Policy Research Working Paper 8851

The Apprenticeship-to-Work Transition

Experimental Evidence from Ghana

Morgan Hardy
Isaac Mbiti
Jamie McCasland
Isabelle Salcher

WORLD BANK GROUP
Social Protection and Jobs Global Practice
May 2019

Policy Research Working Paper 8851

Abstract

This paper examines the effects of a government-sponsored apprenticeship training program designed to address high levels of youth unemployment in Ghana. The study exploits randomized access to the program to examine the short-run effects of apprenticeship training on labor market outcomes. The results show that apprenticeships shift youth out of wage work and into self-employment. However, the loss of wage income is not offset by increases in self-employment profits in the short run. In addition, the study uses the randomized match between apprentices and training providers to examine the causal effect of characteristics of trainers on outcomes for apprentices. Participants who trained with the most experienced trainers or the most profitable ones had higher earnings. These increases more than offset the program's negative treatment effect on earnings. This suggests that training programs can be made more effective through better recruitment of trainers.

This paper is a product of the Social Protection and Jobs Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at morgan.hardy@nyu.edu; imbiti@virginia.edu; jamie.mccasland@ubc.ca; and isabelle.salcher@nyu.edu.

The Apprenticeship-to-Work Transition: Experimental Evidence from Ghana*

Morgan Hardy† Isaac Mbiti‡ Jamie McCasland§ Isabelle Salcher¶

JEL Classification: C93, I21, I25, I26, J24, J46, O12, O15, O17
Keywords: Apprenticeship, Youth Labor Markets, Skills, Ghana

*This project was conducted in cooperation with the Council for Technical and Vocational Education and Training in Ghana. We thank the Executive Director, the management team, and staff for their support and input. We are grateful to Emily Beam, Kathleen Beegle, Alex Eble, Andrew Foster, Louise Fox, Jessica Goldberg, Maddalena Honorati, Dhushyanth Raju, Jay Shimshack, Tavneet Suri and seminar participants at Boston University, the Washington Area Development Economics Symposium, University of Virginia, Brown University, University of Vermont, University of Houston, Penn State University, University of Washington, and New York University, for helpful comments and suggestions. The project would not have been possible without the efforts of numerous individuals including Yani Tyskerud, Charles Sefenu, Lois Aryee, Robert Obenya, Anthony Mansaray, Dina Grossman, Taylor Salisbury, Stephanie Adjovu, Renaud Comba, Eric Protzer, Ishita Ahmed, and Kym Cole. We are grateful for the hard work of the field teams at IPA-Ghana. We are especially grateful to Dr. Daniel Baffour-Awuah for his input and guidance in the early phases of this project. We also acknowledge Kim Lehrer's efforts and intellectual inputs at the start of the project. This research was supported by funding from 3ie, the Center for Economic Policy Research, the International Growth Center, USAID, the Abdul Latif Jameel Poverty Action Lab, the University of Virginia, the University of British Columbia, New York University Abu Dhabi, and the World Bank Jobs Umbrella Trust Fund which is supported by the Department for International Development/UK AID, and the Governments of Norway, Germany, Austria, the Austrian Development Agency, and the Swedish International Development Cooperation Agency. This paper was partly written while Mbiti was visiting the Department of Economics and the Population Studies and Training Center at Brown University. The research protocols were reviewed and approved by the Institutional Review Boards of Innovations for Poverty Action, Southern Methodist University, NYU-Abu Dhabi, the University of British Columbia, and the University of Virginia. The project is registered under the AEA RCT Registry (AEARCTR-0001948). The views expressed in this paper are those of the authors and do not necessarily reflect the the position of any agency or funder.

†NYU- Abu Dhabi; morgan.hardy@nyu.edu
‡University of Virginia, J-PAL, IZA; imbiti@virginia.edu
§University of British Columbia; jamie.mccasland@ubc.ca
¶New York University; isabelle.salcher@nyu.edu

1 Introduction

Youth unemployment and underemployment are pressing policy challenges, particularly in developing regions such as sub-Saharan Africa. For many African youth the transition from school to productive work is slow and often unsuccessful, resulting in youth unemployment rates that are generally twice those of adults (Filmer & Fox, 2014; International Labor Organization Department of Statistics, 2018). Moreover, their employment prospects are generally limited to low-productivity sectors and jobs (Honorati & Johansson de Silva, 2016). The low employability and productivity of youth are often attributed to their inability to obtain marketable (or appropriate) skills (Filmer & Fox, 2014). This is in part because a large number of African youth are locked out of the mainstream education system by school capacity constraints, poor academic performance, or financial constraints.[1]

Job training programs have the potential to provide skills to young people, especially those who are locked out of the mainstream education system. Yet traditional approaches such as the provision of training through public vocational institutions are often criticized for their inability to provide market-ready skills in a cost-effective manner (Johanson & Adams, 2004; Blattman & Ralston, 2015).[2] In contrast, apprenticeship training programs are considered to be promising avenues to deliver skills training to youth, although there is limited empirical evidence on their effectiveness. Apprenticeships are common in West African countries such as Ghana, where they are responsible for training almost four times as many individuals as all other (formal) alternatives combined (Darvas & Palmer, 2014; Filmer & Fox, 2014). By providing on-the-job training, apprenticeships could overcome both the skills mismatch and the lack of relevant employment experience that impede youth in the labor market. Since the training is typically conducted in the informal sector, they are also potentially better placed to prepare youth to transition into that sector, which accounts for about the majority of employment opportunities in many African countries (Filmer & Fox, 2014).[3]. However, there

[1] For example, data from the nationally representative 2014 Ghana Demographic and Health Survey (DHS) show that almost 70 percent of primary-age students were attending school, while only 38 percent of secondary-age students were attending secondary school (Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF International, 2014). The financial constraints in many African countries, including Ghana, will be less binding in the near future due to the introduction of free secondary schooling.

[2] Previous program evaluations have focused extensively on formal vocational training programs, with limited research on apprenticeship programs. These studies have shown mixed results on the effectiveness of training programs in increasing youth employment. In his literature review, McKenzie (2017) finds that the estimated treatment effects from training programs range from no change to an eight percent increase.

[3] For example, 85 percent of jobs in Ghana are in the informal sector (Honorati & Johansson de Silva, 2016).

are concerns that apprenticeships may not actually improve labor market outcomes for youth because of their reliance on traditional (often outdated) technology and the lack of standards and quality assurance (Darvas & Palmer, 2014).

In this paper we conducted a randomized field experiment to examine how apprenticeships for youth affect their labor market participation, earnings, and other life outcomes. We partnered with the Ghanaian government to evaluate the National Apprenticeship Programme (NAP), which placed youth applicants into apprenticeships with small informal sector firms (microenterprises). Because apprenticeships typically require the up-front payment of a training fee, many youth may be locked out of training opportunities by credit constraints (Darvas & Palmer, 2014; Frazer, 2006). The program eliminated these fee barriers and offered youth the opportunity to train in one of five trades: garment making, cosmetology, carpentry, welding, and masonry.[4]

The evaluation was conducted in 32 districts across all regions of Ghana starting in 2012. It featured two, sequential randomizations: the first selected the participants from the pool of applicants, and the second matched the selected participants with training providers (firms). To participate in the program, youth were required to complete an application form indicating their trade of choice and attend an in-person interview with a local (district) selection committee to ascertain their suitability for the program. Each selection committee would either reject applicants outright or categorize them as eligible for the program. Among the eligible applicants, the committees could guarantee access to the program for up to 16 percent by designating them as "priority applicants." The remaining (nonpriority but eligible) applicants were then randomly assigned to the treatment or control group. This randomization was conducted in late 2012. It was stratified by district and trade and designed to fill all available program slots.

Both the selected applicants and potential training providers were invited to a series of matching meetings in mid- to late 2013. Separate meetings were organized for each trade within each of the 32 evaluation districts. In these meetings trainers would make a brief presentation outlining their location, experience, and other characteristics of their business. Choosing among trainers within walking distance, the treatment group participants would then list those they were interested in training with (that is, a preference set). They were then assigned to train with a provider randomly selected from their preference set. Training began shortly after these meetings.

We examine the impact of the NAP training on youth labor market outcomes using the baseline applicant data collected during the application process in 2012, the baseline

[4] In our analysis we pool carpentry, welding, and masonry into one group, which we refer to as construction.

training provider data collected during the matching meetings in 2013, and the endline applicant data collected from mid-2017 through May 2018. On average the endline data were collected about four years after the start of training. Since apprenticeships typically last three years, the endline data would capture short-run returns to training. The endline data focus primarily on labor market outcomes, but they also measure training history and include a trade-specific skills test designed in collaboration with local Ghanaian industry experts. Focusing only on the set of applicants that were randomized into the treatment and control groups, we find that the randomization was balanced across both groups. Typically, evaluations of training programs in developing countries are plagued by high attrition (McKenzie, 2017). In this study, sample attrition was relatively low (10 percent) and balanced across the treatment and control groups. This provides some assurance about the validity of our experiment. Our data show that the training offered under the NAP was essentially the same as that in the traditional apprenticeships available in the market. Given the geographic scope of our evaluation, our results provide insights into the effectiveness of the traditional apprenticeship system in Ghana.

We report three main findings. First, we find that access to the program led to modest increases in the probability of starting an apprenticeship, the probability of completing training, and the duration of training. Youth offered training under the NAP were 13 percentage points more likely to commence training, and 10 percentage points more likely to complete training, than the control group. Because the training completion rate among the control group was relatively low (25 percent) compared with its starting rate (63 percent), this suggests that the program was relatively more effective at encouraging the completion of training than it was the start of training. In addition, the treatment group completed four more months of training than the control group. The data suggest that female participants in cosmetology were the most responsive to the program offer, while male participants in construction were the least responsive (in terms of completion).

