Can Financial Incentives to Firms Improve Apprentice Training? Experimental Evidence from Ghana

Gabriel Brown
University of British Columbia

Morgan Hardy
New York University-Abu Dhabi

Isaac Mbiti
University of Virginia

Jamie McCasland
University of British Columbia

Isabelle Salcher
New York University

December 2022

Abstract

We use a field experiment to test whether financial incentives can improve the quality of apprenticeship training. Trainers (firm owners) in the treatment group participated in a tournament incentive scheme where they received a bonus payment based on their apprentices' rank-order performance on a skills assessment. Trainers in the control group received a fixed payment based on their apprentices' participation in the assessment. Performance on the assessment was higher in the treatment group. Two years later, apprentices scored 0.15σ higher on a low-stakes oral skills test and earned 24% more in total earnings, driven by higher self-employment profits.

JEL Classifications: D22, M53, O12
Word Count: ∼ 6011

1 Introduction

Apprenticeships are viewed as a promising pathway to augment the skills, productivity, and labor market outcomes of youth (Wolter and Ryan, 2011). By combining on-the-job training with practical work experience, apprenticeships can deliver market relevant skills to apprentices by harnessing the knowledge and experience of firms (or firm-owners) in a scalable and potentially cost-effective manner (ILO, 2019). While apprenticeships in developed nations generally occur in the formal sector, in developing countries they are typically conducted within informal sector small-scale enterprises with the firm owner serving as the trainer (Teal, 2016).[1] This type of training is responsible for the majority of skill development in developing countries (Crepon and Premand, 2019; Filmer and Fox, 2014), especially in West African countries such as Ghana, Benin, and Cameroon (Filmer and Fox, 2014; Darvas and Palmer, 2014; ILO, 2011).[2] In addition to providing training, informal apprenticeships are an important and growing form of employment, especially in more urbanized settings in sub-Saharan Africa (Teal, 2016).

Despite the ubiquity of informal apprenticeships, there are concerns that they may not be able to deliver high quality training due to the limited nature of monitoring, formal contracting, standardization (e.g. a syllabus), and skills certification (Filmer and Fox, 2014; Darvas and Palmer, 2014; Teal, 2016; ILO, 2019). Trainers (firm-owners) may prolong the apprenticeship duration to extract additional rents from apprentices who are typically paid much less than regular workers. Further, because apprentices tend to start their own business after completing their training by replicating their trainer's business practices (Frazer, 2006b), trainers may be reluctant to exert effort to train a future potential competitor (Ackah et al., 2015).

In principle, well-designed incentive contracts could potentially encourage trainers to exert higher effort levels toward training their apprentices, potentially increasing the skills, productivity, and livelihood of many youth in developing countries. Experimental evidence from primary school settings in developing countries shows that financial incentives for teachers can be a cost-effective policy option to improve the quality of education (Breeding et al., 2021; Pham et al., 2021; Ganimian and Murnane, 2016; Muralidharan and Sundararaman, 2011). However, it is less

---

[1]See Frazer (2006a) for a detailed description of informal apprenticeships in Ghana.

[2]Informal apprenticeships in Ghana train over four times as many individuals as all formal technical and vocational education and training (TVET) alternatives combined (Darvas and Palmer, 2014).

clear if (modest) incentives for trainers can improve apprentices' outcomes given the potential (opportunity) costs of training and the challenges of recovering training (effort) costs.[3]

In this paper, we report the results of a field experiment in Ghana that provided (tournament-style) financial incentives to firms training apprentices in the context of a government sponsored program. In the absence of the program, aspiring apprentices would be required to find a training firm and pay an upfront training fee. The government program waived the training fee and matched apprentices with trainers (or firm owners). Trainers participating in the program were then randomly assigned to either a treatment group or a control group. Control trainers received a fixed bonus payment for every apprentice that participated in a trade-specific skills assessment, while treatment trainers received bonus payments that varied based on their apprentices' rank-order performance on the same test. This design encouraged all trainers to avail their apprentices for testing. The skills assessment was based on the national vocational skills standards (or curriculum) and consisted of a theoretical and a practical component. Apprentices earned an official skills certification if they passed the test. Because the average bonus payment in both groups was equalized, any observed differences between the treatment and control groups would not simply be driven by income effects.

Approximately 8 months after the start of the incentives intervention, apprentices in the treatment group attained more skills: their practical test scores were 0.13σ (p-value < 0.05) higher and they were more likely to pass the test and earn a certificate. These learning gains persisted: two years later, apprentices from treated firms performed 0.15σ better on a low-stakes craft (or trade) specific skills test and 0.19σ better on an index of sales skills (which are related to soft-skills). The results for the sub-sample of female apprentices (which make of 94% of the overall sample) were of similar magnitude.[4]

We also measure apprentices' labor market outcomes two years after the intervention and find that apprentices in the treatment group were 20% more likely to leave the firm, often to start their own businesses. Apprentices who trained with treatment firms earn 24% more each month on average from all sources, driven by higher self-employment profits. There is suggestive evidence that treatment shifts workers from unskilled to skilled self-employment, but we find no effect of

---

[3]See Acemoglu and Pischke (1999) for a discussion on the importance of recouping training costs in determining a firm's training decisions.
[4]We do not have sufficient sample size to report the results for the sub-sample of males.

treatment on the probability of any self-employment or on overall labor supply.

We make three main contributions. First, we use an experiment to show that trainer incentives can improve the quality of training in informal sector apprenticeships. While there is an extensive literature on the potential for incentives to improve education and health services in developing countries, to the best of our knowledge, we are not aware of any other paper that experimentally tests the impact of financial incentives on firm training behavior.[5] Second, our study contributes to the literature on training within firms (Becker, 1964; Acemoglu and Pischke, 1998a, 1999; Caicedo et al., 2022), where many of the papers highlight the proclivity of firms to under invest in (general) training.[6] Our results show that relatively small incentives, averaging 3.3% of average annual firm profits in the control group, improve training quality offered by firms, which in turn increases the probability that the apprentice exits the firm (a proxy for "graduation"). The increase in apprentice separation suggests that absent intervention, firms strategically slow-walk training to retain apprentices and capture additional rents from their productivity (relative to their low pay).[7] Although trainers improve training quality, we do not find any negative profit effect on firms during the intervention, suggesting that the opportunity costs of training were sufficiently low (relative to the incentive). This could reflect the significant idle (or nonproductive) time among firm owners and employees in these settings—our data suggest that trainers and their employees are idle over 20% of the time.[8] Finally, our results contribute to the broader job (and vocational) training literature in developing countries. While some high-quality programs have large medium-term impacts on earnings (e.g. Alfonsi et al. (2020)), most rigorous evaluations of job (or vocational) training programs find limited impacts on employment and earnings, coupled with high program costs (Mckenzie, 2017; Blattman and Ralston, 2015). The small but growing literature on apprenticeships in developing countries is also mixed, but the evidence suggests that better quality programs can improve trainee labor market outcomes (Alfonsi et al., 2020; Crepon and Premand, 2019; Hardy et al., 2019; Cho et al., 2013; Monk et al., 2008).

Our results highlight that trainer incentives can be a cost-effective approach to improve the

---

[5]Filmer and Fox (2014) and Adoho et al. (2014) argue that there is a need for rigorous evidence examining the potential for trainer incentives to improve the quality of skills training programs.

[6]See Leuven (2005) for a discussion on how factors such as credit constraints among workers and the probability of worker turnover (or poaching by other firms) can lead to inefficient investment in general training.

[7]See Fudenberg and Rayo (2019) and Fudenberg et al. (2021) for theoretical insights on the length of apprenticeships.

[8]Bassi et al. (2022b) and Bassi et al. (2022a) also find that firm owners and employees in informal sector firms in Uganda are idle roughly 25% of the time.

quality of training programs in developing countries—our incentive program pays for itself in returns to apprentices in less than two years. The results suggest that there are significant returns to the technical and sales (or soft) skills gained through apprenticeships. However, consistent with Frazer (2006a), these returns accrue in self-employment rather than in wage employment.

2 Study Design

2.1 Study Background

This study was conducted in collaboration with the Council for Technical and Vocational Education and Training (COTVET), a Ghanaian government agency that develops technical and vocational training policy and programs, and works to improve the informal apprenticeship sector. COTVET recognized that informal apprenticeships suffer from misaligned incentives between apprentices and trainers. They were concerned that trainers slow-walk skill development and retain workers at low wages for excessive periods of time, with most apprenticeships lasting about three years. COTVET was keen on implementing reforms to accelerate the pace of training.

Our partnership with COTVET yielded a nested set of three experiments, embedded in the implementation of the National Apprenticeship Program (NAP)–a nationwide program that aimed to place unemployed youths into apprenticeships with firms (trainers) in the informal sector. From August 2012 to January 2013, a pool of NAP applicants were randomly selected to participate in the program, on a district by district basis. These selected apprentices were (conditionally) randomly assigned to firms and commenced training between October 2013 and January 2014.[9] The sample for the present study is composed of apprentice-firm pairs that were participating in the NAP at time of the intervention in January and February of 2015, about a year after training began.

Hardy and McCasland (forthcoming) and Hardy et al. (2019) provide a detailed discussion of

---

[9]In the absence of the NAP program, apprentices would be required to pay a training fee and also find a trainer. Through the NAP program, the government facilitated search and matching by directly recruiting training firms. In addition, the government program initially promised a fee payment to trainers of 150 Ghana Cedis (GhC), which was never delivered due to a fiscal crisis, meaning the program essentially eliminated the traditional entrance fee. In lieu of a payment from the government and in an effort to foster ongoing goodwill and trust, the research team provided all participating firms with 100 (GhC) per apprentice just before the roll-out of this study in December 2014. The NAP program also provided no government stipend to support apprentices, who were paid (low) wages from their training firms equivalent to those normally received by apprentices in this context. See Hardy and McCasland (forthcoming) and Hardy et al. (2019) for more details on the sister experiments.

the external validity of the sample. Hardy et al. (2019) show the apprentices recruited under the NAP were observably similar to traditional (non-NAP) apprentices along all dimensions (including cognitive ability and schooling) except that NAP apprentices were poorer. Hardy and McCasland (forthcoming) show that firms (trainers) in the NAP program were broadly representative of firms that typically train traditional apprentices in our sample districts in Ghana. Finally, Hardy et al. (2019) find there were no differences in the self-reported training experiences of NAP apprentices, as compared with traditional apprentices.

2.2 Sample Summary Statistics

The NAP program included five trades, three in skilled construction, cosmetology, and garment-making. Due to the unavailability of updated training curricula for the construction trades at the time of the intervention, the incentive program implementation was restricted to garment-making and cosmetology. These were also the two most popular trades, accounting for almost 80% of NAP applicants at baseline (Hardy et al., 2019). This restriction yields a sample of 797 apprentices across 467 training firms in the two focal trades at the start of the intervention. Due to the gender patterns of occupational segregation, 94% of the apprentice sample is female. Consequently, we report results for both the full sample and also for the sample of female apprentices consisting of 751 apprentices training and working at 446 different firms. Due to the small sample size, we cannot report results for the sub-sample of males.

The sample comes from 32 districts around Ghana. Appendix Table A1 reports baseline (in 2012) characteristics of apprentices. Apprentices were 23 years old, with 7.5 years of schooling. Roughly one-third were married and 48% had at least one child. 22% of apprentices in the sample were self-employed and just 4% were wage employed at baseline.

2.3 Experimental Design

Trainers (firms) in the experiment were randomized into treatment and control, stratified by district, trade, and a measure of the number of NAP apprentices at the firm.[10] These three

---

[10]Randomizations were conducted separately for the South of Ghana (24 sample districts across 7 Southern regions of the country as of 2015) and the North of Ghana (8 sample districts across 3 Northern regions of the country as of 2015). In the South, where the average apprentice in our sample trained with 1.3 other NAP apprentices, firms were split into those training a single NAP apprentice and those training more than one NAP apprentice. In the North, where the average apprentice in our sample trained with 1.6 other NAP apprentices, firms were split into those training one or

dimensions generate a set of 107 strata, given the structure of our sample. In January and February of 2015, when apprentices in our sample had been at their firms for about one year, the intervention roll-out was conducted one-on-one and in person. A member of the research team explained the program (either the treatment or control incentive structure), provided the firm owner with a copy of COTVET's occupational standards handbook (from which the tests were designed), and officially enrolled apprentices in the program.

Control firm owners would receive a flat payment of 100 GhC for each apprentice who participated in the skills test. Treatment firm owners would receive payments based on the decile ranking of their performance within their trade and region.[11] Apprentices in the lowest decile earned their firm owners 25 GhC, which increased to 200 GhC in the top decile. The payment schedule is presented in Appendix Table A2. In addition, across every region and trade, the highest performing treatment apprentice in the region by trade group earned their firm owner an additional 250 GhC, in an effort to encourage effort even among the highest performers. Both in expectation and ex-post, treatment firm payouts averaged 100 GhC.[12] Because the average payouts are equalized across groups, we mitigate concerns that outcomes could be driven by income effects. The skills test was implemented a little more than eight months after the intervention roll-out, in August and September of 2015. Incentive payments were made in October of 2015.

