SCREENING AND SIGNALING NON-COGNITIVE SKILLS: EXPERIMENTAL EVIDENCE FROM UGANDA

Vittorio Bassi* Aisha Nansamba†

September 21, 2018

Supplemental Online Appendix

*University of Southern California, INET and IFS. Email: vbassi@usc.edu.
†BRAC NGO. Email: nansamba.aisha@gmail.com.

1 Selection of Vocational Training Institutes

This section describes the selection of Vocational Training Institutes (VTIs) for the intervention, and presents summary statistics, by the VTI participation status. To select the VTIs for the project we contacted 24 large, reputable institutions taken from the list of VTIs registered with the Directorate of Industrial Training. We only considered VTIs training at least 50 workers about to graduate in the sectors considered in our project, and operating in areas where our project was to be implemented. Contacted institutions included VTIs that BRAC had worked with before for other skills training programs, as well as new potential partner VTIs.[1]

VTI principals were told that, if they accepted to participate, BRAC would promote the Job Placement Program with students at their institutions. VTI principals were informed about all components of the program: they were made aware that BRAC would be conducting skills measurements on their trainees using a variety of methods, including in-class observations, and that some of this information could be disclosed to managers during the job placement intervention. The project was well received: only five of the 24 VTIs reported not being interested to participate. We followed up with the interested VTIs and established partnerships with 15 of them.[2] Table S.I reports basic VTI descriptives by whether the VTI took part in the project or not.[3] The table shows that we were successful at identifying large, well established VTIs: the average institution taking part in the project has been operating for about 35 years, offers 11 different types of courses, and has about 350 students currently enrolled. The table further shows that there are no significant differences in means between the institutions that took part in the program, and those that did not, for six of the seven characteristics considered.[4] This suggests that such basic observables do not drive selection into the project, although any conclusions from Table S.I are obviously subject to the caveat of the sample being very small.

2 Skills Assessments

2.1 Skills Assessments of Eligible Workers in the Census

Information on the cognitive skills and Big 5 traits of all eligible trainees was collected at the VTIs during the initial worker census. To measure cognitive skills, we used a 10-item Raven matrices test. To measure the Big 5 traits we used a 10-item questionnaire. Questions were

[1] We contacted 24 VTIs as we expected this would be sufficient to achieve our target sample of workers.

[2] Four of the VTIs interested to participate were not included in the final sample for the project as their trainees were graduating at a time of the year not compatible with our intervention.

[3] This information is missing for 3 VTIs.

[4] Institutions not participating to the program have a higher number of students currently enrolled. However, this difference is driven by one outlier: if we remove the VTI with the highest value of this variable the difference between the two groups is not significant anymore ($p\text{-}value = 0.459$).

translated in Ugandan with the help of a local psychologist from the School of Psychology at Makerere University in Kampala. Panel A of Figure S.I reports the distribution of cognitive skills and of the Big 5 traits.

2.2 Skills Assessments of Workers in the Experimental Sample

We used teacher surveys to measure the soft skills of trainees on the following dimensions: attendance, discipline, communication skills, proactivity and willingness to help other students in class.[5] Teachers were asked to use a 0-10 scale to rate the skills of their trainees, in absolute terms (and so not relative to the class). BRAC field staff conducted in-class observation of worker behavior. This took the form of spot-checks, whereby our field staff showed up unannounced at the VTIs and attended a number of classes, noting down details of student behavior.[6] The main purpose of the spot-checks was to verify the information provided in the teacher surveys. For attendance/time-keeping, BRAC staff were asked to record whether trainees were observed as present and on-time in class; for willingness to help others, we recorded whether they were observed voluntarily helping other students; for proactivity, we recorded whether they were observed asking questions to the teacher; for discipline, we recorded whether they were called up by the teacher for inappropriate behavior in class, such as making noise.[7] For each trainee, we can then construct the percentage of classes in which they were observed by BRAC staff as engaging in a particular type of behavior.

Panel B of Figure S.I shows that, reassuringly, there is a positive and significant correlation between the measures from the teacher-surveys, and the corresponding observed behaviors from the spot-checks. As teacher surveys have fewer missing values, we use information from the teacher surveys to develop our final measures of the following skills: attendance, proactivity, discipline, willingness to help others and communication skills.

To measure creativity, we developed an 8-item questionnaire, together with the help of the psychologist. We construct a creativity index by calculating the unweighted average of answers to the 8 questions. The distribution of the creativity index is reported in Panel C of Figure S.I. The index goes from 1 to 5, with a higher value of the index corresponding to higher creativity. Finally, to measure trustworthiness, we used a version of a standard trust game, played by trainees with real money.[8] Every trainee played exactly the same game. Specifically, trainees

[5]The instructors taking part in the survey had been teaching the trainees in our sample for 12 months on average by the time of the survey. Teachers were not asked to rate the creativity and trustworthiness of trainees as we wanted to keep the teacher surveys focused on those skills that are easier to assess for an external examiner.

[6]Spot-checks were conducted weekly while trainees were still enrolled at the VTIs. The median trainee was observed in six different spot-checks.

[7]BRAC staff were not asked to rate the communication skills of trainees, as it was not straightforward to link communication skills to an objective type of behavior observable during a short time-period in class.

[8]The main reference for the trust game is Berg et al. (1995).

were told that an anonymous sender (not among the other trainees at the VTIs) had decided to send them 1,000 Ugandan shillings (about 30 USD cents), and that BRAC had tripled this amount, so that every trainee had received 3,000 Ugandan shillings (or approximately 90 USD cents). Trainees were then asked how much (if anything) they were willing to send back to the original sender, to reciprocate for the initial transfer received. A higher amount sent back is interpreted as higher trustworthiness. Panel C of Figure S.I reports the distribution of the amounts sent back.

2.2.1 Grading Procedure

To grade the workers on each of the five soft skills, we begin from the skills assessed in the teacher surveys: communication, willingness to help others and attendance, and observe the distribution of grades given by teachers. Teachers were explicitly asked to grade workers using an *absolute* scale, and so not to curve the grades within the sample participating in the study. We then match the grade distribution of creativity and trustworthiness (which are not measured from the teacher surveys) to the average grade distribution of the other three skills.

3 Treatment Assignment and Matching Procedure

We decided to match firms and workers only within sectors, as it is reasonable to assume that workers would mainly look for jobs in the same sector of training. In terms of how many workers to match to each firm, the baseline data reveal that the median firm normally interviews three applicants before a vacancy is filled. Our aim was to replicate a "hiring round" in our sample of firms, and so we tried to keep the number of workers matched to each firm as close as possible to three. However, to keep our intervention logistically feasible, we imposed as a constraint that a worker could not be matched to more than five different firms. Therefore, we developed a matching rule that in each sector fixed the number of workers matched to each firm to exactly three workers, as long as this was compatible with all workers in that sector being matched with at least one firm, and with at most five different firms.

The main challenge with implementing this rule was that the number of workers and firms in our experimental sample varied substantially across sectors. We were able to assign exactly three workers per firm in catering and hairdressing. Firms in tailoring were assigned instead four workers each, and firms in motor-mechanics (the sector with the highest ratio of workers to firms) were assigned six workers each. On the other hand, firms in carpentry and welding (sectors with a relatively low supply of graduates) were matched to two workers each.[9]

[9] All firms in welding were matched to two workers, with the exception of firms in Fortportal, Gulu, Kasese, Mbale, Mbarara, Nyendo, Ojwina which were assigned only one worker.

In practice, we decided to implement the matchings at the submarket level, where a submarket is defined as a BRAC branch-sector combination. So firms in a given submarket were only matched with workers assigned to the same submarket. This was done to reflect the local nature of informal labor market search, appropriate to this context: while workers are mobile in principle, evidence shows that transportation costs limit significantly the ability of workers to move for job search in similar informal contexts (Abebe et al., 2017, Franklin, 2016).[10]

To assign workers to BRAC branches for the matchings, we followed as much as possible their stated preferences at baseline: workers were asked about their first, second and third most preferred urban areas where they would like to look for employment after the end of training. Each trainee was initially assigned to her most preferred urban area. This initial allocation resulted in a few submarkets that had either "excess supply" or "excess demand" of workers, relative to the ranges needed for the allocation rule described above to be implementable. So workers were randomly moved across urban areas for the purpose of the matching assignments, until every submarket had a number of assigned workers and firms falling in the ranges required by the assignment rule. When moving workers across submarkets, their second and third stated preferences were taken into account as much as possible. While some workers had to be matched in an urban area different from their preferred one, such cases were limited: 72% of workers were matched in their preferred urban area, and only 9% were matched in a urban area outside their three most preferred ones.

