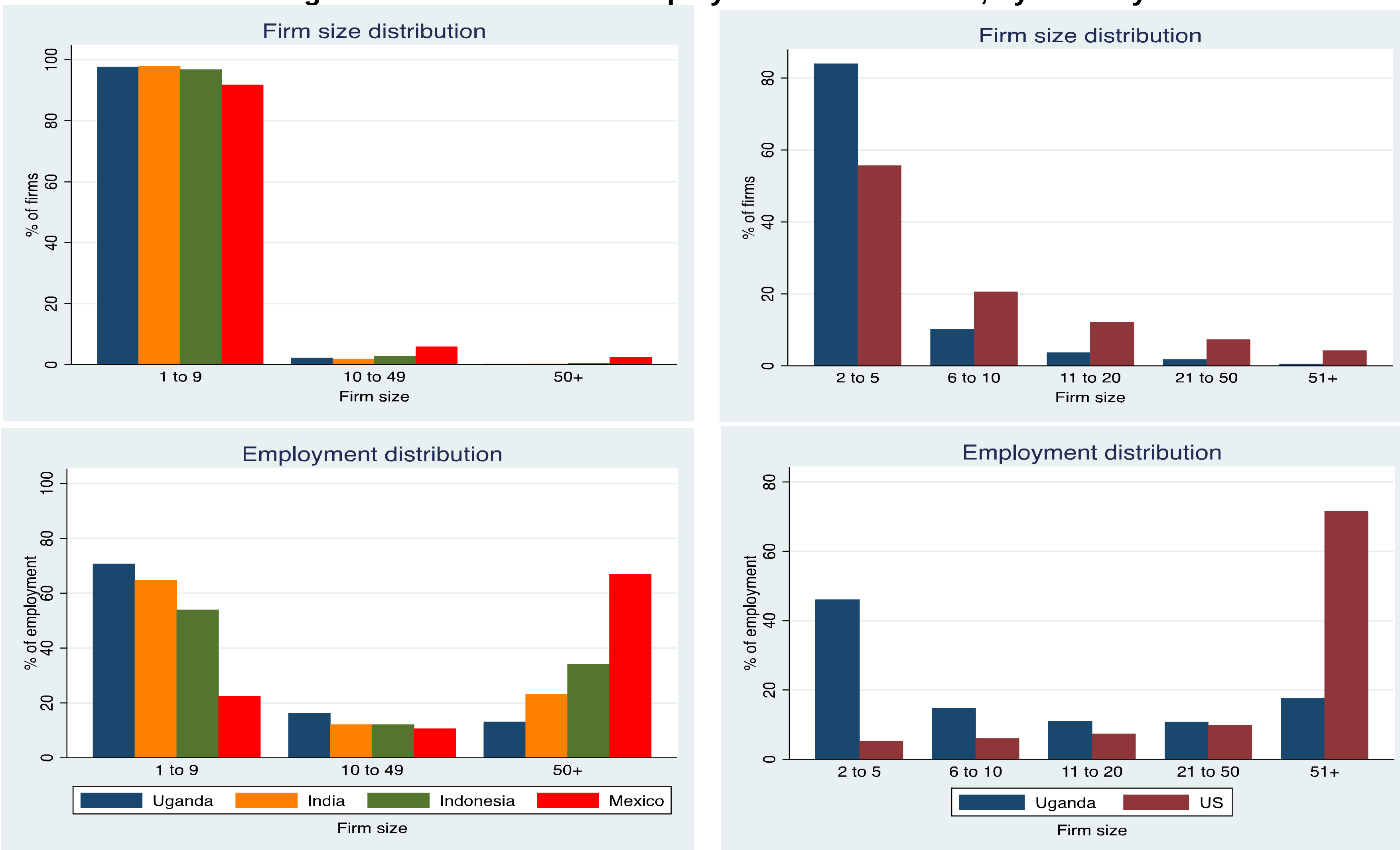
Notes: ***denotes significance at the 1% level, ** at the 5% level, * at the 10% level. The data used is from the baseline and first three follow-up worker surveys. We report OLS regressions, where we use inverse probability weighting (Columns 1 to 8) and robust standard errors are reported in parentheses. We report Lee [2009] bounds in brackets, where we implement a conditional Lee Bounds procedure that is able to condition on strata dummies in Columns 1-3 and 6-9, and to condition on region dummies and a dummy for having a level of education at the median or above at baseline in Columns 4-5. We report p-values adjusted for multiple testing in braces. These are computed using the step-down procedure discussed in Romano and Wolf [2016], with 5000 bootstrap replications. The dependent variable is the Labor Market Index that is computed using the following variables: any paid work in the last month (dummy), any wage employment in the last month (dummy), any casual work in the last month (dummy), hours worked in wage employment last week, hourly wage rate, total earnings in the last month. Hourly earnings and total earnings are set to zero for workers with no earnings. The index is constructed by converting each component into a z-score, averaging these and taking the z-score of the average. z-scores are computed using means and standard deviations from the control group at baseline. In those specifications we are able to control for the baseline value of the index. Manufacturing sectors are: motor-mechanics, plumbing, construction, electrical wiring and welding. Service sectors are: hairdressing, catering and tailoring. Workers are assigned to Manufacturing or Sevice sectors according to stated preferences over their ideal job, reported at baseline. In Column 6 we restrict the sample to labor markets outside of Kampala. All regressions include strata dummies, survey wave dummies, a dummy for the implementation round and dummies for the month of interview. In columns 1 to 7 we also control for the following baseline characteristics of workers: age at baseline, a dummy for whether the worker was married at baseline, a dummy for whether the worker had any children at baseline, a dummy for whether the worker was employed at baseline, and a dummy for whether the worker scored at the median or above on the cognitive test administered at baseline. Columns 1 and 4-9 further control for a complete set of strata dummies. Columns 2 and 3 further control for region dummies, and a dummy for having a level of education at the median or above at baseline. The weights for the IPW estimates are computed separately for attrition at first, second and third follow-up. The instruments for the IPW estimates are whether the worker was an orphan at baseline, a dummy if anyone in the household of the worker reported having a phone at baseline, a dummy for whether the worker reported being willing to work in more than one sector at the time of their original application to the VTIs and dummies for the survey team the worker's interview was assigned to in each of the three follow-up survey rounds. At the foot of each Column we report p-values on the null that the impact of the vocational training is equal to the impact of firm training (T2=T3). All monetary variables are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD.