2.4 Data

Baseline data for apprentices and firms was collected in 2012, and 2013, respectively, prior to randomization. Instruction hours, wages, and profits during the intervention period were collected via a phone survey with firm owners in June 2015, before the assessment. All monetary outcomes for both apprentices and firms are inflated or deflated to October 2015, when firm owners received their bonus payments from the intervention.

The assessment consisted of a theoretical component, delivered verbally by an assessor, and a practical component, which required apprentices to physically demonstrate a series of skills, including a role play with a model client. Both sections were conducted at the firm and scored by

---

two NAP apprentices and those training more than two NAP apprentices. This stratification ensured that we equalized the numbers of apprentices in each experimental group.

[11]Our 32 districts, by design, have representation across all 10 regions as of 2015, generating 20 performance scheme groups. Group size ranges from 10 to 100; mean and median group size is 40 apprentices.

[12]Median payouts were also close to 100 GhC (averaging the fifth and sixth deciles).

the assessor, an expert in the craft. The assesment was developed in collaboration with a team of experts, and conformed to the national standards developed by COTVET. The assessment and scoring protocols were tested and refined in a pilot district. We therefore drop the pilot district from the analysis of the assessment below. Because COTVET mandated that the all the exams had to be anchored to the national standards, the theoretical component of the exam contained a large share of questions that are unlikely to measure remunerable craft skills.[13]

The practical component of the assessment, on the other hand, tested core competencies in each craft. Garment-makers were required to demonstrate several different types of stitches and seams, and asked to measure, cut, and sew a garment for a human model. Cosmetologists could choose to either shampoo and braid a human model or relax, roller-set, and blowout the model's hair. Assessors scored each step, taking the quality of work into account.

We use the scoring schemes that determined whether or not an apprentice passed the exam as our measure of performance on the assesment. From the theoretical section, these scoring schemes dropped questions for which more than 25% of the variation was explained by the assessor. The theoretical score is then the sum of the remaining correct answers. For the practical section, the certificate scoring scheme up-weighted the most technical components of the exam like the quality of the final garment and down-weighted less technical components like a score for greeting the client in a professional manner.[14]

The endline survey, conducted from July 2017 to May 2018, provides data on long-term skills outcomes, earnings, labour supply, apprenticeship completion, and firm exit. The craft skills quiz was developed by the research team, in consultation with local experts, and focused on core competencies in each craft. It was an oral quiz and did not require practical demonstration of skills, but we believe it to be a high-quality measure of craft competencies as it significantly predicts higher earnings cross-sectionally in the endline survey.

We also present results on a measure of sales skills, composed of two indices. First, we asked the respondent to discuss skills training in Ghana, and the surveyor rated the respondent on the audibility, animation, engagement, and coherency of the response. Second, we asked respondents

---

[13]For example, (1) *State three reasons why it is important to create an organized and tidy workspace.*, (2) *State three duties involved in tidying the workplace. For each duty, explain how regularly you should complete it and why?*, (3) *What are the broader societal benefits of the hairdressing and beauty industry in Ghana?*, and (4) *Name three items that should be found in a first aid box.*

[14]Results using IRT models are available upon request.

to pretend to sell a pen. Respondents were scored on standard sales strategies such as describing the attributes of the pen, describing how others like the pen, and using humor to engage the enumerator. Here we take the first principal component of all ratings from both sections. We believe this index can be interpreted as a measure of soft-skills.[15] We also report results on an overall skills index that combines both craft and sales skills.

The key labor market outcomes come from a retrospective panel that asked respondents to relay their wage-employment, self-employment, farming, and apprenticeship earnings (and labor supply) 1 month, 4 months, 7 months, and 10 months prior to the endline interview. Our measure of total earnings is the sum across all four earned income sources.[16] Profits in self-employment follow self-reported question structures from De Mel et al. (2009). We stack the four retrospective observations in all earnings and labor supply specifications. Apprenticeship completion and firm exit data are self-reported.

2.5   Attrition and Balance Tests

Appendix Tables A3 and A4 present apprentice-level balance tests for the two samples that generate variation in our outcome analysis: those apprentices with a follow-up survey and those who participated in the assessment. In both tables, 19 of 20 covariates suggest a balanced sample and F-tests of the joint significance of all 20 covariates fail to reject joint orthogonality.[17] Only our measure of household assets, the first principal component of a set of housing quality and household asset measures, is imbalanced.[18] Because the baseline household asset index predicts some follow-up outcomes of interest (namely earnings and apprentice turnover), we control for it (and a dummy for cases in which it is missing) in all apprentice-level specifications and report results that exclude these controls in the appendix. We test for firm-level balance along baseline firm characteristics in Tables A8 and A9. We find that treatment and control groups are balanced along (baseline) firm-level characteristics such as profits, firm size, firm owner cognitive ability, firm owner experience training apprentices, and several self-reported measures of firm owner instruction.

---

[15]See Adhvaryu et al. (2018) for a recent study on the importance of soft-skills.

[16]Wage- and self-employment earnings are about 87% of total control group earnings.

[17]Appendix Tables A5- A7 present analogous balance tests for the female apprentice sample, with similar conclusions.

[18]Note that an index of food consumption/security and both wage-employment and self-employment earnings are balanced.

Appendix Table A10 presents information on attrition for each of the five data sources. We find that attrition rates are balanced across treatment and control groups in each piece of data we exploit in the paper.

3 Results

3.1 Skill Development

We present results on skills outcomes in Table 1. Intent-To-Treat (ITT) incentivized test score outcomes are estimated using the following least squares regression:

$$Y_i = \beta_0 + \beta_1 * T_i + \gamma_s + \delta_a + \alpha_b * X_i + u_i \tag{1}$$

where γ_s are strata dummies, δ_a are assessor dummies (to control for more and less generous grading), X_i are controls for the baseline household asset index and a dummy for cases in which that variable is missing, T_i is a treatment indicator, and u_i is an error term. Skills outcomes measured in the endline survey are estimated using the same specification, with the exception of the assessor fixed effects (since those are only relevant for the incentivized exam). We focus on ITT estimates because LATE (2SLS) estimates would require very strong assumptions about the exclusion restriction.[19] Errors are clustered at the firm level. We report results for the full sample (Panel A) and also the sample of females (Panel B). The sample size of men is too small to examine a male sub-sample.

Columns (1) and (2) show that about 65% of apprentices in both the treatment and the control group participated in the skills assessment, facilitating a clearer interpretation of our results.[20] The estimated treatment effects on the practical component of the assessment are 0.13σ in the whole sample and 0.18σ for females (Column 3). We find no statistically significant effect of the performance-based incentives on test performance in the theoretical component of the assessment (Column 4), likely due to its content. COTVET requested relatively lenient pass thresholds, which

---

[19]We can estimate 2SLS models using test-taking or skills test-scores as the endogenous variable and use the treatment assignment as the IV. We believe such 2SLS models would be biased because trainers could respond to the incentive and increase effort but their apprentice could miss the assessment for idiosyncratic reasons. IV estimates for our endline skills outcomes are also likely biased because of unmeasured skills that could improve apprentice productivity and labor market outcomes.

[20]Appendix Table A11 shows that there are no statistically significant correlates of participation in the skills assessment.

were collaboratively set at 50% for the garments practical, 45% for the cosmetology practical, and 50% for the theoretical exam, yielding pass rates of 80% in the control group.[21] The observed improvement in performance on the practical exam is associated with a noisy 5 percentage point increase in the probability of passing that exam (Panel A, Column (5)) and in earning a certificate (Panel A, Column (6)). The probability of passing the practical exam and earning a certificate is more precisely estimated among the female sample (Panel B, Column (5) and (6)).[22]

Columns (7) and (9) present results on our measures of craft skills and sales skills collected at endline, and a summary (PCA) index of these skills. The treatment increased craft skills by 0.15σ (Column 7). This suggests that the learning gains persisted more than two years after the intervention. These results validate the results from the assessment, allowing us to rule out potential gaming and concerns about attrition in the assessment sample.[23] The treatment also increased sales skills by 0.19σ. The improvements in verbal and soft skills associated with this measure could be the result of direct sales skill training or indirectly through improvements in confidence or experience with self-employment.

3.2 Labor Market Outcomes

In Table 2, we present our second set of main outcomes. Earnings and labor supply outcomes at endline are a stacked panel of four retrospective rounds, estimated with the following specification:

$$Y_{it} = \beta_0 + \beta_1 * T_i + \gamma_s + \alpha_b * X_i + e_{it} \tag{2}$$

where γ_s are strata dummies, X_i are controls for the baseline household asset index and a dummy for cases in which that variable is missing, T_i is a treatment indicator, and e_{it} is an error term. Errors are clustered at the firm level. Only a small fraction of the apprentice sample had any positive earnings at baseline and earnings data were measured differently. We therefore exclude baseline values of the dependent variable from our outcomes analysis. Given that the four retrospective rounds are 1, 4, 7, and 10 months prior to the endline survey that was collected

---

[21] Extremely strict marking of exams paired with relatively lenient pass thresholds is culturally common in Ghana, as in many low- and middle-income countries.

[22] Although there were a few test-takers who failed the theoretical portion of the exam, the theory section was never binding; test-takers who failed the theory section also failed the practical section.

[23] We observe craft skills at endline for 93% of the analysis sample).

between July 2017 and May 2018, β_1 is an average of the treatment effects five months prior to the survey date, and about two years from the time of the assessment.

ITT treatment effects on earnings are estimated at 10.43 GhC per month, measured in October 2015 GhC (Panel A, Column 1). This is a meaningful increase as it translates to a 24% increase in earnings relative to the control group mean. These earnings effects are largely driven by increases in self-employment profits (Column (2)), rather than by increases in wage-employment earnings (Column (3)). Appendix Table A12 presents effects of treatment on the extensive margin of labor supply. We do not find any significant changes in the probability of working in any sector, working in self employment, or working for a wage. Focusing on self-employment, we find suggestive evidence of a small (3 percentage point) shift out of unskilled self-employment (retail and market vending) and into skilled self-employment (cosmetology and garment-making). This shift is too small given the control group mean to explain the differences in self-employment profits estimated in Column (2) of Table 2. Instead, the earnings effects are likely driven by higher self-employment profits— the treatment increased unconditional self-employment profits by 8.41 GhC, a 35% increase compared to the control group mean (Column (2)), and conditional profits by a noisy 15.15GhC (Column (4)).[24] We find no treatment effects on wage-employment earnings (Columns (3) and (5)).

3.3 Heterogeneity

Although rank-order tournaments elicit quality improvements on average, there could be important differences in the responsiveness of participants of various backgrounds. An important concern is that "weaker" participants may be discouraged from exerting effort if they believe they are competing with "stronger" participants (Dechenaux et al., 2015; Charness and Kuhn, 2011). We test for this possibility by examining the heterogeneity in treatment effects by (baseline) trainer and apprentice characteristics such as cognitive ability. Overall, while we find evidence of heterogeneity along some dimensions for certain outcomes (see Appendix Tables A13-A18), we do not find evidence consistent with discouragement.

The large bonus awarded to the trainer of the best performing apprentice could result in greater effort at the top of the distribution. We test for this possibility using quantile regressions

---

[24]The estimated treatment effects for females in Panel B tend to be larger than the estimates for the overall sample.

in Appendix Figures S1 and S2. Generally, across our skills and earnings outcomes, we find similar treatment effects across the distribution. This provides some reassurance that the gains from the intervention were broad-based, and not driven by increases among the best performing apprentices.

4   Potential Mechanisms

We first explore the role of costs of training in Table 3, using data from a phone survey that was conducted with firm owners after the intervention was launched, but prior to the assessments. We do not find any statistically significant changes in hours worked, instructional hours, or firm profits (Columns (1) to (3)).[25] As the treatment does not negatively impact firm profits, this suggests that the opportunity costs of training are not very large. This may be due to the existence of idle time in many informal sector firms–our estimates suggest that apprentices and trainers are idle over 20% of the workday.[26] In Appendix Figures S3 and S4 we examine heterogeneity in treatment by quintiles of trainer idle time. We find some suggestive evidence consistent with this theory—apprentices training with the most idle firms learned more under the treatment compared to trainees in less idle firms (Panel A). However, we do not find similar patterns for labor market outcomes. In competitive labor markets, improving training quality and skills could yield an increase in the wage bill because apprentices' outside options improve. However, we do not find any statistically significant effect on apprentice wages in Table 3 Column (4).