Second, with respect to labor market outcomes, we find that access to the program shifted participants out of wage work and into self-employment. Since wage work is generally more lucrative than self-employment, this resulted in a reduction of 12 Ghanaian cedis (GHS), or about 13 percent, in average total monthly earnings compared with the control group.[5] This reduction in monthly earnings was most pronounced among male participants in construction (47 GHS). But it was almost negligible among female participants in cosmetology (2 GHS), since reductions in wage earnings among this group

[5] The exchange rate was roughly 1 USD = 4.5 GHS.

were almost fully offset by increases in self-employment earnings.

Finally, using the randomized match between training providers and apprentices, we show that trainer characteristics can have a causal effect on an apprentice's labor market outcomes. In particular, we find that apprentices who trained with the most profitable trainers had greater total monthly earnings compared to their peers who trained with less profitable trainers. We also find similar patterns when apprentices trained with the most experienced trainers, or trainers with the largest wage bills (a proxy for the size and skill of their wage workforce). The magnitude of this increase in earnings more than offset the negative treatment effects of training on earnings for the full sample, suggesting that the effectiveness of training programs can be improved by carefully screening training providers.[6]

We make two main contributions to the literature. First, we add to the very limited set of experimental studies of apprenticeship training in developing countries. There is a larger literature that evaluates vocational training programs as well as business training programs in these contexts. Overall, these studies find that training programs are generally ineffective (Blattman & Ralston, 2015; McKenzie & Woodruff, 2013). We are aware of only two recent randomized control trials on apprenticeships in Africa. Cho, Kalomba, Mobarak, and Orozco (2013) evaluate a three-month apprenticeship program in Malawi and find no improvements in labor market outcomes. The study was plagued by high attrition rates, however. Crépon and Premand (2018) examine a formalized (or improved) apprenticeship training program in Côte d'Ivoire that included an in-classroom training component, a formal certification scheme, and a training wage (or subsidy) for apprentices. They also find limited improvements in labor market outcomes. Our study provides evidence on an existing traditional system of apprenticeships rather than on an improved or formalized system like those examined in the studies in Côte d'Ivoire and Malawi. Further, since the NAP does not offer training wages, it is arguably more cost-effective for the government.

Second, our study is among the very few that use two-sided randomization designs. We are aware of two studies that employ such a design: Crépon and Premand (2018) and Alfonsi et al. (2017). Crépon and Premand (2018) use their design to test whether receiving an apprentice through the randomization affected a firm's subsequent apprenticeship hiring decisions. They find evidence of displacement within firms, such that firms that received apprentices through the program hired fewer apprentices through

[6] This is consistent with evidence from a recent randomized control trial on training in which Alfonsi et al. (2017) argue that their training program was effective in part because they worked with the best trainers in the Ugandan capital city of Kampala.

regular channels. In contrast, Hardy and McCasland (2016) use the data from the NAP experiment in Ghana and find no displacement among firms in our sample. Although they do not examine a typical apprenticeship program, Alfonsi et al. (2017) compare the effectiveness of an on-the-job training program with that of a formal vocational training program in Kampala, Uganda. They find that both forms of training improve labor market outcomes, but individuals assigned to formal vocational training had greater earnings growth because of their acquisition of "transferable skills" rather than "firm-specific skills," highlighting the importance of skills acquisition. Alfonsi et al. (2017) also find evidence that firms are labor constrained. This is consistent with the findings of Hardy and McCasland (2016) for our study context. Rather than focus on firm outcomes, we use the randomized match between trainers and apprentices to identify the characteristics of training providers (firms) that deliver the best outcomes for youth. This novel exercise provides us with a better understanding of the production function for apprenticeship training. To our knowledge, our study is the first to explore this dimension in the context of training programs. This exercise is important, as it can help governments improve the effectiveness of training by introducing a more selective recruitment process for trainers.

Our findings are more pessimistic than those of studies of apprenticeships in European contexts (for example, Acemoglu and Pischke (1998); Fersterer, Pischke, and Winter-Ebmer (2008). These studies have generally found positive effects of the apprenticeship training on individual participants and have also highlighted the potential for firms to benefit from providing such training. Observational studies of apprenticeship training in Ghana have argued that returns to training are greatest for individuals with low levels of education and are not significantly different from zero for those with more than six years of schooling (Monk, Sandefur, & Teal, 2008; Teal, 2016). Our results are consistent with this notion, since the average duration of schooling in our sample exceeded seven years. Further, Frazer (2006) argues that apprenticeships essentially teach apprentices to replicate their trainer's business. This is consistent with our findings showing that the treatment shifts individuals out of wage work and into self-employment. Although these shifts are associated with lower earnings, there could be offsetting nonmonetary benefits from self-employment. This possibility is discussed extensively in the U.S. literature on entrepreneurship (Hamilton, 2000; Hyytinen, Ilmakunnas, & Toivanen, 2013). Given the lack of wage jobs in the Ghanaian context, many individuals likely enter self-employment involuntarily. This could potentially reduce the nonmonetary benefits from self-employment.

2 Context and Program Design

2.1 Context

Youth in Ghana, as elsewhere, often face unique challenges transitioning into the labor market. The overwhelming majority of jobs (almost 90 percent) are in the informal sector (Honorati & Johansson de Silva, 2016). Most of these jobs are low productivity and require limited use of cognitive skills (Honorati & Johansson de Silva, 2016). Thus, obtaining a job with growth and earnings potential is especially challenging. Recent data from Ghana show that youth ages 15-24 are much less likely than adults ages 25-65 to be working: while just over half of young people (52 percent) are working, the majority of other adults (89 percent) are (Ghana Statistical Service, 2014). The lower labor force attachment among youth reflects in part the fact that many young people are still in school. But policy makers and researchers are increasingly concerned by the growing share of young people who are neither in school nor at work.

There are important gender differences in the data. Among those aged 15-24 the female unemployment rate, at 24 percent, is 50 percent higher than the male unemployment rate (World Bank, 2019). In the age group 15-35 the female unemployment rate is 80 percent higher than the male rate (World Bank, 2019). Women in this age group are also less likely than men to work for a wage: while 20 percent of men ages 15-35 are employed in the wage sector, only 11 percent of women in this age group are. Because work in the wage sector is more lucrative than that in other sectors, this gap is an important driver of the lower earnings of women relative to men. These employment challenges stem in part from human capital constraints. Filmer and Fox (2014) argue that human capital is a key facilitator for youth in their efforts toward obtaining productive work. Although Ghana has made significant progress in improving access to education, less than a third of young people ages 15-24 have any senior secondary schooling (29 percent), only slightly higher than the rate for ages 25-34 (25 percent) though much higher than that for the oldest cohort (13 percent for ages 35-65). Skills are rarely reported as the most important obstacle for businesses in Ghana; nonetheless, they are cited as a major obstacle by nearly 20 percent of firms (Honorati & Johansson de Silva, 2016).

The skills deficit in Ghana is also driven in part by the education system, where large numbers of students fail to progress beyond critical junctures such as the end of junior high school. Compulsory education in Ghana consists of six years of primary school and three years of junior high school. Upon completing junior high school, young people can choose to continue their studies by attending a senior high school, a secondary technical school, or a technical institute (Gondwe & Walenkamp, 2011). Access to these institutions is based on performance on the Basic Education Certificate Examination (BECE), taken at the end of junior high school. While the government has made some efforts to increase the number of senior high schools in the country, there are still too few places relative to the number of applicants, and the quality of schooling varies substantially (Ajayi, 2013). The shortage of places is reflected in the gap between primary and secondary completion rates. Recent data show that almost 60 percent of 15- to 24-year-olds had completed primary school, while only a quarter had completed secondary school (World Bank, 2019). Limited capacity at government senior high schools, combined with costly fees in informal training, prevents many young people from furthering their education and improving their skills.[7]

2.2 The National Apprenticeship Programme

The NAP was conceived by the Council for Technical and Vocational Education and Training (COTVET) as a potential policy solution to address the growing numbers of youth unable to progress to secondary school. Because youth who are unable to complete secondary school are typically confined to low-productivity (and low-paying) jobs, the program was designed to provide them with an alternative avenue to acquire skills. By eliminating fee barriers, COTVET hoped the program would enable youth across Ghana to access training opportunities.

The program was implemented in 78 districts across all 10 regions of Ghana to ensure national equity.[8] Because the Northern region of Ghana is disadvantaged and marginalized relative to the southern part of the country, the NAP purposely provided relatively more opportunities for youth in the north. The program was intended to target youth between the ages of 15 and 30. But its decentralized implementation made it difficult for COTVET to enforce these age limits. The NAP offered youth fee-free access to apprenticeship training in masonry, welding, carpentry, garment making, and cosmetology (hairdressing and beauty). The choice of these five trades was determined by COTVET. To our knowledge, this choice was not in response to analysis or predictions of market demand but instead reflected the presence of strong and active trade associations in these fields. Further, COTVET tried to be sensitive to gender equity concerns by including a mix of both female-dominated trades (garment making and cosmetology) and male-dominated ones (masonry, welding, and carpentry). This program design ad-

[7] The recent introduction of free secondary school will likely alleviate the financial barriers. However, academic qualification barriers may still bind for many.