Once the allocation of workers and firms to submarkets was finalized, we randomly assigned workers and firms to a Treatment group and a Control group of approximately equal sizes. So the randomization produced four groups: (i) Treatment workers, (ii) Control workers, (iii) Treatment firms; (iv) Control firms. The randomization was stratified by submarket.

Finally, we implemented the matching allocations: within each submarket and treatment group, firms and workers were matched randomly, according to the assignment rule described above. So, for example, hairdressing firms in the Kasese urban area were randomly matched with exactly three workers who had received hairdressing training and were assigned to be matched in the Kasese area. The three workers were then matched with between one and five hairdressing firms in Kasese.

4 Balance Checks at the Match Level

Tables S.II and S.III report balance checks at the match level for workers and firms, respectively. Starting from Table S.II, Panel A shows that at baseline the sample of scheduled matches is well balanced on worker characteristics across Treatment and Control. Out of the 12 character-

[10]To further limit any concerns related to transportation costs, BRAC paid for the travel costs of workers to the location of interview on the day of the matchings.

istics considered, only the Age variable shows significant differences in means between the two experimental group, with workers in Treatment matches being slightly older. However, column 4 shows that the normalized difference for this variable is small (-.114). Also, we are not able to reject the null hypothesis that all observable worker characteristics are jointly insignificant in predicting Treatment assignment, as shown by the *p-value* on the *F-test* at the bottom of column 2 (*p-value* = .866). Panel B shows that the sample of matches *that take place* remain well balanced on worker characteristics: there are no differences at the 5% level between Treatment and Control in any of the 12 variables considered for the balance checks.

Table S.III shows similar balance checks on firm characteristics, and again the unit of observation is the match. Panel A shows that we achieve good balance at baseline: out of the 10 firm characteristics considered, only the dummy for whether the firm owner attended a VTI shows means that are statistically different between Treatment and Control at the 5% level, with firms in Control being more likely to be managed by owners who attended a VTI in the past. However, column 4 shows that normalized differences remain small, as they are never larger than .13. Reassuringly, all the observable firm characteristics are jointly insignificant in predicting Treatment assignment, as shown by the *p-value* on the *F-test* at the bottom of column 2 (*p-value* = .402). Panel B shows that the sample of matches that take place remain balanced on all variables, apart from the dummy for whether the owner attended a VTI in the past, which continues to show significant differences in means across Treatment and Control. However, normalized differences between Treatment and Control remain relatively small, and there is no evidence that the observable firm characteristics are jointly significant in predicting treatment assignment, as shown by the *p-value* on the *F-test* at the bottom of column 6 (*p-value* = .207). To further alleviate concerns related to lack of balance along this margin, in our main specifications we control for a dummy for whether the owner attended a VTI in the past.

5 Robustness Checks

This section describes the robustness checks mentioned in the paper and not already addressed in the main Appendix.

5.1 Match-Level Impacts on Manager's Beliefs

5.1.1 Robustness to Using the Full Sample of Scheduled Matches

Columns 1-4 of Table S.IV.1 report the results of regressions like (1) but run on the full sample of 1,230 *scheduled* matches, regardless of whether the job interview took place or not. To do so, we assign a value of zero to the outcome of those job interviews that never took place. The results show that our key finding that high ability owners revise upwards their beliefs on the

skills of matched workers is robust to this alternative specification: the coefficient on Treatment in column 2 is positive and significant at the 5% level. In line with the results in Table IV, we do not find significant treatment effects for managers of lower ability. However, column 4 shows that the difference in the treatment effect between manager types is now less precisely estimated ($p\text{-}value = .144$). This is not surprising, as effectively we are assigning a value of zero to the outcome of more than half of the sample, which reduces the precision of the estimates. Taken together, these results show that our key findings on firm owner's beliefs are robust to running the analysis on the full sample of scheduled matches. This is consistent with our evidence that Treatment does not affect selection into the final sample that meets presented in Table A.VI.

5.1.2 Robustness to Missing Values for High Ability Owner Dummy

In columns 5-8 of Table S.IV.1, missing values for the High Ability Owner dummy are imputed using information from the baseline survey of managers. More specifically, the cognitive test score is imputed for firm owners with a missing value of the score, by running an OLS regression of the cognitive score on all the following firm and owner baseline characteristics: dummies for BRAC branch and sector; dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees; and all the variables used in the balance checks in Table A.II in the main Appendix (see Table A.II for the exact variable definition). Columns 5-8 of Table S.IV.1 show that the results are robust to imputing missing values of the High Ability Owner dummy in this way.

5.1.3 Heterogeneous Impacts by Type of Skills

Table S.IV.2 reports heterogeneous effects on manager's beliefs by whether the worker had a Pass/Fail grade: (i) on the skill reported as most important by managers at baseline (creativity), and (ii) on the skill reported as least important (attendance). The results show very similar patterns to Table IV, regardless of whether we use creativity or attendance as our measure of skill. This evidence is in line with high ability managers reacting similarly to information on the different skills, so that any heterogeneity across skills is limited.

5.2 Match-Level Impacts on Sorting

5.2.1 Robustness to Using the Full Sample of Scheduled Matches

Table S.V.1 reports the results of regressions equivalent to (1) but run on the full sample of 1,230 *scheduled* matches, regardless of whether the job interview took place or not. To do so, we assign a value of zero to the outcome of those job interviews that never took place. The results are broadly consistent with the ones reported in Table VII, and confirm that positive treatment

effects of the certificates on hiring are found only for matches with high ability managers. This is shown by the positive treatment effects in columns 2 (for job offers) and column 6 (for hires), which are both significant at the 5% level. However, the difference in treatment effects across manager types is now less precisely estimated, as shown by the *p-values* in columns 4 and 8. This is not surprising though, as effectively we are assigning a value of zero to the outcome of more than half the sample, which reduces the precision of the estimates. Taken together, these results show that our key findings on employment at the matched firms are robust to using the full sample of matches. This is consistent with our evidence that Treatment does not affect selection into the final sample that meets presented in Table A.VI.

5.2.2 Robustness to Running the Analysis at the Worker Level

Table S.V.2 reports the results of regressions like (1) but at the worker level. To do so, the dependent variable is a dummy equal to one if the worker (i) received at least one job offer (columns 1-4) and (ii) was hired by at least one firm (columns 5-8). The results confirm that we are still able to pick up a positive treatment effect in matches with high ability managers. This effect is significant at the 1% level for offers in column 2, and remains significant at the 10% level when we consider hires in column 6. However, the comparison of treatment effects across manager types becomes less precise, as shown by the *p-values* in columns 4 and 8.

5.3 Worker-Level Impacts on Employment

5.3.1 Heterogeneous Effects by Matches with High and Low Ability Managers

Table S.VI reports the results of worker-level regressions like (2) but where treatment effects are allowed to differ by whether the worker was matched to high or low ability managers.[11] We check this dimension of heterogeneity because Table VII shows that there is a positive treatment effect on hires in matches with high ability owners, and the literature has emphasized that even short term initial employment opportunities can have long-lasting effects on workers by providing a "stepping stone" in the labor market (Pallais, 2014). We consider our main worker-level outcomes from Tables V, VIII and IX. The *p-values* on the test of equality of treatment effects reported at the bottom of the table show that impacts never vary significantly by whether the worker was initially matched to a high or low ability manager.

[11]The heterogeneous analysis is confined to workers who were matched to either only high ability managers or only low ability managers (those 23% of workers matched to both high and low ability managers are excluded).