Table 4 presents evidence of an additional salient cost to firms that provide high-quality training: accelerated apprentice exit. Although training completion and promotion are sometimes fluid concepts in this setting (39% of certificate-holders self-report non-completion) and retrospective data on completion and exit timing in the endline survey naturally generate some data quality issues, treatment increased completion rates. Apprentices from treated firms were 22% (7 percentage points) more likely to self-report completion after the start of the program in 2015 (Column (1)) and 67% (10 percentage points) more likely to self-report paying a non-zero ceremony or exit fee to formally celebrate the completion of the apprenticeship (Column (2)).[27] Treatment increased

---

[25]The results on instructional time in Column (2) make the conservative assumption that, in firms with multiple NAP trainees, no instruction time was simultaneous (summing the instruction time reported for all NAP trainees in a firm).

[26]See Appendix Tables A19 and A20. Bassi et al. (2022a) and Bassi et al. (2022b) report similar findings.

[27]These ceremonies sometimes are just parties financed by the trainee. Other times they have traditional elements as well.

the probability of completion and exit, while having no impact on the probability of completion with retention (Column (3) and (4)). Apprentices at treated firms are 20% (8 percentage points) more likely to exit the firm in the two years after the assessment, as reported by apprentices in the endline survey and presented in Column (5).

Using plausible assumptions detailed in Appendix B, we estimate that trainers gained an additional 7GhC from the incentive on average and apprentice exit cost 40 GhC. Comparing 7GhC paid out in October 2015 to losses of 40 GhC incurred about two years later (or within two years), firm owners would need an approximately 7.5% monthly discount rate to be time-consistent exponential discounters, much higher than standard estimates in the literature of about 5% per year (Engen et al., 1994a,b; Hubbard et al., 1994), but much lower than implied discount rates calibrated in other studies of working populations in low-income countries. For example, Kaur et al. (2015) find that data entry workers in India display present bias in effort, as time-consistent calibrations yield discount rates in their study of 4% per day. Our estimates could suggest present bias. Alternatively, firm owners could be overestimating their expected payouts (e.g. they believe they will receive the large bonuses for the best performers) or underestimating the costs (and/or probability) of apprentice turnover.[28]

5   Analysis of Program Costs and Benefits

The largest costs associated with the implementation of this demonstration program were (1) the cost of the financial incentives, and (2) the cost of developing and administering the assessment. Auxiliary costs such as program enrollment and certificate distribution were subsumed into the larger administrative and survey costs associated with the research study and ongoing administrative follow-up through COTVET.

We make some simplifying assumptions in our analysis of costs. First, although financial incentives were only actually paid out for the 65% of the sample who participated in the exam, we assume an incentive cost of 100 GhC per trainee for 100% of the trainees in the program. Second, we exclude the cost of developing the exam, since it is a fixed cost and instead focus on variable costs per apprentice. Third, we treat the auxiliary administrative costs mentioned above as fixed

---

[28]Appendix Tables A21-A25 replicate our main four tables and outcomes for extensive margin labor supply excluding baseline controls, showing all results are robust to this exclusion.

costs and also exclude them from our analysis.[29]

After these restrictions and simplifications, we are left with 100 GhC per apprentice in financial incentive costs and the variable costs associated with testing. We spent about 40,000 GhC on the testing program, with assessor salaries, per diems, and travel as the largest costs. Allocated on a per apprentice basis across all 797 apprentices at the start of this study, this amounts to about 50 GhC per apprentice, yielding a total cost per apprentice of roughly 150 GhC. Taking monthly returns from Table 2, an average across all 748 apprentices for whom we have endline data, the 12-month return to the program is estimated at about 125 GhC. If we assume a 10% discount rate we can show that the program costs would be covered by 20 months of apprentice post-intervention earnings. Although these calculations are quite rough, they suggest a highly cost-effective program in terms of earnings outcomes.

Following Dhaliwal et al. (2013) and Kremer et al. (2013), we also conduct a more traditional cost-effectiveness analysis focusing on the learning outcomes measured in the skills assessment. We find the program increases learning 0.325σ per 100 USD spent. This is less cost-effective compared to the other programs studied in the broader teacher incentive literature, due to the larger costs of testing apprentices at their workplace.[30] However, it is likely that these costs could be reduced if the program was scaled up.

While the program we study in this paper targets on-the-job training *quality* conditional on apprenticeship rather than the overall impacts of training, it may still be useful to compare the per-trainee cost of the program to other cost-effective programs. For example, restricting the sample to compliers and estimating returns over a 15 year productive life, Alfonsi et al. (2020) calculate a benefit-to-cost ratio of 2.69 for their firm-provided training program which costs 368 USD per trainee. By contrast, the incentive program we study here costs about 40 USD (using October 2015 conversion rates).

---

[29]Note that we also exclude the flat cash payment that was paid to firm owners immediately prior to the program, since it was largely the result of political issues unrelated to the performance-based incentive and testing scheme.

[30]For example, the annual cost of the teacher incentive programs studied in Mbiti et al. (2019) were between 7 and 9 USD per student, compared to 40 USD per apprentice in this study. We use October 2015 exchange rates to covert GhC to USD.

6 Conclusion

Skills-upgrading through apprenticeship training offers a potentially low-cost, high-access alternative to institutional vocational training. However, the quality of training provided by these programs could be low due to lack of formalization, and monitoring. In this paper, we show that trainer incentives can encourage firms to invest more in training, which can yield large earnings gains likely associated with substantial welfare improvements in people's lives.

We explore potential determinants of trainer choice to provide quality training and find little empirical evidence supporting immediate training cost considerations, with no detectable impacts on training time, apprentice wages, nor firm productivity during training. Instead, we find significant increase in apprentice turnover, echoing existing literature from high-income country contexts that suggest firms strategically withhold training of general skills and are more willing to provide training in monopsonistic labor markets (Acemoglu and Pischke, 1998b). Because of the relatively small-scale of our experiment, it is unclear if these results would change if the program was scaled up. This remains a promising area for future research.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998a. "Why Do Firms Train? Theory and Evidence." *Quarterly Journal of Economics*.

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998b. "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics* 113 (1): 79–119.

Acemoglu, Daron, and Jörn-Steffen Pischke. 1999. "Beyond Becker: Training in Imperfect Labour Markets." *The Economic Journal*.

Ackah, Charles, Abena Oduro, and Richard Boateng. 2015. "Gender and Enterprise Development in Africa: The Case of Ghana."

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2018. "The Skills to Pay the Bills: Returns to On-the-job Soft Skills Training."

Adoho, Franck, Dala T. Chakravarty, Shubha, Mattias Lundberg, and Afia Tasneem. 2014. "The impact of an adolescent girls employment program : the EPAG project in Liberia."

Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2020. "Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda." *Econometrica*.

Bassi, Vittorio, Jung Hyuk Lee, Alessandra Peter, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2022a. "Self-employment within the firm." *Working paper*.

Bassi, Vittorio, Raffaela Muoio, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2022b. "Achieving scale collectively." *Econometrica* 90 (6): 2937–2978.

Becker, Gary. 1964. "Human Capital."

Blattman, Christopher, and Laura Ralston. 2015. "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs." *Available at SSRN 2622220*.

Breeding, Mary E., Tara Beteille, and David K. Evans. 2021. "Teacher Pay-for-Performance: What Works? Where? and How?." http://documents.worldbank.org/curated/en/183331619587678000/Teacher-Pay-for-Performance-What-Works-Where-and-How.

Caicedo, Santiago, Miguel Espinosa, and Arthur Seibold. 2022. "Unwilling to Train?—Firm Responses to the Colombian Apprenticeship Regulation." *Econometrica* 90 (2): 507–550.

Charness, Gary, and Peter Kuhn. 2011. "Lab labor: What can labor economists learn from the lab?" In *Handbook of labor economics*, Volume 4. 229–330, Elsevier.

Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, and Victor Orozco-Olvera. 2013. "Gender differences in the effects of vocational training: Constraints on women and drop-out behavior." *World Bank Policy Research Working Paper* (6545): .

Crepon, Bruno, and Patrick Premand. 2019. "Direct and Indirect Effects of Subsidized Dual Apprenticeships." *Working Paper*.

Darvas, Peter, and Robert Palmer. 2014. "Demand and Supply of Skills in Ghana: How Can Training Programs Improve Employment and Productivity?".

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2009. "Measuring Microenterprise Profits: Must We Ask How The Sausage Is Made?" *Journal of Development Economics* 88, pages 19-31.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M Sheremeta. 2015. "A survey of experimental research on contests, all-pay auctions and tournaments." *Experimental Economics* 18 (4): 609–669.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. 2013. "Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education." *Education Policy in Developing Countries* 285–338.

Engen, Eric, William Gale, and John Scholz. 1994a. "Do Saving Incentives Work?".

Engen, Eric, William Gale, and Cori Uccello. 1994b. "The Adequacy of Household Saving."

Filmer, Deon, and Louise Fox. 2014. "Youth Employment in Sub-Saharan Africa."

Frazer, Garth. 2006a. "Learning the Master's Trade: Apprenticeship and Human Capital in Ghana." *Journal of Development Economics* 81 259–298.

Frazer, Garth. 2006b. "Learning the master's trade: apprenticeship and human capital in Ghana." *Journal of Development Economics* 81 (2): 259–298.

Fudenberg, Drew, George Georgiadis, and Luis Rayo. 2021. "Working to learn." *Journal of Economic Theory* 197 105347.

Fudenberg, Drew, and Luis Rayo. 2019. "Training and effort dynamics in apprenticeship." *American Economic Review* 109 (11): 3780–3812.

Ganimian, Alejandro J, and Richard J Murnane. 2016. "Improving education in developing countries: Lessons from rigorous impact evaluations." *Review of Educational Research* 86 (3): 719–755.

Hardy, Morgan, Isaac Mbiti, Jamie McCasland, and Isabelle Salcher. 2019. "The Apprenticeship-to-Work Transition." *World Bank Policy Research Working Paper 8851*.

Hardy, Morgan, and Jamie McCasland. forthcoming. "Are Small Firms Labor Constrained? Experimental Evidence from Ghana." *American Economic Journal: Applied Economics*.

Hubbard, Glenn, Jonathan Skinner, and Stephen Zeldes. 1994. "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving."

ILO. 2011. "Upgrading informal apprenticeship systems."

ILO. 2019. "Quality Apprenticeships: Addressing skills mismatch and youth unemployment."

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. "Self-control at work."

Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The challenge of education and learning in the developing world." *Science* 340 (6130): 297–300.

Leuven, Edwin. 2005. "The economics of private sector training: a survey of the literature." *Journal of economic surveys* 19 (1): 91–111.

Mbiti, Isaac, Mauricio Romero, and Youdi Schipper. 2019. "Designing effective teacher performance pay programs: experimental evidence from Tanzania."Technical report, National Bureau of Economic Research.

Mckenzie, David. 2017. "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence." *World Bank Research Observer*.

Monk, Courtney, Justin Sandefur, and Francis Teal. 2008. "Does Doing an Apprenticeship Pay Off? Evidence from Ghana." *CSAE Working Paper Series 2008-08* Centre for the Study of African Economies, University of Oxford.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): pp. 39–77, `http://www.jstor.org/stable/10.1086/659655`.

Pham, Lam D, Tuan D Nguyen, and Matthew G Springer. 2021. "Teacher merit pay: A meta-analysis." *American Educational Research Journal* 58 (3): 527–566.

Teal, Francis. 2016. "Are apprenticeships beneficial in sub-saharan africa?" *IZA World of Labor*.

Wolter, Stefan C, and Paul Ryan. 2011. "Apprenticeship." In *Handbook of the Economics of Education*, Volume 3. 521–576, Elsevier.