[8] Over the study period, there were 10 administrative regions in Ghana. This was expanded to 16 in 2019. We use the pre-2019 administrative boundaries and definitions in this paper.

hered to the very strong patterns of occupational segregation in Ghana. Data from a nationally representative survey show that female apprenticeships were limited almost exclusively to garment making and cosmetology. In contrast, just over a third of male apprenticeships were in construction trades (Ghana Statistical Service, 2014).

Participants were matched with a training provider, often referred to as a master craftsperson (MCP). Participants would work in the MCP's firm and obtain skills through learning by doing in an unstructured environment similar to a traditional apprenticeship program. The NAP training period was supposed to last one year, but in practice trainers generally kept their apprentices for 18 months to almost 4 years, depending on the district and trade. The length of training was ultimately decided by each trainer. Because most trainers considered one year to be too short, they pushed back on COTVET's suggested duration. Since the program was decentralized, COTVET could not enforce the one-year training term. The program was meant to pay trainers 150 GHS to train an apprentice, an amount equivalent to the traditional apprenticeship entrance fee. As a result of the government's fiscal crisis, however, COTVET was unable to pay this fee. But the research team was able to secure donor funds to pay 100 GHS to each participating trainer. The program was also supposed to provide participants with a tool kit relevant to their trade (for example, a sewing machine for garment makers). But most tool kits were never delivered. While the program provided no subsidy to apprentices, firm owners typically paid apprentices small wages or "chop money" (about 20 GHS a month in our midline surveys of firm owners), which increased with seniority and varied with firm productivity or revenues. Thus the program essentially functioned as a subsidized version of a traditional apprenticeship with training timelines of around three years and limited government monitoring or additional support.

3 Research Design

3.1 Participant Recruitment and Randomization Procedure

We use a randomized control trial to rigorously evaluate the effectiveness of the NAP in Ghana. Starting with the entire list of 78 program districts, we chose a set of 32 districts for the evaluation using population-weighted random sampling, stratifying by north and south. This resulted in a representative set of evaluation districts across all 10 regions of the country. Starting in July 2012, COTVET announced the NAP on radio stations. Program materials and fliers were distributed by local officials, including the Technical and Vocational Education Training (TVET) coordinators of the Ghana Education Service (GES). In our evaluation districts these dissemination efforts were supplemented by teams of enumerators who worked with community leaders and religious institutions to bolster the outreach efforts.

To apply to the program, applicants submitted a formal application to the district office and attended an interview with a panel of district officials. The interview panel assessed all applicants and determined whether they were eligible for the program. Due to political considerations, district officials were given the opportunity to "handpick" applicants to fill about 16 percent of the slots. The remaining eligible applicants were then placed in the random lottery. The randomization was stratified by choice of training and district and was conducted electronically but announced locally in conjunction with district officials. In our sample districts 3,928 youth applied to the program. The selection committees selected a total of 329 youth as priority applicants. The randomization assigned 2,031 applicants to the treatment group and 1,568 to the control group. The treatment group was larger than the control group because the selection process had to ensure that that there were no unused training slots. This selection process was completed by the end of 2012.

The 2012 elections resulted in a change in the political regime, and this delayed the implementation of the program. The program was finally launched in late 2013. Treatment group applicants were informed by phone and were invited to a series of "matching meetings," where prospective training providers introduced themselves and their firms. Potential trainers described the location of their shops, their experience in training apprentices, a summary of their firm, and any trade specializations. Potential apprentices then completed a preference sheet, identifying the set of trainers they could feasibly train with based on distance. For clarity we suggested that potential apprentices focus on listing "trainers within walking distance."[9] Apprentices were then randomly assigned to a provider in their feasible set. Training began in late 2013 and lasted between two and three years.

3.2 Data and Balance

We collected baseline data from applicants while they waited to be interviewed by the district selection committees. The baseline apprentice survey covered a broad range of information, including educational attainment, family background, cognitive and noncognitive assessments, and labor market outcomes. Baseline data on trainers were

[9]Apprentices could also list trainers who were further away as long as they had reliable means of transportation.

collected in 2013 during the matching process. These data include information on the trainer's background; the number of workers and apprentices employed; assets, sales, and profits; and management practices. We also conducted cognitive and noncognitive assessments with the trainers.

The endline survey of apprentices was launched in August 2017 and continued through May 2018. The endline apprentice survey covered topics similar to those in the baseline but included more details about labor market outcomes. For example, for self-employed workers the survey captured firm management practices. The survey also used survey questions comparable to those in other impact evaluations conducted on youth labor markets (e.g. Hicks, Kremer, Mbiti, and Miguel (2013)), as well as those in large-scale labor market surveys in Ghana such as the World Bank STEP survey and the Ghana Living Standard Survey.[10]

The timeline of program and evaluation activities is summarized in Figure 1. In our analysis we use data primarily from the endline survey collected in 2017-18, complemented by baseline measures for heterogeneity and balance analysis. Note that apprentice placement occurred between October 2013 and January 2014, between 42 and 52 months before the endline survey data collection.

The baseline characteristics of program applicants as well as the estimated differences between the treatment and control groups are reported in Table 1. On average applicants were 23 years old at the baseline and had completed just over seven years of schooling. The education levels of both mothers and fathers were lower than the schooling of our primary respondents, and mothers had almost 2.5 years less education than fathers. Among measures of labor market attachment, a quarter of the sample had ever started an apprenticeship and just over 40 percent were working. Only 5 percent of the sample worked for a wage, and just under 20 percent were self-employed. Applicants were working about nine hours a week and earning 15 GHS a month from all sources. Garment making and cosmetology were the two most popular trades, which is unsurprising given the gender composition of our sample. Of the applicants, 44 percent expressed interest in an apprenticeship in garment making, 35 percent were interested in cosmetology, and the rest were interested in construction trades (welding, masonry, and carpentry).

In order to provide evidence for the internal validity of our randomization, we test whether our treatment and control groups are similar on observable characteristics measured at baseline. For a given characteristic we run an ordinary least squares (OLS)

[10] Additional details of the World Bank STEP survey can be found at http://microdata.worldbank.org/index.php/catalog/step/about.

regression with assignment to treatment as the independent variable to test whether the difference in means between treatment and control groups is statistically significant. In these regressions we also include district $\times$ trade fixed effects, the stratification unit of our randomization. In addition, we perform an F-test to test whether characteristics of individuals assigned to treatment are jointly different from characteristics of the control group. The results in Table 1 show that baseline characteristics are indeed balanced. We reject the null hypothesis of equal means between the treatment and control groups in only 2 of 20 cases, which is consistent with random selection. Mothers of individuals assigned to treatment tend to have a third fewer years of schooling, while individuals assigned to treatment on average tend to have scored higher on the vocabulary test. Each of these differences is small, however. Moreover, our joint F-test shows that we cannot reject the null hypothesis that characteristics of the treatment and control groups are the same. Thus we take these results as evidence for the internal validity of our randomization. Appendix Tables A.1-A.3 repeat this exercise for each trade offered by the NAP. For clarity we group the construction trades together, and because there were only a handful of male participants training in garment making, we exclude these male participants from the analysis. This restriction allows us to better interpret the results because each trade grouping is homogeneous with respect to gender (i.e. male participants in construction, female participants in garments, female participants in cosmetology). Overall, within each trade group we find that baseline characteristics are balanced between the treatment and control groups.

Table 2 reports the survey follow-up rates between baseline (2012) and endline (2017-18) for the analysis sample (excluding the priority applicants). More than 90 percent of participants in the baseline were surveyed at endline, with no differences in the follow-up rate between the treatment and control groups (Column 1). This is a much lower attrition rate than those in previous studies on training in developing countries (McKenzie, 2017). We also test whether the survey follow-up rates were balanced between the treatment and control groups within each trade group (Columns 2-4). Within each trade group our survey teams achieved follow-up rates of more than 90 percent, with no statistically significant differences between the treatment and control groups except in garment making. The relatively low levels of survey attrition provide reassurance that our results are not confounded by imbalances in the survey follow-up rates. There is a small and marginally statistically significant imbalance among the female garment-making trade group. We were about 2.5 percentage points less likely to follow up with the treatment group in this subsample, which corresponds to about 17 individuals. Given the slight imbalance in attrition in the garment-making trade group, we also compute Lee bounds to examine whether our results are robust to attrition.[11]

3.3 Empirical Specifications

3.3.1 Intent-to-Treat Analysis

We estimate the effect of NAP a variety of outcomes using the following OLS equation:

$$Y_i = \delta_0 + \delta_1 Treat_i + \delta_2 X_i + \gamma_d \times \gamma_j + \gamma_t + \varepsilon_i. \quad (1)$$

In this specification Y_i is our set of outcomes, including labor market outcomes and other ancillary measures of well-being. $Treat_i$ is a binary variable for the assignment to the NAP treatment. X_i is a set of baseline controls. To maintain a parsimonious specification, we control only for mother's schooling and vocabulary scores, in order to correct for the small imbalances in these characteristics. We also include our stratification variables, $\gamma_d \times \gamma_j$. These are a set of district (d) $\times$ trade-group (j) fixed effects. γ_t denotes survey month fixed effects that are included to capture any temporal differences. ε_i is an error term. Because we randomize at the individual level, we use robust standard errors. Our coefficient of interest is δ_1, which captures the average differences in outcomes between those assigned to the treatment group and those assigned to the control group (or the intent-to-treat estimator). We focus on the intent-to-treat estimate because it is well identified by the experiment and is policy relevant.[12]

We estimate Equation 1 on the full sample and also present results for each of our three trade-groups. To ameliorate concerns about multiple hypothesis testing we include adjusted p-values following Westfall and Young (1993).

