Econometrica, Vol. 88, No. 6 (November, 2020), 2369–2414

TACKLING YOUTH UNEMPLOYMENT: EVIDENCE FROM A LABOR MARKET EXPERIMENT IN UGANDA

LIVIA ALFONSI
Department of Agricultural and Resource Economics, UC Berkeley

ORIANA BANDIERA
Department of Economics, IGC and STICERD, LSE

VITTORIO BASSI
Department of Economics, USC

ROBIN BURGESS
Department of Economics, IGC and STICERD, LSE

IMRAN RASUL
Department of Economics, UCL and IFS

MUNSHI SULAIMAN
Research and Evaluation Division, BRAC

ANNA VITALI
Department of Economics, UCL

We design a labor market experiment to compare demand- 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 compare the effect of offering workers either vocational training (VT) or firm-provided training (FT) for six months in a common setting where youth unemployment is above 60%. Relative to control workers, we find that, averaged over three post-intervention years, FT and VT workers: (i) enjoy large and similar upticks in sector-specific skills, (ii) significantly improve their employment rates, and (iii) experience marked improvements in an index of labor market outcomes. These averages, however, mask differences in dynamics: FT gains materialize quickly but fade over time, while VT gains emerge slowly but are long-lasting, leading VT worker employment and earning profiles to rise above those of FT workers. Estimating a job ladder model of worker search reveals the key reason for

Livia Alfonsi: livia.alfonsi@berkeley.edu
Oriana Bandiera: o.bandiera@lse.ac.uk
Vittorio Bassi: vbassi@usc.edu
Robin Burgess: r.burgess@lse.ac.uk
Imran Rasul: i.rasul@ucl.ac.uk
Munshi Sulaiman: munshi.slmn@gmail.com
Anna Vitali: anna.vitali.16@ucl.ac.uk

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, Victor Quintas-Martinez, Mark Rosenzweig, Johannes Schmieder, T. Paul Schultz, John Strauss, Miguel Urquiola, Gerard van der Berg, Eric Verhoogen, Alessandra Voena, Chris Woodruff, Fabrizio Zilibotti, and many seminar participants for comments. All errors are our own.

this: VT workers receive significantly higher rates of job offers when unemployed, thus hastening their movement back into work. This likely stems from the fact that the skills of VT workers are certified and therefore can be demonstrated to potential employers. Tackling youth unemployment by skilling youth using vocational training pre-labor market entry therefore appears to be more effective than incentivizing firms through wage subsidies to hire and train young labor market entrants.

KEYWORDS: Vocational training, on-the-job training, human capital, youth unemployment.

1. INTRODUCTION

YOUTH UNEMPLOYMENT is a major challenge in the developing world. A growing mass of young workers are failing to find work in manufacturing and service sectors consisting mainly of small-scale firms. This raises two questions. On the supply side, why don't workers acquire the skills that can help them secure jobs? On the demand side, what prevents firms 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 employed in 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 formal sector youth employment rates are below 30%, and youth are mostly engaged in insecure and informal casual work.

To do this, we design a two-sided experiment involving 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. As the vocational training and firm-provided training interventions are fielded in the same setting, we can directly compare their impacts on workers. To investigate mechanisms, we use the experimental results to estimate a job ladder model with treatment-specific transition parameters. This is our core contribution.[1]

Both interventions are designed to improve skills, but they do so by relaxing different constraints. On the supply side, subsidized vocational training may help workers overcome credit market imperfections which prevent them from investing in skills or imperfect knowledge regarding the return to different skills (Jensen (2010)). Moreover, vocational training formally provided by vocational training institutes (VTIs) gives workers *certified* skills, showing which sector-specific skills they were trained in. This ameliorates adverse selection and can enhance the labor mobility of vocationally trained workers as long as

[1]Earlier studies have often evaluated a combination of in-class vocational and on-the-job training, for example, JTPA in the U.S. and the YTS in the UK. In low-income settings, Card, Ibarran, Regalia, Rosas-Shady, and Soares (2011) and Attanasio, Kugler, and Meghir (2011) both evaluated the impacts of combining three months of vocational training followed by three-month apprenticeships, in the Dominican Republic and Colombia, respectively. 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. The justifications for such approaches are twofold: (i) to reduce employer screening costs (Autor (2001), Hardy and McCasland (2017)); (ii) to provide workers some labor market experience that can have persistent impacts (Pallais (2014)).

there are firms willing and able to hire them.[2] If these do not exist, only a policy that relaxes firms' hiring constraints will increase employment rates.

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, or of learning about the ability and match quality of inexperienced workers. However, firms cannot avoid the time costs associated with training a new worker, which may be large in a context where firms are small and where much of the training must be done by the firm owner.

Our research design provides evidence on these elements. Workers in our study are disadvantaged youth entering the labor market. On the demand side, we have small and medium size enterprises (SMEs) in both manufacturing and service sectors, a core segment of the Ugandan economy. We track 1700 workers and 1500 firms over four years, after randomly assigning workers to either control, vocational training (VT) for six months, or firm training (FT) for six months.

To evaluate treatment effects on skills, we develop a sector-specific skills test together with skills assessors in Uganda. Our first finding is that, two to three years post-intervention, workers who have received training have accumulated sector-specific skills (equivalent to a 30% or 0.4sd increase over control workers). The magnitude of the improvement is almost identical across both VT and FT workers ($p = 0.902$). This is important because it helps to shut down one potential difference between treatments.[3]

Our second finding is that there is substantial divergence in compliance: 68% of workers assigned to VT start this training, but only 24% of workers assigned to FT do. This gap is driven by firm, rather than worker, characteristics. In common with earlier studies, firm interest is a key limiting factor (Groh, Krishnan, Mckenzie, and Vishwanath (2016)), here due to the fact that training was a time- and resource-costly requirement for firm owners as their involvement in the training of apprentices was monitored and enforced. This feature of the experiment is policy-relevant when thinking through supply- and demand-side policies to tackle youth unemployment. It also shapes how we estimate treatment effects, the structural model, and how we conduct the IRR analysis. Throughout, we focus on workers' labor market outcomes and estimate both the ITT and the ATE for compliers: ITT, because by averaging over compliers and non-compliers, they reflect likely binding challenges to scaling-up VT and FT interventions in the same context, or of exporting them to other contexts; ATE, because by focusing on compliers, they map closely to theories of training and enable us to examine the channels via which VT and FT effects differ for trained workers.

Our third finding is that both treatments improve an index of worker labor market outcomes, that combines employment, total labor supply, and earnings. Due to differences in compliance, the ranking of the two treatments depends on whether we look at ITT or ATE, but in both cases, we fail to reject the null of equality. Indeed, the ATE estimates show increases by 0.473sd for FT workers and by 0.272sd for VT workers ($p = 0.202$),

[2]Evidence of the value of certification in labor markets has been provided by Pallais (2014), MacLeod, Riehl, Saavedra, and Urquiola (2015), Abebe, Caria, Fafchamps, Falco, Franklin, and Quinn (2020), Bassi and Nansamba (2020), Abel, Burger, and Piraino (2019), and Carranza, Garlick, Orkin, and Rankin (2019).

[3]Our setting departs from the standard Beckerian framework in two ways: (i) we subsidize the apprenticeships through the wage subsidy, making firms more willing to provide skills that are not firm-specific; (ii) firms are contractually required to provide sector-specific skills to workers. This form of apprenticeship—where firms do not bear the full cost of training but are contractually obliged to train workers and monitored in doing so—is a policy that can be replicated and scaled-up by government.

while ITT show increases by 0.105sd for FT workers and by 0.170sd for VT workers ($p = 0.169$).[4]

These similarities mask differences in dynamic treatment effects. FT workers find employment more quickly than VT workers, but over time, their employment rate converges to the control group, while employment rates for VT workers increase over time. This reversal of fortune between FT and VT workers is also found for earnings—FT workers do well initially, but then over time, their earnings fall behind those of VT workers. VT workers steadily increase their earnings and diverge away from the control group. In other words, the similarity in ATE treatment effects between FT and VT workers when averaged over post-intervention survey waves is driven by the earlier quarters in which FT workers were hired by firms incentivized through wage subsidies. Subsequent to that, the patterns of employment and earnings differ across VT and FT workers, with FT workers having employment profiles similar to those of control workers.[5]

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 estimate a job ladder model of worker search. This emphasizes three mechanisms driving labor market outcomes: (i) arrival rates of job offers when unemployed or when employed; (ii) job separation rates; (iii) skills. The model allows workers to be heterogeneous in two dimensions: their training (treatments), and their type-ε that determines their productivity on-the-job. We operationalize worker types by linking them to the measurable skills of workers. This form of observed worker heterogeneity provides an avenue for using the model for counterfactual analysis.

The job ladder model estimates reveal that: (i) VT workers have significantly higher steady-state rates of unemployment-to-job (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 become unemployed. At the same time, accepted earnings conditional on employment are similar between FT and VT workers, consistent with them having similar skills, but VT workers can more easily certify their skills and climb back onto the job ladder if unemployed. In steady state, unconditional annual earnings of complier VT workers rise by 55% over controls, while the earnings of complier FT workers rise by just over half of that, 31%.

Combining these results gives us a precise interpretation to what drives the dynamic treatment effects: vocational trainees pull away from FT workers in their employment rates and earnings because they are more likely to get back onto the job ladder if they fall into unemployment. These dynamics are not due to greater job-to-job mobility, suggesting the returns to skills certifiability are higher when unemployed than when employed. Moreover, compliers across training routes move as far up the job ladder as each other—wages *conditional* on employment are similar for complier VT and FT workers because their skills are similar. The key distinction is that VT workers are more likely to get back onto the job ladder if they fall off it. Tackling youth unemployment by skilling youth using vocational training pre-labor market entry therefore appears to be more effective than incentivizing firms through wage subsidies to hire and train young labor market entrants.

[4]In relation to earlier studies that have evaluated a combination of vocational and on-the-job training, Card et al. (2011) found no evidence of employment impacts; Attanasio, Kugler, and Meghir (2011) found a 7% increase in employment rates for women and a 20% earnings increase. Galasso, Ravallion, and Salvia (2004), Levinsohn, Rankin, Roberts, and Schoer (2014), and Groh et al. (2016) evaluated wage subsidy interventions.

[5]Comparing two supply-side interventions, Abebe et al. (2020) also found that the effect of subsidies—in their case, to workers to fund transport—is short-lived, while certification has lasting impacts.

To understand how the results derived from the partial equilibrium model map to general equilibrium impacts, we exploit the two-sided experimental design allowing us to compare training routes from the dual perspectives of workers *and* firms. As in Crepon and Premand (2019), the firm-side of the experiment shows there are no employment displacement (or crowding-in) effects on other workers—in either the short run when wage subsidies are in place, or in the long run long after apprentices have left the firm.

We use the model parameters to conduct two counterfactual simulations: (i) to assess the relative importance of the mechanisms at the heart of the model in explaining steady-state outcomes; (ii) to simulate treatment effects if the training interventions were targeted to other workers in the economy; in particular, by drawing on data from other studies in the same context, we simulate what would have been the impacts of targeting our treatments to workers that are able to self-finance vocational training, and that are already employed in SMEs in manufacturing sectors similar to those in this current study.

We then combine program accounting costs with steady-state earnings benefits to derive the internal rate of return from each treatment. Assuming gains last 15 years, the IRR to vocational training is 22%, while the IRR for firm training is negative. It does *not* pay for the social planner to replicate the kind of subsidized apprenticeship offered in the FT treatment. However, the reason for this negative IRR is the low compliance in the FT treatment: only 24% of such workers end up being hired by firms they are matched to.[6] However, in these labor markets, we do observe workers paying firms for apprenticeships using the kind of payment structure we set up in the FT treatment. To see why this is so, we redo the IRR calculations based on the steady-state earnings for compliers—namely, those that acquire firm-provided or vocational training. Among this group, the IRR for VT workers rises to 33%, and for FT workers, the IRR is 25%. The rise in IRR for FT workers highlights the high social returns from being able to overcome firms' constraints in taking on and training young workers. Under these IRRs, both training routes pay for themselves.

This of course raises the question of why workers are not availing themselves of these returns by paying for VT or FT themselves? A key reason may be credit constraints: the costs of vocational training, or of self-financing apprenticeships, run into hundreds of dollars and so are both orders of magnitude higher than young workers' earnings at baseline ($6/month).

Despite their popularity, the evidence base for training programs, or in contrasting alternative training programs in the same context, is thin. The meta-analyses of Blattman and Ralston (2015), McKenzie (2017), and Card, Kluve, and Weber (2018) 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. A key explanation is that our training interventions were big push: lasting longer and delivered more intensely than some earlier studies. Other factors, such as the selection of workers and vocational training institutes worked with, also play a role. On each dimension, we make suggestions for future research.

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

[6]This low compliance is driven by a lack of firms taking up the offer of the wage subsidy and the matched-to worker (workers are as likely to accept offers from firms as vocational trainees are to accept the offer of training from VTIs). Moreover, we document that less profitable firms are more likely to take on workers through the FT treatment. This suggests that these firms are financially constrained in hiring young job seekers and that this is an important demand-side constraint.

outcomes. Section 4 develops the job ladder model. Section 5 presents model estimates. Section 6 presents counterfactuals. Section 7 presents the IRR estimates and discusses external validity. Section 8 concludes. Robustness checks and further estimation details are in the Supplemental Material (Alfonsi et al. (2020)).

2. SETTING, DESIGN, AND TREATMENTS

Our study is a collaboration with the NGO BRAC, who implemented all treatments, and five reputable vocational training institutes (VTIs). The VTI sector in Uganda is well established, with hundreds in operation. Each could offer standard six-month training courses in eight sectors: welding, motor mechanics, electrical wiring, construction, plumbing, hairdressing, tailoring, and catering. These sectors constitute a source of stable wage employment for young workers in Uganda: around 25% of employed workers aged 18–25 work in them.[7]

2.1. *Setting*

Workers. Individuals were recruited into our evaluation throughout Uganda. We advertised an offer of potentially receiving six months of sector-specific vocational training at one of the VTIs we collaborated with. The eligibility criteria targeted disadvantaged youth. We received eligible applications from 1714 individuals whose characteristics are shown in Table A.I: 44% are women, they are aged 20 on average, and the vast majority have never received vocational training.[8]

The first row of Table I shows baseline labor market outcomes for our workers: unemployment rates are over 60% for these youth (Columns 2 and 3) with insecure casual work being the most prevalent labor activity. Unconditionally, average monthly earnings are \$6, corresponding to around 10% of the Ugandan per capita income at the time. Hence, these individuals are unlikely to be able to self-finance investment into vocational training (that costs over \$400), or to pay a firm for an apprenticeship. Table A.II compares our sample to those aged 18–25 in the Uganda National Household Survey from 2012/2013. The intervention 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 who report being labor market active.[9]

Our experiment uses an oversubscription design. This is informative of the impact of marginally expanding such training. Given Ugandan demographics, there is no shortage of the kind of disadvantaged youth that applied to our offer.

[7]The VTIs we worked with: (i) were founded decades earlier; (ii) were mostly for-profit; (iii) trained hundreds of workers with an average student-teacher ratio of 10; (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 2012/2013 Uganda National Household Survey.

[8]The program was advertised 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; (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 geographic-specific threshold score was used to select eligibles.

[9]Unemployment rates are often difficult to define in low-income contexts. It is thus perhaps more accurate to speak of rates of non-employment, and for expositional ease, this is what we will implicitly have in mind when referring to unemployment.