Table A9: Spell Descriptives

	Steady State: November 2015 (Data from Second and Third FUP)
Unemployment rate (%)	52.27%
% of initial spells that are LEFT CENSORED	50.58%
% of initial spells that are RIGHT CENSORED	54.66%
% of initial UNEMPLOYMENT spells that are RIGHT CENSORED	60.03%
% of initial EMPLOYMENT spells that are RIGHT CENSORED	36.16%
Average duration of initial UNEMPLOYMENT spell [Months]	20.99
Average duration of initial EMPLOYMENT spell [Months]	15.38

Notes: The descriptive statistics use data from the second and third worker follow-up. The dataset is a cross-section of workers, and for each worker it contains information on: spell type (employment, unemployment), spell duration (in months), earnings in employment spells (in USD), dates of transitions between spells, and type of transition: (i) job to unemployment, (ii) unemployment to job, or (iii) job to job. Wages are deflated and expressed in terms of August 2012 prices, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted into August 2012 USD. The statistics are computed for at most two spells (and one transition) per individual. The initial spell is identified as the (employment or unemployment) spell that was ongoing in November 2015. Spells are right censored at the date of interview. Spells are left censored at 1 August 2014. Casual and agricultural occupations are coded as unemployment. Self-employment is coded as employment (but self-employment spells are assigned a separate spell). The unemployment rate is defined as the percentage of the sample that was unemployed in their first spell.