3.3.2 Heterogeneity by Trainer Characteristics

In addition to randomizing access to the NAP training, our research design (partially) randomized the match between apprentices and training providers or master craftspeople (MCPs). This two-sided randomization allows us to credibly estimate the effect of a trainer's characteristics on an apprentice's labor market outcomes. During the matching meetings apprentices were asked to indicate which MCPs they would like to train with.

[11] The full set of Lee bounds for garment making is available on request.

[12] Our research design also included an incentive design for MCPs who were training NAP apprentices. This treatment rewarded MCPs with a bonus based on the performance of their apprentices on a skills test. We include controls for this treatment as a test of robustness. The findings presented in this paper are robust to the inclusion of this control.

They were instructed to list only those they could feasibly train with, considering distance and travel costs. We also encouraged apprentices to list as many MCPs as possible. Apprentices were randomly assigned to train with one of the MCPs on their list. Because there were close to 1,000 MCPs in our study and apprentices listed as many as 10 MCPs each, data limitations mean that we cannot include fixed effects for all combinations of MCPs. In addition, each MCP has different attributes that we need to incorporate into the analysis. We focus on four MCP attributes that would arguably influence the quality of training: math test scores (a proxy for education), previous training experience, profitability of the business, and their wage bill (a proxy for the scale of the business). We rank the MCPs within their district and trade on each attribute and estimate the following regression:

$$Y_i = \beta_0 + \beta_1 TopRanked_{ik} + X_{ik}\Gamma + \gamma_d \times \gamma_j + \varepsilon_i. \quad (2)$$

In this specification Y_i denotes an apprentice's outcomes. $TopRanked_{ik}$ is a binary variable equal to one if the assigned MCP is ranked either first or second in the corresponding district and trade for a particular attribute k (e.g. math scores). Because we are focusing on the rank order of MCPs by attributes, we can include the full set of fixed effects for each apprentice's choices. These fixed effects will be in terms of the rank order of the MCP's attribute (k) in each district (d) and trade (j). All these fixed effects are captured in X_{ik}. Since MCP rankings will change depending on the attribute, the set of fixed effects will also change to reflect the differences in the rank ordering.[13] We also include district $\times$ trade fixed effects to account for the distinct district-by-trade matching meetings that were held. Our analysis includes only individuals who listed two or more MCPs in the meetings. This empirical strategy mirrors those commonly used in the education literature examining the effects of gaining admission to an elite (or magnet) school.[14]

[13] For the cases where there are more than two MCPs with the top-ranked characteristic, we simply redefine the treatment variable to include the entire group of top-ranked MCPs.

[14] Examples of papers that examine the effects of going to a better-ranked school include Abdulkadiroğlu, Angrist, and Pathak (2014); Jackson (2010); Lucas and Mbiti (2014); Pop-Eleches and Urquiola (2013). These studies condition on student choice sets and then use a regression-discontinuity framework, with test scores as the running variable, to identify the impact of getting admitted to a better school.

4 Results

4.1 First Stage: Take-up, Completion, and Training Duration

We first use Equation 1 to examine the level of compliance with our treatment assignment. We use three separate measures of compliance: starting an apprenticeship, completing an apprenticeship, and the duration (in months) of training. The results are reported in Table 3. Because the NAP training was essentially a traditional apprenticeship in practice, our compliance measures treat the NAP and traditional apprenticeships as equivalent. This assumption is tested in Appendix Tables A.4 -A.5, where we compare the characteristics of apprenticeships under the NAP with those of the traditional programs available in the market. Overall, the characteristics of the programs are very similar. The main difference is the lower fees charged by NAP trainers, consistent with the program design.

Over the full sample the NAP increased the probability of starting an apprenticeship by 13 percentage points (Table 3, Column 1). Since 63 percent of the control group also started an apprenticeship, the NAP treatment increased the probability of training by 21 percent. Being offered an NAP apprenticeship increased the probability of apprenticeship completion by almost 10 percentage points (Column 2). Because 25 percent of the control group completed an apprenticeship, this translates into a 40 percent increase in completion rate. Finally, a NAP treatment offer increased the duration of apprenticeship training by approximately four months, which represents a 20 percent increase relative to the control group (Column 3).

We also examine whether our first-stage results differ by trade and gender for each of our three compliance measures, again showing results in Table 3. The availability and attractiveness of labor market options other than an apprenticeship might differ based on the market for each trade. In addition, unobserved factors inherent to each trade (such as how trainers treat their apprentices) could influence the probability of starting or completing an apprenticeship as well as the duration of training.

We find some evidence of heterogeneity in compliance rates by trade. The estimated treatment effects of the program on the probability of starting an apprenticeship are similar across all three trade groups (Column 1). But differences emerge in training completion (Column 2). We find that the treatment did not affect the completion rate for male participants in construction. However, the treatment increased the probability of completion by 13 percentage points for female participants in cosmetology (a 45 percent increase relative to the control group). The treatment also increased the probability of completion for female participants in garment making by 9 percentage points (a 40 percent increase relative to the control group). We find similar patterns when we examine the duration of training (Column 3). We do not find a significant increase in duration among male participants in construction. In contrast, the treatment induced female participants to do longer apprenticeships- about five months longer in cosmetology and about four months longer in garment making. Relative to the control group, the treatment increased the duration of training by roughly a third for female participants.

One of the primary justifications for the NAP intervention was that many youth would be locked out of training opportunities by credit constraints. To examine this further, we test for heterogeneity in compliance rates by an applicant's poverty status, which we measure using an asset index (Appendix Table A.6). In general poorer applicants had lower rates of starting training (Column 1), lower rates of completing training (Column 2), and had shorter training periods (Column 3). However, since the interaction between treatment and poverty is positive, it suggests that program was more effective at lowering the barriers to training for poorer applicants. In particular, among the treatment group, a 1σ increase in the poverty index increased the likelihood of starting training by 6 percentage points (p-value < 0.01) and increased the duration of training by 2.6 months (p-value < 0.01). We also test for heterogeneity in compliance by baseline measures of ability and social network connections (Panel B and Panel C). However, the results show that compliance with treatment does not vary with these measures of ability and social network connections.

4.2 Intent to Treat Estimates

We use Equation 1 to examine the impact of the NAP treatment offer on labor market outcomes, assets, and consumption. This analysis is guided by our pre-analysis plan. Following the structure of the analysis presented earlier, we estimate treatment effects for our full sample as well as separately by trade groups. We use the standard p-values for inference in our discussion and also note cases where the inference is not robust to adjustments for multiple testing following Westfall and Young (1993) and Jones, Molitor, and Reif (2018).

4.2.1 Labor Market Outcomes: Labor Supply and Earnings

We first examine the effect of the NAP intervention on the extensive margin of labor supply, showing results in Table 4. In addition to documenting overall labor market participation (Column 1), we also measure any potential sectoral shifts by examining wage work (Column 2), self-employment (Column 3), work on a participant's own farm

(Column 4), apprenticeship (Column 5), and unpaid work (Column 6).[15] Results for the full sample show that those offered an NAP apprenticeship were almost 3 percentage points (4 percent relative to the control mean) less likely to be working compared with the control group (p-value < 0.1). This decline in the probability of working was driven in part by a 4 percentage point (25 percent) reduction in wage employment compared to the control group (p-value < 0.01). There was a limited (and insignificant) offsetting increase in the probability of self-employment. The treatment offer also encouraged participants to transition out of farm work, further contributing to the overall decline in the probability of working. The offer reduced the probability of youth working on their own farms, typically in low-productivity subsistence farming, by 2 percentage points (25 percent) compared with the control group (p-value < 0.05).

We also explore the program's effects on the extensive margin of labor supply in our three trade groups. Among male participants in construction trades there was no statistically significant change in the probability of working in any sector, in the wage sector, in self-employment, or in unpaid work. The treatment offer did lead to a 9 percentage point (40 percent) reduction in farming compared with the control group (p-value < 0.05), although this is not robust to multiple testing adjustments. In addition, the treatment offer increased the probability that male participants in construction were in apprenticeship training by 13 percentage points (100 percent) relative to the control group (p-value $<$ 0.01). This is consistent with the first-stage results showing that the treatment offer did not increase the probability of training completion for male participants in construction.

Among female participants in cosmetology we do not find a significant change in the probability of working. The treatment offer reduced the probability of wage work by 5 percentage points (33 percent) relative to the control group (p-value < 0.01). But this was offset by a 7 percentage point (22 percent) increase in self-employment compared with the control group (p-value < 0.05). There was a marginally significant 2 percentage point reduction in own farm work resulting from the treatment offer, but this is not robust to multiple testing adjustments. We also do not find any changes in the probability of apprenticeship or unpaid work.

Among female participants in garment making we do not find any statistically significant changes in any measure of labor supply (on the extensive margin). Further, all the coefficients are negative except for the estimated effect on working as an apprentice, which is positive but insignificant.

Taken together, the broad patterns suggest that the treatment shifts youth from wage

[15] We treat aggregate labor supply measures (working in any job and total hours worked) as summary measures and do not adjust the p-value for multiple testing.

work toward self-employment. In some trades the (positive) offsetting transition into self-employment is slower, leading to an overall reduction in labor market participation. Even though our endline data were collected almost four years after the start of training, the results on working in apprenticeship show that there is variance in the length of training across trades. Cosmetologists are most able to move on from their apprenticeships, followed by garment makers and then construction workers. This is consistent with the training completion rates reported in Table 3. The results on the intensive margin of labor supply or number of hours worked are generally in line with our extensive margin estimates. For brevity, these results are reported in Appendix Table A.7.