TABLE I

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[a]

	Number of Workers (1)	Currently Working (2)	Has Worked in the Last Month (3)	Has Done Any Wage Employment in the Last Month (4)	Any Self Employment in the Last Month (5)	Has Done Any Casual Work in the Last Month (6)	Total Earnings in the Last Month [USD] (7)	F-Test of Joint Significance (8)
All Workers	1714	0.360	0.383	0.130	0.046	0.257	5.92	
		(0.045)	(0.044)	(0.023)	(0.013)	(0.044)	(1.11)	
T1: Control	451	0.381	0.401	0.120	0.038	0.296	5.11	
		(0.049)	(0.048)	(0.025)	(0.015)	(0.047)	(1.27)	
T2: Firm Trained	283	0.369	0.387	0.103	0.064	0.266	6.42	{0.968}
		(0.035)	(0.035)	(0.023)	(0.017)	(0.032)	(1.34)	
		[0.979]	[0.985]	[0.520]	[0.096]	[0.592]	[0.243]	
T3: Vocationally Trained	390	0.358	0.389	0.149	0.034	0.253	7.29	{0.811}
		(0.032)	(0.032)	(0.023)	(0.013)	(0.029)	(1.26)	
		[0.763]	[0.990]	[0.188]	[0.802]	[0.265]	[0.063]	
T4: Vocationally Trained + Matched	307	0.320	0.360	0.149	0.050	0.205	5.25	{0.758}
		(0.033)	(0.034)	(0.026)	(0.015)	(0.030)	(1.20)	
		[0.316]	[0.747]	[0.229]	[0.266]	[0.070]	[0.757]	
T5: Untrained, Matched	283	0.364	0.367	0.127	0.057	0.251	5.56	{0.996}
		(0.033)	(0.034)	(0.025)	(0.016)	(0.031)	(1.25)	
		[0.707]	[0.386]	[0.821]	[0.207]	[0.210]	[0.721]	
F-test of joint significance		{0.882}	{0.908}	{0.301}	{0.214}	{0.433}	{0.380}	

[a]All data are 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 in parentheses 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 7. 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 includes 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.

Firms. To draw a sample of firms, we conducted a firm census in 15 urban labor markets. We selected firms: (i) operating in one of the eight manufacturing and service sectors in which we offered sector-specific vocational training; (ii) having between one and 15 employees (plus a firm owner). The first criterion limits skills mismatch in our study. The second restriction excludes micro-entrepreneurs and ensures we focus on SMEs that are central to employment generation in Uganda. We end up with a sample of 1538 SMEs, employing 4551 workers in total at baseline. On constraints to expansion: (i) 65% of firm owners reported the terms of available finance limiting their growth; (ii) 67% reported access to skilled labor as a constraint; (iii) 52% reported the inability to screen workers as a constraint. Wage subsidies might help relax demand-side constraints on SMEs related to hiring young labor market entrants.

Returns to Vocational Training. Table A.III 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 a standard Mincerian wage regression of log wages. The first row pools across all sectors and documents that at baseline, 31% of 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 workers with such training. The Mincerian returns to vocational training are over 50%, and this holds in each sector. Of course, the Mincerian returns are upward-biased due to selection into employment. Our experimental results help quantify this selection bias. This evidence shows there is demand for, and potentially high returns to, vocational training in the sectors that 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, Kluve, and Weber (2018)).

Use of Apprenticeships. Firm-sponsored training is another route through which workers accumulate human capital. Apprenticeships are a common labor contract throughout Sub-Saharan Africa. Table A.IV provides evidence on such contracts from our sample of SMEs. Panel A shows that half the workers employed in control 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, some are paid, and others pay for their training. For those paid during apprenticeships, they report an average monthly wage of $39. Firm owners were asked about the skills composition of apprentices. Self-financed apprentices are more likely to be reported to have sector-specific rather than firm-specific skills.[10]

Panel C shows that the main opportunity cost for taking on new hires is the firm owner's time: they are predominantly tasked to train apprentices. This is especially so for self-financed apprentices: 56% report being trained exclusively by the firm owner, and none report being trained only by employees. Firm owners have the skills to train employees: they have significantly more years of education than workers, and are significantly more likely to have received vocational training. As mentioned above, most SMEs report an

[10]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). (ii) For 29% of workers that pay for their apprenticeship, the average total payment is over $500. Whichever way we calculate it, the expected cost of an apprenticeship is above the baseline annual earnings of our sample workers.

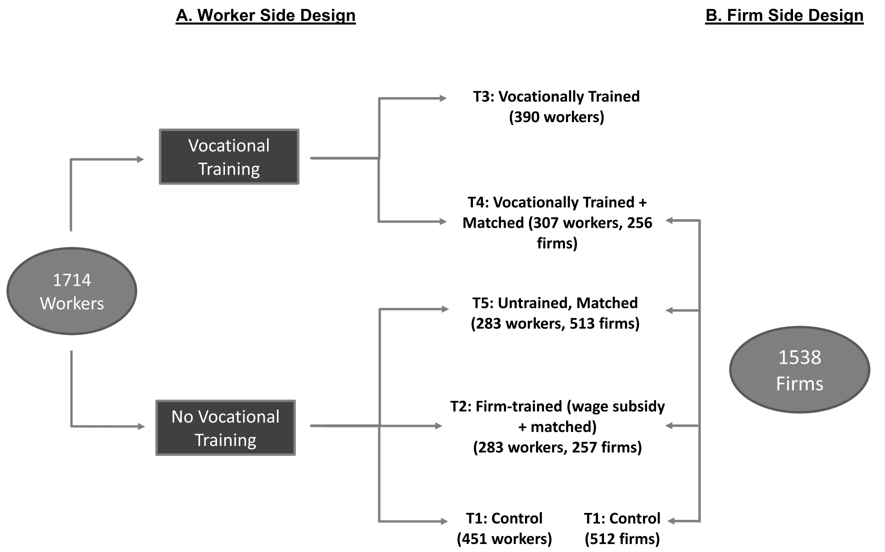

FIGURE 1.—Experimental design. *Note*: Numbers in parentheses refer to the number of applicants originally assigned to each treatment, and the number of firms assigned to each treatment.

inability to screen workers as constraining expansion. Hence, if SMEs are credit constrained, it is these kinds of up-front screening costs, or firm owners' opportunity costs of training new hires, that are reduced in our apprenticeship treatment.[11]

2.2. *Design*

The left-hand side of Figure 1 presents the design from a worker's perspective. One thousand seven hundred fourteen workers applied to the offer of vocational training. Those randomly assigned to vocational training were split into two treatments. The first group completed their six months of 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 vocationally trained workers, upon graduation, were matched to firms operating in the same sector as the worker had been trained in, and located in the same region as the worker. On graduation, all trainees leave their VTI with a certificate stating which VTI was attended, and the six-month training course taken.

Workers not offered vocational training were randomly split into three groups: (i) to be matched to firms; (ii) to be matched to firms and those firms offered a wage subsidy to hire and train them on-the-job for six months (i.e., as an apprentice); (iii) held as a control. This design allows us to thus compare and contrast supply- and demand-side interventions designed to raise skills and reduce youth unemployment, and to understand the nature of constraints on workers and firms that prevent such human capital investments being undertaken.

[11]Firm owners' role in training workers is well recognized. In the firm-side surveys, we interviewed employees in our SMEs pre-intervention. We asked them about the role of the 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.

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. 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 three years of follow-up data.

We assign workers to treatment using a stratified randomization where strata are region of residence, gender, and education. Table I shows the labor market characteristics of workers in each treatment. Table A.I shows other background characteristics. In both cases, the samples are well balanced, and normalized differences in observables are small.

The right-hand side of Figure 1 shows the design from a firm's perspective: firms were randomly assigned to either be matched with: (i) vocationally trained workers; (ii) untrained workers; (iii) untrained workers and given a wage subsidy to hire and train them; (iv) or held as a control.[12]

Timeline. Figure 2 shows the study timeline: the baseline worker survey took place from June to September 2012 just after workers applied for vocational training. Eligible workers were tracked in surveys fielded 24, 36, and 48 months after baseline (12, 24, and 36 months after the end of vocational training/apprenticeships).[13] Only 13% of workers attrit by the 48-month endline. Supplemental Material Appendix A.1 describes correlates of attrition, confirming attrition is uncorrelated to treatment.

The lower part of Figure 2 shows the timeline of firm surveys. We use these data to compare firm outcomes between those offered the apprenticeship and the control group. In particular, we estimate short- and long-run employment displacement effects of the FT treatment. As we describe later, these results on (a lack of) displacement help us to extrapolate results from the structural model—that is, in partial equilibrium—to understand the general equilibrium impacts of firm-provided training in these urban labor markets.

2.3. *Treatments*

Vocational Training. Vocational training provides workers six months of sector-specific training in one of eight sectors. In treatment arms involving vocational training (T3, T4), BRAC entirely covered training costs, at $470 per trainee. Lessons were held Monday–Friday, for six hours per day; 30% of course content was dedicated to theory, 70% to practical work covering sector-specific skills and managerial/business skills. VTIs signed contracts with BRAC to deliver these standard training courses to workers. They were monitored by regular and unannounced visits by BRAC staff to ensure workers were

[12]In current work in progress, we are conducting a comprehensive analysis of the firm-side impacts of all these treatments (and other treatments), 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 (but not by endline), and we account for this by weighting observations using inverse probability weights.

[13]We surveyed those randomized out of vocational training just as vocational trainees were transitioning into the labor market. The tracker survey had a 23% attrition rate. The work status of respondents were as follows: 19% were currently involved in some work activity, 11% had been involved in a work activity in the last six months (but not on survey date), and 70% had not worked in the last six months.

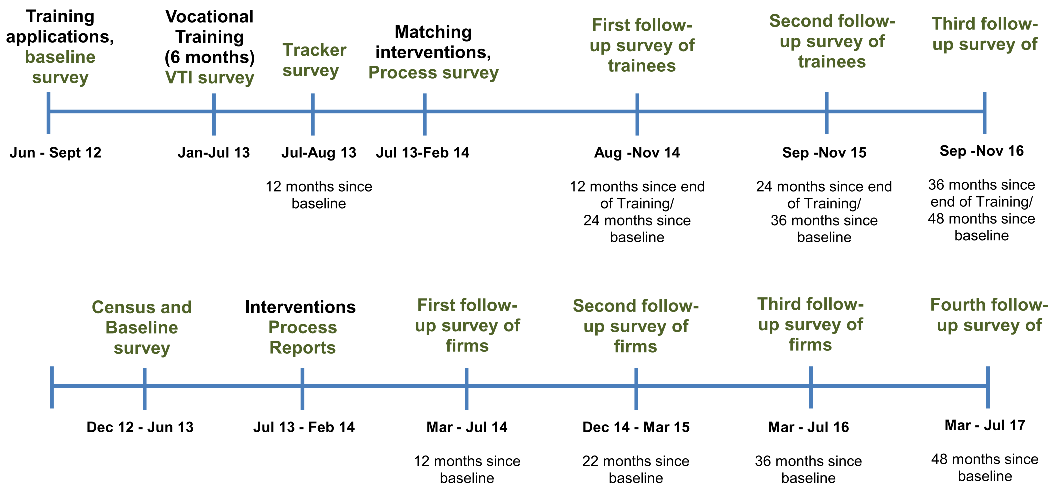

FIGURE 2.—Timeline. *Notes*: The timeline highlights the dates relevant for the main batch of worker applications and baseline surveys. A second smaller round of applications and baseline surveys were conducted in May and June 2013. The majority of trainees from the first round of applicants started 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.

present and being trained. For each worker, VTIs were paid half the training fee at the start of training, and half at the end, conditional on them having trained the worker (this staggered timing of payments ensured VT workers nearly always completed the full course of training conditional on enrolling).[14]

Firm Training. In the firm-provided training treatment, we offered firms to meet untrained workers and receive $50 a month for six months if they hired and trained one such worker on-the-job. This was an inflexible wage subsidy: $12/month was to be retained by the owner, and $38/month was to be paid to the worker. This differs from the standard Beckerian apprenticeship model in that the firm does not bear the full training cost. As such, the skills provided to workers might differ from in other apprenticeship structures. As Table A.IV shows, workers do sometimes pay for their training in these labor markets, and this is associated with more intensive involvement of firm owners in training workers.

We assess whether the level of the wage subsidy is reasonable using two anchors: (i) Table A.IV shows that during apprenticeships, if workers were paid, their mean wage

[14]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 for travel and accommodation ($70). The staggered incentive contract solved drop-out problems associated with training programs in low-income settings (Blattman and Ralston (2015)). There was no additional stipend paid to trainees during training, and no child care offered (recall that around 10% of our worker sample have at least one child).

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 111% (Figure A.1a shows the distribution of unskilled wages at baseline among those employed in our SMEs). This is high: for example, de Mel, Mckenzie, and Woodruff (2019) evaluated a wage subsidy program with a 50% subsidy rate.

This FT treatment was designed as a formal training program, backed by an explicit contractual agreement between firm owners and our implementing partner, BRAC. This was intended to mirror the vocational training program as far as possible, but with training being conducted at a firm rather than at a VTI. The contract between BRAC and firm owners in the FT treatment is shown in Figure A.2. The contract is succinct and clearly states: (i) firm owners agree to train the worker in a specific trade for six months; (ii) firm owners would pay back the entire subsidy if the trainee was found not to be receiving training, or was not showing up to the firm regularly. There was no explicit training curriculum: firm owners were free to train workers as they saw fit, as long as sector-specific training was provided (as VTIs were also tasked to do).

As with VTIs, monitoring checks were used to ensure these agreements were adhered to. Firm owners and trainee workers were monitored during the six months of training in two ways: (i) each firm was provided an attendance register, and every day both the firm owner and the worker had to sign it, providing the worker's time of arrival and departure; (ii) BRAC staff conducted monthly unannounced spot checks at firm premises to verify the worker was showing up and being trained. Payments were disbursed monthly at the local BRAC office, and both the firm owner and the worker had to be present at disbursement, where they were asked to sign an additional register to certify the worker was showing up regularly at the firm and receiving training. Figure A.1b shows worker and firm reports on the wage subsidy being received by the worker, with a clear spike at \$38/month as intended.

Matching. 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. 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 in the same region as the worker was located.

3. TREATMENT EFFECTS

3.1. *Compliance*

Workers may not comply with their treatment, and for treatments involving worker-firm matches, there can also be non-compliance driven by firms because worker-firm matches only occur if both a worker and the firm express a willingness to meet. Table A.VI shows worker and firm take-up rates by treatment. For 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) 68% of workers take up the offer of vocational training and complete the training (Column 2); conditional on enrollment, over 94% of them completed the training.

For workers assigned to the FT treatment, 51% are actually offered a meeting with a firm (Column 3). In common with earlier studies, firm interest is a key limiting factor on worker-firm matches occurring (Groh et al. (2016)). The explanation is that the FT treatment required firm owners to provide time- and resource-costly training. We provide additional evidence on this from process reports collected during the intervention roll-out. Firm owners who were not interested in taking on a worker in the FT treatment were asked why; 62% said the subsidy was not large enough to cover training costs, in line with demand-side credit constraints binding.

However, conditional on the worker-firm match, 80% of meetings take place (Column 4), 90% of interviewed workers are offered a job (Column 5), and two thirds of job offers are accepted (Column 6). This offer acceptance rate to firm-provided training is very close to the 68% compliance rate for vocational training. The difference in compliance rates between FT and VT is not driven by the share of workers taking up training offers, but by the share receiving *offers from firms* to provide them training. As a result, only 24% of workers initially assigned to FT end up being employed and trained at the matched firm.

Firms' lack of interest in meeting workers is most severe in the treatments involving matching (T4, T5): only 13% and 19% of workers end up being offered a meeting with a firm in these treatments (Column 3). This is not surprising in this context: given youth unemployment rates of around 60,% firms should have little difficulty in meeting untrained workers, and as Table A.III shows, 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.

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 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 vocationally trained (VT) workers (T3 and T4) and firm-trained (FT) workers.