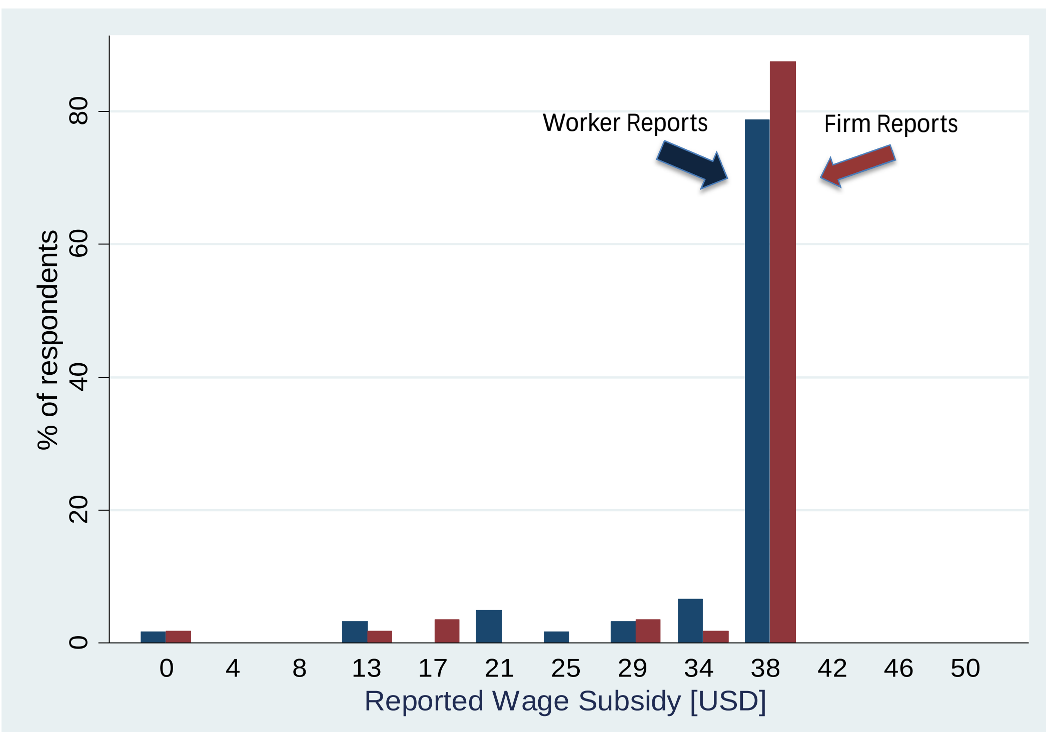
Figure A1: Firm Size and Employment Distributions, by Country



Notes: The data used to produce these figures is from the following sources: Uganda – 2010 Census of Business Establishments collected by the Uganda Bureau of Statistics; US - 2010 Business Dynamic Statistics collected by the US Census Bureau; India, Indonesia and Mexico - data are from Hsieh and Olken (2014). Data are from 2011 for India, 2006 for Indonesia and 2008 for Mexico. The firm size distribution reflects the % of firms that employ 0 to 9, 10 to 49, and 50+ employees. The employment distribution reflects the percentage of workers employed in firms that employ 0 to 9, 10 to 49, and 50+ employees.

Figure A2a: Wage Distribution of Unskilled Workers at Baseline

Figure A2b: Worker-Firm Wage Subsidy Splits



Notes: The top graph shows the distribution of unskilled workers' wages at baseline. The solid line is drawn in correspondence to the total amount of wage subsidy under the Firm Training treatment, and the dashed line indicates the median (unskilled) wage at baseline. A Kernel density estimate of the distribution of wages is also shown. The lower histogram shows the reported monthly earnings of workers hired through the Firm Training treatment, where the first bar is always the worker's self-reported wage, and the second bar is what the firm reports paying the worker.

Figure A3: Worker Spells Data