We examine the impact of the program on earnings from the previous month (relative to the survey date) in Table 5. We present results on all sources of income from the previous month including total earnings (Column 1), income from wage jobs (Column 2), income (profits) from self-employment (Column 3), income (profits) from farming (Column 4), and any income from apprenticeship (Column 5). All monthly earnings are "unconditional on working," where we assign zero income for those who did not report working in a given sector in the previous month.

For the full sample we find that the treatment offer reduced total monthly labor earnings by 11.5 GHS (13 percent) relative to the control group (p-value < 0.05). This was driven primarily by the estimated 15 GHS (35 percent) reduction in wage earnings compared with the control group (p-value < 0.01), a result consistent with the previously documented labor supply shifts out of wage employment. We do not find any significant changes in earnings from self-employment, farming, or apprenticeship. Without any offsetting increases in earnings, total monthly earnings declined as a result of the program offer.

The reduction in total earnings was more pronounced for male participants in construction trades. Intention-to-treat estimates for this group show that total monthly earnings fell by 47 GHS (24 percent) relative to the control group (p-value < 0.1). This was driven by the 59 GHS (46 percent) reduction in wages compared with the control group (p-value < 0.05). This reduction in wage earnings was not offset by any statistically significant (or meaningful) increases in earnings from self-employment, farming, or apprenticeship.

Among female participants in cosmetology we find a small and statistically insignificant reduction in total monthly earnings. The estimated 11 GHS (33 percent) reduction in monthly wage earnings compared with the control group (p-value < 0.05) was somewhat offset by increases in earnings from self-employment and farming. Neither of the intention-to-treat estimates on earnings from self-employment or farming is statistically significant, however.

The treatment offer reduced total monthly earnings among female participants in garment making by almost 11 GHS (15 percent) compared with the control group, although the p-value is just outside the threshold for marginal significance. The reduction in total earnings was driven by noisy reductions in wage earnings and self-employment income. Although the estimated coefficients on farming and apprenticeship income are positive, the magnitudes are negligible and statistically insignificant.

4.2.2 Material Well-Being

We present results on durable household assets, personal consumption expenditure on a limited set of items (such as phone calls), savings, loans, and migration in Table 6. These indicators enable us to examine broad measures of individual welfare. Our durable household assets index measured whether the household in which the respondent lived owned any of the following items: a radio, television, car, motorbike, and refrigerator, with the restriction that these items had to be functioning at the time of the endline survey (Column 1). Previous studies constructed similar indices based on household-owned durable assets for measuring poverty and conducting welfare analysis (Booysen, Van Der Berg, Burger, Von Maltitz, & Du Rand, 2008; Filmer & Scott, 2008). Although the overall (full-sample) treatment effect on durable household assets is positive, it is not statistically significant. For male participants in construction, however, the intention-to-treat estimate is negative and insignificant. Among female participants in cosmetology the program offer increased assets by 0.1σ (p-value < 0.05), but this is not very robust to multiple testing adjustments. The estimated effect among female participants in garment making is small (0.03σ) and insignificant. Overall, these results could be driven by the greater propensity for cosmetologists to finish training and shift into self-employment, leading to greater asset accumulation.

Because of time and budget constraints, we could not field a full detailed consumption module. Rather, we measured personal consumption expenditure as the sum of the respondent's expenditures on phone credit, personal items, and eating out during the week before the endline survey (Column 2). We do not find any evidence that the program altered spending patterns on these items. Further, we do not find any evidence that the program had an effect on savings behavior or borrowing activity (Columns 3 and 4).

Finally, we examine whether the program led to greater migration (Column 5). Migration can be viewed as an investment, where migrants move to pursue economic opportunities. For the full sample we find that the program offer increased the probability of migration by 4.4 percentage points (11 percent) compared with the control group. The estimated effects are similar across all trade groups, ranging from 3.4 to 4.7 percentage points, although these sub-sample effects are less precisely estimated.

5 Do Trainer Characteristics Matter?

We exploit the randomized match between apprentices and trainers to identify the effect of trainer characteristics on apprentice outcomes. This analysis can help provide insights into the specification of the training production function. We estimate Equation 2 on the sample of respondents who attended a matching meeting. Thus our empirical strategy compares two trainees who expressed the same preferences for trainers but were randomly assigned to trainers with different characteristics. In our analysis we focus on the following trainer characteristics: math test scores as a measure of cognitive ability, profits as a measure of business performance, the number of apprentices trained in the past as a measure of training experience, and the wage bill as a measure of business size. We chose these characteristics because they capture different dimensions of a trainer. Previous research has shown that math skills are important for firm behavior (Kremer, Lee, Robinson, & Rostapshova, 2013). Because Frazer (2006) argues that apprentices copy their training provider's business, we examine the extent to which trainers' profitable practices can be replicated by their apprentice. More experienced trainers may be better able to instruct. Finally, trainers with larger wage bills may be able to rely on other workers to instruct the apprentices. Apprentices can also learn by observing these workers.

Focusing on these dimensions, we first compare the characteristics of the two top-ranked trainers with those of lower-ranked trainers as a specification check. Overall, compared with trainers outside the top two, the top-ranked trainers scored on average 1.36σ higher in math and had trained almost 26 more apprentices (or more than four times as many). Compared with lower-ranked trainers, the two highest-ranked trainers earned almost 700 GHS more in monthly profits (almost four times as much) and incurred about 330 GHS more in monthly wage expenses (five times as much). All these differences are highly statistically significant (results not shown). The differences in trainer characteristics were clear and apparent to the apprentices. For all observable characteristics, trainers with the best characteristics were generally the most popular choices in apprentices' preference sets. This suggests that apprentices recognized the differences in quality and reputation of the trainers in the matching meetings. In addition, we find that apprentice characteristics are balanced between top-ranked and lower-ranked trainers in Appendix Table A.8. This provides us with additional reassurance about the validity of this empirical exercise. Table 7 examines the effect of trainer characteristics on the first-stage outcomes. Because the majority of individuals who attended a matching meeting started training, it is not surprising that we find only limited effects on any of these first-stage outcomes.

We examine the transmission of skills in Table 8. We measure craft skills using a short test designed in collaboration with industry experts. Innovation is measured using a series of self-reports such as the number of new designs created. Job skills are measured based on the skill content of a job following the World Bank STEP surveys. A higher value of this index reflects a shift away from jobs that are intensive in physical tasks. This index is only relevant for individuals who are working. Managerial skills are measured by the number of improved management practices adopted. These are relevant only for individuals who run their own business.

Apprentices assigned to trainers with the best math skills did not obtain better skills (Column 1). Nor did training with the most profitable trainers affect apprentice skills (Column 2). Training with the most experienced trainers (Column 3) improved test scores by about 0.25σ (p-value < 0.1) compared with trainees assigned to less experienced providers. These trainers increased the skill content of jobs by a similar magnitude (p-value < 0.1). Neither of these effects is robust to multiple testing adjustments, however. Apprentices paired with trainers with the highest wage bill (Column 4) scored almost 0.2σ higher on the skills test than those assigned to trainers with smaller wage bills, although this effect was just outside the standard thresholds for statistical significance. Training with these providers also increased innovation (or creativity) by 0.18σ (p-value < 0.05).

Finally, we examine the effect of trainer characteristics on labor market outcomes for apprentices in Table 9. For brevity, we focus only on wage and self-employment outcomes as well as on aggregate outcomes such as total earnings. Apprentices assigned to trainers with the best math scores did not have better labor market outcomes (Column 1). Those assigned to trainers with the highest profits (Column 2) were 16 percentage points more likely to be working (p-value < 0.01). This was driven primarily by the almost 10 percentage point increase in self-employment (p-value < 0.1). This match increased total monthly earnings by 63 GHS (p-value < 0.05). This was driven mostly by increased wage earnings, but the estimate is imprecise. We find that there were limited (or imprecisely estimated) labor supply effects when apprentices trained with the most experienced trainers (Column 3). But apprentices paired with these trainers had 65 GHS more in total monthly earnings (p-value < 0.01) than their peers assigned to less experienced trainers. This effect was driven primarily by the 42 GHS increase in wage earnings (p-value < 0.05). Finally, we find that apprentices assigned to trainers with the highest wage bills had 45 GHS more in total monthly earnings (p-value < 0.1) than their peers assigned to lower-ranked trainers (Column 4). Although neither of the treatment effects on wages or profits is statistically significant, the point estimates suggest that this increase was likely driven by self-employment profits.

6 Discussion

Overall, we find limited evidence that the apprenticeship training improved average labor market outcomes in the short run. One key mechanism for explaining negative, large, and significant earnings effects for male participants is that 33 percent of the compliers (those who took up treatment) were still in their low-paid apprenticeships. Additional evidence for this mechanism can be seen in Table 3, which shows that male participants in the treatment group were not significantly more likely to have completed an apprenticeship despite first-stage magnitudes for starting an apprenticeship similar to those for female participants in cosmetology and garment making. This could simply reflect the differences in training requirements among trades, where construction requires the most hours of training. An alternative possibility is that the work in construction trades is more lumpy and intermittent. This would limit training possibilities and extend the duration of training. In contrast, hairdressing and garment-making businesses generally have more clients and thus provide apprentices more opportunities to acquire skills in a shorter time. In addition, many women already have some basic skills in hairdressing and garment making, which can further speed up the training process.