3.1.1. *Differential Compliance Between Vocational and Firm-Provided Training*

There is a divergence in compliance between VT and FT treatments: 68% of workers assigned to vocational training start this training, and 24% of workers assigned to firm-provided training are hired by firms they are matched to. To understand whether outcome differences between VT and FT can be due to lower compliance in FT, we first establish whether FT compliance relates to worker traits, which might in turn determine returns to training. The experimental design has two features that help rule this out. First, eligibility requirements mechanically ensure that individuals in the sample are relatively homogeneous. Second, by design, firms in the FT treatment could only train one of two workers assigned to them. Hence, their ability to choose on worker traits was limited among a relatively homogeneous pair of workers presented to them.

Table A.VII provides regression evidence on the correlates of compliance in the FT treatment. Column 1 controls for worker characteristics, and shows that 10 out of 11 of these do not predict take-up (with the other being marginally significant). This is in line with there being limited variation in worker traits presented to firms. Column 2 shows this is robust when we add program-related characteristics. In contrast, when firm characteristics are added, these are jointly significant in predicting compliance ($p = 0.002$). This is driven by firms with significantly lower profits per worker taking on FT workers. Hence, firms hiring workers when given a wage subsidy incentive appear to be negatively selected.

Finally, Column 4 shows that when we control for firm fixed effects (that is possible given multiple workers were presented to the same firm), worker characteristics remain jointly insignificant predictors of compliance ($p = 0.976$).[15]

This all suggests compliers are similar in FT and VT treatments because lower compliance in FT is explained by firm, not worker, characteristics. We thus consider take-up to be exogenous, given that compliance is not predicted by worker characteristics. This allows us to compare ATE estimates of both training routes, narrowing the interpretation of differences between them as stemming from either the skills imparted, or the certifiability of those skills. To explore this further, we next provide descriptive evidence on skill accumulation among VT and FT workers. In the next section, we estimate treatment effects on various dimensions of skills accumulation.

3.1.2. *Descriptive Evidence on Skill Accumulation Among Compliers*

Vocational Training. We surveyed VT workers towards the end of the training and asked about their satisfaction with it: 76% were extremely happy/very happy with the experience; 86% were extremely happy/very happy with the skills gained; 96% reported skills acquisition as being better than or as expected, and 56% reported that six months of training was enough time for them to learn the skills they had wanted to.

Firm-Provided Training. We provide two pieces of evidence to confirm workers hired under the FT treatment were trained. From the first firm follow-up survey, which was deliberately fielded around the end of the six-month apprenticeship period, firm owners were asked to indicate for each employee hired in the last six months: (i) their productivity (on a 1 to 5 scale) when hired; (ii) their productivity at the time of the survey (or when the worker left the firm).

Panel A of Figure A.3 shows the productivity growth of hired FT workers alongside that for: (i) hired workers at control firms, (ii) hired workers at control firms who received on-the-job training, (iii) hired workers at control firms who did not receive on-the-job training; (iv) hired workers in FT firms excluding those hired through our FT treatment. We see that hired FT workers had an increase in productivity of 2.24 points during the wage subsidy period, higher than for all other comparison groups. Hence, workers hired under the FT treatment received training over and above what these firms would normally provide, in line with their contractual requirements with BRAC, and reinforcing the notion that this was a resource- and time-intensive treatment from firms' perspective.

We also use data from the first worker follow-up survey, conducted around six months *after* the wage subsidy expired. For each job spell in the previous year, workers were asked to report: (i) their ability to perform a typical sector-specific task at the start of the spell (e.g., if the job spell was in motor-mechanics, they were asked whether they were able to mend a tire tube); (ii) their ability to perform the same task at the end of the job spell (or on survey date if the spell was ongoing). The same questions were asked about an important firm-specific task at the firm where the worker was employed (the worker indicated this task). Firm owners were asked the same questions about each employee in the first firm follow-up. We compare the rate of learning on these tasks at the matched

[15]Evidence from the process reports further bolsters this. In less than 4% of cases did firms report turning down a worker in the FT treatment because of worker characteristics. Firms interested in taking on a trainee reported they were happy to take on any of the workers assigned to them, and were not searching for specific trainee characteristics. Moreover, most firms that did not take on a worker did not even meet any worker matched to them, again consistent with worker characteristics not mattering in firm selection.

firm for workers hired under the FT treatment. This comparison is in Panel B of Figure A.3. Reassuringly, worker and firm reports are well aligned, and both show substantial learning for both sector- and firm-specific tasks.

3.2. *Estimation*

We present impacts on workers' labor market outcomes and estimate both the ITT and the ATE for compliers. The former are useful from a policymaker's perspective because they reflect likely binding challenges to scaling-up the training interventions in the same context, or to export them to other contexts. We present ATE estimates because these map closely to theories of training (that typically do not model non-compliance), and to show channels through which VT and FT differ for trained workers. Our ITT estimates are based on the following ANCOVA specification for worker i in strata s in survey wave $t = 1, 2, 3$:

$$y_{ist} = \sum_j \beta_j T_{ij} + \gamma y_{i0} + \delta \mathbf{x}_{i0} + \lambda_s + \vartheta_t + u_{ist},$$

where y_{ist} is the outcome of interest, T_{ij} denotes worker i being randomly assigned to treatment j (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, we use robust standard errors, and we weight ITT estimates using inverse probability weights (IPWs) to account for attrition. In the Supplemental Material, we 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.[16]

The ATE specification replaces treatment assignment with treatment take-up (with the same controls), where take-up is defined as a dummy equal to 1 if the worker: (i) started firm training in FT, or (ii) started vocational training in VT. The earlier results showed compliance is not driven by worker characteristics. We use treatment assignment as an IV for treatment take-up and report 2SLS regression estimates, which measure the effect of treatment on the compliers. We bootstrap standard errors using 1000 replications, and we report unadjusted p-values alongside Romano and Wolf (2016) p-values accounting for multiple hypothesis testing.

The coefficient of interest is β_j: the treatment effect of T_{ij} as *averaged* over the three post-intervention survey waves. To transition from the treatment effects to the structural estimates, we then estimate dynamic treatment effects. To do so, we convert our data to a worker job spells data set, which is possible because, in each survey, workers were asked to provide their monthly labor market history since the previous survey. These shed light on the evolution of treatment effects, as well as on whether workers are in steady state

[16]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.

towards the end of our study period, underpinning the job ladder model we develop and estimate.

β_j measures the causal effect of treatment on outcomes under SUTVA. This will not hold if treatment displaces control workers because treated workers are 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, or change sector, to find work.[17] 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 5% of all workers. Hence, we do not expect the control group to be contaminated by treated workers in the same market as they are unlikely to be competing for the same exact job.[18]

3.3. *Skills*

If training routes are to impact labor market outcomes, they should first impact worker skills. We present results on three aspects of skills acquisition. It is natural to focus on the ATEs, namely, on those that attended vocational training, or were hired as apprentices.

The first dimension of skills relates to 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 II show that for both outcomes, workers hired by firms in the FT treatment are between 56 and 77pp more likely than the control group to be firm trained.[19] More surprisingly: (i) vocationally trained workers are no more likely than the control group to report being trainees in their first employment spell; (ii) workers assigned to firm training are significantly more likely to report having received training or view themselves as trainees than vocationally trained workers ($p = 0.000$ in Columns 1 and 2).

This suggests firms are less willing to train workers who have already been vocationally trained in sector-specific skills. We thus find no evidence of a complementarity between firm-provided skills and skills provided by vocational training institutes. The finding is consistent with both training routes providing workers similar skills, but also with firms anticipating VT workers to be more mobile than others because their skills are certifiable.

The second dimension we consider is a sector-specific skills test we developed in conjunction with skills assessors and modulators of written and practical occupational tests in Uganda. Each test comprises seven questions: Figure A.4 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 11. The test was conducted on all workers (including those assigned to the control

[17]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. On workers' sectoral mobility, at baseline 96% of workers reported being willing to work in more than one sector. Moreover, only 15% of all main job spells of workers in the control group at first follow-up are in the same sector as the ideal sector mentioned at baseline.

[18]Crepon, Duflo, Gurgand, Rathelot, and Zamora (2013) provided 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.

[19]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 A.IV where 50% of workers employed in the SMEs at baseline reported having been apprentices in the firm.

TABLE II

ATE ESTIMATES, TRAINING AND SKILLS. 2SLS REGRESSION COEFFICIENTS, BOOTSTRAPPED STANDARD ERRORS IN PARENTHESES. BOOTSTRAP p-VALUES IN BRACES: UNADJUSTED p-VALUES (LEFT) AND ROMANO AND WOLF (2016) ADJUSTED p-VALUES (RIGHT)[a]

	Being Trained by Firms		Sector-Specific Skills		Skills Transferability
Treatment Effects on:	Received On the Job Training	Position is "Trainee"	Any Skills (0/1)	Test Score (0–100)	
Measured at:	First Job	First Job	Two-Three Years After Training	Two-Three Years After Training	Three Years After Training, Conditional on Employment
	(1)	(2)	(3)	(4)	(5)
Firm Trained	0.570	0.767	0.422	9.67	−0.072
	(0.179)	(0.123)	(0.100)	(5.29)	(0.341)
	{0.001; 0.022}	{0.001; 0.002}	{0.001; 0.011}	{0.087; 0.292}	{0.831; 0.841}
Vocationally Trained	−0.048	0.015	0.407	10.3	0.253
	(0.056)	(0.029)	(0.032)	(1.70)	(0.104)
	{0.426; 0.815}	{0.591; 0.841}	{0.001; 0.001}	{0.001; 0.002}	{0.049; 0.136}
Mean (SD) Outcome in Control Group	0.402	0.041	0.596	30.1 (22.9)	–
Control for Baseline Value	No	No	No	No	No
p-values on tests of equality:					
Firm Trained = Vocationally Trained	[0.000]	[0.000]	[0.863]	[0.902]	[0.264]
N. of observations	789	797	1818	1818	650

[a]The data used are from the baseline and three worker follow-up surveys. We report 2SLS regression estimates, where treatment assignment is used as IV for treatment take-up. Treatment take-up is defined as a dummy equal to 1 if the worker (i) started firm training in FT or (ii) started vocational training in VT. Bootstrap standard errors are calculated using 1000 replications and reported in parentheses. We also report unadjusted p-values (left) and p-values adjusted for multiple testing (right) in braces. These are computed using the step-down procedure discussed in Romano and Wolf (2016), with 1000 bootstrap replications. The number of hypotheses being tested simultaneously in the five columns is ten—for each outcome, we test the impact of two treatments, FT and VT, and we do so for five outcomes. Therefore, the p-values in Columns 1 to 5 are adjusted for testing on ten hypotheses. All regressions include 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. 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, the dependent variable is a dummy for whether the respondent reports having any sector-specific skills or not at second and third follow-up. In Column 4, the dependent variable is the skills test score, from the test administered to workers in the second and third worker follow-up. In Column 5, 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 3 and 4, 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 3–4. 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.

group) at second and third follow-up, so this outcome measures the persistence of skills accumulation in the VT and FT treatments. There is no differential attrition by treatment into the test.[20]

Before administering the test, we asked a filtering question to workers on whether they had *any* skills relevant for sectors in our study. The dependent variable in Column 3 of Table II is a dummy equal to 1 if the worker reported having skills for a sector. The ATE estimates show that VT workers and FT workers all report being significantly more likely to have relevant skills than control workers. As reported at the foot of the table, 60% of controls report having skills for some sector, and reassuringly, this rises to 100% for FT workers that were hired by firms, and for VT workers that attended a VTI. This is consistent with the descriptive evidence in Figure A.3 suggesting complier workers in VT and FT treatments received actual training.

All workers that reported having sectoral skills took the test; others (mostly in the control group) were assigned a score of 11 assuming they would answer the test at random. Column 4 shows that VT and FT workers significantly increase their measurable sector-specific skills, as recorded two and three years after the training. Relative to controls, VT workers increase sector-specific skills by 34% (or 0.4sd of test scores). FT workers increase sector-specific skills by 32%. Strikingly, both training routes cause persistent skills accumulation, although there is no significant difference in sector-specific skills accumulation between VT and FT workers ($p = 0.902$).

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, Dearden, Meghir, and Sianesi (1999)), especially once the endogenous selection of workers into training is corrected for (Leuven and Oosterbeek (2008)).

The fact that FT workers are provided sector-specific skills goes against a standard Beckerian framework of firm-sponsored training. Our setting departs from this in two ways: (i) BRAC subsidizes the apprenticeships through the wage subsidy, making firms more willing to provide skills that are not firm-specific; (ii) firms are contractually required by BRAC to provide sector-specific skills to workers (Figure A.2). As the subsidy likely remains below the full cost of training, this reinforces the notion that a key reason for low take-up by firms is that the FT treatment imposed a costly requirement on firm owners to provide training, that was monitored and enforced.

The final dimension we consider probes whether VT and FT workers differ in the firm-specificity of their skill set at endline. It is hard to directly measure firm-specific skills for our study sectors. We thus approach this issue using data from the endline survey where we asked employed workers whether they considered their skills to be transferable across

[20]A few earlier papers have also used data from skills tests (Berniell and de la Mata (2016), Adhvaryu, Kala, and Nyshadham (2019)). 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).

firms. As Column 5 shows, relative to control workers, VT workers are significantly more likely than controls to report having transferable skills, although there is no statistical difference to FT workers ($p = 0.264$). This is again consistent with the earlier descriptive evidence in Figure A.3, that showed substantial learning for both sector-specific and firm-specific tasks among FT workers.

The key implication from these results is the similarity in long-run sector-specific skill accumulation in VT and FT treatments. This largely shuts down a channel through which differences in outcomes across treatments could have been generated, and reinforces the idea that such differences stem from the greater certifiability of skills obtained through vocational training.

3.4. *Employment and Earnings*

Table III presents ITT estimates for labor market outcomes, starting with the extensive margin of being in paid employment. Column 1 shows that, averaged over the three post-intervention survey waves, both forms of worker training raise employment probabilities: workers assigned to FT and VT treatments are 6pp and 9pp more likely to be employed, corresponding to 14% and 21% impacts over controls, whose unemployment rate is 56%. Hence, these ITT impacts of both training routes on youth unemployment rates are economically significant.

On the total effect margin, Column 2 shows VT and FT workers significantly increase the months worked in the year by 0.88 and 0.52, respectively, corresponding to 19% and 11% increases over controls. Hence, through the offer of either training route, young workers increase their labor market attachment. For VT workers, Column 3 shows this is further enhanced by them significantly increasing their weekly hours worked. This is evidence that the VT treatment has a stronger impact on employment.

Column 4 combines extensive margin and total effect margin effects to derive ITT impacts on total monthly earnings. Averaged over all post-intervention waves, the ITT earnings impact for VT workers is an increase of 25% over the control group. In contrast, there is no ITT earnings impact for FT workers, and the difference in earnings between VT and FT is statistically significant ($p = 0.048$). Hence, from a social planner's point of view, the use of wage subsidies attached to workers does not—on average—lead to earnings gains for those workers.

Column 5 combines these multiple labor market outcomes into one index following Anderson (2008), so accounting for the covariance structure in components, and we normalize by the standard deviation of the index in the control group to ease interpretation. The labor market index rises significantly for both FT and VT workers, with the magnitude being slightly larger for VT workers, because such workers have improved outcomes along all four components. However, all the findings point to both groups of workers increasing their labor market attachment: by being more likely to work, and by supplying more labor over time. This notion is reinforced by the fact that for both FT and VT workers, increases in employment are driven by increases in wage, not casual, employment. Following up on this, we consider whether workers are employed in their sector of training (for VT workers), or the firm sector they were matched to (for FT workers), or their first or second preferred sectors of employment (for controls). Column 6 shows: (i) for VT workers, this likelihood rises by 167%; (ii) for FT workers, this rises by 67%; (iii) the difference between VT and FT workers is significant ($p = 0.000$).