6 A Two-Sided Sample Selection Model

This section presents the estimates of a two-sided sample selection model to check the robustness of running the analysis on the selected sample of matches that took place. Table A.VI in the main text shows that there is no evidence of Treatment affecting the *selection* of workers and firms in the final sample of job interviews that took place. At the same time, Panel B of Tables S.II and S.III show that the sample of matches that took place remains balanced on observables. These two results justify our focus on the selected sample of matches that took place for our preferred specifications reported in the main text. As discussed in the main text, since conditional on the match taking place there was close to full compliance with Treatment, we interpret our estimates of the Treatment effects on the selected sample as the Average Treatment Effect (ATE) for the population of firms and workers who met. The reason why we also estimate a two-sided model of sample selection is that this allows us to investigate whether the estimated ATE on the selected sample can be considered to be *representative* for the ATE of the population *assigned* to meet, which is our original population of interest.

6.1 Timing of Events and Information Set

For all workers and firms we know whether they were interested in participating in the scheduled job interviews, regardless of whether the job interviews they were assigned to actually took place. This is due to the specific process we followed to implement the matching intervention. Specifically, the matching intervention was implemented in two steps:

1. **Step 1**: A few days before the matching intervention, BRAC enumerators contacted *all* workers and firms initially assigned to meet, to inform them that the intervention was about to start. Workers and firms were asked to confirm their interest in participating in the job interviews, and this was recorded by BRAC enumerators. Importantly, at this stage firms and workers were *not* given any information about whom they had been matched with, nor about the Treatment group they had been assigned to.

2. **Step 2**: Job interviews were carried out between the workers and firms that had confirmed their interest, following the initial matching allocations. The certificates were shown to both the worker and the matched firm only conditional on the job interview taking place.

Conditional on meeting at least one worker in Step 2, only 7% of firms did not meet one of the other matched workers because of loss of interest in conducting the additional interviews. Also, only 3% of the workers that met at least one firm did not meet one of the other matched firms because of lack of interest in doing the additional interviews. This confirms that selection into the final sample of matched workers and firms operates mainly through the *extensive margin*

decision of whether to participate at all, rather than through the intensive margin decision of which interviews to take part in, conditional on participation.

6.2 Selection Equations

The decision of worker i and firm j to participate in the job interviews can be modelled as follows:

$$m_i = \eta \mathbf{Z}_i + \nu_i \tag{1}$$
$$m_j = \theta \mathbf{Z}_j + \nu_j \tag{2}$$

$$D_i = 1 \quad if\ m_i > 0; \quad D_i = 0 \quad if\ m_i \leq 0 \tag{3}$$
$$D_j = 1 \quad if\ m_j > 0; \quad D_j = 0 \quad if\ m_j \leq 0 \tag{4}$$

where m_i is a latent variable representing the propensity of worker i to participate. This is a linear projection of observable worker characteristics $\mathbf{Z}_i$ and unobserved (to the econometrician) worker component ν_i. Note that the participation decision of workers does not depend on (observed or unobserved) firm characteristics. This reflects a feature of the implementation described in Step 1 above. The latent variable representing the participation decision of firms, m_j, is modelled in a similar way, and so does not depend on worker characteristics.

Workers are willing to participate in the job interviews when the value of m_i is greater than 0. In this case, $D_i = 1$. Similarly, for firms $D_j = 1$ only if $m_j > 0$. For the job interview to take place, *both* worker i and firm j must be willing to participate. Therefore, interview outcomes y_{ij} are observed only if *both* $D_i = 1$ and $D_j = 1$. Following the notation in Section 3.3 of the main text, we have:

$$y_{ij} = \beta_0 + \beta_1 Treat_{ij} + \gamma \mathbf{X}_i + \delta \mathbf{X}_j + \theta \mathbf{Strata}_{ij} + \alpha Int_{ij} + \epsilon_i + \epsilon_j + \epsilon_{ij} \quad if\ D_i = 1\ and\ D_j = 1 \tag{5}$$

$$y_{ij} = not\ observed \quad if\ D_i = 0\ or\ D_j = 0 \tag{6}$$

where the error term in (5) is decomposed into a worker-specific component (ϵ_i), a firm-specific component (ϵ_j) and a job interview-specific component (ϵ_{ij}). Under the assumption that the decision of firms (workers) to participate in the intervention does not depend on the characteristics of the workers (firms) they are matched to (which is justified by the discussion around

Step 2 above), we obtain:

$$E[y_{ij}|\mathbf{C}_{ij}, D_i = 1, D_j = 1] = \beta_0 + \beta_1 Treat_{ij} + \gamma \mathbf{X}_i + \delta \mathbf{X}_j + \theta \mathbf{Strata}_{ij} + \alpha Int_{ij} + E[\epsilon_i|\nu_i > -\eta \mathbf{Z}_i] + E[\epsilon_j|\nu_j > -\theta \mathbf{Z}_j] \quad (7)$$

where $\mathbf{C}_{ij}$ defines the set of worker, firm, strata and interview controls. Equation (7) makes clear that the presence of selection bias is driven only by the potential correlation between ϵ_i and ν_i on the worker side, and ϵ_j and ν_j on the firm side.

6.3 Using a Control Function Approach to Control for Selection

We can correct for sample selection in the estimation of (5) by controlling for the selection terms highlighted in (7). If joint normality of the unobservables in the selection and main equation is assumed, then it is easy to show that (7) can be rewritten as:

$$E[y_{ij}|\mathbf{C}_{ij}, D_i = 1, D_j = 1] = \beta_0 + \beta_1 Treat_{ij} + \gamma \mathbf{X}_i + \delta \mathbf{X}_j + \theta \mathbf{Strata}_{ij} + \alpha Int_{ij} + \rho_w \sigma_w \lambda(\eta \mathbf{Z}_i) + \rho_f \sigma_f \lambda(\theta \mathbf{Z}_j) \quad (8)$$

where: ρ_w and ρ_f are the correlation coefficients between the unobservables in the main equation and in the selection equation; σ_w and σ_f are the standard deviations of the unobservables in the main equation, relative to the standard deviations of the unobservables in the selection equation (which are normalized to 1); and $\lambda(\cdot)$ is the inverse Mille ratio: $\lambda(\cdot) = \frac{\phi(\cdot)}{\Phi(\cdot)}$. If one is willing to assume joint normality of the errors, then a two-step procedure similar to the one developed by (Heckman 1974; 1977; 1979) can be used to recover consistent estimates of the parameters for the original population of workers and firms *assigned* to meet.

Equation (8) can be written in more general form, without invoking joint normality:

$$E[y_{ij}|\mathbf{C}_{ij}, D_i = 1, D_j = 1] = \beta_0 + \beta_1 Treat_{ij} + \gamma \mathbf{X}_i + \delta \mathbf{X}_j + \theta \mathbf{Strata}_{ij} + \alpha Int_{ij} + g(\eta \mathbf{Z}_i) + h(\theta \mathbf{Z}_j) \quad (9)$$

where $g(\cdot)$ and $h(\cdot)$ are non-parametric functions. We follow an approach similar to Newey (2009) and replace $g(\cdot)$ and $h(\cdot)$ with series approximations, as follows:

$$E[y_{ij}|\mathbf{C}_{ij}, D_i = 1, D_j = 1] = \beta_0 + \beta_1 Treat_{ij} + \gamma \mathbf{X}_i + \delta \mathbf{X}_j + \theta \mathbf{Strata}_{ij} + \alpha Int_{ij} + + \zeta_1 \lambda(\eta \mathbf{Z}_i) + \zeta_2 \lambda(\eta \mathbf{Z}_i) \cdot (\eta \mathbf{Z}_i) + \zeta_3 \lambda(\theta \mathbf{Z}_j) + \zeta_4 \lambda(\theta \mathbf{Z}_j) \cdot (\theta \mathbf{Z}_j) \quad (10)$$

where $\lambda(\cdot)=\frac{\phi(\cdot)}{\Phi(\cdot)}$. This is more robust than simply assuming joint normality of the unobservables. In the next subsection, we go through the details of estimating equation (10). The estimates from this model recover the ATE for the sample of *scheduled* matches, under the assumption that the model is correctly specified.