Another factor that may explain the limited results is the lack of contract enforcement mechanisms. In the absence of the NAP, apprenticeships would be organized through social network ties (Frazer, 2006; Velenchik, 1995). These social ties would provide a mechanism for monitoring and enforcing contractual and training obligations between the apprentice and the trainer. Since the NAP recruits both apprentices and trainers and matches them, this could displace or crowd-out this traditional contract enforcement mechanism. Further, the limited government oversight in the NAP could lead to some exploitative situations in which little training is provided and the duration of training is extended so that trainers can benefit from cheap labor.

Finally, our results could reflect the numerous weaknesses in the informal apprenticeship system that previous researchers have documented (Darvas & Palmer, 2014; Frazer, 2006; Palmer, 2009). For example, trainers may have an incentive to withhold information if they are concerned about their apprentice eventually competing against them (Frazer, 2006). The lack of training materials and syllabi as well as reliance on older technology could also hinder apprentices' learning and earning outcomes (Darvas & Palmer, 2014). Further, since apprentices copy their trainer's business, their outcomes could be limited by their trainer's knowledge and practices (Darvas & Palmer, 2014; Frazer, 2006). Our results using the randomized match between apprentices and trainers show that whom you train with matters. Training with the most profitable or the most experienced firms increased apprentices' total monthly earnings. Because training opportunities with these trainers were limited, apprentices assigned to other trainers were relegated to less qualified trainers who delivered worse outcomes. Thus the overall results suggest that trainer quality is a constraint to the effectiveness of apprenticeships.

7 Conclusion

In this paper we present findings from our experimental evaluation of the NAP in Ghana, a program that because of its implementation process is equivalent to a traditional apprenticeship. Among the applicant pool, opportunities to train under the NAP were randomized. This allows us to compare the outcomes of participants offered training to those who were not offered the opportunity to train. Close to 75 percent of participants who were offered the opportunity to train under the NAP commenced training. However, since 62 percent of the control started an apprenticeship outside of the program, we find a relatively modest first stage effect on starting an apprenticeship. This could suggest that the program recruitment was not particularly well targeted. The implementation delays caused by political transition were likely an important contributing factor to the modest first stage effect on starting an apprenticeship. We find a larger difference in the apprenticeship completion rates between individuals who received a NAP training offer and those who did not. This suggests that apprenticeship fees present a barrier for many youth, especially when it comes to completion of training. Because many youth in this context are likely to be credit constrained, this provides some justification for government subsidies.

With respect to labor market outcomes, apprenticeships move participants out of wage employment, but there is a slower transition into self-employment. This reduces earnings in wage employment, with limited effects on self-employment earnings, leading to lower overall earnings. But we find that these reductions can be more than offset by the higher earnings generated if an apprentice trains with a high-quality trainer. Because the number of high-quality trainers is limited, this finding suggests that policy measures to screen trainers, as well as to augment their skills and productivity, can potentially improve the overall effectiveness of apprenticeship training.

References

Abdulkadiroğlu, A., Angrist, J., & Pathak, P. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, *82*(1), 137–196.

Acemoglu, D., & Pischke, J.-S. (1998). Why Do Firms Train? Theory and Evidence. *The Quarterly Journal of Economics*, *113*(1), 79–119.

Ajayi, K. (2013). School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana. *Boston University Working Paper*.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., & Vitali, A. (2017). Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda. *STICERD-Development Economics Papers*.

Blattman, C., & Ralston, L. (2015). Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs. *Available at SSRN 2622220*.

Booysen, F., Van Der Berg, S., Burger, R., Von Maltitz, M., & Du Rand, G. (2008). Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries. *World Development*, *36*(6), 1113–1130.

Cho, Y., Kalomba, D., Mobarak, A. M., & Orozco, V. (2013). Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-out Behavior. *World Bank Working Paper*.

Crépon, B., & Premand, P. (2018). Creating New Positions? Direct and Indirect Effects of a Subsidized Apprenticeship Program. *World Bank Working Paper*.

Darvas, P., & Palmer, R. (2014). *Demand and Supply of Skills in Ghana: How Can Training Programs Improve Employment and Productivity?* The World Bank.

Fersterer, J., Pischke, J.-S., & Winter-Ebmer, R. (2008). Returns to Aapprenticeship Training in Austria: Evidence from Failed Firms. *Scandinavian journal of economics*, *110*(4), 733–753.

Filmer, D., & Fox, L. (2014). *Youth Employment in sub-Saharan Africa*. The World Bank.

Filmer, D., & Scott, K. (2008). *Assessing Asset Indices*. The World Bank.

Frazer, G. (2006). Learning the Master's Trade: Apprenticeship and Human Capital in Ghana. *Journal of Development Economics*, *81*(2), 259–298.

Ghana Statistical Service. (2014). *Ghana Living Standard Survey (GLSS) Round 6.*

Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF International. (2014). *Ghana Demographic and Health Survey 2014.* Rockville, Maryland, USA.

Gondwe, M., & Walenkamp, J. (2011). Alignment of Higher Professional Education with the Needs of the Local Labour Market: The Case of Ghana. *The Hague: NUFFIC and The Hague University of Applied Sciences*.

Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political economy*, *108*(3), 604–631.

Hardy, M., & McCasland, J. (2016). Are Small Firms Labor Constrained? Experimental Evidence from Ghana. *UBC Working Paper*.

Hicks, J. H., Kremer, M., Mbiti, I., & Miguel, E. (2013). Vocational Education in Kenya: Evidence from a Randomized Evaluation among Youth. *Working Paper*.

Honorati, M., & Johansson de Silva, S. (2016). *Expanding Job Opportunities in Ghana*. The World Bank.

Hyytinen, A., Ilmakunnas, P., & Toivanen, O. (2013). The Return-to-Entrepreneurship Puzzle. *Labour Economics*, *20*, 57–67.

International Labor Organization Department of Statistics. (2018). *ILOSTAT Database.*

Jackson, C. K. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. *The Economic Journal*, *120*(549), 1399–1429.

Johanson, R. K., & Adams, A. V. (2004). *Skills Development in Sub-Saharan Africa*. The World Bank.

Jones, D., Molitor, D., & Reif, J. (2018). *What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study* (Tech. Rep.). National Bureau of Economic Research.

Kremer, M., Lee, J., Robinson, J., & Rostapshova, O. (2013). Behavioral Biases and Firm Behavior: Evidence from Kenyan Retail Shops. *American Economic Review: Papers and Proceedings*, *103*(3), 362-368.

Lucas, A. M., & Mbiti, I. M. (2014). Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya. *American Economic Journal: Applied Economics*, *6*(3), 234–63.

McKenzie, D. (2017). How Effective are Active Labor Market Policies in Developing Countries? a Critical Review of Recent Evidence. *The World Bank Research Observer*, *32*(2), 127–154.

McKenzie, D., & Woodruff, C. (2013). What are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World? *The World Bank Research Observer*, *29*(1), 48–82.

Monk, C., Sandefur, J., & Teal, F. (2008). Does Doing an Apprenticeship Pay Off? Evidence from Ghana. *CSAE Working Paper*.

Palmer, R. (2009). Formalising the Informal: Ghana's National Apprenticeship Programme. *Journal of Vocational Education and Training*, *61*(1), 67–83.

Pop-Eleches, C., & Urquiola, M. (2013). Going to a better School: Effects and Behavioral responses. *American Economic Review*, *103*(4), 1289–1324.

Teal, F. (2016). Are Apprenticeships Beneficial in sub-Saharan Africa? *IZA World of Labor*.

Velenchik, A. D. (1995). Apprenticeship Contracts, Small Enterprises, and Credit Markets in Ghana. *The World Bank Economic Review*, *9*(3), 451–475.

Westfall, P. H., & Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for P-value Adjustment* (Vol. 279). John Wiley & Sons.

World Bank. (2019). *Ghana Jobs and Skills Project Appraisal Document*. The World Bank.