Of course, some of the differences in ITT impacts might be driven by differential compliance between VT and FT treatments. To account for this and also map to theories of

TABLE III

ITT ESTIMATES, LABOR MARKET OUTCOMES. OLS IPW REGRESSION COEFFICIENTS AND ROBUST STANDARD ERRORS IN PARENTHESES. BOOTSTRAP p-VALUES IN BRACES: UNADJUSTED p-VALUES (LEFT) AND ROMANO AND WOLF (2016) ADJUSTED p-VALUES (RIGHT)[a]

	Any Paid Work in the Last Month (1)	Number of Months Worked in the Last Year (2)	Hours Worked in the Last Week (3)	Total Earnings in the Last Month [USD] (4)	Labor Market Index (5)	Worked in Sector of Training/Matching in the Last Month (6)
Firm Trained	0.063	0.518	−0.196	1.89	0.105	0.045
	(0.025)	(0.259)	(2.27)	(2.20)	(0.051)	(0.015)
	{0.016; 0.046}	{0.049; 0.126}	{0.945; 0.945}	{0.408; 0.601}	{0.043; 0.043}	{0.005; 0.005}
Vocationally Trained	0.090	0.879	3.76	6.10	0.170	0.112
	(0.020)	(0.207)	(1.84)	(1.80)	(0.041)	(0.013)
	{0.001; 0.001}	{0.001; 0.001}	{0.043; 0.126}	{0.001; 0.005}	{0.001; 0.001}	{0.001; 0.001}
Mean Outcome in Control Group	0.438	4.52	28.2	24.7	0.003	0.067
Control for Baseline Value	Yes	No	Yes	Yes	Yes	Yes
p-values on tests of equality:						
Firm Trained = Vocationally Trained	[0.255]	[0.134]	[0.059]	[0.048]	[0.169]	[0.000]
N. of observations	3256	3256	2057	3115	3256	3256

[a]The data used are from the baseline and three worker follow-up surveys in all columns apart from Column 3. Since hours worked in the last month are not available at first follow-up (24 months after baseline), Column 3 uses data from the baseline, and the second and third follow-up surveys implemented 36 and 48 months after baseline, respectively. We report OLS IPW regression estimates in all columns, together with robust standard errors in parentheses. We also report unadjusted p-values (left) and p-values adjusted for multiple testing (right) in braces. These are computed using the step-down procedure discussed in Romano and Wolf (2016), with 1000 bootstrap replications. The number of hypotheses being tested simultaneously in the first four columns is eight—for each outcome, we test the impact of two treatments, FT and VT, and we do so for four outcomes. Therefore, the p-values in Columns 1 to 4 are adjusted for testing on eight hypotheses. In addition, we adjust the p-values in Column 5 for two hypotheses, that is, for testing two treatments at the same time. We again adjust the p-values for testing on two hypotheses in Column 6. 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. 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. 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 Labor Market Index in Column 5 is a standardized index of the variables in Columns 1–4, where we use the procedure discussed in Anderson (2008) to construct the index. In Column 6, the dependent variable is a dummy equal to 1 if the individual conducted any work in the last month in: the sector of training (for VT); the sector of matching (for FT); either the first or second most preferred sectors of employment, as indicated in the baseline survey (for Control). The weights for the Inverse Probability Weights (IPW) 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 = 1 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 IV

ATE ESTIMATES, LABOR MARKET OUTCOMES. 2SLS REGRESSION COEFFICIENTS, BOOTSTRAPPED STANDARD ERRORS IN PARENTHESES. BOOTSTRAP p-VALUES IN BRACES: UNADJUSTED p-VALUES (LEFT) AND ROMANO AND WOLF (2016) ADJUSTED p-VALUES (RIGHT)[a]

Dependent Variable:	Any Paid Work in the Last Month (1)	Number of Months Worked in the Last Year (2)	Hours Worked in the Last Week (3)	Total Earnings in the Last Month [USD] (4)	Labor Market Index (5)	Worked in Sector of Training/Matching in the Last Month (6)
Firm Trained	0.246	2.31	4.13	11.9	0.473	0.245
	(0.085)	(0.917)	(7.56)	(8.08)	(0.176)	(0.062)
	{0.004; 0.023}	{0.013; 0.029}	{0.622; 0.622}	{0.145; 0.241}	{0.009; 0.009}	{0.001; 0.001}
Vocationally Trained	0.135	1.38	7.12	10.3	0.272	0.190
	(0.028)	(0.302)	(2.61)	(2.65)	(0.059)	(0.019)
	{0.001; 0.001}	{0.001; 0.001}	{0.013; 0.026}	{0.001; 0.001}	{0.001; 0.001}	{0.001; 0.001}
Mean Outcome in Control Group	0.438	4.52	28.2	24.7	0.003	0.067
Control for Baseline Value	Yes	No	Yes	Yes	Yes	Yes
p-values on tests of equality:						
Firm Trained = Vocationally Trained	[0.141]	[0.255]	[0.661]	[0.830]	[0.202]	[0.343]
N. of observations	3256	3256	2057	3115	3256	3256

[a]The data used are from the baseline and three worker follow-up surveys in all columns apart from Column 3, which uses data from baseline and second and third follow-up, since hours worked in the last month are not available at first follow-up. We report 2SLS regression estimates in Columns 1–6, where treatment assignment is used as IV for treatment take-up. Treatment take-up is defined as a dummy equal to 1 if the worker (i) started firm training in FT or (ii) started vocational training in VT. Bootstrap standard errors are calculated using 1000 replications and reported in parentheses. We also report unadjusted p-values (left) and p-values adjusted for multiple testing (right) in braces. These are computed using the step-down procedure discussed in Romano and Wolf (2016), with 1000 bootstrap replications. The number of hypotheses being tested simultaneously in the first four columns is eight—for each outcome, we test the impact of two treatments, FT and VT, and we do so for four outcomes. Therefore, the p-values in Columns 1 to 4 are adjusted for testing on eight hypotheses. In addition, we adjust the p-values in Column 5 for two hypotheses, that is, for testing two treatments at the same time. We again adjust the p-values for testing on two hypotheses in Column 6. 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. 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. 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 Labor Market Index in Column 5 is a standardized index of the variables in Columns 1–4, where we use the procedure discussed in Anderson (2008) to construct the index. In Column 6, the dependent variable is a dummy equal to 1 if the individual conducted any work in the last month in: the sector of training (for VT); the sector of matching (for FT); either the first or second most preferred sectors of employment, as indicated in the baseline survey (for Control). 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.

training, Table IV presents ATE estimates for the same outcomes. The treatment effects are similar between VT workers that started vocational training, and FT workers that were hired and trained by firms incentivized by wage subsidies. The overall labor market index shows no difference in ATEs between FT and VT training ($p = 0.202$). Indeed, taking into account the differential compliance, the point estimate on the labor market index is actually higher for FT workers. This is driven by FT workers having a higher likelihood of any paid work, and a higher number of months worked on average over the three-year follow-up period—reversing the ranking from the ITT estimates.

The ATE estimates on monthly earnings in Column 4 show high experimental returns to vocational training: 42%, averaged over the post-intervention period. This begs the question of why workers do not themselves invest in vocational training given such returns. One explanation is credit constraints: as documented earlier, worker monthly earnings at baseline are \$6, while the vocational training costs over \$400.

An alternative explanation is that workers have incorrect beliefs about the returns to vocational training. In Supplemental Material Appendix A.2, we assess this using information collected from workers at baseline over their expected probability of finding work, and their expected earnings conditional on employment, if they received vocational training. We find that workers expect the returns to vocational training to be nearly 200%, many times more than the ATE estimate of returns, at 42% (Table A.VIII). In short, workers are overly optimistic with regard to the returns to vocational training, and such expectations do not explain their lack of investment in their own human capital.

Supplemental Material Appendix A.3 presents robustness checks on our baseline ITT findings.

3.5. *Dynamics*

3.5.1. *Retention*

Examining dynamic responses across treatments allows us to bridge to the job ladder model that assumes workers are in steady state. To begin with, we consider retention rates among apprentices. In each survey wave, we asked workers hired under the FT treatment if they were still employed at the same firm they were originally matched to. Figure A.5 plots the survival function for them: among those actually hired, 57% are employed for at least 6 months. Yet, their tenure does not last much longer: the average duration of employment at the matched firm, conditional on being strictly higher than 6 months, is 9 months. Crucially, by endline, almost none of these workers remain in the firm they were originally matched to. The fact that apprentices have relatively short employment spells at their matched-to firm suggests the FT treatment provided them skills that could be less firm-specific than if firms had borne the cost of training or not been contractually obliged by BRAC to provide training. The fact that FT workers transition away from firms they were trained by after the wage subsidy expires further limits any additional degree of firm-specific skills they accumulate relative to VT workers. Figure A.5 also shows survival functions for the first employment spell among VT and control workers: both groups have considerably longer first employment spells than workers hired in the FT by firms incentivized through a wage subsidy.

3.5.2. *Dynamic Treatment Effects*

Figure 3 shows quarterly dynamics of: (i) number of months worked (Panel A); (ii) total earnings (Panel B); (iii) average hourly wage in wage employment (conditional on being employed). For each outcome, the left-hand panel shows the descriptive evolution of the outcome, and the right-hand panel provides dynamic treatment effect estimates for these quarterly outcomes.

On the dynamics of quarterly employment (Panel A), we see that FT workers find employment more quickly than VT workers, but over time their employment rate converges to the control group, while employment rates for VT workers increase over time. This reversal of fortune between FT and VT is shown clearly in the dynamic ITT effects in the right-hand panel. The dynamic treatment effects on employment show that transition rates are stable by the fourth quarter of 2015. This underpins the assumption workers are in steady state for the structural model.

On the dynamics for quarterly earnings (Panel B), again FT workers do well initially, but then over time, their earnings fall behind those of VT workers. In contrast, VT workers steadily increase their earnings (the gradient is near linear) and diverge away from the control group over time.

Hence, the earlier documented similarity in treatment effects between FT and VT when averaged over post-intervention survey waves is driven by the earlier quarters in which FT workers were hired by firms incentivized through wage subsidies. Subsequent to that, the patterns of employment and earnings differ across VT and FT workers, with FT workers having similar employment profiles as control workers.

Panel C shows quarterly hourly wages (*conditional* on employment). Conditional hourly earnings for FT and VT workers rise relative to controls, and are not different from each other. For FT workers, hourly wages rise by 12% relative to controls, and for VT workers, they rise by 11%. To understand these wage impacts, we note the earlier results that sector-specific skills accumulation is similar among complier FT and VT workers. Hence, on-the-job productivity and wages should be similar for young workers entering the labor market through either training route.

4. JOB LADDER MODEL

4.1. *Set-up*

The labor market features a continuum of measure 1 of risk-neutral workers. Time is continuous and we assume workers have reached their steady-state labor market trajectories by the end of our study period, as suggested by Figure 3. Workers are heterogeneous in two dimensions: their training (treatment) status, T (where $T = VT$, FT, or C), and the amount of effective labor ε they supply each period given their training status T. We later relate a worker's effective labor to their sector-specific skills. Given the treatment effects on skills, we assume that relative to control workers, ε is increasing in training T, and is fully transferable across firms. Conditional on T, ε is assumed fixed so there is no human capital accumulation over and above that provided by training. We can thus think of ε as characterizing a worker's fixed productivity or type. The cross-sectional distribution of ε conditional on training is denoted $H(\varepsilon|T)$, with density $h(\varepsilon|T)$.

Following van den Berg and Ridder (1998) and Barlevy (2008), we assume firms post wage contracts indexed to worker type-ε, namely, they post a piece rate r paying a constant price per unit of effective labor. This fits our context, where the majority of workers are paid piece rates in manufacturing and service sectors. A worker of type-ε employed at a firm posting a piece rate r earns a wage $w = r\varepsilon$. We assume the offered piece rate comes from a distribution $F(r)$ with density $f(r)$, and denote the lower (upper) bound of the support of $F(r)$ as $\underline{r}$ ($\overline{r}$). All workers sample piece rates from this same distribution. The firm commits to pay w each period until the worker is laid off or quits.

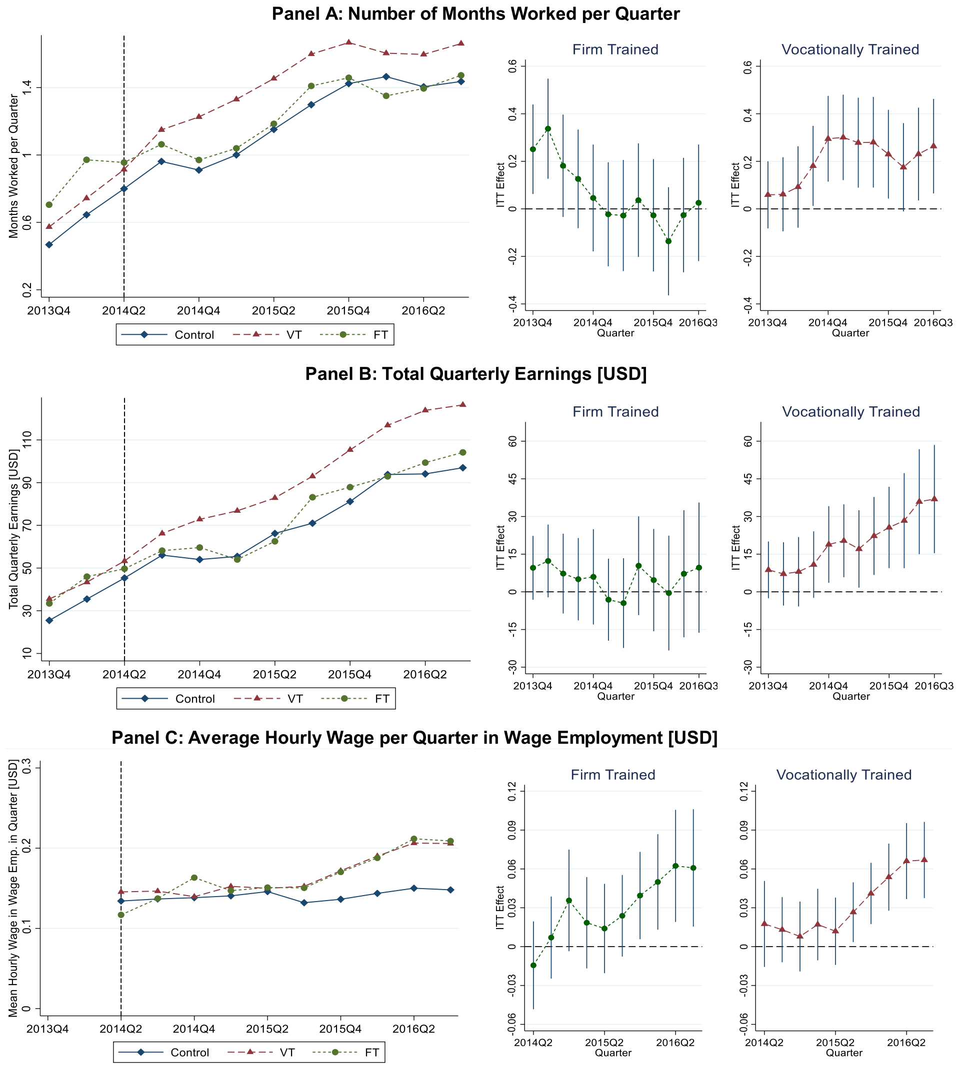

FIGURE 3.—Dynamics of employment, earnings, and wages. *Notes*: Data is from the first, second and third follow-up surveys. We use information on all employment and self-employment job spells reported by the workers in the twelve months prior to each survey. So the period considered goes from the fourth quarter of 2013, which is the first quarter covered in the first follow-up survey, until the third quarter of 2016, which is the last quarter covered in the third follow-up survey. Figures on the left of each panel report average number of months, earnings, and hourly wages for each quarter. The vertical dashed line corresponds to the end of the Firm Training intervention. Figures on the right of each panel report quarterly ITT treatment effects of Firm Training and Vocational Training on various outcomes, with 95% confidence intervals. All coefficients reported in each panel are estimated from the same dynamic treatment effects regression, where the FT and VT treatment indicators are interacted with dummies for each quarter considered, with robust standard errors. All regressions further include strata dummies, dummies for quarters, and a dummy for the implementation round. 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. Casual and agricultural occupations are coded as unemployment. Wages and earnings 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 earnings and wages are excluded.