Table 1: Skills

	Assessment						Endline survey		
	=1 if took incentivized test	=1 if took incentivized test	Practical component (z - score)	Theoretical component (z - score)	Passed practical	Earned certificate	Craft skills quiz (z - score)	Sales skills (z - score)	Craft-Sales skills (z - score)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Full sample									
Treatment	-0.01	-0.01	0.13*	-0.08	0.05	0.05	0.15**	0.19***	0.22***
	(0.04)	(0.04)	(0.07)	(0.07)	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)
Observations	797	763	488	488	488	488	743	748	743
Mean of Dep Variable T=0	0.65	0.65	0.00	0.00	0.80	0.80	0.00	0.00	-0.09
Panel B: Women only									
Treatment	-0.00	-0.01	0.18**	-0.05	0.07*	0.07*	0.15**	0.19**	0.22***
	(0.04)	(0.04)	(0.08)	(0.07)	(0.04)	(0.04)	(0.07)	(0.08)	(0.08)
Observations	751	717	457	457	457	457	703	706	703
Mean of Dep Variable T=0	0.64	0.64	0.00	0.00	0.81	0.81	0.00	0.00	-0.10

Notes: This table shows the results from estimating equation 1 for different skill measures. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Specifications from the assessment include assessor dummies. Columns 2 through 6 exclude test-takers from the first district, because the assessment (being implemented in collaboration with our government partners) differed from the test given in all other districts. Standard errors, clustered at the firm level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Earnings

	Unconditional			Conditional	
	Total earnings (GHC)	Self-employment profits (GHC)	Wage employment earnings (GHC)	Self-employment profits (GHC)	Wage employment earnings (GHC)
	(1)	(2)	(3)	(4)	(5)
Panel A: Full sample					
Treatment	10.43*	8.41*	3.41	15.15	-4.10
	(5.56)	(4.70)	(3.40)	(12.23)	(19.55)
Observations	2992	2992	2992	869	301
Mean of Dep Variable T=0	44.34	24.18	12.98	83.58	136.89
Panel B: Women only					
Treatment	12.89**	8.93*	3.02	22.22*	0.21
	(5.46)	(4.75)	(3.57)	(12.67)	(21.40)
Observations	2824	2824	2824	806	287
Mean of Dep Variable T=0	41.52	22.99	13.10	80.62	135.34

Notes: This table shows the results from estimating equation 2 for different labor market outcomes. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. Total earnings includes earnings from wage-employment, self-employment, farming, and apprenticeship. Profits in self-employment follow self-reported question structures from De Mel et al. (2009). All earnings are in October 2015 Ghana Cedis. Standard errors, clustered at the firm level, are in parentheses. $* p < 0.10, ** p < 0.05, *** p < 0.01$

Table 3: Firm owner (trainer) effort

	Firm level			Apprentice level	
	Firm owner hours last week	Firm owner instruction hours per day	Firm profits last month (GHC)	Firm owner instruction per day	Worker wages last month (GHC)
	(1)	(2)	(3)	(4)	(5)
Panel A: Full sample					
Treatment	-1.24	-0.09	12.60	0.10	1.55
	(1.85)	(0.21)	(31.11)	(0.09)	(3.08)
Observations	339	341	335	595	566
Mean of Dep Variable T=0	52.33	3.42	251.39	1.29	22.17
Panel B: Women only					
Treatment	-0.74	-0.04	9.64	0.08	1.63
	(1.76)	(0.22)	(32.72)	(0.09)	(3.28)
Observations	323	325	319	556	528
Mean of Dep Variable T=0	52.46	3.33	237.20	1.32	21.85

Notes: All specifications include strata fixed effects. Specifications in Columns (4) and (5) include controls for the baseline household asset index of the apprentice (including a dummy for where the variable is missing). Standard errors in Columns (4) and (5) are clustered at the firm level. Firm profits follow self-reported question structures from De Mel et al. (2009). Wages for each apprentice are reported by the firm owner. Both profits and wages are in October 2015 Ghana Cedis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Apprentice completion and turnover

	Apprentice completion				Apprentice turnover	
	Completed	Paid exit/ceremony	Completed and exited	Completed and retained	Exited	Missed months
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full sample						
Treatment	0.07**	0.10***	0.08**	-0.01	0.08**	1.50*
	(0.04)	(0.03)	(0.04)	(0.01)	(0.04)	(0.84)
Observations	748	748	748	748	748	748
Mean of Dep Variable T=0	0.31	0.15	0.28	0.03	0.39	7.73
Panel B: Women only						
Treatment	0.07*	0.11***	0.08**	-0.00	0.08**	1.63*
	(0.04)	(0.03)	(0.04)	(0.01)	(0.04)	(0.86)
Observations	706	706	706	706	706	706
Mean of Dep Variable T=0	0.32	0.16	0.30	0.02	0.40	7.99

Notes: All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Exited in Column (5) includes those who exit and complete and those who self-report exiting without completing. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A   Online Appendix

Table A1: Balance: Apprentice Characteristics (Full Sample - Men and Women)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	797	0.56	0.54	-0.18	(0.17)
North	797	0.34	0.35		
Program Apprentices (incl. Respondent)	797	2.51	2.45		
Female	797	0.92	0.96	0.04	(0.10)*
Age	774	22.87	22.96	0.10	(0.81)
Married (0/1)	772	0.37	0.34	-0.03	(0.38)
Number of Children	781	0.74	0.79	-0.01	(0.85)
Lives with Parent(s) (0/1)	778	0.49	0.45	-0.01	(0.81)
Household Asset score (z-score)	766	0.00	0.16	0.13	(0.04)**
Years Schooling	760	7.50	7.38	-0.17	(0.44)
Ability Index (z-score)	780	0.00	0.01	0.01	(0.94)
Soft Skills Index (z-score)	778	0.00	0.08	0.09	(0.19)
Food Security (z-score)	776	0.00	-0.03	-0.06	(0.40)
Self-Reported Health (z-score)	779	0.00	0.01	-0.01	(0.94)
Mother Years Schooling	647	3.28	3.53	0.32	(0.35)
Father Years Schooling	567	5.55	5.66	0.08	(0.86)
Prior Apprenticeship Experience (0/1)	780	0.26	0.27	0.01	(0.74)
Wage-Employed (0/1)	781	0.04	0.04	0.00	(0.92)
Wage-Employment Earnings (GhC)	781	1.44	1.39	0.15	(0.82)
Self-Employed (0/1)	781	0.22	0.18	-0.02	(0.59)
Self-Employment Profits (GhC)	781	11.74	10.59	0.36	(0.90)
Total Weekly Hours Worked	772	18.54	16.61	-0.70	(0.72)
Desires Self-Employment (0/1)	781	0.55	0.52	-0.04	(0.22)

Notes: *North* indicates that the apprentice lives in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 781 of 797 apprentices participated in the baseline survey between August and December of 2012. This table includes all available data from the baseline survey. *Household Asset* is the first principal component of a set of housing quality (floor material, roof material, wall material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of the normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the first principal component of two food consumption measures: meals eaten per day and meals per day that included meat (excluding eggs). *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 20 covariates (from female to desires self-employment) in predicting treatment yields of p-value of 0.39. * p < 0.10, ** p < 0.05, *** p < 0.01

25

Table A2: Pay for Performance Scheme

Group (Decile)	Payment to MCP (Ghana Cedis)
1	200
2	140
3	130
4	120
5	105
6	90
7	80
8	70
9	35
10	25

Notes: The above lists the payouts by apprentice performance decile for each apprentice in the treatment group. In addition, firm owners in the treatment group whose apprentice performed best in the region and craft earned their trainer 250 GhC. Firm owners in the control group received 100 GhC for every apprentice who took the test regardless of performance. Payouts for both control and treatment firms averaged 100 GhC per apprentice, both in expectation and ex-post.

Table A3: Balance: Apprentice Characteristics (Endline Sample - Men and Women)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	748	0.56	0.53		
North	748	0.34	0.35		
Program Apprentices (incl. Respondent)	748	2.52	2.46	-0.16	(0.23)
Female	748	0.92	0.96	0.03	(0.13)
Age	727	22.91	22.98	0.04	(0.92)
Married (0/1)	725	0.36	0.35	-0.02	(0.57)
Number of Children	734	0.74	0.80	-0.02	(0.82)
Lives with Parent(s) (0/1)	732	0.49	0.44	-0.01	(0.77)
Household Asset score (z-score)	720	-0.01	0.14	0.15	(0.03)**
Years Schooling	713	7.49	7.40	-0.11	(0.62)
Ability Index (z-score)	733	-0.02	0.00	0.02	(0.82)
Soft Skills Index (z-score)	731	0.00	0.08	0.07	(0.30)
Food Security (z-score)	730	-0.01	-0.02	-0.04	(0.58)
Self-Reported Health (z-score)	732	0.00	-0.01	-0.02	(0.85)
Mother Years Schooling	611	3.34	3.44	0.10	(0.76)
Father Years Schooling	536	5.60	5.74	0.00	(1.00)
Prior Apprenticeship Experience (0/1)	733	0.26	0.28	0.02	(0.50)
Wage-Employed (0/1)	734	0.04	0.04	0.00	(0.86)
Wage-Employment Earnings (GhC)	734	1.46	1.36	0.12	(0.86)
Self-Employed (0/1)	734	0.22	0.19	-0.02	(0.61)
Self-Employment Profits (GhC)	734	12.23	11.03	0.10	(0.98)
Total Weekly Hours Worked	726	19.06	17.32	-0.41	(0.84)
Desires Self-Employment (0/1)	734	0.56	0.52	-0.05	(0.17)

Notes: *North* indicates that the apprentice lives in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 781 of 797 apprentices participated in the baseline survey between August and December of 2012. 748 apprentices participated in the endline survey. 734 participated in both and thus have baseline covariates against which to test for balance and endline survey data, from which most of our treatment effects are measured. This table restricts the sample to those who participated in the endline survey. *Household Asset* is the first principal component of a set of housing quality (floor material, roof material, wall material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of the normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the first principal component of two food consumption measures: meals eaten per day and meals per day that included meat (excluding eggs). *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 20 covariates (from female to desires self-employment) in predicting treatment yields of p-value of 0.56. * p < 0.10, ** p < 0.05, *** p < 0.01

27

Table A4: Balance: Apprentice Characteristics (Assessment Sample - Men and Women)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	488	0.59	0.57		
North	488	0.47	0.48		
Program Apprentices (incl. Respondent)	488	2.65	2.51	-0.24	(0.16)
Female	488	0.91	0.96	0.03	(0.23)
Age	474	22.69	22.71	-0.07	(0.89)
Married (0/1)	473	0.41	0.42	-0.01	(0.80)
Number of Children	480	0.75	0.88	0.01	(0.90)
Lives with Parent(s) (0/1)	479	0.50	0.48	0.02	(0.60)
Household Asset score (z-score)	472	-0.12	0.06	0.21	(0.01)**
Years Schooling	463	7.16	7.42	0.29	(0.33)
Ability Index (z-score)	480	-0.05	0.02	0.03	(0.79)
Soft Skills Index (z-score)	478	-0.07	-0.05	0.07	(0.50)
Food Security (z-score)	477	-0.04	-0.09	-0.03	(0.76)
Self-Reported Health (z-score)	478	0.01	-0.04	0.00	(0.96)
Mother Years Schooling	416	2.75	2.86	0.15	(0.71)
Father Years Schooling	367	4.59	5.11	0.40	(0.47)
Prior Apprenticeship Experience (0/1)	479	0.30	0.26	-0.02	(0.65)
Wage-Employed (0/1)	480	0.04	0.03	0.00	(0.89)
Wage-Employment Earnings (GhC)	480	1.28	1.28	0.18	(0.84)
Self-Employed (0/1)	480	0.24	0.20	-0.03	(0.52)
Self-Employment Profits (GhC)	480	12.12	11.79	1.98	(0.62)
Total Weekly Hours Worked	476	21.54	17.24	-2.69	(0.33)
Desires Self-Employment (0/1)	480	0.53	0.51	-0.03	(0.55)

Notes: *North* indicates that the apprentice lives in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 781 of 797 apprentices participated in the baseline survey between August and December of 2012. 488 apprentices participated in the assessment (excluding the pilot district). 480 participated in both and thus have baseline covariates against which to test for balance and assessment data, from which some of our skills treatment effects are measured. This table restricts the sample to those who participated in the assessment. *Household Asset* is the first principal component of a set of housing quality (floor material, roof material, wall material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of the normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the first principal component of two food consumption measures: meals eaten per day and meals per day that included meat (excluding eggs). *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 20 covariates (from female to desires self-employment) in predicting treatment yields of p-value of 0.62. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Balance: Apprentice Characteristics (Full Sample – Women Only)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	751	0.53	0.52		
North	751	0.34	0.33		
Female	751	1.00	1.00		
Program Apprentices (incl. Respondent)	751	2.44	2.48	-0.15	(0.24)
Age	729	22.62	22.86	0.29	(0.43)
Married (0/1)	727	0.38	0.34	-0.05	(0.17)
Number of Children	736	0.75	0.79	-0.03	(0.72)
Lives with Parent(s) (0/1)	733	0.48	0.45	-0.01	(0.69)
Household Asset score (z-score)	722	0.00	0.17	0.15	(0.03)**
Years Schooling	715	7.40	7.30	-0.17	(0.45)
Ability Index (z-score)	735	0.00	0.03	0.02	(0.80)
Soft Skills Index (z-score)	733	0.00	0.09	0.09	(0.18)
Food Security (z-score)	731	0.00	-0.02	-0.05	(0.49)
Self-Reported Health (z-score)	734	0.00	0.03	0.01	(0.85)
Mother Years Schooling	611	3.37	3.59	0.26	(0.47)
Father Years Schooling	531	5.60	5.67	0.10	(0.84)
Prior Apprenticeship Experience (0/1)	735	0.25	0.25	0.01	(0.86)
Wage-Employed (0/1)	736	0.02	0.03	0.01	(0.48)
Wage-Employment Earnings (GhC)	736	0.57	1.11	0.56	(0.19)
Self-Employed (0/1)	736	0.21	0.18	-0.01	(0.68)
Self-Employment Profits (GhC)	736	11.43	9.03	-1.38	(0.60)
Total Weekly Hours Worked	728	18.27	15.85	-0.95	(0.62)
Desires Self-Employment (0/1)	736	0.55	0.52	-0.05	(0.21)