6.4 Estimation Using a Two-step Procedure

We have direct information on D_i and D_j from the matching surveys: we define $D_i = 1$ for workers that met at least one firm, or that reported being interested to participate when contacted by BRAC enumerators just before the start of the intervention. The variable D_j is defined in a similar way on the firm side.[12]

To achieve identification of the parameters in equation (10), the observables $\mathbf{Z}_i$ and $\mathbf{Z}_j$ in the selection equations should contain at least one continuous regressor excluded from the main equation (5). For this, we use the characteristics of the BRAC enumerators that were assigned to contact the workers and firms for the job interviews. In particular, for the worker side we use the following instruments: age, gender and number of previous surveys conducted with BRAC by the enumerator assigned to contact the matched worker, as well as dummies for the month when the worker was contacted by the enumerator. The firm side instruments are: age, gender and number of previous surveys conducted with BRAC by the enumerator assigned to contact the matched firm. Workers and firms were assigned randomly to BRAC enumerators, so these variables are exogenous. The two-step estimation procedure is implemented as follows:

1. Regress D_i on $\mathbf{Z}_i$ by Probit using the *full sample* of workers initially assigned to be matched. $\mathbf{Z}_i$ includes the instruments and the same worker controls as in (5). We store the predicted values $\hat{\eta}\mathbf{Z}_i$ and compute an estimate of the inverse Mills ratio for workers $\lambda(\hat{\eta}\mathbf{Z}_i)$.

2. Do the same for firms, by regressing D_j on $\mathbf{Z}_j$ by Probit. $\mathbf{Z}_j$ contains the instruments and the same firm-side control variables included in (5).

3. Estimate equation (10) by OLS, using the estimated inverse Mills ratios and predicted values from the previous step to correct for sample selection bias. Standard errors are calculated using 2000 bootstrap replications, with resampling clustered by firm, to account for the fact that firms were assigned to meet more than one worker.

[12]Note that we cannot simply recover the selection indicators from whether a scheduled meeting took place.

6.5 Estimation Results

We apply this two-step procedure to the analysis of our main match level outcomes: the beliefs of managers on the skills of the matched workers, and whether a job interview turned into a hire. These are observed only if the interview between the worker and the firm took place. Tables S.VII.1 and S.VII.2 report the estimates of the selection equations for workers and firms, respectively. In line with the results in Table A.VI of the main text, we see that Treatment does not predict selection into the sample of workers and firms interested to participate in the job interviews. We notice that our excluded regressors are significant predictors in both the worker and firm selection equations: the joint *F-test* of significance of the instruments in the worker selection equation has *p-value* = .027. On the firm side, the joint *F-test* of the instruments has *p-value* = .039.

Table S.VII.3 reports the results of the two-step model. In columns 1-3 the dependent variable is a dummy equal to one if the matched worker was reported by the firm as being more skilled than the usual applicant (this is the same dependent variable as in Table IV in the main text). In columns 4-6 the dependent variable is a dummy equal to one if the job interview turned into a hire (this is the same dependent variable as in Table VII in the main text). We only focus on the split between matches with high and low ability managers, as this is the margin where we found heterogeneous effects of the intervention in the main text.

There are two points to note about the results from Table S.VII.3: first, as shown in the lower half of the table, the inverse Mills ratios and their interaction with the predicted values from the selection equations are largely not significant. This is an indication that sample selection bias is not substantial in these regressions. We further evaluate the results from the two-step model by comparing them with the estimates of equation (5) on the sample of matches that took place, without controlling for selection. Comparing Table S.VII.3 with the results in Table IV and VII in the main text shows that controlling for selection does not change substantially the point estimates of the treatment effect: for example, focusing on hires, the point estimate in column 4 of Table S.VII.3 is .148, while this is .134 in column 6 of Table VII. This can be appreciated graphically in Figure S.II, where the two point estimates are reported next to each other. The stability of coefficients across the two specifications provides additional evidence that selection into the final sample of matches is not creating substantial bias in these regressions, so that our estimates of the ATE for the selected sample that met are valid estimates for the ATE of the sample that was *scheduled* to meet. Controlling for selection substantially increases the standard errors of the estimates however, as the control function approach is clearly inefficient.

References

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn (2017): "Anonymity or Distance? Job Search and Labor Market Exclusion in a Growing African City," mimeo.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122–142.

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

Heckman, J. J. (1974): "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42, 679–694.

——— (1977): "Sample Selection Bias as a Specification Error (with an Application to the Estimation of Labor Supply Functions)," NBER Working Paper No. 172.

——— (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Newey, W. K. (2009): "Two-Step Series Estimation of Sample Selection Models," *The Econometrics Journal*, 12, 217–229.

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

Table S.I: Basic VTI descriptives

Means, standard deviations in parentheses

P-value from test of equality of means in brackets

	VTI took part in the project	VTI did not take part in the project	P-value
	(1)	(2)	(3)
Number of VTIs	14	7	
A. VTI characteristics			
Years since establishment	35.2	24.7	[.388]
	(30.3)	(9.86)	
Public institution	.214	.143	[1.00]
	(.426)	(.378)	
Number of training courses offered	10.9	9.57	[.561]
	(5.46)	(3.64)	
Number of students currently enrolled	349	604	[.084]*
	(201)	(447)	
B. Region			
Kampala	.500	.571	[1.00]
	(.519)	(.535)	
North	0	0	
	(0)	(0)	
East	.214	.286	[1.00]
	(.426)	(.488)	
West	.286	.143	[.624]
	(.469)	(.378)	

Notes: The table uses data from the 21 VTIs that were contacted about the intervention and for which information is available. There are two additional VTIs that did not take part in the project for which this information is missing, and one additional VTI that took part in the project for which this information is missing. Column 3 reports p-values from two-sided t-tests for all variables apart from the dummy variables (region dummies and private vs public dummy), for which the p-value is from Fisher exact test.

Table S.II: Match-level worker balance at baseline and conditional on meeting

Means, standard deviations in parentheses

P-value on t-test of equality of means with control group in brackets

P-value on F-test that worker characteristics do not jointly predict treatment assignment in braces

	Panel A: Balance at baseline				Panel B: Balance conditional on meeting			
	Control Matches	Treatment Matches	P-value	Normalized Differences	Control Matches	Treatment Matches	P-value	Normalized Differences
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of matches	614	616			257	258		
A. Background characteristics at baseline								
Age [Years]	20.2	20.6	[.045]	-.114	20.3	20.9	[.255]	-.131
	(2.32)	(2.80)			(2.48)	(3.21)		
Female	.515	.477	[.436]	.053	.506	.457	[.553]	.069
	(.500)	(.500)			(.501)	(.499)		
Completed prior education [Years]	10.2	10.3	[.474]	-.044	10.1	10.1	[.782]	.006
	(1.83)	(1.88)			(1.84)	(1.99)		
Course duration [Years]	1.44	1.45	[.941]	-.007	1.44	1.36	[.597]	.066
	(.918)	(.863)			(.930)	(.850)		
Ever employed	.179	.214	[.309]	-.063	.156	.209	[.216]	-.098
	(.384)	(.411)			(.363)	(.408)		
Monthly expected earnings [USD]	119	124	[.502]	-.047	114	122	[.342]	-.088
	(68.8)	(69.4)			(57.6)	(67.5)		
B. Skills at baseline								
Attendance [1-5 scale]	3.45	3.43	[.789]	.014	3.54	3.29	[.432]	.148
	(1.14)	(1.14)			(1.18)	(1.13)		
Communication skills [1-5 scale]	3.30	3.32	[.898]	-.013	3.31	3.20	[.645]	.068
	(1.09)	(1.16)			(1.12)	(1.16)		
Creativity [1-5 scale]	3.39	3.49	[.244]	-.067	3.42	3.38	[.917]	.026
	(1.11)	(1.08)			(1.13)	(1.06)		
Trustworthiness [1-5 scale]	3.51	3.51	[.925]	.002	3.56	3.45	[.483]	.077
	(.984)	(.996)			(.921)	(1.03)		
Willingness to help others [1-5 scale]	3.39	3.38	[.856]	.007	3.49	3.26	[.344]	.148
	(1.14)	(1.08)			(1.17)	(1.11)		
Cognitive test score [0-10 scale]	5.12	5.08	[.717]	.011	5.33	4.87	[.073]	.136
	(2.41)	(2.38)			(2.40)	(2.44)		
F-test of joint significance		{.866}				{.605}		

Notes: Data is from the 787 workers included in the final research sample. The t-stats are from OLS regressions of the variable of interest on a constant, treatment dummy and stratification variables (dummies for BRAC branch and sector). The regressions are at the match level and standard errors are clustered at the level of both the firm and the worker. The F-stats are from OLS regressions where the dependent variable is the treatment dummy, and the independent variables are all the variables considered for the balance checks in the table as well as stratification variables (dummies for region and sector). The regressions are at the match level and standard errors are clustered at the level of both the firm and the worker. Panel A considers balance at baseline, and so the sample includes all the 1230 scheduled matches. Panel B considers balance conditional on the meeting taking place, and so the sample includes those 515 matches that were actually carried out. Expected earnings are constructed as follows: respondents were asked to report: (i) their minimum expected earnings; (ii) their maximum expected earnings; (iii) the probability that they could earn at least the midpoint. We use this information to fit a triangular probability distribution of expected earnings for each respondent. All monetary amounts are deflated and expressed in terms of the price level in January 2015, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted in January 2015 USD. The top 1% values of expected earnings are excluded.