Two points are of note. First, the distribution from which piece rates are drawn $F(\cdot)$ does not depend on treatment T. Once hired, worker productivity/skill is realized and so higher type-ε workers are paid a higher wage (at the same piece rate r) because $w = r\varepsilon$. We relax this assumption in a later extension, allowing for $F(r|T)$: this enables us to investigate, in a very reduced form way, whether, for example, workers in different treatments search differently across firms, who might then offer piece rates from different underlying distributions. An alternative interpretation of this extension is a set-up in which even once a worker is hired, their skills are not perfectly observable to the firm, as in a model of statistical discrimination where skill certificates are just a signal of unobserved worker ability.

Second, firms play no role in the model. Hence, to understand how the results derived from this partial equilibrium framework map to general equilibrium impacts, we later present evidence from the firm side of the experiment. This identifies how firms react to treatments along one key margin: the displacement (or crowding in) of other workers in the economy, which is central to understanding general equilibrium effects of our training interventions.

4.1.1. *Value Functions*

Workers can be unemployed or employed each period. Unemployed workers earn zero income each period of unemployment. $\lambda_0(T)$ is the arrival rate of job offers for an unemployed worker with 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 discount rate ρ, the value of unemployment for a type-ε worker with training status T is

$$\rho U(\varepsilon, T) = \lambda_0(T) \int_{R(\varepsilon,T)}^{\bar{r}} \big[V(x, \varepsilon, T) - U(\varepsilon, T)\big]\, dF(x). \tag{1}$$

Employed workers of type-ε earn $w = r\varepsilon$ at their firm in each period. They face an exogenous job destruction rate $\delta(T)$, depending on their training status. This captures both the quality of jobs and the expected duration of the employment relation. On-the-job search is allowed. $\lambda_1(T)$ is the arrival rate of job offers for an employed worker with training status T. She takes up this opportunity if the expected value of the job offer exceeds the current job value.[21] A type-ε worker with training status T has the following valuation of a job paying piece rate r:

$$\begin{aligned}\rho V(r, \varepsilon, T) &= r\varepsilon + \delta(T)\big[U(\varepsilon, T) - V(r, \varepsilon, T)\big] \\ &\quad + \lambda_1(T) \int_{r}^{\bar{r}} \big[V(x, \varepsilon, T) - V(r, \varepsilon, T)\big]\, dF(x).\end{aligned} \tag{2}$$

Combining (1) and (2), the key endogenous choice for workers—their reservation wage—solves the following for a type-ε worker with training status T:

$$R(\varepsilon, T) = \big[\lambda_0(T) - \lambda_1(T)\big] \int_{R(\varepsilon,T)}^{\bar{r}} \frac{\bar{F}(x)}{\rho + \delta(T) + \lambda_1(T)\bar{F}(x)}\, dx,$$

[21]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, Robin, and van den Berg (2000), Moscarini and Postel-Vinay (2018)).

where $\bar{F}(x) = 1 - F(x)$. Unemployed workers accept piece rates above $R(\varepsilon, T)$. We assume $\underline{r} \geq R(\varepsilon, T)$, so unemployed workers accept any job offer. Employed workers accept piece rates drawn from $F(r)$ that are higher than their current one irrespective of ε.

It is important to be clear on the distinct roles that two sources of worker heterogeneity play in the model: worker types-ε (as later proxied by their skills) determine wages conditional on employment, but do not play a role for labor market transitions. An alternative way to view this is that we assume skills are immediately observable to firms upon job offers being made to workers. Treatment status instead is allowed to impact both worker types and the labor market transition parameters, $\lambda_0(T)$, $\lambda_1(T)$, and $\delta(T)$. In particular, worker types are affected by skills accumulated during training, and that makes workers more productive when employed. Transitions are impacted by treatment if this makes it easier for workers to receive job offers, for instance, if treatment facilitates job search by allowing workers to better demonstrate their skills to employers.[22]

Finally, note that since transition parameters do not depend on ε, then the reservation wage for a type-ε worker does not depend on ε, so $R(\varepsilon, T) = R(T)$.

We do not explicitly model search effort. Rather, this is encompassed within the labor mobility parameters (λ_0, λ_1) that capture both job search when unemployed, on-the-job search effort, and underlying factors that drive search effort in these states and differ across treatments such as the certifiability of skills.

4.1.2. *Steady State*

We close the model by deriving steady-state conditions where, for expositional ease, we omit conditioning on T. The following steady-state relationship characterizes when outflows and inflows for unemployment are equal for workers of type-ε:

$$\lambda_0 u(\varepsilon) = \delta\big[h(\varepsilon) - u(\varepsilon)\big],$$

$$u(\varepsilon) = \frac{\delta}{\delta + \lambda_0} h(\varepsilon),$$

where $u(\varepsilon)/h(\varepsilon)$ is the unemployment rate for type-ε workers. This is independent of type-ε, which is unsurprising as worker labor market mobility is independent of ε, and depends only on T (through $\delta(T)$ and $\lambda_0(T)$). Hence, this is also the population unemployment rate.

As employed workers can search on-the-job, the cross-sectional distribution of *observed* piece rates for type-ε workers $G(r|\varepsilon)$ differs from the offer sampling distribution $F(r)$. This is because observed piece rates are those accepted by workers. For type-ε employed workers with piece rate $\leq r$, the steady-state relationship for employment is

$$\big[\delta + \lambda_1 \bar{F}(r)\big]\big[h(\varepsilon) - u(\varepsilon)\big]G(r|\varepsilon) = \lambda_0 F(r) u(\varepsilon). \tag{3}$$

The LHS of (3) is the outflow from the stock of type-ε workers employed at a piece rate less than r.[23] The RHS of (3) is the inflow into employment from unemployment.

[22] By not allowing worker characteristics, except treatment assignment, to impact these transition parameters, the model does not deal with additional forms of selection into and out of employment.

[23] To see this, note that of the $(h(\varepsilon) - u(\varepsilon))G(r|\varepsilon)$ workers employed at a piece rate $\leq r$, a fraction δ have the job exogenously terminated, while a fraction $\lambda_1 \bar{F}(r)$ receive and accept an offer greater than r.

Using these steady-state relationships, we can derive the link between $G(\cdot)$ and $F(\cdot)$:

$$F(r) = \frac{(\delta + \lambda_1)G(r|\varepsilon)}{\delta + \lambda_1 G(r|\varepsilon)}, \tag{4}$$

$$G(r|\varepsilon) = \frac{\delta F(r)}{\delta + \lambda_1 \bar{F}(r)}, \tag{5}$$

where $G(r|\varepsilon) = G(r)$ is independent of ε (given worker mobility is independent of ε). We see that $G(r)$ FOSD $F(r)$ unless there are no job-to-job transitions ($\lambda_1 = 0$), that is, because on-the-job search leads to outside offers, there exists a wedge between offered and accepted piece rates.

4.2. *Estimation*

4.2.1. *Data*

In each survey, we asked respondents to provide their monthly labor market history since the last survey. We use this 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, one year after the end of vocational training, to our endline in November 2016. 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 between job and unemployment states (τ_{JU_i}, τ_{JJ_i}, τ_{UJ_i}) using a maximum of two spells since the steady state has been reached. Hence, the model is estimated off the last two survey waves.

4.2.2. *Identification*

Our survey records i's wage in each employment spell j, w_{ij}. However, the distributions of worker productivity $H(\varepsilon)$, piece rates $G(r)$, and piece rate offers $F(r)$ are not observed. We tackle these identification issues as follows. First, the data include a proxy for worker productivity: their sector-specific skills test score, s. We assume measurable skills relate to true worker productivity as follows: $\varepsilon = s^\alpha$. Taking the natural logarithm of wages, we obtain the following expression for worker i in spell j: $\ln(w_{ij}) = \ln(r_{ij}) + \alpha \ln(s_i) + \epsilon_{ij}$, where r_{ij} denotes the piece rate paid to worker i in spell j and where ϵ_{ij} captures idiosyncratic measurement error, which we allow for to bring the model to the data. As r_{ij} is unobserved, to identify α using OLS we need r_{ij} to be independent of s_i. However, in our model, training simultaneously impacts worker type-ε and observed wages w_{ij} (because $G(r|\varepsilon)$ depends on $\delta(T)$ and $\lambda_1(T)$). To correct for this omitted variable bias, we run the following regression of wages on skills, controlling for treatment status:

$$\ln(w_{ij}) = \gamma_0 + \alpha \ln(s_i) + \sum_k \gamma_k T_{ik} + u_{ij}, \tag{6}$$

where T_{ik} is a dummy equal to 1 if worker i is assigned to treatment k, and $u_{ij} = \ln(r_{ij}) + \epsilon_{ij}$. We estimate this for workers transitioning from unemployment into employment because of the assumption that workers accept any job offer when unemployed.

Table A.X shows estimates of α, using the same survey waves from which we estimate the model parameters. Our baseline estimate is $\widehat{\alpha} = 0.263$ from Column 1. This falls slightly to $\widehat{\alpha} = 0.245$ when we condition on worker characteristics and strata dummies.[24] Given $\widehat{\alpha}$, i's measured sector-specific skills (s_i), and i's wages in employment spell j (w_{ij}), we recover the estimated piece rate for each worker-spell as $\hat{r}_{ij} = w_{ij}/s_i^{\hat{\alpha}}$. We thus recover $G(r)$ and use the steady-state conditions to estimate the distribution of piece rate offers, $F(r)$. These functions are recovered for each group of workers (controls, non-compliers, compliers), so we do not impose a common $F(\cdot)$ across treatments.[25]

4.2.3. *Parameterization*

We assume the model parameters have the following parametric form:

$$\lambda_0 = \lambda_{00} + \sum_k \lambda_{0k} T_k, \tag{7}$$

$$\lambda_1 = \lambda_{10} + \sum_k \lambda_{1k} T_k, \tag{8}$$

$$\delta = \delta_0 + \sum_k \delta_k T_k, \tag{9}$$

where T_k denotes worker's treatment status. In line with SUTVA, we assume workers across treatments do not interact with each other in the labor market. Given differences in compliance between FT and VT treatments, we treat compliers and non-compliers as separate groups and simultaneously estimate different parameters (λ_0, λ_1, δ) and $h(\varepsilon)$ for each of the five groups (control, non-compliers in FT and VT, compliers in FT and VT). From the treatment effect estimates on skills, we expect compliers and non-compliers to have different $h(\varepsilon)$ distributions. In the Supplemental Material, we detail the construction of the likelihood function. We estimate the parameters using maximum likelihood using the two-step procedure in Bontemps, Robin, and van den Berg (2000), recovering asymptotic standard errors for the parameters.

4.2.4. *Supportive Evidence*

We provide three pieces of evidence supporting the model structure. First, a key assumption is that labor mobility does not depend on worker type-ε. We test this in the sample used to estimate the model parameters. Table A.XI shows estimates from an OLS regression of a dummy equal to 1 if the individual is employed in November 2015, on the worker's sector-specific skills test score, controlling for treatment status, where the unit of observation is the job spell. Across worker samples in Columns 1 to 3, there is zero correlation between skills (our proxy for ε) and the likelihood the worker's initial steady-state spell is in employment. Column 4 confirms this to be the case when we allow skills to differentially impact this employment probability by treatment.

Second, we test the model prediction that wage growth occurs between, not within, job spells. Decomposing workers' wage growth into that occurring within and between job spells, 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.[26]

[24]Column 3 adds interactions of the logarithm of the skills test score with treatment dummies: we see both point estimates are close to zero, suggesting there is no need to include a $T_{ik} \times s_i$ interaction in (6).

[25]In not imposing a common $F(\cdot)$ across treatments, we are implicitly assuming that differences across treatments arise from measurement error.

[26]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

Third, we take the sample used to estimate the model parameters and estimate ITT regressions on outcomes closely linked to labor mobility, a key mechanism. The results in Table V show that VT workers experience more job spells (Column 1), and this leads them to gain over two months of extra employment relative to the control group (Column 2). On the other hand, and consistent with the dynamic evidence, in steady state, FT workers do not experience more job spells and are not employed for longer than the control group. These differences between VT and FT are significant at the 1% level.

We further see that while VT workers are more likely to be employed in the first spell in steady state (Column 3), VT workers who are unemployed in the first spell are significantly more likely than control workers to transition to employment (Column 4). The estimate in Column 4 relates to λ_0, and shows the positive treatment effect on the number of work spells for VT workers in Column 1 is not only driven by them being more likely to be initially employed in the first spell. Again, we find no impact of FT on transitions away from unemployment, and the difference between FT and VT workers is significant at the 5% level.

Finally, in Columns 5 and 6, we estimate treatment effects relating to λ_1 and δ, respectively. Column 5 shows no significant impact of treatments on the number of transitions conditional on being employed in the first spell. Column 6 shows that, conditional on being employed in the first spell, FT workers have significantly fewer transitions into unemployment, but we are unable to reject this estimate is different for VT workers ($p = 0.274$).

While suggestive of the mechanisms that might be at play, these estimates do not map exactly into structural parameters, and so cannot be easily interpreted. We therefore move to presenting the full set of model parameter estimates.

5. MODEL ESTIMATES

5.1. *Parameters*

Table VI presents the baseline results. Panel A shows the mean worker type-$\overline{\varepsilon}$ and parameter estimates ($\widehat{\delta}$, $\widehat{\lambda}_0$, $\widehat{\lambda}_1$) for controls (Column 1), non-compliers in each training arm (Columns 2 and 3), and compliers in each training arm (Columns 4 and 5). This confirms the distribution of worker types-ε has higher means for compliers from either training route (2.65, 2.58), with both being at least 10% higher than the mean type-ε among controls and non-compliers (2.31, 2.28, 2.35). We next note that job destruction rates δ are identical for FT and VT workers (0.023). To the extent that δ captures job quality, this is consistent with the similar skills acquired through both training routes leading to similar job qualities.

On labor mobility, for the arrival rate of job offers when unemployed (λ_0): (i) controls and non-compliers have similar estimates; (ii) remarkably, the arrival rate of job offers when unemployed is almost identical for complier firm-trained workers and control workers ($\widehat{\lambda}_0 = 0.019, 0.020$): the additional skills and labor market experience gained

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 jobs 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).