Notes: *North* indicates that the apprentice lives in the the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. All apprentices in this table are female. Each treatment coefficient in this table are female. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 736 of the 751 women in the analysis sample participated in the baseline survey between August and December of 2012. *Household Asset* is the first principal component of a set of housing quality (floor material, wall material, roof material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of the normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the first principal component of two food consumption measures: meals eaten per day and meals per day that included meat (excluding eggs). *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 19 covariates (from age to desires self-employment) in predicting treatment yields of p-value of 0.24. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

29

Table A6: Balance: Apprentice Characteristics (Endline Sample - Women Only)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	706	0.53	0.52		
North	706	0.33	0.34		
Female	706	1.00	1.00		
Program Apprentices (incl. Respondent)	706	2.45	2.50	-0.14	(0.29)
Age	686	22.66	22.89	0.21	(0.58)
Married (0/1)	684	0.37	0.35	-0.04	(0.35)
Number of Children	693	0.75	0.80	-0.03	(0.73)
Lives with Parent(s) (0/1)	691	0.48	0.44	-0.01	(0.73)
Household Asset score (z-score)	680	-0.01	0.16	0.15	(0.02)**
Years Schooling	672	7.40	7.33	-0.10	(0.67)
Ability Index (z-score)	692	-0.02	0.03	0.04	(0.63)
Soft Skills Index (z-score)	690	0.01	0.08	0.07	(0.36)
Food Security (z-score)	689	-0.01	0.00	-0.04	(0.63)
Self-Reported Health (z-score)	691	0.00	0.01	0.01	(0.95)
Mother Years Schooling	578	3.43	3.52	0.07	(0.83)
Father Years Schooling	502	5.71	5.74	-0.03	(0.95)
Prior Apprenticeship Experience (0/1)	692	0.25	0.26	0.01	(0.70)
Wage-Employed (0/1)	693	0.02	0.03	0.01	(0.38)
Wage-Employment Earnings (GhC)	693	0.60	1.18	0.67	(0.16)
Self-Employed (0/1)	693	0.22	0.18	-0.01	(0.65)
Self-Employment Profits (GhC)	693	11.83	9.35	-2.09	(0.44)
Total Weekly Hours Worked	686	18.84	16.62	-0.72	(0.72)
Desires Self-Employment (0/1)	693	0.56	0.52	-0.06	(0.14)

Notes: *North* indicates that the apprentice lives in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. All workers in this table are female. Each treatment coefficient comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 736 of the 751 women in the analysis sample participated in the baseline survey between August and December of 2012. 706 participated in the endline survey. 700 participated in both and thus have baseline covariates against which to test for balance and endline survey data, from which most of our treatment effects are measured. *Household Asset* is the first principal component of a set of housing quality (floor material, roof material, wall material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of two normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 19 covariates (from age to desires self-employment) in predicting treatment yields of p-value of 0.23. $* p < 0.10, ** p < 0.05, *** p < 0.01$

Table A7: Balance: Apprentice Characteristics (Assessment Sample - Women Only)

	(1) Observations	(2) Mean Control	(3) Mean Treatment	(4) Treatment Coefficient	(5) Coefficient p-value
Garments	457	0.57	0.55		
North	457	0.46	0.46		
Female	457	1.00	1.00		
Program Apprentices (incl. Respondent)	457	2.61	2.55	-0.22	(0.22)
Age	443	22.40	22.61	0.16	(0.75)
Married (0/1)	442	0.43	0.41	-0.04	(0.45)
Number of Children	449	0.76	0.87	0.00	(0.99)
Lives with Parent(s) (0/1)	448	0.50	0.47	0.01	(0.88)
Household Asset score (z-score)	441	-0.12	0.07	0.23	(0.01)***
Years Schooling	432	6.96	7.32	0.36	(0.24)
Ability Index (z-score)	449	-0.07	0.06	0.07	(0.47)
Soft Skills Index (z-score)	447	-0.07	-0.05	0.05	(0.63)
Food Security (z-score)	446	-0.04	-0.07	-0.02	(0.81)
Self-Reported Health (z-score)	447	0.01	-0.04	0.00	(0.98)
Mother Years Schooling	391	2.82	2.93	0.15	(0.73)
Father Years Schooling	342	4.65	5.12	0.43	(0.47)
Prior Apprenticeship Experience (0/1)	448	0.28	0.24	-0.03	(0.53)
Wage-Employed (0/1)	449	0.02	0.02	0.00	(1.00)
Wage-Employment Earnings (GhC)	449	0.40	0.88	0.40	(0.46)
Self-Employed (0/1)	449	0.23	0.19	-0.03	(0.51)
Self-Employment Profits (GhC)	449	11.17	9.02	-1.12	(0.73)
Total Weekly Hours Worked	446	21.03	16.25	-3.42	(0.20)
Desires Self-Employment (0/1)	449	0.53	0.52	-0.02	(0.69)

Notes: *North* indicates that the apprentice lives in the the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and program apprentices (including respondent) to be balanced by construction. All workers in this table are female. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Errors are clustered at the firm level. 736 of the 751 women in the analysis sample participated in the baseline survey between August and December of 2012. 457 participated in the assesment (excluding the women from the pilot district where the assessment differed from other districts). 449 participated in both and thus have baseline covariates against which to test for balance and assessment data, from which some of our key skills treatment effects are measured. *Household Asset* is the first principal component of a set of housing quality (floor material, roof material, wall material, sanitation access, water source, lighting source, primary cooking implement, and number of people who sleep in the same room) and asset measures (mattress, tv, radio, refrigerator, car, motorbike, bicycle, working mobile phones, and shoes). *Ability Index* is the normalized sum of the normalized scores on each of four cognitive tests: Digit Span Recall, four math questions we developed ourselves, Ravens Matrices Group B, and a fifteen word oral English vocabulary recognition test. *Soft Skills Index* is the normalized sum of the normalized scores on each of two (adapted to context) non-cognitive tests: the Rosenberg Self-Esteem scale and the Rotter Locus of Control scale. *Food Security* is the first principal component of two food consumption measures: meals eaten per day and meals per day that included meat (excluding eggs). *Self-Reported Health* is the normalized response to *All in all, how would you describe your state of health these days?* on a four-point Likert scale. The top 0.5% of wage job earnings and self-employment profits have been winsorized; these variables are measured in October 2015 Ghana Cedis. *Desires Self-Employment* is an indicator variable for the respondent reporting that the reason they are interested in apprenticeship training is that it will lead to self-employment. An F test of the joint significance of all 19 covariates (from age to desires self-employment) in predicting treatment yields of p-value of 0.26. * p < 0.10, ** p < 0.05, *** p < 0.01

Table A8: Balance: Firm Characteristics - Firms Training Men and Women

	(1)	(2)	(3)	(4)	(5)
	Observations	Mean Control	Mean Treatment	Treatment Coefficient	Coefficient p-value
Garments	467	0.55	0.56		
North	467	0.31	0.32		
Program Apprentices	467	1.70	1.71	-0.10	(0.16)
Female	467	0.86	0.90	0.03	(0.34)
Age	457	35.47	35.05	-0.29	(0.69)
Years Schooling	462	8.52	8.48	-0.09	(0.78)
Completed Apprenticeship (0/1)	461	0.97	0.97	0.01	(0.62)
Ability Index (z-score)	421	0.00	0.02	0.01	(0.95)
Firm Age (Years)	463	11.22	10.72	-0.37	(0.59)
Revenues Last Month (GhC)	419	372.00	373.40	11.36	(0.78)
Profits Last Month (GhC)	419	187.70	199.10	21.48	(0.33)
Firm Size (Incl Owner)	421	4.43	4.52	-0.05	(0.86)
Wage Bill (GhC)	421	55.11	56.16	-3.84	(0.72)
Firm Assets (GhC)	421	2,870.00	3,029.00	249.05	(0.32)
Trade Association Member (0/1)	421	0.85	0.86	0.00	(0.94)
Past Completed Apprentices	420	12.39	10.29	-1.94	(0.28)
Customer Service Instruction (0/1)	406	0.98	0.97	0.00	(0.98)
Inventory Management Instruction (0/1)	406	0.59	0.65	0.05	(0.38)
Financial Management Instruction (0/1)	406	0.54	0.56	0.01	(0.90)
Business Confidence Instruction (0/1)	406	0.60	0.65	0.00	(0.97)
Provides More Instruction to Slow Apprentices (0/1)	402	0.30	0.30	-0.01	(0.84)
Months Until Apprentices Work with Clients	406	9.96	9.62	-0.14	(0.87)

Notes: *North* indicates that the firm is located in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and number of program apprentices to be balanced by construction. This table includes all firms in our sample. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. *Ability Index* is the normalized sum of the normalized scores on two cognitive tests: Digit Span Recall and four math questions we developed ourselves. All financial variables are in October 2015 GhC. An F test of the joint significance of all 19 covariates (from female to months until apprentices work with clients) in predicting treatment yields of p-value of 0.98. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Balance: Firm Characteristics - Firms Training Women Only

	(1)	(2)	(3)	(4)	(5)
		Mean	Mean	Treatment	Coefficient
	Observations	Control	Treatment	Coefficient	p-value
Garments	446	0.53	0.54	-0.08	(0.22)
North	446	0.31	0.31	0.01	(0.61)
Program Apprentices	446	1.71	1.73	0.02	(0.72)
Female	446	0.91	0.92	0.01	(0.65)
Age	436	35.37	35.20	-0.11	(0.88)
Years Schooling	441	8.46	8.44	-0.07	(0.83)
Completed Apprenticeship (o/1)	440	0.97	0.97	0.01	(0.67)
Ability Index (z-score)	403	0.00	0.03	0.02	(0.86)
Firm Age (Years)	442	11.20	10.80	-0.24	(0.73)
Revenues Last Month (GhC)	401	378.10	377.80	14.23	(0.73)
Profits Last Month (GhC)	401	191.80	201.00	20.67	(0.37)
Firm Size (Incl Owner)	403	4.44	4.52	-0.03	(0.92)
Wage Bill (GhC)	403	55.18	54.70	-5.80	(0.59)
Firm Assets (GhC)	403	2,849.00	3,054.00	240.90	(0.35)
Trade Association Member (o/1)	403	0.85	0.87	0.01	(0.84)
Past Completed Apprentices	402	12.81	10.68	-1.98	(0.30)
Customer Service Instruction (o/1)	389	0.98	0.97	0.00	(0.97)
Inventory Management Instruction (o/1)	389	0.59	0.67	0.07	(0.20)
Financial Management Instruction (o/1)	389	0.53	0.56	0.02	(0.77)
Business Confidence Instruction (o/1)	389	0.62	0.65	-0.02	(0.66)
Provides More Instruction to Slow Apprentices (o/1)	385	0.32	0.29	-0.03	(0.58)
Months Until Apprentices Work with Clients	389	9.68	9.38	-0.11	(0.90)

Notes: *North* indicates that the firm is located in the the northern regions of the country as of 2015, which are socio-culturally distinct from the 7 southern regions of Ghanas of 2015. The randomization was stratified by district, trade, and a measure of the number of NAP apprentices training at the firm, so we expect garments, north, and number of program apprentices to be balanced by construction. This table includes all firms training female workers in our sample. Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. *Ability Index* is the normalized sum of the normalized scores on two cognitive tests: Digit Span Recall and four math questions we developed ourselves. All financial variables are in October 2015 GhC. An F test of the joint significance of all 19 covariates (from female to months until apprentices work with clients) in predicting treatment yields of p-value of 0.97. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A10: Attrition

| | Targeted Sample | = 1 if Surveyed | | Treatment Coefficient | Coefficient p-value |
		Control	Treatment		
Panel A: Full Sample - Men and Women					
Apprentice Baseline Survey	751.00	0.98	0.98	0.00	(0.88)
Apprentice Incentivized Test	797.00	0.65	0.63	-0.01	(0.87)
Apprentice Endline Survey	797.00	0.94	0.94	-0.01	(0.71)
Firm Baseline Survey	467.00	0.90	0.91	0.01	(0.72)
Firm Midline Survey	467.00	0.72	0.74	0.01	(0.81)
Panel B: Analysis Sample - Women Only					
Apprentice Baseline Survey	751	0.98	0.98	0.00	(0.88)
Apprentice Incentivized Test	751	0.64	0.63	0.00	(0.97)
Apprentice Endline Survey	751	0.94	0.94	-0.01	(0.66)
Firm Baseline Survey	446	0.90	0.90	0.00	(0.96)
Firm Midline Survey	446	0.72	0.73	0.01	(0.85)