Table S.III: Match-level firm balance at baseline and conditional on meeting

Means, standard deviations in parentheses

P-value on t-test of equality of means with control group in brackets

P-value on F-test that firm characteristics do not jointly predict treatment assignment in braces

	Panel A: Balance at baseline				Panel B: Balance conditional on meeting			
	Control Matches	Treatment Matches	P-value	Normalized Differences	Control Matches	Treatment Matches	P-value	Normalized Differences
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of matches	614	616			257	258		
A. Owner characteristics								
Owner is female	.384	.393	[.428]	-.012	.381	.357	[.868]	.036
	(.487)	(.489)			(.487)	(.480)		
Owner age [Years]	36.0	36.4	[.700]	-.030	36.9	35.5	[.229]	.112
	(8.63)	(8.83)			(8.91)	(8.66)		
Owner completed years of education	10.7	10.1	[.095]	.129	10.3	9.65	[.192]	.131
	(3.18)	(3.33)			(3.51)	(3.15)		
Owner has received training from a VTI	.402	.316	[.049]	.128	.440	.324	[.006]	.170
	(.491)	(.465)			(.497)	(.469)		
Owner scored at median or above on cognitive test	.512	.531	[.922]	-.026	.550	.471	[.137]	.112
	(.500)	(.500)			(.499)	(.500)		
B. Firm characteristics								
Business is registered	.912	.909	[.969]	.007	.907	.930	[.242]	-.061
	(.283)	(.288)			(.292)	(.255)		
Number of employees	3.11	2.89	[.419]	.056	3.03	2.90	[.647]	.036
	(2.95)	(2.66)			(2.40)	(2.73)		
Age of business [Years]	6.80	7.09	[.611]	-.035	7.56	7.64	[.927]	-.009
	(4.99)	(6.49)			(5.04)	(6.42)		
Average monthly revenues [USD]	503	491	[.603]	.015	501	486	[.934]	.019
	(570)	(617)			(561)	(565)		
Average monthly profits [USD]	205	202	[.679]	.011	195	189	[.481]	.016
	(247)	(216)			(237)	(198)		
F-test of joint significance from column regression		{.402}				{.207}		

Notes: Data is from the 422 firms included in the final research sample. The t-stats are from OLS regressions of the variable of interest on a constant, treatment dummy and stratification variables (dummies for BRAC branch and sector). The regressions are at the match level and standard errors are clustered at the level of both the firm and the worker. The F-stats are from OLS regressions where the dependent variable is the treatment dummy, and the independent variables are all the variables considered for the balance checks in the Table as well as stratification variables (dummies for BRAC branch and sector). The regressions are at the match level and standard errors are clustered at the level of both the firm and the worker. Panel A considers balance at baseline, and so the sample includes all the 1230 scheduled matches. Panel B considers balance conditional on the meeting taking place, and so the sample includes those 515 matches that were actually carried out. All monetary amounts are deflated and expressed in terms of the price level in January 2015, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted in January 2015 USD. Firms in the top 1% of the profit and revenues distribution at baseline are excluded from the balance checks on such variables.

Table S.IV.1: Impacts on firm owner beliefs about matched workers - Robustness

OLS regression coefficients, standard errors adjusted for two-way clustering in parentheses

P-values on t-test of equality of coefficients for High and Low Ability Owners in brackets

Dependent variable:	**Matched worker reported as MORE SKILLED than usual applicant [Yes=1]**							
Robustness check:	**Assigning value zero to the outcome of matches that did not take place**				**Imputing missing values of the High Ability dummy**			
Sample of firm owners:	**All**	**High Ability Owners**	**Low Ability Owners**	**P-value (2) = (3)**	**All**	**High Ability Owners (Imputed)**	**Low Ability Owners (Imputed)**	**P-value (6) = (7)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Treatment	.002	.040**	-.006	[.144]	.001	.084**	-.049	[.028]
	(.012)	(.020)	(.024)		(.025)	(.042)	(.037)	
Mean of dep. var. in Control group	.041	.037	.047		.041	.037	.047	
Number of observations (matches)	1,229	532	489		515	270	245	

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Results from the matching surveys are reported. Standard errors are adjusted for two-way clustering (at the level of both the firm and the worker), following the procedure in Cameron et al. [2011]. In columns 1-3 the dependent variable takes value zero if the match was not carried out (so if the firm and the worker never met). All regressions control for stratification variables (dummies for BRAC branch and sector) as well as for dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on all five soft skills measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. All regressions also control for the following firm characteristics measured at baseline: dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. In column 1 and 5 the sample includes all matches. In columns 2 and 6 it is restricted to matches with High Ability firm owners. In columns 3 and 7 it is restricted to matches with Low Ability firm owners. The p-values in columns 4 and 8 are from similar OLS regressions estimated on the full sample of matches and where each independent variable is interacted with the High Ability owner dummy. In columns 5-7, the cognitive test score is imputed for firm owners with a missing value of the score, by running an OLS regression of the cognitive score on all the following firm and owner baseline characteristics: dummies for BRAC branch and sector; dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees; and all the variables used in the balance checks in Table A.II in the main Appendix (see Table A.II for the exact variable definition). All regressions further control for dummies for missing values in each of the independent variables.

Table S.IV.2: Heterogeneous impacts on firm owner beliefs about matched workers

OLS regression coefficients, standard errors adjusted for two-way clustering in parentheses

P-values on t-test of equality of coefficients for High and Low Ability Owners in brackets

Dependent variable:	**Matched worker reported as MORE SKILLED than usual applicant [Yes=1]**							
Sample of firm owners:	**All**	**High Ability Owners**	**Low Ability Owners**	**P-value (2) = (3)**	**All**	**High Ability Owners**	**Low Ability Owners**	**P-value (6) = (7)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Pass grade on Creativity	-.013	-.035	.000	[.551]				
	(.031)	(.045)	(.039)					
Fail grade on Creativity X Treatment	-.055	-.014	-.060	[.687]				
	(.045)	(.077)	(.076)					
Pass grade on Creativity X Treatment	.018	.149***	-.032	[.019]				
	(.029)	(.055)	(.046)					
Pass grade on Attendance					-.044	.027	-.056	[.359]
					(.035)	(.045)	(.073)	
Fail grade on Attendance X Treatment					-.038	.059	-.006	[.592]
					(.046)	(.055)	(.100)	
Pass grade on Attendance X Treatment					.016	.124**	-.044	[.015]
					(.028)	(.055)	(.037)	
Mean of dep. var. in Control group	.041	.037	.047		.041	.037	.047	
P-value Fail = Pass	[.163]	[.096]	[.716]		[.308]	[.352]	[.709]	
Number of observations (matches)	515	232	222		515	232	222	

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Results from the matching surveys are reported. Standard errors are adjusted for two-way clustering (at the level of both the firm and the worker), following the procedure in Cameron et al. [2011]. All regressions control for stratification variables (dummies for BRAC branch and sector) as well as for dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on Creativity (columns 1-4) or Attendance (columns 5-8), as measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. All regressions also control for the following firm characteristics measured at baseline: dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. In columns 1 and 5 the sample includes all scheduled matches. In columns 2 and 6 it is restricted to matches with High Ability firm owners. In columns 3 and 7 it is restricted to matches with Low Ability firm owners. The p-values in columns 4 and 8 are from similar OLS regressions estimated on the full sample of matches and where each independent variable is interacted with the High Ability owner dummy. All regressions further control for dummies for missing values in each of the independent variables.