TABLE V

ITT ESTIMATES, LABOR MOBILITY. OLS REGRESSION COEFFICIENTS, ROBUST STANDARD ERRORS IN PARENTHESES[a]

	Number of Non-Casual Work Spells	Total Months of Non-Casual Work	Non-Casual Work in First Spell	Number of UJ Transitions	Number of JJ Transitions	Number of JU Transitions
Sample of Workers:	All	All	All	Unemployed in First Spell	Employed in First Spell	Employed in First Spell
	(1)	(2)	(3)	(4)	(5)	(6)
Firm Trained	−0.043	−0.155	0.004	−0.013	0.008	−0.197
	(0.062)	(0.735)	(0.033)	(0.055)	(0.076)	(0.112)
Vocationally Trained	0.167	2.35	0.094	0.122	0.074	−0.092
	(0.052)	(0.599)	(0.028)	(0.046)	(0.099)	(0.090)
Mean Outcome in Control Group	1.03	10.6	0.180	0.734	0.338	0.569
p-value FT = VT	0.000	0.000	0.005	0.011	0.461	0.274
Observations	1158	1158	1158	894	264	264

[a]We report OLS regression coefficients, and robust standard errors in parentheses. The data set is a cross-section of workers built using data from the second and third worker follow-up surveys. 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. 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 the third follow-up interview (which ended in December 2016). 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 sample is restricted to individuals who were unemployed in their first spell in Column 4, and individuals who were employed in their first spell in Columns 5 and 6. The outcomes in Columns 4, 5, and 6 are respectively the number of unemployment-to-job, job-to-job, and job-to-unemployment transitions the worker experienced between November 2015 and the date of the interview. All regressions include strata dummies and a dummy for the implementation round. 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. 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 VI

BASELINE ESTIMATES OF THE JOB LADDER SEARCH MODEL. TWO-STEP ESTIMATION PROCEDURE IN BONTEMPS, ROBIN, AND VAN DEN BERG (2000). ASYMPTOTIC STANDARD ERRORS IN PARENTHESES. STEADY STATE: NOVEMBER 2015 (DATA FROM SECOND AND THIRD FOLLOW UP)[a]

			Non-Compliers		Compliers	
		Control	Firm Trained	Vocationally Trained	Firm Trained	Vocationally Trained
		(1)	(2)	(3)	(4)	(5)
Panel A: Parameter Estimates (Monthly)						
Average units of effective labor [USD]		2.31	2.28	2.35	2.65	2.58
Job destruction rate, δ		0.027	0.027	0.026	0.023	0.023
		(0.003)	(0.006)	(0.005)	(0.007)	(0.004)
Arrival rate of job offers if UNEMPLOYED, λ_0		0.019	0.019	0.018	0.020	0.028
		(0.002)	(0.003)	(0.003)	(0.005)	(0.003)
Arrival rate of job offers if EMPLOYED, λ_1		0.038	0.042	0.054	0.032	0.039
		(0.010)	(0.019)	(0.022)	(0.022)	(0.013)
Panel B: Competition for Workers and Unemployment						
Interfirm competition for workers		1.41	1.54	2.08	1.41	1.68
	% Impact:		**8.7%**	**47%**	**−0.5%**	**19%**
Unemployment rate		0.59	0.59	0.59	0.53	0.46
	% Impact:		**0.26%**	**0.70%**	**−9.9%**	**−23%**
Unemployment duration (months)		52.8	53.1	56.2	50.0	35.9
	% Impact:		**0.54%**	**6.5%**	**−5.2%**	**−32%**
Employment duration (months)		36.8	36.7	38.5	44.2	42.9
	% Impact:		**−0.10%**	**4.7%**	**20%**	**17%**
Panel C: Wages and Earnings						
Average monthly OFFERED wages [USD]		43.1	42.6	43.9	49.4	48.2
Average monthly ACCEPTED wages [USD]		62.6	63.2	70.3	71.8	73.1
Impact on annual earnings [USD]			**1.66**	**34.1**	**95.4**	**169**
	% Impact:		**0.54%**	**11%**	**31%**	**55%**

[a]The data set 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 data set contains at most two spells (and one transition) per individual. The data come 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 the third follow-up interview (which ended in December 2016). 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 nonparametrically 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 conducted using information from both the first and second spells for each individual to recover the parameter estimates. As shown in Panel A, we estimate separate parameters for Control and Treatment groups, and, within treatments, for compliers and non-compliers. Outputs in Panel B are derived from the model and computed as functions of the estimated parameters: (i) interfirm competition for workers $= \lambda_1/\delta$; (ii) unemployment rate $= \delta/(\delta + \lambda_0)$; (iii) unemployment duration $= 1/\lambda_0$; employment duration $= 1/\delta$. In Panel C, average monthly offered and accepted wages are computed as the product of average offered and accepted piece-rates, and average units of effective labor. We assume workers draw piece-rates from the same offer distribution $F(r)$. $F(r)$ is the kernel density estimate of a weighted average of the distributions of offered piece-rates across treatments—$F(r|T)$—where such distributions are obtained from their steady-state relationship with nonparametrically estimated $G(r|T)$. Weights are equal to the share of individuals in each treatment. For each treatment, we then re-invert $F(r)$ using estimated parameters and steady-state relationships to obtain $G(r|T)$ under the assumption that workers draw piece-rates from the same offer distribution.

by firm-trained workers count for little if they fall off the job ladder into unemployment; (iii) complier VT workers have transition rates 40% higher than complier FT workers. These λ_0 estimates are statistically different from each other ($p = 0.082$); $\widehat{\lambda}_0(\text{VT})$ is also statistically higher than for the control group ($p = 0.004$). On job offer arrival rates when employed (λ_1) among compliers, we see that VT workers have arrival rates 22% higher than FT workers, although this is not statistically different ($p = 0.365$).

Among trained workers, the difference in labor mobility of vocationally trained workers relative to firm-trained workers is the higher rate of UJ transitions: when unemployed, VT workers get back onto the job ladder more quickly. This pattern of mobility is in line with VT workers having more certifiable skills than FT workers, and this certifiability having especially high returns when workers are unemployed. When employed, the certifiability of skills matters less because potential employers have a signal of a worker's skills because they are already employed. Hence, JJ transitions play less of a role in explaining differences between training routes.[27]

Finally, we note the overall similarity in parameter estimates between controls and non-compliers on nearly all dimensions in Panel A: this reinforces the idea that compliance is essentially exogenous, in line with the earlier discussion.[28]

5.2. *Unemployment, Wages, and Earnings*

Panel B shows impacts on key labor market statistics for workers. Beneath each, we report the percentage impact relative to controls. The first row measures the intensity of interfirm competition for workers (labor market tightness): this is the number of outside offers received before being laid off: $\frac{\lambda_1}{\delta}$. Interfirm competition for complier VT workers is slightly higher than for FT workers, driven by VT workers receiving slightly more outside job offers when employed. Relative to controls, interfirm competition for complier VT workers rises by 19%. This is in line with the earlier ATE estimates on skills accumulation, which showed VT workers reporting more transferable skills across firms. In contrast, the steady-state interfirm competition for FT workers is very similar to that for control workers.

Both training routes substantially reduce youth unemployment rates for compliers relative to controls. FT compliers have 10% lower unemployment rates in steady state; for VT compliers, the reduction is 23%. Both impacts are of economic significance given the high levels of youth unemployment in Uganda. Unemployment durations fall for both sets of compliers, but this fall is far larger for VT workers (32% vs. 5.2%): this is because of the significantly greater unemployment-to-job mobility of VT workers, so they more quickly get back onto the job ladder if they fall into unemployment (recall job destruction rates are identical across compliers).

Ultimately, what we are concerned with is whether workers transition to better paying jobs. These results are shown in Panel C. To derive these earnings impacts, we take the appropriately weighted mean of the kernel density estimate of each group specific $F(r)$, to impose a common $F(r)$. We invert this using the steady-state relationships (4) and (5) to obtain $G(r|\cdot)$ and $G(w|\cdot)$ distributions for each group of workers.

[27]These results are also consistent with a world in which job interviews are effective but costly, and skill certificates reduce the cost of conducting an interview. Hence, FT workers and VT workers have the same probability of hire (and the same wage offer) conditional on being interviewed (because they have the same skills), but VT workers are more likely to get an interview because they have certificates.

[28]The p-values from F-tests of the joint significance of the $(\varepsilon, \widehat{\delta}, \widehat{\lambda}_0, \widehat{\lambda}_1)$ parameters between VT and FT non-compliers and Controls are 0.314, 0.971, 0.932, and 0.774, respectively.

Panel C shows that when unemployed, the mean offered wages for control and non-complier workers are similar. This is in line with the assumption unemployed workers accept any job offer, and with the estimated mean type-ε being similar across controls and non-compliers. In contrast, steady-state wages offered to complier FT and VT workers are substantially higher; this is because these workers are more skilled, so even if firms draw from the same piece rate distribution $F(r)$, this translates into higher wages for workers because of the complementarity between piece rates and skills ($w = r\varepsilon$). Of course, workers only accept job offers if the value of the offered job is greater than their current one. The mean accepted wage for complier VT and FT workers is thus much higher than offered wages, but not much different between the two (73.1 vs. 71.8). This is in line with the skills of both trained workers being similar, so earnings conditional on employment are similar. The impact of training on skills also explains why accepted wages for compliers are higher than for controls/non-compliers. The final row in Panel C shows that the steady-state annual earnings (unconditional) of complier VT workers rise by 55% over controls, while the earnings of complier FT workers rise by just over half of that, 31%.

Combining these results precisely explains the dynamic treatment effects: vocational trainees pull away from FT workers in their employment rates and earnings (Panels A and B in Figure 3) because they are more likely to get back onto the job ladder if they fall into unemployment. These dynamics are not so much due to any greater job-to-job mobility, suggesting the returns to skills certifiability are higher when unemployed. Moreover, compliers across training routes move as far up the job ladder as each other—wages conditional on employment are similar for complier VT and FT workers because their skills are similar (Panel C in Figure 3). The key distinction is that VT workers are more likely to get back onto the job ladder if they fall off it.

Two further points are of note. First, it is useful to contrast the experimental and model estimates of the returns to training. To be clear, these are not estimated from the same samples, but the former informs us about dynamics to the steady state, and the latter informs on the nature of the steady state. The estimated returns to vocational training in steady state for compliers (55%) are higher than the ATE estimate (42%). In contrast, the steady-state returns to firm-provided training (31%) are lower than the experimental returns (48%). This contrast arises because the steady-state calculations account for the lower UJ transition rates of FT workers. In steady state, they get back on the job ladder at the same rate as controls, slowly closing the gap between them in terms of employment rates.[29]

Second, controls and non-compliers have the highest job arrival rates when employed (Panel A, Columns 1 to 3). These selected groups of less skilled workers that find employment churn between jobs at a high rate, but they do not progress up the job ladder. To see this, note that, within treatment, the accepted wage for non-compliers is on average lower than for compliers; hence, they do not move far up the job ladder despite higher JJ mobility and interfirm competition (especially for VT non-compliers). This is because unskilled workers do not improve their type over time, say through skills accumulation. Hence, as shown in the final row of Table VI, ultimately the steady-state earnings impact

[29] 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 (2018)). Dynamic impacts are then driven by intertemporal complementarity in workers' capacity to learn. We partially explore this hypothesis by estimating whether workers' cognitive abilities, and other preference parameters, change differentially by training route. We find no evidence of such mechanisms.

is almost identical between controls and non-compliers (remaining well below earnings impacts for either complier groups).

In Supplemental Material Appendix A.4, we present robustness checks where we: (i) allow piece rates to be drawn from treatment-specific functions $F(r|T)$; (ii) show how our results vary when using the highest and lowest $\widehat{\alpha}$'s across specifications in Table A.X.

6. EXTENSIONS

6.1. *Employment Displacement*

To understand how these results can map to general equilibrium impacts, we present evidence from the firm side of the experiment, focusing on one key margin: the displacement (or crowding in) of other workers. The right-hand side of Figure 1 summarizes the design from firms' perspective. We focus on the comparison between firms assigned to the wage subsidy offer and the control group of firms.[30] We estimate the effect on firms of being offered to meet an untrained worker and a wage subsidy to hire and train that worker, using the following ITT specification for firm f in randomization strata s:

$$y_{fst} = \beta \text{Firm-Trained}_f + \gamma y_{f0} + \delta \mathbf{x}_{f0} + \lambda_s + \vartheta_t + u_{fst}.$$

y_{fst} is the firm outcome of interest in post-intervention survey wave t, Firm-Trained$_f$ is a dummy equal to 1 if the firm is in the FT treatment. We also estimate ATE impacts, where we instrument hiring a worker the firm is matched to with treatment assignment. y_{f0} is the 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-BRAC branch, and account for attrition using IPWs, and we also present Lee bounds. Finally, we examine dynamic impacts by estimating effects: (i) in the short run, only using the first firm follow-up survey conducted towards the end of the six-month FT training intervention; (ii) in the long run, averaging treatment effects over survey waves two to four, that run to years after wage subsidies have expired, and long after any initially hired workers have left FT firms (Figures 2 and A.5).[31]

The results are in Table VII. Column 1 shows no evidence of other hires being crowded out by FT workers in the short run when the wage subsidy is in place. The change in the number of employees hired almost equals the number of post-intervention hires (Column 2), and there is no evidence of more workers being fired post-intervention among FT firms (Column 3). There are also no long-run impacts on employment, hires, or fires for FT firms relative to control firms (Columns 5 to 7). The (scaled-up) similarity in the pattern of results between the ITT and ATE estimates suggest impacts are driven by firms that hire a worker they are matched to.

Overall, we see little evidence of employment displacement of control workers in the long run (a result robust to using only the final year of data), and there are no net employment effects of wage subsidies in the long run. Although we can never be certain of the impacts on workers outside of our evaluation sample, this lack of employment crowd-in

[30] Measuring employment displacement effects of hiring VT workers was part of our original design with the VT + match treatment, T4. However, as described earlier, take-up rates in this treatment are too low to say anything about outcomes (even at first follow-up) for firms that VT + match workers were hired by.

[31] $\mathbf{x}_{f0}$ 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 dummies for whether the respondent provided a phone number at baseline, and for whether he/she was an employee of the firm (rather than the firm owner or the manager), the number of network firms, and dummies for interviewers at baseline.

TABLE VII

EMPLOYMENT AND OTHER FIRM OUTCOMES. IPW REGRESSION COEFFICIENTS, STANDARD ERRORS CLUSTERED BY SECTOR-BRANCH IN PARENTHESIS, LEE BOUNDS IN BRACKETS[a]

	Short Run (First Follow-Up)				Long Run (Second to Fourth Follow-Ups)			
	Number of Employees (1)	Number of Post-Intervention Hires (2)	Number of Post-intervention Fires (3)	Log (Average Monthly Profits) (4)	Number of Employees (5)	Number of Post-intervention Hires (6)	Number of Post-intervention Fires (7)	Log (Average Monthly Profits) (8)
PANEL A: ITT Estimates								
Firm Trained	0.350	0.370	−0.118	0.011	−0.116	−0.054	−0.093	0.113
	(0.205)	(0.137)	(0.160)	(0.114)	(0.154)	(0.077)	(0.150)	(0.050)
	[0.553; 1.16]	[0.430; 0.668]	[−0.272; 0.111]	[−0.089; 0.219]	[−0.133; 0.237]	[−0.087; 0.176]	[−0.007; 0.435]	[−0.069; 0.192]
PANEL B: ATE Estimates								
Firm Trained	1.34	1.42	−0.453	0.036	−0.358	−0.127	−0.182	0.313
	(0.770)	(0.441)	(0.623)	(0.375)	(0.431)	(0.229)	(0.451)	(0.169)
Mean outcome in control firms	2.43	0.650	1.56	211	2.32	0.893	1.41	182
Number of observations	569	569	569	444	1611	1606	1611	1178

[a]The data used are from the firm follow-up data surveys. Panel A reports OLS IPW regression estimates together with standard errors adjusted for heteroscedasticity and clustered at the branch-trade level in parentheses. We report Lee (2009) bounds in brackets, where we implement a conditional Lee bounds procedure that is able to condition on dummies for the interview round and baseline trade. Underlined bounds are significantly different from zero at the 95% condfidence level. Panel B reports 2SLS IPW regression estimates, where treatment assignment is used as IV for treatment take-up. Treatment take-up is defined as a dummy equal to 1 if the firm hired one of the workers it was matched with. All regressions control for the value of outcome at baseline (when available), and include branch and trade fixed effects, survey wave dummies, and dummies for the month of interview. Baseline controls also include the owner's sex, business age (measured as number of years since the business was established) and business age squared, firm size, and owner's years of education. The weights for the Inverse Probability Weights (IPW) are computed separately for attrition at second, third, and fourth follow-up. The instruments for the IPW estimates are dummies for whether the respondent provided a phone number at baseline, and for whether he/she was an employee of the firm (rather than the firm owner or the manager), the number of network firms, and dummies for interviewers 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 (1 USD = 2385 UGX). Monthly profits and revenues are truncated at 99th percentile.

or crowd-out in our control group of firms in the long run suggests the partial and general equilibrium employment effects of the model coincide.[32]

6.2. *Counterfactuals*

6.2.1. *Understanding the Relative Importance of Mechanisms*

We assess the relative importance of the mechanisms at the heart of the model to explain steady-state unemployment rates, earnings conditional on employment, and unconditional earnings. These mechanisms relate to differences in: (i) arrival rates of job offers (λ_0, λ_1); (ii) separation rates (δ); (iii) skills (s) that drive worker type-ε. For counterfactual analysis, we hold two sets constant and allow only the third to vary with treatment. We thus assess the qualitative importance of each mechanism (due to the parameter interactions and nonlinearities in the model, these are not exact decompositions).[33]

The results are in Table VIII. Panel A shows the baseline level of each outcome across controls, FT, and VT compliers (applying to each counterfactual scenario considered). Panel B compares FT and VT compliers to controls. When equating parameters, we set them all to the value in the control group. Panel C compares compliers in FT and VT to each other. When equating parameters, we set them all equal to the value for VT compliers. We do not conduct this comparison for earnings conditional on employment as Panel A shows minimal differences along this margin.