Notes: Each treatment coefficient in column (4) comes from a separate regression that includes strata fixed effects. Standard errors in apprentice-level regressions are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Correlates of attrition

	Full sample (1)	Women only (2)
Treatment	-0.00	0.01
	(0.04)	(0.04)
Age	-0.01	-0.01
	(0.00)	(0.00)
Married (0/1)	0.01	0.01
	(0.04)	(0.04)
Number of Children	-0.01	-0.01
	(0.02)	(0.02)
Household Asset score (z-score)	0.01	0.02
	(0.02)	(0.02)
Years Schooling	0.01	0.01
	(0.01)	(0.01)
Ability Index (z-score)	-0.00	-0.00
	(0.02)	(0.02)
Soft Skills Index (z-score)	-0.02	-0.02
	(0.02)	(0.02)
Food Security (z-score)	-0.02	-0.01
	(0.02)	(0.02)
Self-Reported Health (z-score)	-0.01	-0.01
	(0.02)	(0.02)
Desires Self-Employment (0/1)	0.00	0.01
	(0.03)	(0.03)
Observations	797	751

Notes: The dependent variable is a dummy of test take up (Take exam = 1). Standard errors, clustered at the firm level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A12: Labor supply

	Working	Self-employment	Skilled self-employment	Unskilled self-employment	Wage-employment	Skilled wage-employment	Unskilled wage-employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full sample							
Treatment	0.03	0.01	0.04	-0.03*	0.00	-0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	2992	2992	2992	2992	2992	2992	2992
Mean of Dep Variable T=0	0.62	0.29	0.16	0.13	0.10	0.05	0.05
Panel B: Women only							
Treatment	0.03	0.00	0.04	-0.03*	-0.00	-0.00	0.00
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	2824	2824	2824	2824	2824	2824	2824
Mean of Dep Variable T=0	0.62	0.29	0.16	0.13	0.10	0.05	0.05

Notes: This Table shows the results from estimating Equation 2 for different labor market outcomes. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Standard errors are clustered at the firm level. Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

36

Table A13: Heterogeneity in treatment effects (Practical Test)

	Full Sample					Women only				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)

Panel A: Apprentice characteristics

	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	-.091	.31**	-.027	-.12**	.039	-.043	.29**	.0013	-.12*	.082
	(.072)	(.14)	(.068)	(.062)	(.15)	(.074)	(.14)	(.067)	(.062)	(.14)
N. of obs.	488	463	480	478	480	457	432	449	447	449

Panel B: Firm owner characteristics

	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	-.071	.1	-.025	-.092	-.0041	.1	.084	-.0027	-.069	-.0064
	(.18)	(.078)	(.031)	(.17)	(.0056)	(.17)	(.082)	(.03)	(.17)	(.0055)
N. of obs.	486	464	464	462	464	455	434	434	432	434

Notes: The outcome variable is the practical component z-score. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 – 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects, assessor dummies and controls for the baseline household asset index (including a dummy for where the variable is missing). All specifications exclude pilot district. Standard errors, clustered at the firm level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A14: Heterogeneity in treatment effects (Craft skills (z-score))

Panel A: Apprentice characteristics

	Full Sample					Women only				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	.091	.27*	-.12	-.043	-.17	.07	.29*	-.13	-.024	-.18
	(.077)	(.16)	(.077)	(.072)	(.16)	(.079)	(.16)	(.082)	(.075)	(.17)
N. of obs.	743	710	730	728	731	703	671	691	689	692

Panel B: Firm owner characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	.23	-.025	-.029	-.022	.0052	.22	-.0014	-.031	-.069	.0065
	(.18)	(.09)	(.032)	(.17)	(.0041)	(.18)	(.086)	(.032)	(.17)	(.0041)
N. of obs.	738	693	693	691	692	698	654	654	652	653

Notes: The outcome variable is the craft skills quiz z-score. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 - 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Standard errors, clustered at the firm level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A15: Heterogeneity in treatment effects (Sales skills (z-score))

	Full Sample					Women only				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Apprentice characteristics										
	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	.14*	-.012	.1	.16**	-.13	.17**	.061	.092	.13*	-.16
	(.076)	(.16)	(.072)	(.075)	(.15)	(.077)	(.16)	(.076)	(.079)	(.16)
N. of obs.	748	713	733	731	734	706	672	692	690	693
Panel B: Firm owner characteristics										
	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	-.16	.2**	.037	.026	.017***	-.16	.2**	.036	.044	.018***
	(.16)	(.086)	(.032)	(.18)	(.0059)	(.17)	(.091)	(.033)	(.18)	(.0059)
N. of obs.	743	696	696	694	695	701	656	656	654	655

Notes: The outcome variable is the sales skills quiz z-score. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 - 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Standard errors, clustered at the firm level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A16: Heterogeneity in treatment effects (Total earnings)

Panel A: Apprentice characteristics

	Full Sample					Women only				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	16*	-31	4.1	14**	19	18*	-27	-.66	14**	17
	(9.4)	(23)	(5.3)	(6.4)	(16)	(9.7)	(24)	(5.3)	(6.6)	(16)
N. of obs.	2,992	2,852	2,784	2,924	2,936	2,824	2,688	2,624	2,760	2,772

Panel B: Firm owner characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	-34***	4.1	-.61	38***	.052	-29**	-.66	.87	26**	-.1
	(13)	(5.3)	(2.3)	(12)	(.4)	(13)	(5.3)	(2.2)	(11)	(.4)
N. of obs.	2,972	2,784	2,784	2,776	2,780	2,804	2,624	2,624	2,616	2,620

Notes: The outcome variable is the total earnings. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 - 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. Total earnings includes earnings from wage-employment, self-employment, farming, and apprenticeship. Standard errors, clustered at the firm level, are in parentheses. Earnings are in October 2015 Ghana Cedis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

40

Table A17: Heterogeneity in treatment effects (Self-employment business profits (UC))

| | Full Sample | | | | | Women only | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Apprentice characteristics	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	11	-31	-4.7	14**	18	12	-30	-5.2	14**	21
	(8.2)	(23)	(4.3)	(5.9)	(14)	(8.5)	(24)	(4.4)	(6.2)	(14)
N. of obs.	2,992	2,852	2,784	2,924	2,936	2,824	2,688	2,624	2,760	2,772
Panel B: Firm owner characteristics	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	-.95	-4.7	.11	15	-.044	-.63	-5.2	.84	13	-.13
	(10)	(4.3)	(1.7)	(10)	(.38)	(11)	(4.4)	(1.6)	(10)	(.39)
N. of obs.	2,972	2,784	2,784	2,776	2,780	2,804	2,624	2,624	2,616	2,620

Notes: The outcome variable is the unconditional self-employment business profits. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 - 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. Total earnings includes earnings from wage-employment, self-employment, farming, and apprenticeship. Standard errors, clustered at the firm level, are in parentheses. Earnings are in October 2015 Ghana Cedis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

41

Table A18: Heterogeneity in treatment effects (Self-employment business profits (Conditional))

	Full Sample					Women only				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)

Panel A: Apprentice characteristics

	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment	Asset score	Less than grade 8	Ability	Soft skills	Desires self employment
Treatment*Covariate	7.7	-18	7.2	25**	52**	12	-29	3.1	24*	64**
	(12)	(20)	(11)	(11)	(25)	(12)	(21)	(11)	(12)	(26)
N. of obs.	869	831	787	856	860	806	771	727	796	800

Panel B: Firm owner characteristics

	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained	Less than grade 8	Ability	Total workers	Above median sales	Apprentices trained
Treatment*Covariate	-9.1	7.2	-5.6	30	.0023	-19	3.1	-6.1	23	-35
	(24)	(11)	(4.4)	(31)	(1.2)	(26)	(11)	(4.9)	(31)	(1.2)
N. of obs.	864	787	787	785	787	801	727	727	725	727

Notes: The outcome variable is the conditional self-employment business profits. Each column shows the heterogeneous treatment effect by different worker (Panel A) and firm (Panel B) characteristics. For both panels, Columns 1 - 5 use the full apprentices sample and Columns 6-10 only the female apprentices sample. For apprentices characteristics: "Asset score" is the z-score of the apprentice's household asset score; "Less than grade 8" is a dummy variable that equals 1 if the apprentice has less than eight years of formal schooling; "Ability" is the z-score of the apprentice's ability index; "Soft skills" is the z-score of the apprentice's soft-skills index; "Desires self employment" is a dummy variable if the apprentice answered yes to the prompt. For firm owner characteristics: "Less than grade 8" is a dummy variable that equals 1 if the firm owner has less than eight years of formal schooling; "Ability" is the z-score of the firm owner's ability index; "Total workers" is the baseline firm size (including owner); "Above median sales" is a dummy variable that equals 1 if the firm earned more than median revenues last month; "Apprentices trained" is the total number of past completed apprentices. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. Total earnings includes earnings from wage-employment, self-employment, farming, and apprenticeship. Standard errors, clustered at the firm level, are in parentheses. Earnings are in October 2015 Ghana Cedis. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$

Table A19: Summary statistics of time use (MCP)

	Mean	SD	Min	Max	N
Panel A: Percentage of total work hours					
Instruction	0.06	0.09	0.00	1.00	1060.00
Supervising	0.04	0.09	0.00	0.69	1060.00
Trade association	0.01	0.05	0.00	0.44	1060.00
Learning new skills	0.00	0.03	0.00	0.38	1060.00
Work in shop	0.62	0.23	0.00	1.00	1060.00
Other	0.01	0.04	0.00	0.44	1060.00
Idle time	0.26	0.19	0.00	1.00	1060.00
Panel B: Percentage of idle hours					
Errands outside shop	0.28	0.29	0.00	1.00	1006.00
Break from work	0.57	0.34	0.00	1.00	1006.00
Waiting for customers	0.15	0.29	0.00	1.00	1006.00

Table A20: Summary statistics of time use (Apprentices)

	Mean	SD	Min	Max	N
Panel A: Percentage of total work hours					
Instruction MCP	0.10	0.13	0.00	1.00	882.00
Instruction workers	0.01	0.04	0.00	0.88	882.00
Observation	0.06	0.12	0.00	0.75	882.00
Practice	0.04	0.10	0.00	1.00	882.00
Work in shop	0.57	0.23	0.00	1.00	882.00
Other	0.01	0.05	0.00	1.00	882.00
Idle time	0.21	0.12	0.00	1.00	882.00
Panel B: Percentage of idle hours					
Errands outside shop	0.45	0.30	0.00	1.00	843.00
Break from work	0.45	0.33	0.00	1.00	843.00
Non shop tasks	0.10	0.26	0.00	1.00	843.00

Table A21: Skills - Robustness

	Assessment						Endline survey		
	= 1 if took incentivized test	= 1 if took incentivized test	Practical component (z - score)	Theoretical component (z - score)	Passed practical	Earned certificate	Craft skills quiz (z - score)	Sales skills (z - score)	Craft-Sales skills (z - score)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Full sample (Baseline controls)									
Treatment	-0.01	-0.01	0.13*	-0.08	0.05	0.05	0.15**	0.19***	0.22***
	(0.04)	(0.04)	(0.07)	(0.07)	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)
Observations	797	763	488	488	488	488	743	748	743
Mean of Dep Variable T=0	0.65	0.65	0.00	0.00	0.80	0.80	0.00	0.00	-0.09
Panel B: Full sample (No controls)									
Treatment	-0.01	-0.01	0.13*	-0.06	0.05	0.05	0.16**	0.19**	0.22***
	(0.04)	(0.04)	(0.07)	(0.07)	(0.03)	(0.03)	(0.07)	(0.07)	(0.07)
Observations	797	763	488	488	488	488	743	748	743
Mean of Dep Variable T=0	0.65	0.65	-0.00	-0.00	0.80	0.80	-0.00	-0.00	-0.09
Panel C: Women only (No controls)									
Treatment	0.00	-0.00	0.18**	-0.04	0.07*	0.07*	0.16**	0.20***	0.23***
	(0.04)	(0.04)	(0.07)	(0.07)	(0.04)	(0.04)	(0.07)	(0.08)	(0.08)
Observations	751	717	457	457	457	457	703	706	703
Mean of Dep Variable T=0	0.64	0.64	0.00	0.00	0.81	0.81	0.00	0.00	-0.10

Notes: This Table shows the results from estimating Equation 1 for different skill measures. All specifications include strata fixed effects. Specifications from the assessment include assessor dummies. All standard errors are clustered at the firm level. Columns 2 through 6 exclude test-takers from the first district, because the assessment (being implemented in collaboration with our government partners) differed from the test given in all other districts. The first panel mirrors specifications in Table 1, including controls for the baseline value of the household asset index (and a dummy for where that variable is missing) in the full sample of 797 apprentices (including both men and women). The second and third panels test for robustness to the exclusion of the baseline controls in the full sample (men and women) and the analysis sample (women only). Columns 2-6 exclude pilot district. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