Table S.V.1: Impacts on sorting - Including zeros

OLS regression coefficients, standard errors adjusted for two-way clustering in parentheses

P-values on t-test of equality of coefficients for High and Low Ability Owners in brackets

Dependent variable:	Worker was made a job offer by the matched firm [Yes=1]				Worker was hired by the matched firm [Yes=1]			
Sample of firm owners:	All	High Ability Owners	Low Ability Owners	P-value (2) = (3)	All	High Ability Owners	Low Ability Owners	P-value (6) = (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	.004	.057**	.023	[.382]	-.001	.041**	.005	[.196]
	(.019)	(.023)	(.033)		(.013)	(.019)	(.021)	
Mean of dep. var. in Control	.079	.053	.093		.041	.031	.038	
Number of observations (matches)	1,126	473	449		1,126	473	449	

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Results from the worker followup surveys are reported. Standard errors are adjusted for two-way clustering (at the level of both the firm and the worker), following the procedure in Cameron et al. [2011]. The dependent variable takes value zero if the match was not carried out (so if the firm and the worker never met). All regressions control for stratification variables (dummies for BRAC branch and sector) as well as for dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on all five soft skills measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. All regressions also control for the following firm characteristics measured at baseline: dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. In columns 1 and 5 the sample includes all scheduled matches. In columns 2 and 6 it is restricted to matches with High Ability firm owners. In columns 3 and 7 it is restricted to matches with Low Ability firm owners. The p-values in columns 4 and 8 are from similar OLS regressions estimated on the full sample of matches and where each independent variable is interacted with the High Ability owner dummy. All regressions further control for dummies for missing values in each of the independent variables.

Table S.V.2: Impacts on sorting - Worker level

OLS regression coefficients, robust standard errors in parentheses
P-values on t-test of equality of coefficients for High and Low Ability Owners in brackets

Dependent variable:	Worker received at least one job offer by the matched firms [Yes=1]				Worker was hired by at least one matched firm [Yes=1]			
Sample of workers match to:	All	High Ability Owners	Low Ability Owners	P-value (2) = (3)	All	High Ability Owners	Low Ability Owners	P-value (6) = (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	.006	.112***	.030	[.172]	.007	.067*	-.004	[.113]
	(.023)	(.042)	(.043)		(.017)	(.037)	(.025)	
Mean of dep. var. in Control	.104	.065	.096		.053	.037	.044	
Number of observations (workers)	709	210	248		709	210	248	

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Results from the first worker follow-up survey are reported. Standard errors are adjusted for heteroskedasticity in these regressions. All regressions control for stratification variables (dummies for BRAC branch and sector) as well as for dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on all five soft skills measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. In columns 1 and 5 the sample includes all workers. In columns 2 and 6 it is restricted to workers matched only with High Ability firm owners. In columns 3 and 7 it is restricted to workers matched only with Low Ability firm owners. The p-values in columns 4 and 8 are from similar OLS regressions estimated on the full sample of workers and where each independent variable is interacted with a dummy for whether the worker was matched with High Ability owners only. All regressions further control for dummies for missing values in each of the independent variables.

Table S.VI: Heterogeneous impacts on post-intervention worker outcomes - Matched to High or Low Ability Owner

OLS regression coefficients, standard errors clustered at the worker level in parentheses

Dependent variable:	Monthly expected earnings [USD]	Expected probability of employment in the next six months (0 to 10 scale)	Any work as employee in the last week	Weekly hours worked in last job as employee	Any paid work in the last month	Total earnings in the last month [USD]	
Sample of workers:	**All**	**All**	**All**	**All**	**All**	**All**	**Positive earnings**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Matched to High Ability Owners	3.71	.180	.010	-.191	-.010	-5.91	-7.07
	(5.69)	(.237)	(.051)	(3.76)	(.046)	(5.93)	(6.51)
Matched to Low Ability Owners X Treatment	8.91	.520**	.039	-1.12	-.040	2.28	6.71
	(5.83)	(.218)	(.051)	(3.65)	(.046)	(5.97)	(6.49)
Matched to High Ability Owner X Treatment	3.23	.109	.026	1.57	.006	-1.48	-.884
	(5.93)	(.234)	(.054)	(3.65)	(.047)	(5.67)	(6.83)
Mean of dep. var. in Control group	114.5	5.53	.428	37.0	.750	47.2	63.1
Pvalue Matched High = Matched Low	[.501]	[.221]	[.870]	[.612]	[.508]	[.661]	[.428]
Controls for baseline value of outcome	Yes	Yes	Yes	No	Yes	Yes	Yes
Uses data from first and second followup	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	868	879	879	877	870	864	631

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Results from the worker follow-up surveys are reported. Standard errors are clustered at the worker level. All regressions control for stratification variables (dummies for region and sector), a dummy for second follow-up and dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on all five soft skills measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. In column 1 the dependent variable is constructed as follows: respondents were asked to report: (i) their minimum expected earnings; (ii) their maximum expected earnings; (iii) the probability that they could earn at least the midpoint. We use this information to fit a triangular probability distribution of expected earnings for each respondent. The top 1% values of expected earnings are excluded. Since at baseline all workers were enrolled at vocational training institutes and only 1% of them were currently doing in any paid work, for the employment outcomes we consider as baseline value of the outcome the expected probability of employment in the six months after graduation, as reported in the baseline survey. So we control for such baseline expected probability of employment in columns 3 and 5. The dependent variable in columns 6 and 7 is total labor earnings from all activities in the month prior to the survey. Since at baseline all workers were enrolled at vocational training institutes and only 1% of them were currently doing in any paid work, we consider as baseline value of the outcome expected earnings at baseline. The top 1% values of labor market earnings are excluded. All monetary amounts are deflated and expressed in terms of the price level in January 2015, using the monthly consumer price index published by the Uganda Bureau of Statistics. Deflated monetary amounts are then converted in January 2015 USD. The variable Matched to high Ability Owners takes value one if the worker was matched only to High Ability owners, and zero otherwise. The variable Matched to Low Ability owners takes value one if the worker was matched only to Low Ability owners, and zero otherwise. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. All regressions further control for dummies for missing values in each of the independent variables.

Table S.VII.1: Two-step selection model - Worker selection equation

Probit marginal effects, robust standard errors in parentheses

Dependent variable:	Worker was interested in being matched during implementation [Yes=1]
	(1)
Treatment	-.013
	(.104)
Pass grade on all skills	-.124
	(.114)
Age	.155
	(.139)
Age squared	-.003
	(.003)
Female	-.124
	(.228)
Completed prior education	-.039
	(.034)
Course duration	-.237***
	(.078)
Ever employed	.029
	(.139)
Mean of dependent variable	.740
P-value on F-test of joint significance of excluded variables	[.027]
P-value on F-test of joint significance of sector dummies	[.004]
P-value on F-test of joint significance of BRAC branch dummies	[.000]
R-squared	.139
Number of observations (workers)	784

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. The table uses data from the baseline and matching surveys. Standard errors are adjusted for heteroskedasticity. The worker-side instruments in the first step are: age, gender and number or previous surveys conducted with BRAC by the enumerator assigned to contact the matched worker, as well as dummies for the month when the worker was contacted by the enumerator. All regressions further control for dummies for missing values in each of the independent variables.

Table S.VII.2: Two-step selection model - Firm selection equation

Probit marginal effects, robust standard errors in parentheses

Dependent variable:	Firm was interested in being matched during implementation [Yes=1]
	(1)
Treatment	-.023
	(.149)
Number of employees	.054
	(.036)
Owner is female	-.010
	(.243)
Owner attended a VTI	.119
	(.170)
Age of owner	-.013
	(.055)
Age of owner squared	.000
	(.001)
High ability owner	-.064
	(.172)
Mean of dependent variable	.739
P-value on F-test of joint significance of excluded variables	[.039]
P-value on F-test of joint significance of sector dummies	[.024]
P-value on F-test of joint significance of BRAC branch dummies	[.000]
R-squared	.026
Number of observations (firms)	422

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. The table uses data from the baseline and matching surveys. Standard errors are adjusted for heteroskedasticity. The firm-side instruments in the first step are: age, gender and number or previous surveys conducted with BRAC by the enumerator assigned to contact the matched firm. All regressions further control for dummies for missing values in each of the independent variables.