On unemployment rates: (i) Panel B shows the impact of firm training over controls is mostly driven by lower separation rates; (ii) Panel C shows that the differential impact on unemployment of VT over FT is nearly all due to differences in job offer rates (explaining 110% of the gap). This confirms the central importance of skills certification for youth unemployment in this setting. On earnings conditional on employment, Panel B shows: (i) the impact of FT over controls is mostly due to skill differentials, with separation rates being qualitatively around one third as important; (ii) the impact of VT over controls is mostly due to skill differentials; separation rates are of less but still non-negligible importance in explaining this difference.

Combining both outcomes, on unconditional earnings we see that: (i) skill and separation rate differences are equally important in explaining the gap between FT and controls; (ii) all three mechanisms—skills, separation rates, and job offer arrival rates—explain the gap between VT and controls; (iii) the gap between VT and FT workers is overwhelmingly due to differences in job offer arrival rates.

---

[32]Our two-sided experimental design adds to a nascent literature examining impacts of wage subsidy programs on firms (McKenzie, Assaf, and Cusolito (2016), Hardy and McCasland (2017), de Mel, Mckenzie, and Woodruff (2019)).

[33]To set the initial conditions for the simulations, we use the parameter estimates in Panel A of Table VI and kernel density estimates of $F(r)$ and $h(\varepsilon)$ for each group of workers. We construct a common $F(r)$ function by taking an appropriately weighted average of the group-specific $F(r)$ functions. We then simulate a panel of 50,000 workers observed over 48 months in steady state, where workers are randomly assigned to treatment in the same proportions as in our experiment. As we showed earlier that compliance is uncorrelated with worker observables, workers are also randomly assigned to take up their treatment in the same proportion as in the experiment. The simulation allows ($h(\varepsilon)$, δ, λ_0, λ_1) to vary across treatments, and in line with the baseline model, we assume all piece rate offers are drawn from the same $F(r)$ distribution across treatments. In each simulation, the average $G(r)$ is calculated as the mean piece rate in the population of employed workers across the 48 months of the simulation. The average of $F(r)$ is calculated as the mean piece rate received by workers transitioning from unemployment to employment. Final statistics are computed as the average results across 10 simulations.

TABLE VIII

COUNTERFACTUAL ANALYSIS ON RELATIVE IMPORTANCE OF MECHANISMS[a]

	Unemployment			Earnings Conditional on Employment			Unconditional Earnings		
	Different Arrival Rates (1)	Different Separation Rates (2)	Different Skills (3)	Different Arrival Rates (4)	Different Separation Rates (5)	Different Skills (6)	Different Arrival Rates (7)	Different Separation Rates (8)	Different Skills (9)
Panel A: Baseline Levels									
Control		0.589			64.0			26.3	
Firm Trained		0.531			73.4			34.4	
Vocationally Trained		0.456			74.4			40.5	
Panel B: $FT = VT = Control$									
Firm Trained	***21%***	***76%***	***0%***	***−39%***	***33%***	***100%***	***−10%***	***56%***	***54%***
Vocationally Trained	***72%***	***29%***	***0%***	***3%***	***27%***	***74%***	***51%***	***30%***	***29%***
Panel C: $FT = VT$									
Vocationally Trained	***110%***	***−9%***	***0%***	–	–	–	***137%***	***−11%***	***−15%***

[a]The table reports OLS estimates from simulated data generated from the model. We run 10 simulations of the behavior of 50,000 workers followed over a period of 48 months. In each simulation, we randomly assign individuals to treatment in the same proportions as in our experiment. Workers are also randomly assigned to take up their treatment in the same proportion as in the experiment. In each simulation, we calculate treatment effects as the average monthly impact of FT and VT on employment and earnings across the 48 months from OLS regressions. We then aggregate estimates across the different simulations. Panel A shows mean unemployment rate, conditional and unconditional earnings in the baseline simulations, when we allow arrival rates λ_0 and λ_1, separation rates δ, and the distribution of effective units of labor $h(\varepsilon)$ to vary across Control and treatment groups. Panel B shows percentage changes in treatment effects beween the baseline and the counterfactual simulations when we set the parameters indicated at the top of the table for individuals in the FT, VT groups to be the same as for the Control group. In Panel C, we set the parameters of FT workers to be equal to those of VT workers. So, in Panel C, the parameters of individuals in VT and Control remain the same as in the baseline simulation. In Columns 1, 4, and 7, we set arrival rates λ_0 and λ_1 to be equal across treatments. In Columns 2, 5, and 8, we set separation rates δ to be equal across treatments. In Columns 3, 6, and 9, we set the distribution of effective units of labor $h(\varepsilon)$ to be equal across treatments. The percentages in Panel B are calculated as the percentage change in FT and VT coefficients between baseline and counterfactual simulation. The percentages in Panel C are instead calculated as the percentage change in the difference between the VT and FT coefficients in the baseline and counterfactual simulations.

6.2.2. *Extending Training to Other Workers in the Economy*

We construct counterfactuals considering if the training interventions were targeted to other workers. The issue is relevant because individuals were recruited into our sample based on the potential offer of vocational training, and eligibility criteria targeting disadvantaged youth. As Table A.II shows, relative to labor market active youth in Uganda, our sample is worse off in terms of labor market outcomes at baseline. Workers in our sample might be especially selected on two traits relative to other labor market entrants: ability and patience.

The first is relevant because our workers are unemployed at baseline with worse labor market histories. The second trait is relevant because only those prepared to forgo the opportunity cost of labor market offers during six months of training would have been willing to apply for our offer. To build counterfactual scenarios targeting youth with different ability or patience than in our sample, we exploit the fact that the job ladder model has built in observed heterogeneity of workers type-ε, depending on their skill. To begin with, we examine heterogeneous skills accumulation in FT and VT treatments by ability and patience. In our sample, worker ability is measured by Raven matrices tests, and worker patience is measured using answers to questions about their willingness to wait to receive (hypothetical) monetary rewards. For each trait, we classify a worker of being high/low type if they are above/below the sample median.[34]

Table A.XIII shows the results on heterogeneous skills accumulation. Column 1 highlights that there are different levels of skill accumulation by high/low ability. This difference is statistically significant in the FT treatment ($p = 0.072$) and marginally so within the VT treatment ($p = 0.116$). Hence, there is a complementarity between firm-provided training and underlying ability in the accumulation of sector-specific skills. As expected, Column 2 shows such complementarities also exist between firm-provided training and high/low patience workers, but these are less precisely estimated.

We use these estimates to construct counterfactuals assuming: (i) the distribution of worker types, $H(\varepsilon)$, varies by workers with above/below the median trait in controls, complier VT workers, and complier FT workers (so among six groups overall); (ii) there is a share θ of high-trait (ability or patience) workers in the economy; (ii) the parameters (δ, λ_0, λ_1) are the same as in the baseline model for controls and compliers in each treatment (*irrespective* of their ability/patience, so traits only impact outcomes through skill accumulation). We then simulate counterfactual impacts in the economy varying the share of high-trait workers θ.

Recall that in the model, worker types-ε determine wages conditional on employment, but play no role for labor market transitions, so that unemployment rates do not vary with traits (or hence θ) by assumption. Therefore, we focus on simulated impacts on earnings conditional on employment, and unconditional earnings. These are in Figure 4 (along with 95% confidence intervals). Panels A and B vary the share of high-ability workers treated. Panels C and D vary the share of high-patience workers treated.

Panel A shows earnings impacts when employed for FT workers rise steeply in the share of high-ability workers treated. At the extreme, if firm training were only taken up by low-ability workers, treatment effects on earnings conditional on employment would be 5%, and if only high-ability workers were targeted, they would be closer to 20%. At this extreme, the point estimate treatment effect of FT would actually be higher than for

[34]Patience is measured at baseline. Cognitive ability is measured at first follow-up. We verify that there are no direct treatment effect impacts on these traits.

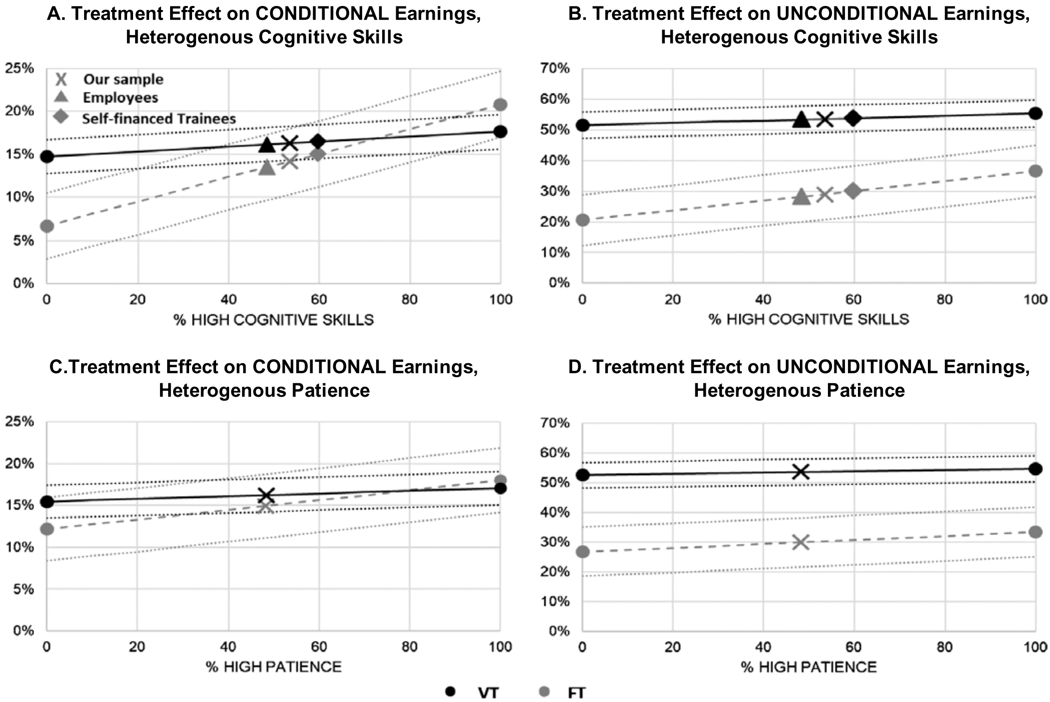

FIGURE 4.—Counterfactuals. *Notes*: This shows the percentage impact of FT and VT on conditional and unconditional earnings from OLS estimates run on simulated data generated from the model. The dashed lines show 95% confidence intervals. We run 10 simulations of the behavior of 50,000 workers followed over a period of 48 months. In each simulation, we randomly assign individuals to treatment in the same proportions as in our experiment. Workers are also randomly assigned to take up treatment in the same proportion as in the experiment. In each simulation we calculate treatment effects as the average monthly impact of FT and VT on employment and earnings across the 48 months from OLS regressions. We then aggregate estimates across the different simulations. We only show estimated impacts on compliers. Each panel shows the treatment effects when we vary the share of individuals with high cognitive skills and high patience in the population (from 0 to 100%). To do that, we first divide workers in our data into high/low Raven matrices using their score on the Raven Matrices test implemented at first follow-up. Workers are assigned to the High Raven group if they scored on or above the median of the Raven Matrices test. Similarly, workers are divided into high/low Patience using their answers to a series of questions about their willingness to wait to receive (hypothetical) monetary rewards at baseline. Workers are assigned to the High Patience group if they had a value of Patience on or above the median. We then obtain kernel density estimates of the distribution of effective units of labor $h(\varepsilon)$ for each of these groups. In the simulations individuals are randomly assigned to be High/Low Raven and High/Low Patience in the proportion indicated on the figures, and draw their effective units of labor from the corresponding distribution. The crosses indicate the exact percentages of High Cognitive Skills and High Patience individuals in our sample. These are respectively 53.5% and 48.3%. The triangles and diamonds indicate the percentage of High Cognitive Skills individuals in two other Ugandan samples from related studies. The first sample—“self-financed trainees”—are youth analyzed in Bassi and Nansamba (2020) and includes trainees that have self-financed their own training. The second sample—“employees”—are employees in a representative sample of firms in welding, furniture making and grain milling from Bassi et al. (2020). The share of High Cognitive Skills individuals is 59.8% in the “self-financed trainees” sample, and 48.5% in the “employees” samples.

VT. Panel B shows that for unconditional earnings, the impact varies from 20% to 38% as θ varies from zero to 1.

Panels C and D show less pronounced impacts for patience, but treatment effect impacts for both training routes are increasing in θ. This is especially the case for earnings conditional on employment for FT: if only high-patience workers were targeted, then the treatment effects of both training routes converge to be almost identical at 17%.

These counterfactuals offer an explanation for why studies in the literature might differ in their estimated returns to firm-provided apprenticeships: as θ varies across samples in the literature, the impacts of training vary, and their relative ranking can also reverse.

We use this range of simulated estimates to discuss implied treatment impacts on two alternative groups of worker from our same context. To do so, we draw on worker samples from related studies with comparable information on the cognitive abilities of workers. First, Bassi and Nansamba (2020) surveyed 1000 young workers currently receiving training in similar sectors and at similar VTIs in Uganda. These trainees have self-financed their vocational training. The same 10-question Raven matrices test was used to measure their cognitive ability. In their sample, the share of high-ability workers is 60% (higher than in our sample). Second, Bassi, Muoio, Porzio, Sen, and Tugume (2020) surveyed over 2000 employees in a representative sample of firms in welding, furniture making, and grain milling sectors, operating in urban areas in Eastern, Central, and Western Uganda. Cognitive ability was measured using a subset of the 10-question Raven matrices we used. Using the overlapping matrices, we find the share of high-ability workers in this employee sample to be 49%.

Panels A and B in Figure 4 superimpose the simulated impact our FT and VT treatments would have had they been targeted to workers self-financing their vocational training, or to workers currently employed in similar manufacturing sectors, under the assumption that the difference between samples is the share of high-cognitive-ability workers in each. Two important comparisons are as follows. First, for self-financed VTI attendees, the impacts of VT are only marginally higher than for our workers. This suggests factors associated with not being credit constrained (and so allowing workers to self-finance vocational training as in Bassi and Nansamba (2020)) are not much correlated to the returns to vocational training. Second, for currently employed workers, the impacts of firm-provided training are slightly lower than for apprenticeships in our sample. This suggests firms might not be hiring workers optimally, in line with evidence presented earlier that 52% of SME firm owners reported the inability to screen workers as a constraint.