Figure 1: Timeline

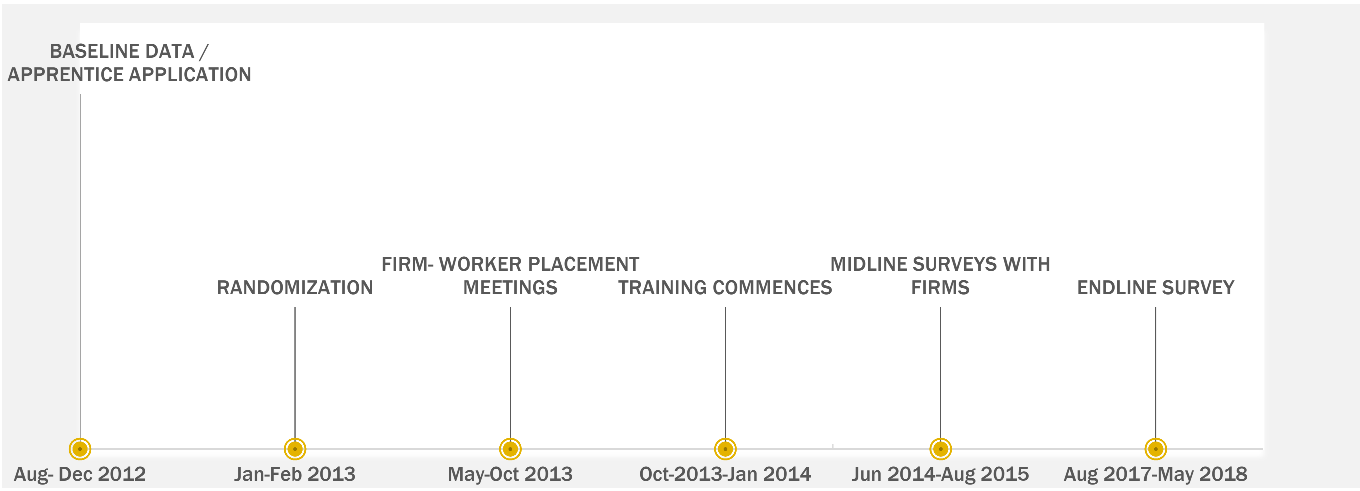

Figure 2: Research Design

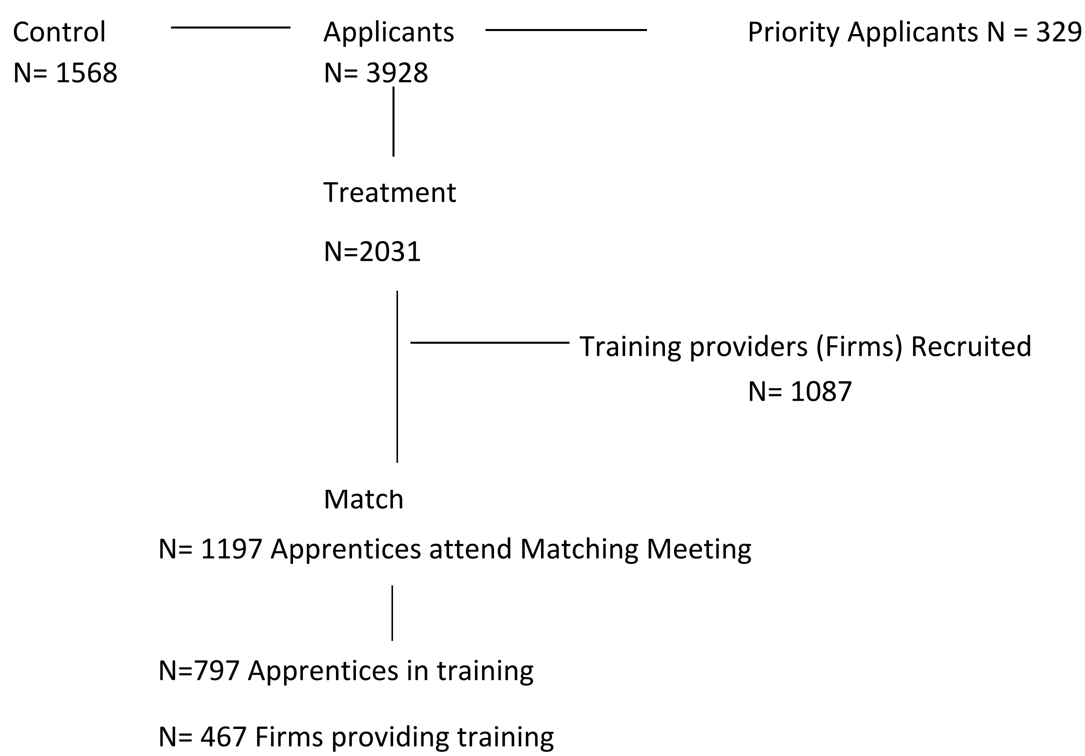

Table 1: Balance of Baseline Characteristics by Treatment/Control - Full Sample TEST

		(1) **N**	(2) **Mean Control**	(3) **Treatment**
Demographics				
(1)	Age (yrs)	3,468	23.14	0.045
(2)	Years of schooling	3,387	7.25	0.092
(3)	HH size (adults+children)	3,299	6.70	0.083
(4)	Mother: years of schooling	2,900	3.83	-0.339*
(5)	Father: years of schooling	2,596	6.23	-0.216
Labor				
(6)	Started an apprenticeship (0/1)	3,600	0.25	-0.002
(7)	Working (0/1)	3,600	0.43	0.011
(8)	Wage empl. (0/1)	3,600	0.05	-0.003
(9)	Self-empl. (0/1)	3,600	0.18	0.019
(10)	Total hours (hrs)	3,600	8.97	0.625
(11)	Wage empl. (hrs)	3,600	2.29	-0.082
(12)	Self-empl. (hrs)	3,600	6.68	0.707
(13)	Total earnings (GHC)	3,600	14.92	3.249
(14)	Wage empl. (GHC)	3,600	2.39	-0.443
(15)	Self-empl. (GHC)	3,600	8.52	-0.254
Skills				
(16)	Vocabulary score (z-score)	2,556	0.00	0.080*
(17)	Math score (z-score)	3,346	0.00	0.018
(18)	Digits score (z-score)	3,490	0.00	0.034
(19)	Ravens score (z-score)	3,486	0.00	0.018
Other				
(20)	Asset score (z-score)	3,345	0.00	0.028
(21)	Married (0/1)	3,600	0.31	-0.006
(22)	Children (0/1)	3,600	0.45	-0.013
(23)	Close family works in Govt/GES/DA (0/1)	3,600	0.30	-0.009
(24)	Urban (0/1)	3,326	0.77	0.002
(25)	Top 10 Metro (0/1)	3,473	0.14	0
(26)	Top 10 + District Capitals (0/1)	3,473	0.53	0.021
	F-test statistic	1,457		0.600

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1

Balanced baseline covariates are tested via OLS regressions for a sample of 3,600 individuals. Each row corresponds to such a regression. District x Trade Fixed Effects have been included. F-test statistic reported.

Table 2: Attrition: Endline Survey Completion Rates

	(1) **Full Sample**	(2) **Males Construction**	(3) **Females Cosmetology**	(4) **Females Garment-making**
Treatment	0.002	0.006	0.022	-0.024*
	(0.010)	(0.030)	(0.017)	(0.014)
Mean Completion Rate	0.909	0.926	0.907	0.918
Mean Control Group	0.906	0.914	0.897	0.930
Mean Treatment Group	0.911	0.929	0.917	0.906
Observations	3,600	740	1,240	1,438
Controls	No	No	No	No
Strata FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$

Estimation via OLS with treatment assignment as the independent variable and control group as the omitted category.

Outcome variable: completed endline survey (0/1).

Table 3: First Stage

	(1) **Started apprenticeship?** (0/1)	(2) **Completed apprenticeship?** (0/1)	(3) **Apprenticeship duration** (months)
	Full Sample		
Treatment	0.133***	0.099***	4.088***
	(0.017)	(0.017)	(0.742)
Adjusted p-value	0.000	0.000	0.000
Mean Control	0.626	0.249	18.608
Observations	3,270	3,270	3,270
	Males in Construction		
Treatment	0.161***	0.010	1.426
	(0.050)	(0.049)	(2.875)
Adjusted p-value	0.004	0.851	0.847
Mean Control	0.572	0.252	26.985
Observations	685	685	685
	Females in Cosmetology		
Treatment	0.140***	0.131***	5.033***
	(0.027)	(0.028)	(1.059)
Adjusted p-value	0.000	0.001	0.000
Mean Control	0.630	0.288	16.030
Observations	1,129	1,129	1,129
	Females in Garment-making		
Treatment	0.120***	0.086***	4.683***
	(0.025)	(0.024)	(1.058)
Adjusted p-value	0.000	0.001	0.000
Mean Control	0.630	0.216	17.694
Observations	1,327	1,327	1,327
Controls	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Estimation via OLS with treatment assignment as the independent variable. Imbalanced baseline covariates included as controls.

Table 4: Labor Supply - Extensive Margin

	(1) **Working** (0/1)	(2) **Wage empl.** (0/1)	(3) **Self empl.** (0/1)	(4) **Own farm** (0/1)	(5) **App' ship** (0/1)	(6) **Unpaid work** (0/1)
	Full Sample					
Treatment	-0.029*	-0.041***	0.026	-0.023**	0.021*	-0.005
	(0.017)	(0.013)	(0.017)	(0.010)	(0.012)	(0.011)
Adjusted p-value	0.084	0.006	0.257	0.084	0.245	0.633
Mean Control	0.713	0.158	0.297	0.089	0.118	0.094
Observations	3,270	3,270	3,270	3,270	3,270	3,270
	Males in Construction					
Treatment	-0.056	-0.059	-0.044	-0.091**	0.131***	-0.012
	(0.041)	(0.048)	(0.038)	(0.043)	(0.042)	(0.030)
Adjusted p-value	0.166	0.472	0.472	0.123	0.009	0.668
Mean Control	0.849	0.296	0.189	0.220	0.132	0.094
Observations	685	685	685	685	685	685
	Females in Cosmetology					
Treatment	-0.015	-0.053***	0.069**	-0.021*	-0.002	0.006
	(0.029)	(0.020)	(0.029)	(0.013)	(0.017)	(0.017)
Adjusted p-value	0.596	0.035	0.079	0.257	0.908	0.908
Mean Control	0.670	0.156	0.317	0.057	0.082	0.075
Observations	1,129	1,129	1,129	1,129	1,129	1,129
	Females in Garment-making					
Treatment	-0.032	-0.024	-0.005	-0.003	0.025	-0.009
	(0.026)	(0.017)	(0.026)	(0.014)	(0.019)	(0.016)
Adjusted p-value	0.207	0.579	0.978	0.978	0.579	0.924
Mean Control	0.706	0.121	0.313	0.072	0.135	0.111
Observations	1,327	1,327	1,327	1,327	1,327	1,327
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Working comprises wagejob, own business, own farm, unpaid work and apprenticeship. Imbalanced baseline covariates by strata included as controls.