Table S.VII.3: Estimates from two-step selection model

OLS regression coefficients, bootstrap standard errors in parentheses

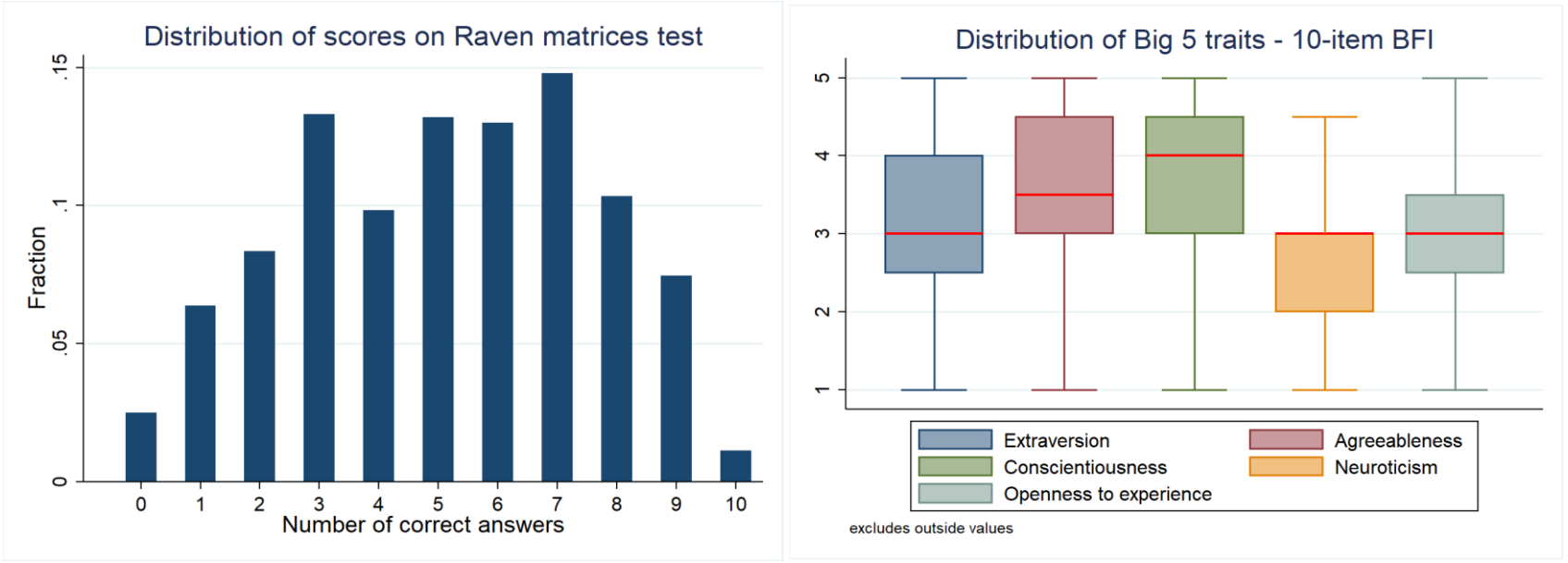
P-values on t-test of equality of coefficients for High and Low ability owners in brackets

Dependent variable:	Worker reported as more skilled than usual applicant [Yes=1]			Worker was hired by the matched firm [Yes=1]		
Sample of firm owners:	High Ability Owner	Low Ability Owner	P-value (1) = (2)	High Ability Owner	Low Ability Owner	P-value (4) = (5)
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	.100	-.042	[.058]	.148	-.013	[.137]
	(.082)	(.090)		(.119)	(.086)	
Mean of Dep. Var. in Control	.079	.115		.080	.105	
P-value from joint F-test on worker Inv. Mills Ratio and interactions with predicted values	[.814]	[.315]		[.782]	[.055]	
P-value from joint F-test on firm Inv. Mills Ratio and interactions with predicted values	[.679]	[.817]		[.695]	[.690]	
P-value from joint F-test on worker and firm Inv. Mills Ratio and interactions with predicted values	[.898]	[.625]		[.888]	[.168]	
Number of observations (matches)	232	222		173	182	
First Stage						
P-value from F-test on excluded variables in Worker selection equation			[.027]			
P-value from F-test on excluded variables in Firm selection equation			[.039]			

Notes: *** (**) (*) denotes significance at the 1% (5%) (10%) level. Columns 1-3 use data from the matching surveys; columns 4-6 use data from the worker follow-up survey. The table reports estimates from a two-sided two-step selection model. Standard errors are bootstrapped, where the re-sampling is clustered by firm, with 2000 replications. All regressions control for stratification variables (dummies for region and sector) as well as for dummies for month of interview. In addition, all regressions control for the following worker characteristics measured at baseline: a dummy for whether the worker had a pass grade (C or above) on all five soft skills measured in the baseline assessments; age and age squared; dummy for female; years of formal education; duration (in years) of the vocational training program the worker was attending at baseline; dummy for any past work experience. All regressions also control for the following firm characteristics measured at baseline: dummy for female owner; age and age squared of the owner; dummy for whether owner attended a VTI in the past; number of employees. Firm owners who scored on or above the median on a cognitive test administered at baseline are assigned to the High Ability group; owners who scored below the median are assigned to the Low Ability group. In columns 1 and 4 the sample is restricted to workers matched only with High Ability firm owners. In columns 2 and 5 it is restricted to workers matched only with Low Ability firm owners. The p-values in columns 3 and 6 are from similar regressions estimated on the full sample of workers and where each independent variable is interacted with a dummy for whether the worker was matched with High Ability owners only. The worker-side instruments in the first step are: age, gender and number or previous surveys conducted with BRAC by the enumerator assigned to contact the matched worker, as well as dummies for the month when the worker was contacted by the enumerator. The firm-side instruments in the first step are: age, gender and number or previous surveys conducted with BRAC by the enumerator assigned to contact the matched firm. All regressions further control for dummies for missing values in each of the independent variables.

Figure S.I: Skills measurement on workers

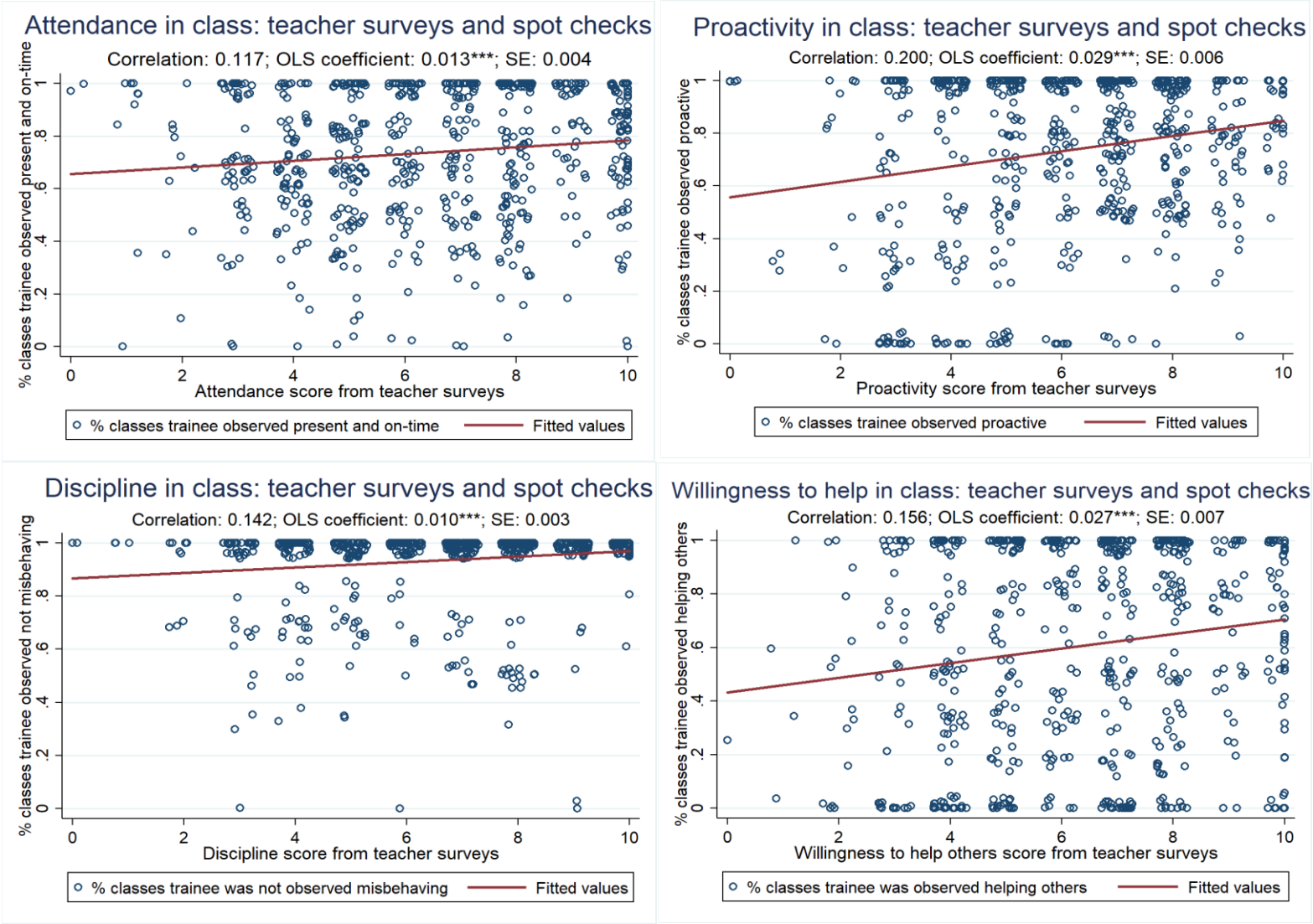
Panel A: Distribution of cognitive skills and of Big-5 traits for trainees in initial census