44

Table A22: Earnings - Robustness

		Unconditional		Conditional	
	Total earnings (GHC) (1)	Self-employment profits (GHC) (2)	Wage employment earnings (GHC) (3)	Self-employment profits (GHC) (4)	Wage employment earnings (GHC) (5)
Panel A: Full sample (Baseline controls)					
Treatment	10.43*	8.41*	3.41	15.15	-4.10
	(5.56)	(4.70)	(3.40)	(12.23)	(19.55)
Observations	2992	2992	2992	869	301
Mean of Dep Variable T=0	44.34	24.18	12.98	83.58	136.89
Panel B: Full sample (No controls)					
Treatment	11.82**	9.36**	3.92	14.87	-12.65
	(5.43)	(4.58)	(3.44)	(12.09)	(24.53)
Observations	2992	2992	2992	869	301
Mean of Dep Variable T=0	44.34	24.18	12.98	83.58	136.89
Panel C: Women only (No controls)					
Treatment	14.84***	10.54**	3.42	22.79*	-10.60
	(5.34)	(4.65)	(3.62)	(12.58)	(26.86)
Observations	2824	2824	2824	806	287
Mean of Dep Variable T=0	41.52	22.99	13.10	80.62	135.34

Notes: This Table shows the results from estimating Equation 2 for different labor market outcomes. All standard errors are clustered at the firm level. Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. Total earnings includes earnings from wage-employment, self-employment, farming, and apprenticeship. Profits in self-employment follow self-reported question structures from De Mel et al. (2009). All earnings are in October 2015 Ghana Cedis. The first panel mirrors specifications in Table 2, including controls for the baseline value of the household asset index (and a dummy for where that variable in missing) in the full sample of 797 apprentices (including both men and women). The second and third panels test for robustness to the exclusion of the baseline controls in the full sample (men and women) and the analysis sample (women only). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A23: Firm owner (trainer) effort - Robustness

	Firm level			Apprentice level	
	Firm owner hours last week	Firm owner instruction hours per day	Firm profits last month (GHC)	Firm owner instruction per day	Apprentice wages last month (GHC)
	(1)	(2)	(3)	(4)	(5)
Panel A: Full sample (Baseline controls)					
Treatment	-1.24	-0.09	12.60	0.10	1.55
	(1.85)	(0.21)	(31.11)	(0.09)	(3.08)
Observations	339	341	335	595	566
Mean of Dep Variable T=0	52.33	3.42	251.39	1.29	22.17
Panel B: Full sample (No controls)					
Treatment	-1.24	-0.09	12.60	0.11	1.53
	(1.85)	(0.21)	(31.11)	(0.09)	(3.04)
Observations	339	341	335	595	566
Mean of Dep Variable T=0	52.33	3.42	251.39	1.29	22.17
Panel C: Women only (No controls)					
Treatment	-0.74	-0.04	9.64	0.08	1.94
	(1.76)	(0.22)	(32.72)	(0.09)	(3.20)
Observations	323	325	319	556	528
Mean of Dep Variable T=0	52.46	3.33	237.20	1.32	21.85

Notes: All specifications include strata fixed effects. Specifications from the assessment include assessor dummies. Standard errors in Columns (4) and (5) are clustered at the firm level. Firm profits follow self-reported question structures from De Mel et al. (2009). Wages for each apprentice are reported by the firm owner. Both profits and wages are in October 2015 Ghana Cedis. The first panel mirrors specifications in Table 3, including controls for the baseline value of the household asset index (and a dummy for where that variable in missing) in Columns (4) and (5) in the full sample of 797 apprentices (including both men and women). The second and third panels test for robustness to the exclusion of the baseline controls in the full sample (men and women) and the analysis sample (women only). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A24: Apprentice completion and turnover - Robustness

	Apprentice completion				Apprentice turnover	
	Completed	Paid exit/ceremony	Completed and exited	Completed and retained	Exited	Missed months
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full sample (Baseline Controls)						
Treatment	0.07**	0.10***	0.08**	-0.01	0.08**	1.50*
	(0.04)	(0.03)	(0.04)	(0.01)	(0.04)	(0.84)
Observations	748	748	748	748	748	748
Mean of Dep Variable T=0	0.31	0.15	0.28	0.03	0.39	7.73
Panel B: Full sample (No controls)						
Treatment	0.07*	0.10***	0.08**	-0.01	0.08**	1.47*
	(0.04)	(0.03)	(0.04)	(0.01)	(0.04)	(0.83)
Observations	748	748	748	748	748	748
Mean of Dep Variable T=0	0.31	0.15	0.28	0.03	0.39	7.73
Panel C: Women only (No controls)						
Treatment	0.07*	0.11***	0.07**	-0.00	0.08**	1.60*
	(0.04)	(0.03)	(0.04)	(0.01)	(0.04)	(0.86)
Observations	706	706	706	706	706	706
Mean of Dep Variable T=0	0.32	0.16	0.30	0.02	0.40	7.99

Notes: All specifications include strata fixed effects. All standard errors are clustered at the firm level. The first panel mirrors specifications in Table 4, including controls for the baseline value of the household asset index (and a dummy for where that variable in missing) in the full sample of 797 apprentices (including both men and women). The second and third panels test for robustness to the exclusion of the baseline controls in the full sample (men and women) and the analysis sample (women only). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

47

Table A25: Labor supply - Robustness

| employment	Working - employment	Self- employment	Skilled self - employment	Unskilled self - employment	Wage- employment	Skilled wage- employment	Unskilled wage- employment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full sample (Baseline controls)							
Treatment	0.03	0.01	0.04	-0.03*	0.00	-0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	2992	2992	2992	2992	2992	2992	2992
Mean of Dep Variable T=0	0.62	0.29	0.16	0.13	0.10	0.05	0.05
Panel B: Full sample (No controls)							
Treatment	0.04	0.01	0.04	-0.03	0.00	-0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	2992	2992	2992	2992	2992	2992	2992
Mean of Dep Variable T=0	0.62	0.29	0.16	0.13	0.10	0.05	0.05
Panel B: Women only (No controls)							
Treatment	0.03	0.01	0.04	-0.03	-0.00	-0.00	0.00
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	2824	2824	2824	2824	2824	2824	2824
Mean of Dep Variable T=0	0.62	0.29	0.16	0.13	0.10	0.05	0.05

Notes: This Table shows the results from estimating Equation 2 for different labor market outcomes. All specifications include strata fixed effects and controls for the baseline household asset index (including a dummy for where the variable is missing). Standard errors are clustered at the firm level. Outcomes are stacked across a four-round retrospective panel that asked respondents to report on earnings 1 month ago, 4 months ago, 7 months ago, and 10 months ago. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$

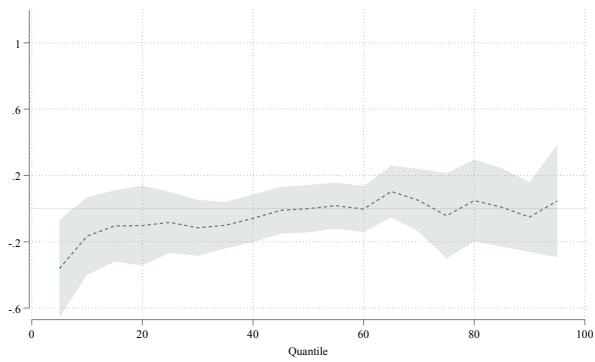# Figure S1: Quantile treatment effects (All)
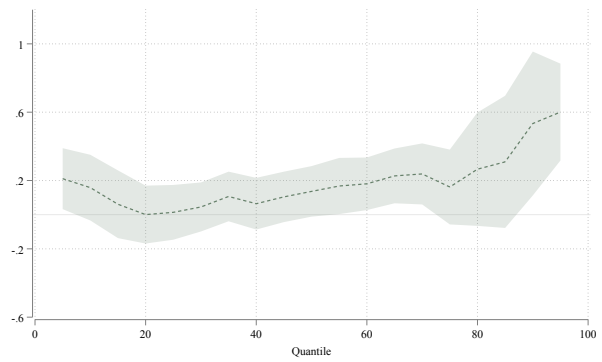
(a) Craft skills (z-score)
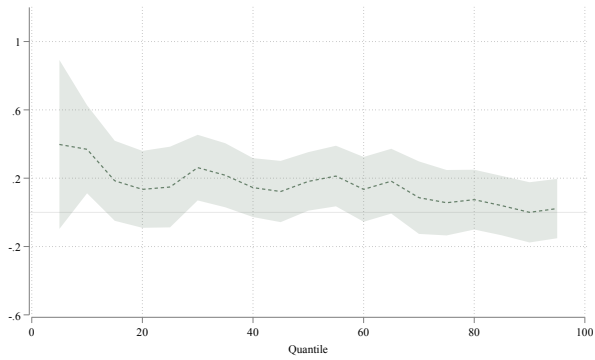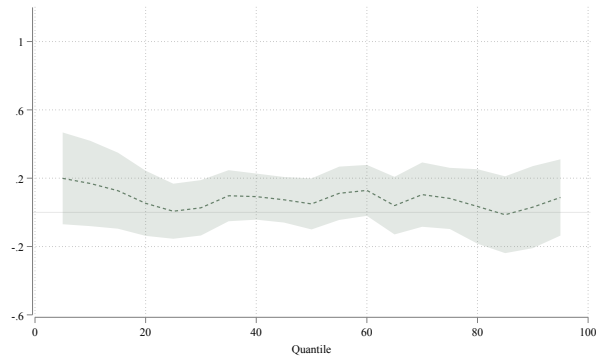
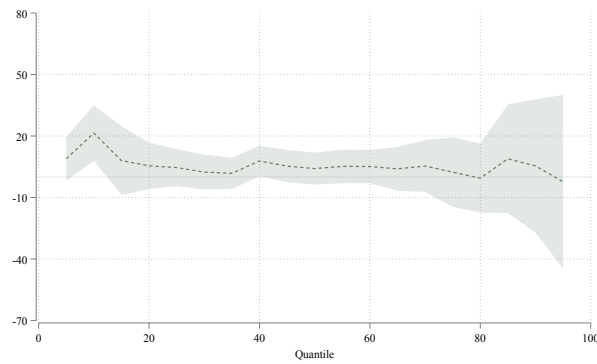(b) Practical component (z-score)

(c) Theoretical component (z-score)
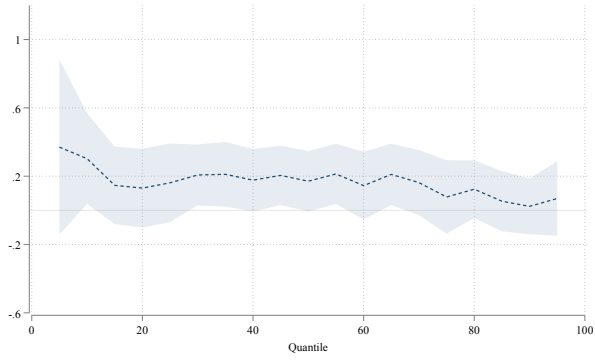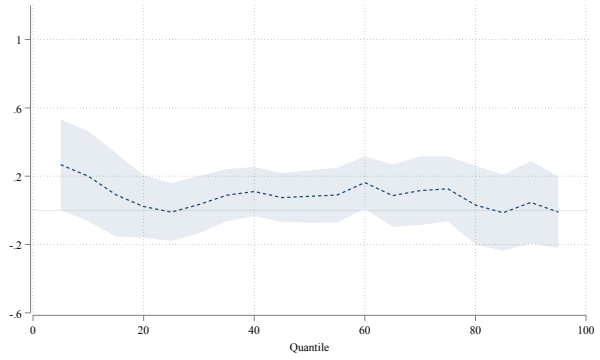
(d) Sale skills (z-score)

(e) Total earnings

Note: This figures plots the coefficients from a quantile regression on several outcomes and their 95% confidence intervals. Standard errors are clustered at the firm level.