7. IRR AND EXTERNAL VALIDITY

7.1. *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 planner implementing either policy. Table IX presents IRR calculations for each treatment, where our benchmark case assumes a social discount rate of 5%, and that the steady-state earnings gains to workers last 15 years. We assume no employment displacement effects of either training route. For FT workers, this is based on the earlier firm-side results. For VT workers, this is an implicit assumption we make. We present IRR calculations based on steady-state earnings impacts on: (i) all workers; (ii) compliers. The former is most appropriate from a social planner's point of view. The latter provides a sense of the private returns to workers if they could overcome constraints to make these kinds of human capital investment themselves.[35]

[35]Table A.XIV shows the model estimates when we pool non-compliers and compliers in each treatment. Panel A shows that because of the high compliance in VT, it remains the case that VT workers have higher

TABLE IX

INTERNAL RATE OF RETURN[a]

	All Workers		Compliers	
	Firm Trained (1)	Vocationally Trained (2)	Firm Trained (3)	Vocationally Trained (4)
Social discount rate = 5%				
Remaining expected productive life of beneficiaries	15 years	15 years	15 years	15 years
Panel A. External parameters				
Total cost per individual at year 0 [USD]:	**368**	**510**	**368**	**510**
(i) Training costs (for 6 months)	302	470	302	470
(ii) Program overheads costs	31	4	31	4
(iii) Foregone earnings (for 6 months)—average at baseline	36	36	36	36
Panel B. Estimated total earnings benefits				
1) NPV change in steady state earnings (from model estimates)	222	1246	990	1753
2) Benefits/cost ratio	**0.604**	**2.44**	**2.69**	**3.44**
3) Internal Rate of Return (IRR)	**−0.017**	**0.224**	**0.250**	**0.327**
Panel C. Sensitivity				
Sensitivity to different expected remaining productive life of beneficiaries				
Remaining expected productive life = 10 years	**−0.088**	**0.196**	**0.225**	**0.309**
Remaining expected productive life = 5 years	**−0.308**	**0.057**	**0.093**	**0.196**
Sensitivity to different earnings				
Forgone earnings = 90th percentile at baseline (120 USD)	**−0.040**	**0.187**	**0.196**	**0.277**
Forgone earnings = double 90th percentile at baseline (241 USD)	**−0.065**	**0.146**	**0.144**	**0.225**
Forgone earnings = max earnings at baseline (794 USD)	**−0.127**	**0.047**	**0.031**	**0.102**
Forgone earnings = double max earnings at baseline (1588 USD)	**−0.169**	**−0.017**	**−0.035**	**0.027**
Panel D. Program costs for IRR to equate social discount rate				
5 Total cost per individual at year 0 [USD]	–	**1246**	**990**	**1753**
Sensitivity to different discount rates/time horizons				
Social discount rate = 10%	–	913	726	1285

[a]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 computation of the IRR uses as input for the benefit the ITT and ATE impacts of Firm and Vocational Training 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.

Panel A in Table IX shows the per intended beneficiary 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 (\$50.3/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.[36]

Panel B shows the NPV of 15 years of earning gains. Focusing on the impacts for all workers, we see that the gains to those assigned to FT are around 18% of those assigned to VT. However, the benefit-cost ratio is below 1 for FT and the IRR is negative. It does *not* pay for the social planner to replicate the kind of subsidized apprenticeship offered in the FT treatment. However, the reason for this negative IRR is the low compliance in the FT treatment. This low compliance is driven by a lack of firms taking up the offer of the wage subsidy and the matched-to worker.

However, in these labor markets, we do observe workers paying firms for an apprenticeship. To see why this is so, we redo the IRR calculations but based on the steady-state earnings for compliers—namely, those that acquire firm-provided or vocational training. For those workers that are hired and trained by firms under the FT treatment, Column 3 shows the benefit-cost ratio is well above 1 (2.69) and the IRR is 25%. The rise in IRR for FT workers highlights the high social returns from being able to overcome firm constraints in taking on and hiring young workers. A core problem remains to design such interventions; this might mean offering higher subsidy rates, or some other incentive.

For the VT treatment, Column 2 shows that based on steady-state gains for all workers assigned to this treatment, the benefit-cost ratio is 2.44 and the IRR is 22%. Even with take-up rates of 68%, vocational training generates high returns, and certainly compares favorably to a menu of other anti-poverty policies focusing on human capital accumulation. When using the benefits for those workers that actually take up the vocational training treatment, Column 4 shows the benefit-cost ratio rises to 2.69 and the IRR rises to 33%.

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. We see that the IRR for FT drops off more quickly with shorter productive lives, while for VT, it remains at 5.7% or above under the alternative scenarios. This is as expected given the different wage profiles to the interventions. The VT intervention always pays for itself within a decade. On foregone earnings, only under very extreme assumptions does the IRR for VT ever fall below 5%.

These calculations are based on the cost structure of the NGO BRAC that we collaborated with. Their overhead costs represent the *marginal* cost of extending their activities in Uganda to the training program evaluated. To get a sense of the 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, focusing on the scenarios where

---

mean ε than controls ($p = 0.000$). The arrival rate of job offers when unemployed is still higher for VT than for FT ($p = 0.069$), and FT and controls have the same $\widehat{\lambda}_0$ estimate. Panel B shows qualitatively similar patterns on unemployment rate impacts and durations as the baseline model, and Panel C shows overall steady-state earnings to be 7%, and 39% higher for all workers assigned to FT and VT with respect to those in Control.

[36]These cost structures are per intended beneficiary and do not change across the two sets of IRR calculations based on all workers or only those that comply with their treatment. This is because we take the view that such costs are incurred by the social planner ex ante, prior to compliance being observed.

the baseline IRR is positive to begin with. For vocational training, in Column 2 we see total costs per beneficiary would have to increase almost threefold for the intervention to break even. The final row performs the same calculation assuming a 10% social discount rate. In this case, the costs for vocational training would still need to nearly double for the social planner not to intervene.

7.2. *External Validity*

In meta-analyses of training interventions in low-income settings, Blattman and Ralston (2015) and McKenzie (2017) documented that most interventions have a very low IRR. Figure A.6 compares our ITT treatment impacts relative to the 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 five reasons why our returns are high relative to other studies, each of which opens up avenues for future work.

First, our treatments are intensive and of a 'big push' variety. Specifically, both treatments last six months, the wage subsidy rate is higher than some other studies, and in the wage subsidy treatment firms are contractually obliged to train hired workers.

Second, we worked with a limited set of VTIs, pre-selected 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 relax credit constraints for workers, it is not obvious the results would replicate 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. This might explain why similar programs providing vouchers to workers redeemable at any training provider have had more limited success (Galasso, Ravallion, and Salvia (2004), Groh et al. (2016)).

Third, there are design issues: our experiment separates out in-class vocational training from a wage subsidy program. It also 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 one of these sectors. Workers were not given free reign over which sector to train in: they had to choose among sectors with substantial demand for skilled workers. This limited scope for mismatch between worker skills and firms they were offered to.

Fourth, only 13% of workers attrit over our four-year evaluation, comparing favorably to other studies. Indeed, in the meta-analysis of McKenzie (2017), all but one study has attrition rates above 18%. As Figure A.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.

Finally, workers selected into our sample given the oversubscription design might differ from other young workers. We exploited this fact earlier to build counterfactual treatment effect estimates of targeting our interventions to other workers. Given youth unemployment rates of 60%, the allocation of talent in the economy might improve if we think of the large pool of unemployed workers as heterogeneous, and those attracted to the sample through the offer of vocational training as being positively selected relative to the average unemployed youth. It is exactly these kinds of motivated job seekers that the economic gains from matching to jobs might be highest for.

8. CONCLUSION

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 lifetime 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 forms of tertiary education as a means of transitioning young people into good jobs.

We document that both types of training, when provided over an extended period, can have highly positive effects on employment and earnings within disadvantaged youth transitioning into the labor market. This is in sharp contrast with workers who receive neither type of training and who remain largely unemployed or employed in casual work, as is common among unskilled workers across the developing world. What is even more revealing is that the steady-state effects on employment and earnings for VT workers are almost twice as large as those for FT workers. This result speaks directly to the value of the certifiability of skills, which is a key difference between skills gained through vocational training and those gained via firm-provided apprenticeships. Estimating a job ladder model of worker search reveals labor market mobility as the main mechanism for the divergence in employment and earnings profiles between VT and FT workers.

Two final implications of our findings are noteworthy. First, as shown in the second counterfactual, the complementarity between cognitive ability and skills differs between FT and VT treatments. This hints at training being imparted differently at firms and VTIs: at firms, higher-ability workers learn more, while VTIs appear to ensure more workers gain skills. This is in line with the objectives of firms and VTIs, but highlights the importance for policy to account for differential targeting by worker ability by skills providers in the economy. Second, the cost of non-certifiable skills is greater in lower-income settings because the firm size distribution is highly skewed: young workers are reliant on hiring by SMEs that have limited potential for promotion *within* firms, which would be a natural alternative to labor mobility between firms as the way to climb the job ladder. Hence, policies to relax constraints on firm size and deepen worker hierarchies within firms are a natural counterpart to policies promoting labor mobility across firms.

These implications open up a rich set of research possibilities for analyzing how vocational education might be best organized in these countries, how governments might intervene to incentivize firms to provide apprenticeships, to certify skills, and to unlock constraints on firm growth. Few areas of research are more important for determining the development trajectory of low-income countries. With overwhelmingly young populations and inadequate education systems, the training that young workers obtain as they transition into the labor force will be pivotal in determining whether we end up with a sea of workers in unskilled, informal work or with a growing share in stable, skilled jobs. At a time when labor markets are undergoing rapid structural change, we need to take up the challenge set by Becker (1964) and Schultz (1981) of working out how best to invest in the human capital of young workers so that they can secure meaningful jobs.

REFERENCES

ABEBE, G., S. CARIA, M. FAFCHAMPS, P. FALCO, S. FRANKLIN, AND S. QUINN (2020): "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Review of Economic Studies* (forthcoming). [2371,2372]

ABEL, M., R. BURGER, AND P. PIRAINO (2019): "The Value of Reference Letters: Experimental Evidence from South Africa," *AEJ: Applied Economics*, 3, 44–71, https://www.aeaweb.org/articles?id=10.1257/app.20180666&&from=f. [2371]

ADHVARYU, A., N. KALA, AND A. NYSHADHAM (2019): "Returns to On-the-Job Soft Skills Training," Report, http://static1.1.sqspcdn.com/static/f/884336/28307195/1591029269300/PACE_dec2019.pdf?token=onKcEtpcq1IYmC%2F3dj4Eg%2BbAbdk%3D, Michigan. [2386]

ALFONSI, L., O. BANDIERA, V. BASSI, R. BURGESS, I. RASUL, M. SULAIMAN, AND A. VITALI (2020): "Supplement to 'Tackling Youth Unemployment: Evidence From a Labor Market Experiment in Uganda'," *Econometrica Supplemental Material*, 88, https://doi.org/10.3982/ECTA15959. [2374]

ANDERSON, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103 (484), 1481–1495. [2387-2389]

ATTANASIO, O., A. KUGLER, AND C. MEGHIR (2011): "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence From a Randomized Trial," *AEJ: Applied Economics*, 3, 188–220. [2370,2372]

AUTOR, D. H. (2001): "Why Do Temporary Help Firms Provide Free General Skills Training," *Quarterly Journal of Economics*, 116, 1409–1448. [2370]

BARLEVY, G. (2008): "Identification of Search Models Using Record Statistics," *Review of Economic Studies*, 75, 29–64. [2391]

BASSI, V., AND A. NANSAMBA (2020): "Screening and Signaling Non-Cognitive Skills: Experimental Evidence From Uganda," USC-INET Research Paper No. 19-08, Available at SSRN: https://ssrn.com/abstract=3268523 or http://dx.doi.org/10.2139/ssrn.3268523. [2371,2407,2408]

BASSI, V., R. MUOIO, T. PORZIO, R. SEN, AND E. TUGUME (2020): "Achieving Scale Collectively," CEPR Discussion Paper No. DP15134, Available at SSRN: https://ssrn.com/abstract=3674922. [2407,2408]

BECKER, G. S. (1964): *Human Capital*. Chicago: University of Chicago Press. [2412]

——— (1994): "Human Capital Revisited," in *Human Capital: A Theoretical and Empirical Analysis With Special Reference to Education*. Chicago: University of Chicago Press. [2412]

BERNIELL, L., AND D. DE LA MATA (2016): "Starting on the Right Track: Experimental Evidence From a Large-Scale Apprenticeship Program," Report, CAF, https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=EEAESEM2016&paper_id=1244. [2386]

BLATTMAN, C., AND L. RALSTON (2015): "Generating Employment in Poor and Fragile States: Evidence From Labor Market and Entrepreneurship Programs,". Report, Chicago, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2622220. [2373,2379,2411]

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. [2386]

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, 305–358. [2393,2396,2399]

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. [2370,2372]

CARD, D., J. KLUVE, AND A. WEBER (2018): "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations," *JEEA*, 16, 894–931. [2373,2376]

CARRANZA, E., R. GARLICK, K. ORKIN, AND N. RANKIN (2019): "Job Search, Hiring, and Matching With Two-Sided Limited Information About Workseekers' Skills," Report, Duke University, https://events.barcelonagse.eu/live/files/2640-kateorkin67266pdf. [2371]

CREPON, B., AND P. PREMAND (2019): "Creating New Positions? Direct and Indirect Effects of an Apprenticeship Program,". Report, CREST, https://www.vwi.unibe.ch/unibe/portal/fak_wiso/b_dep_vwl/a_inst_vwl/content/e195818/e195991/e195992/e788116/PREMAND_Crepon_Premand_Apprenticeship_ger.pdf. [2373]

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, 531–580. [2384]

DE MEL, S., D. MCKENZIE, AND C. WOODRUFF (2019): "Labor Drops: Experimental Evidence on the Return to Additional Labor in Microenterprises," *AEJ: Applied*, 11, 202–235. [2380,2404]

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, 128–142. [2372,2411]

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. [2371,2372,2381,2411]

HARDY, M., AND J. MCCASLAND (2017): "Are Small Firms Labor Constrained? Experimental Evidence From Ghana," Report, UBC, https://pedl.cepr.org/sites/default/files/Hardy%20%26%20McCasland_Are%20Small%20Firms%20Labor%20Constrained.pdf. [2370,2404]

JENSEN, R. (2010): "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 125, 515–548. [2370]

LEE, D. S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102. [2403]

LEUVEN, E., AND H. OOSTERBEEK (2008): "An Alternative Approach to Estimate the Wage Returns to Private-Sector Training," *Journal of Applied Econometrics*, 23, 423–434. [2386]

LEVINSOHN, J., N. RANKIN, G. ROBERTS, AND V. SCHOER (2014): "Wage Subsidies and Youth Employment in South Africa: Evidence From a Randomised Control Trial,". Report, Yale, https://ideas.repec.org/p/sza/wpaper/wpapers207.html. [2372]

MACLEOD, W. B., E. RIEHL, J. E. SAAVEDRA, AND M. URQUIOLA (2015): "The Big Sort: College Reputation and Labor Market Outcomes," *AEJ: Applied Economics*, 9, 223–261. [2371]

MCKENZIE, D. (2017): "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence," *World Bank Research Observer*, 32, 127–154. [2373,2411]

MCKENZIE, D., N. ASSAF, AND 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. [2404]

MOSCARINI, G., AND F. POSTEL-VINAY (2018): "The Cyclical Job Ladder," *Annual Review of Economics*, 10, 165–188. [2393]

NEAL, D. (2018): *Information, Incentives, and Education Policy*. Cambridge: Harvard University Press. [2401]

PALLAIS, A. (2014): "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104, 3565–3599. [2370,2371]

PISSARIDES, C. A. (1994): "Search Unemployment With On-the-Job Search," *Review of Economic Studies*, 61, 457–475. [2412]

ROMANO, J. P., AND M. WOLF (2016): "Efficient Computation of Adjusted P-Values for Resampling-Based Stepdown Multiple Testing," *Statistics and Prob. Letters*, 113, 38–40. [2383,2385,2388,2389]

SCHULTZ, T. W. (1981): *Investing in People: The Economics of Population Quality*. California: University of California Press. [2412]

VAN DEN BERG, G. J., AND G. RIDDER (1998): "An Empirical Equilibrium Search Model of the Labor Market," *Econometrica*, 66, 1183–1221. [2391]

Co-editor Charles I. Jones handled this manuscript.

Manuscript received 20 December, 2017; final version accepted 26 June, 2020; available online 2 July, 2020.