Notes: The sample for both panels includes all trainees surveyed in the initial census. The left panel reports the frequency histogram of the scores on a 10-item Raven matrices test. The right panel reports Box and Whisker plots of the distribution of the Big-5 traits, measured using a standard 10-item questionnaire, with explanations and examples adapted to the Ugandan context. The red line is the median.

Figure S.I: Skills measurement on workers (continued)

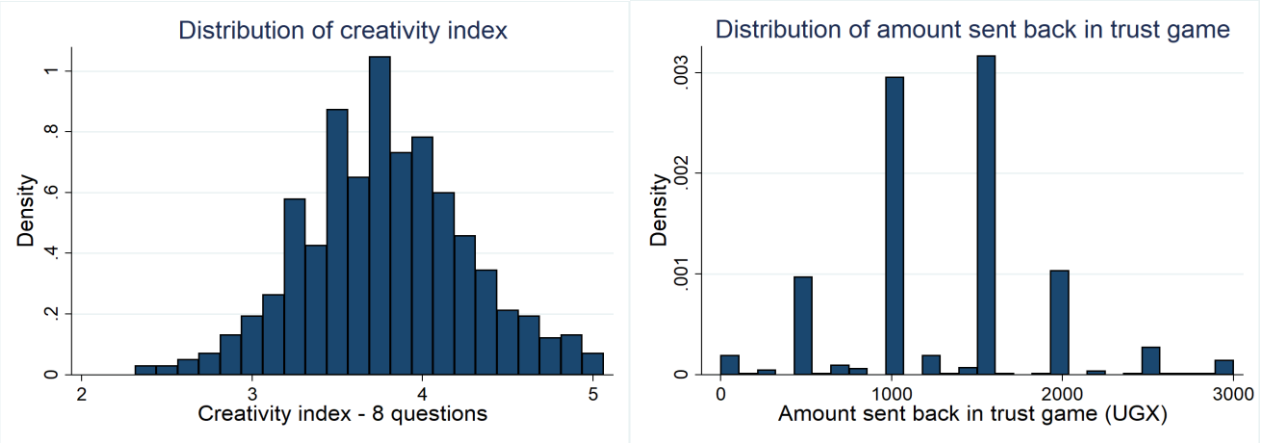
Panel B: Correlation between teacher surveys and spot-checks



Notes: The sample in all panels includes all workers interested in the matching intervention and so included in the baseline survey. The graphs are two-way scatter plots with jittering. The red lines are fitted values from a univariate OLS regression of the variable on the y-axis on the variable on the x-axis. Spot-checks were conducted weekly and the median trainee is observed in six spot-checks.

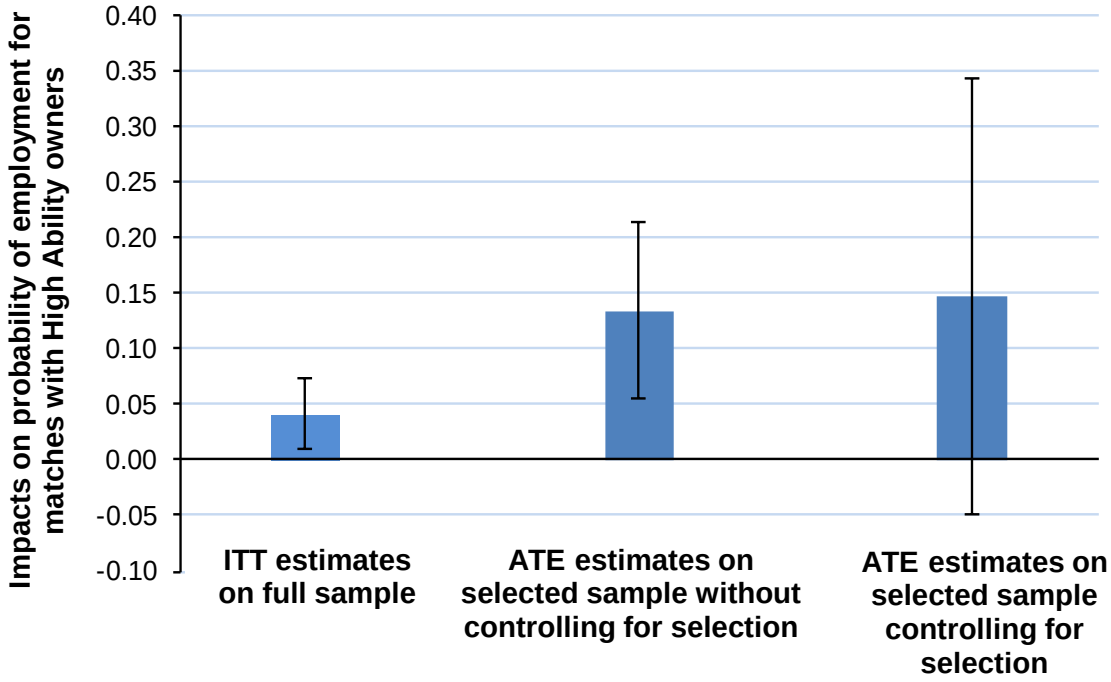
Figure S.I: Skills measurement on workers (continued)

Panel C: Creativity and Trustworthiness



Notes: The sample in all panels includes all 787 workers interested in the matching intervention and so included in the baseline survey. The left panel reports the distribution of the creativity index. This is constructed as the unweighted average of answers to an 8-item battery of questions designed to measure creativity. Creativity is increasing in going from 1 to 5. The right panel reports the distribution of the amount (in Ugandan shillings) that trainees sent back in the trust game they played. A higher amount is interpreted as an indication of higher trustworthiness.

Figure S.II: Controlling for sample selection



Notes: The figure reports the estimated treatment effects on the probability of employment for the sample of matches with High Ability owners. The ITT estimates are taken from column 6 of Table S.V.1, and use the full sample of matches, by assigning a value of zero to those matches that were not carried out (so that all matches are included in the estimation). The ATE estimates on the selected sample without controlling for selection are taken from column 6 of Table VII in the main text, and use only the sample of matches that took place, without any further adjustment for sample selection. The ATE estimates on the selected sample controlling for selection are taken from column 4 of Table S.VII.3 and use only the sample of matches that took place, but additionally controlling for selection using a control function approach. The black bars correspond to 90% CIs around the point estimates.