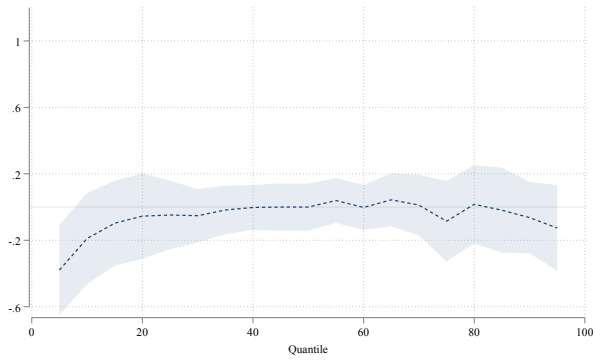Figure S2: Quantile treatment effects (Women only)
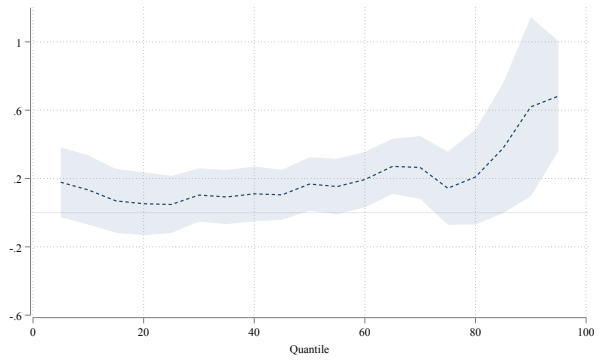
(a) Craft skills (z-score)
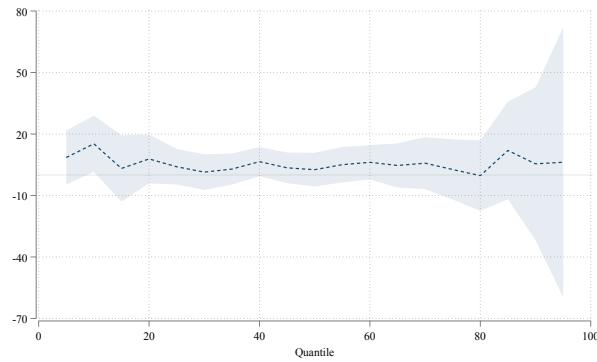
(b) Practical component (z-score)

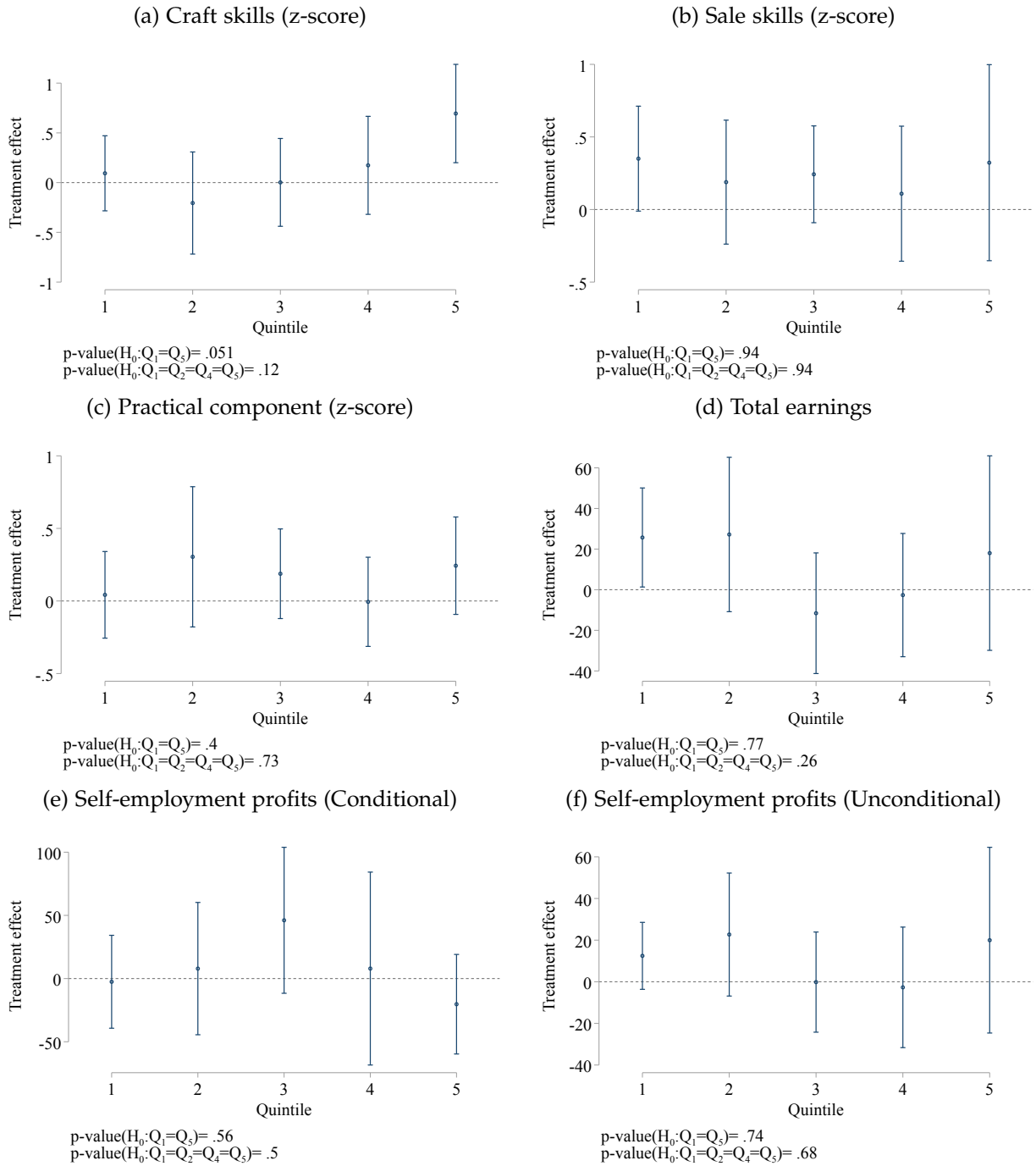(c) Theoretical component (z-score)

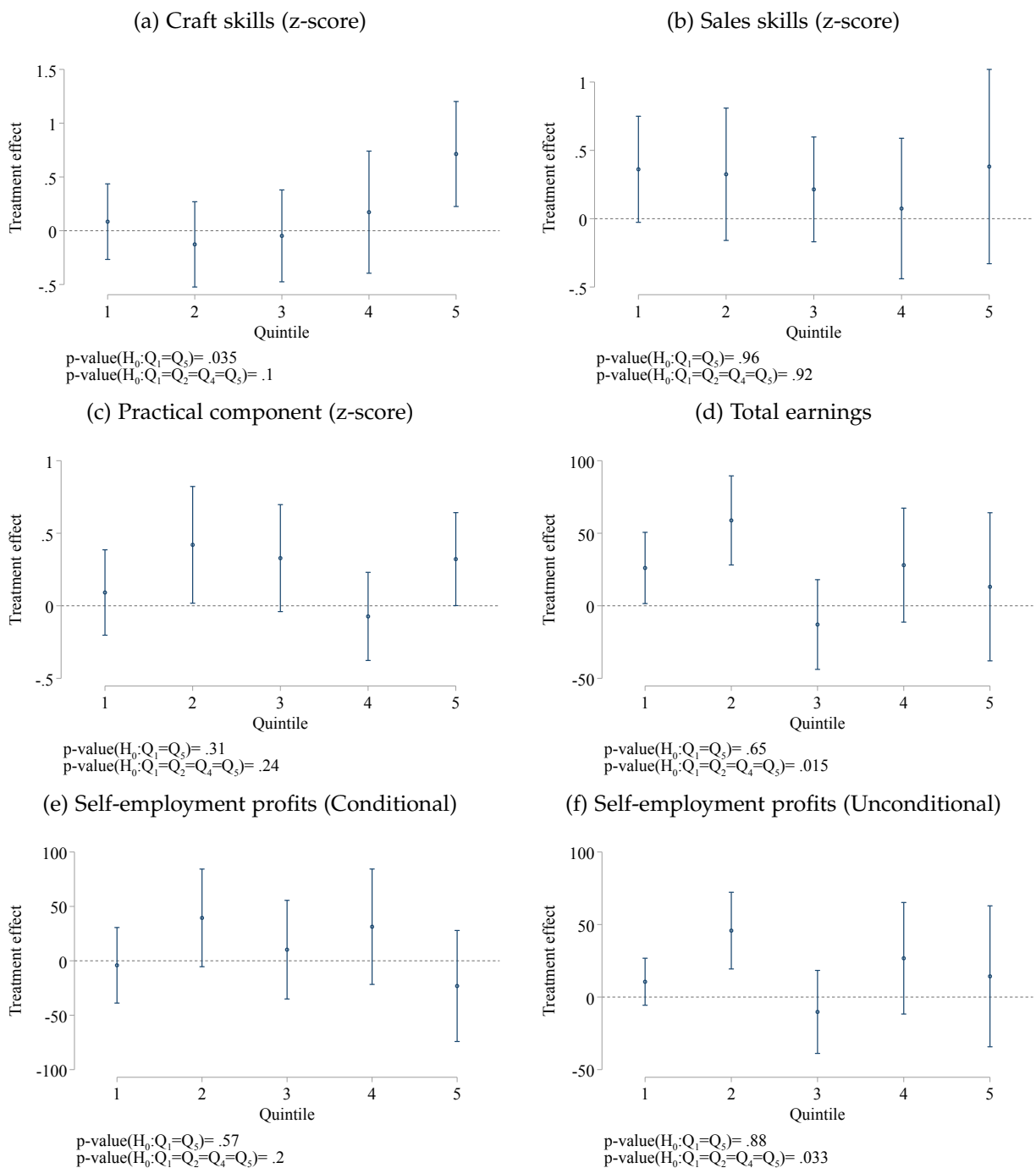(d) Sale skills (z-score)

(e) Total earnings

Note: This figures plots the coefficients from a quantile regression on several outcomes and their 95% confidence intervals. Standard errors are clustered at the firm level.

Figure S3: Outcome heterogeneity by idle time (All)

(a) Craft skills (z-score)

p-value(H_0:Q_1=Q_5)= .051
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .12

(b) Sale skills (z-score)

p-value(H_0:Q_1=Q_5)= .94
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .94

(c) Practical component (z-score)

p-value(H_0:Q_1=Q_5)= .4
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .73

(d) Total earnings

p-value(H_0:Q_1=Q_5)= .77
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .26

(e) Self-employment profits (Conditional)

p-value(H_0:Q_1=Q_5)= .56
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .5

(f) Self-employment profits (Unconditional)

p-value(H_0:Q_1=Q_5)= .74
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .68

Note: This figure shows the treatment effect (and their 95% confidence interval) in a series of outcomes (y-axis) by MCPs' quintile in idle time (x-axis). We also show p-values for testing whether the treatment effect for MCPs' in the first quintile (least idle time) are the same as the effects for those in the last quintile (most idle time), and for testing whether the treatment effect is the same across all five quintiles.

Figure S4: Outcome heterogeneity by idle time (Women)

(a) Craft skills (z-score)

p-value(H_0:Q_1=Q_5)= .035
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .1

(b) Sales skills (z-score)

p-value(H_0:Q_1=Q_5)= .96
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .92

(c) Practical component (z-score)

p-value(H_0:Q_1=Q_5)= .31
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .24

(d) Total earnings

p-value(H_0:Q_1=Q_5)= .65
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .015

(e) Self-employment profits (Conditional)

p-value(H_0:Q_1=Q_5)= .57
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .2

(f) Self-employment profits (Unconditional)

p-value(H_0:Q_1=Q_5)= .88
p-value(H_0:Q_1=Q_2=Q_4=Q_5)= .033

Note: This figure shows the treatment effect (and their 95% confidence interval) in a series of outcomes (y-axis) by MCPs' quintile in idle time (x-axis). We also show p-values for testing whether the treatment effect for MCPs' in the first quintile (least idle time) are the same as the effects for those in the last quintile (most idle time), and for testing whether the treatment effect is the same across all five quintiles.

B Calculating the Cost of Apprentice Exit and the Marginal Benefit of Effort

To attempt to quantify the costs of apprentice exit to the firm, we estimate the number of months between January 2015 and the endline survey (mostly in late 2017) that firm owners missed out on due to accelerated apprentice exit, based on the apprentice's self-reported month and year of exit from the firm (if any) and the timing of the endline survey, assuming zero missed months for apprentices still at the firm. In Table 4 Column (6), we estimate apprentices exit 1.5 months earlier in the treatment group than the control group. Hardy and McCasland (forthcoming) estimate that firm profits increase by 11% per apprentice, in response to access to apprentices in this sample (suggesting that in this labor market firms capture a large share of apprentice productivity). Applying 11% to the profit measures in the firm midline survey in Table 3 (251 GhC per month in the control group), we estimate a loss due to exit of 27.6 GhC per month. Multiplied by 1.5 months, our back-of-the-envelope calculation suggests firm owners lost a little over 40 GhC in the medium-term due to higher apprentice turnover rates.

The performance incentive was designed to yield the same average payout in the treatment and control groups. To estimate gains associated with higher quality training, we regress the payout on the normalized practical test score in the subsample of apprentices from treated firms who participated in the assessment, controlling for strata fixed effects and clustering at the firm level as we do throughout the analysis. We estimate that a 1 standard deviation increase in the practical score is associated with earning 55 GhC more in the performance-based financial incentive among treatment firms. Applying this to the 0.13 standard deviation impact estimate from Table 1 yields a benefit estimate of about 7 GhC.