Table 5: Labor Earnings

	(1) **Total** (GHC)	(2) **Wage empl.** (GHC)	(3) **Self empl.** (GHC)	(4) **Own farm** (GHC)	(5) **App'ship** (GHC)
			Full Sample		
Treatment	-11.54**	-15.35***	-0.94	2.13	0.82
	(5.73)	(4.84)	(4.34)	(2.03)	(0.95)
Adjusted p-value	0.044	0.007	0.816	0.643	0.643
Mean Control	89.19	42.17	41.52	3.21	3.97
Observations	3,270	3,270	3,270	3,270	3,270
			Males in Construction		
Treatment	-47.35*	-59.36**	-16.40	11.12	5.69
	(28.56)	(27.39)	(21.86)	(10.31)	(5.53)
Adjusted p-value	0.098	0.116	0.594	0.594	0.594
Mean Control	197.65	126.97	67.74	0.18	11.76
Observations	685	685	685	685	685
			Females in Cosmetology		
Treatment	-2.25	-11.23**	7.43	1.77	-0.40
	(7.69)	(5.30)	(6.10)	(2.08)	(0.63)
Adjusted p-value	0.769	0.186	0.581	0.664	0.664
Mean Control	73.21	33.62	36.14	1.78	1.87
Observations	1,129	1,129	1,129	1,129	1,129
			Females in Garment-making		
Treatment	-10.95	-8.09*	-5.19	0.67	0.82
	(6.70)	(4.24)	(5.38)	(2.83)	(0.79)
Adjusted p-value	0.102	0.204	0.686	0.829	0.686
Mean Control	71.89	25.25	39.84	4.86	2.05
Observations	1,327	1,327	1,327	1,327	1,327
Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Outcome variable: Unconditional monthly earnings in Ghana Cedi. Total earnings comprise wagejob, own business, own farm, and apprenticeship.

Table 6: Material Well-Being

	(1) **Asset score** (z-score)	(2) **Personal consumption** (GHC)	(3) **Savings** (GHC)	(4) **Ever loan** (0/1)	(5) **Migrated** (0/1)
	Full Sample				
Treatment	0.047	1.70	24.41	0.016	0.044**
	(0.031)	(2.21)	(15.55)	(0.014)	(0.018)
Adjusted p-value	0.372	0.447	0.372	0.433	0.065
Mean Control	0.000	34.80	125.50	0.185	0.406
Observations	3,270	3,263	3,270	3,270	3,270
	Males in Construction				
Treatment	-0.144	-0.08	38.98	0.019	0.034
	(0.096)	(10.98)	(85.01)	(0.037)	(0.046)
Adjusted p-value	0.471	0.995	0.928	0.928	0.907
Mean Control	0.000	60.49	242.97	0.157	0.252
Observations	685	682	685	685	685
	Females in Cosmetology				
Treatment	0.114**	0.20	17.19	-0.008	0.042
	(0.050)	(3.19)	(20.59)	(0.025)	(0.030)
Adjusted p-value	0.102	0.953	0.784	0.941	0.516
Mean Control	0.000	33.06	123.60	0.216	0.443
Observations	1,129	1,126	1,129	1,129	1,129
	Females in Garment-making				
Treatment	0.030	3.77	27.62	0.039*	0.047*
	(0.045)	(2.71)	(18.12)	(0.021)	(0.027)
Adjusted p-value	0.523	0.350	0.350	0.296	0.299
Mean Control	0.000	29.44	95.51	0.169	0.424
Observations	1,327	1,326	1,327	1,327	1,327
Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Asset score computed using PCA. Personal consumption expenditure.

Table 7: First Stage: Match Sample

	(1) Math Score (z-score)	(2) Profits (GHC)	(3) Apprentices Trained (#)	(4) Wage Bill (GHC)
	Started Apprenticeship (0/1)			
Matched with 1st/2nd MCP	-0.036	0.033	0.090*	0.023
	(0.046)	(0.048)	(0.051)	(0.053)
Adjusted p-value	0.685	0.851	0.215	0.689
	Completed Apprenticeship (0/1)			
Matched with 1st/2nd MCP	-0.007	0.004	0.059	0.057
	(0.066)	(0.062)	(0.066)	(0.063)
Adjusted p-value	0.909	0.947	0.619	0.689
	Apprenticeship Duration (months)			
Matched with 1st/2nd MCP	-3.068	0.946	1.042	2.013
	(2.558)	(2.774)	(2.803)	(2.682)
Adjusted p-value	0.540	0.913	0.724	0.689
Observations	567	567	567	567
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1. P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Characteristics of MCPs ordered within districtxtrade. Independent variable: being assigned 1st or 2nd MCP according to this ordering. Controlling for choice set and average characteristics of choice set. Different columns correspond to different MCP characteristics. Match treatment sample are apprentices assigned to treatment group who attended the match meeting where they ranked two or more MCPs.

Table 8: Skills Outcomes: Match Sample

	(1) Math Score (z-score)	(2) Profits (GHC)	(3) Apprentices Trained (#)	(4) Wage Bill (GHC)
	Craft Skills (z-score)			
Matched with 1st/2nd MCP	-0.067	0.057	0.235*	0.208
	(0.121)	(0.134)	(0.135)	(0.131)
Observations	567	567	567	567
Adjusted p-value	0.789	0.861	0.332	0.363
	Innovation (z-score)			
Matched with 1st/2nd MCP	-0.069	0.095	0.074	0.180**
	(0.073)	(0.065)	(0.079)	(0.071)
Observations	567	567	567	567
Adjusted p-value	0.789	0.452	0.636	0.072
	Job Skills (z-score)			
Matched with 1st/2nd MCP	0.129	0.098	0.241*	-0.097
	(0.129)	(0.142)	(0.138)	(0.132)
Observations	489	489	489	489
Adjusted p-value	0.789	0.861	0.332	0.518
	Managerial Skills (z-score)			
Matched with 1st/2nd MCP	-0.317	-0.217	0.181	-0.484
	(0.356)	(0.344)	(0.471)	(0.458)
Observations	201	201	201	201
Adjusted p-value	0.789	0.861	0.750	0.518
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1. P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Characteristics of MCPs ordered within districtxtrade. Independent variable: being assigned 1st or 2nd MCP according to this ordering. Controlling for choice set and average characteristics of choice set. Different columns correspond to different MCP characteristics. Match treatment sample are apprentices assigned to treatment group who attended the match meeting where they ranked two or more MCPs. Skill scores standardized using the mean and standard deviation of the full sample control group.

Table 9: Labor Market Outcomes: Match Sample

	(1) Math Score (z-score)	(2) Profits (GHC)	(3) Apprentices Trained (#)	(4) Wage Bill (GHC)
	Working (0/1)			
Matched with 1st/2nd MCP	-0.064	0.164***	0.015	0.043
	(0.061)	(0.060)	(0.065)	(0.062)
Adjusted p-value	0.302	0.007	0.816	0.490
	Wage Employment (0/1)			
Matched with 1st/2nd MCP	-0.010	0.099*	0.080	0.007
	(0.044)	(0.052)	(0.050)	(0.053)
Adjusted p-value	0.957	0.095	0.209	0.887
	Self-Employment (0/1)			
Matched with 1st/2nd MCP	-0.008	0.072	0.075	0.061
	(0.060)	(0.058)	(0.063)	(0.061)
Adjusted p-value	0.957	0.200	0.217	0.529
	Total Earnings (GHC)			
Matched with 1st/2nd MCP	-13.101	62.738**	65.106***	45.553*
	(21.411)	(25.270)	(22.080)	(26.680)
Adjusted p-value	0.541	0.013	0.003	0.089
	Wage Earnings (GHC)			
Matched with 1st/2nd MCP	-8.978	38.237	42.521**	14.131
	(19.906)	(25.986)	(17.517)	(25.163)
Adjusted p-value	0.867	0.287	0.037	0.598
	Business Profits (GHC)			
Matched with 1st/2nd MCP	7.098	18.523	13.830	25.306
	(15.147)	(15.008)	(15.347)	(17.012)
Adjusted p-value	0.867	0.287	0.349	0.291
Observations	567	567	567	567
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1. P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Characteristics of MCPs ordered within districtxtrade. Independent variable: being assigned 1st or 2nd MCP according to this ordering. Controlling for choice set and average characteristics of choice set. Different columns correspond to different MCP characteristics. Outcomes one month prior to endline survey.

A Appendix

Table A.1: Balance of Baseline Characteristics by Treatment/Control - Males in Construction

		(1) **N**	(2) **Control Mean**	(3) **Treatment**
Demographics				
(1)	Age (yrs)	721	24.46	-0.006
(2)	Years of schooling	713	7.95	0.377
(3)	HH size (adults+children)	688	7.96	0.296
(4)	Mother: years of schooling	612	2.80	0.495
(5)	Father: years of schooling	599	5.95	-0.867
Labor				
(6)	Started an apprenticeship (0/1)	727	0.42	-0.051
(7)	Working (0/1)	727	0.61	-0.113**
(8)	Wage empl. (0/1)	727	0.13	-0.012
(9)	Self-empl. (0/1)	727	0.23	0.031
(10)	Total hours (hrs)	727	13.49	1.431
(11)	Wage empl. (hrs)	727	5.13	0.424
(12)	Self-empl. (hrs)	727	8.36	1.007
(13)	Total earnings (GHC)	727	47.05	13.940
(14)	Wage empl. (GHC)	727	9.43	-2.647
(15)	Self-empl. (GHC)	727	19.09	-6.015
Skills				
(16)	Vocabulary score (z-score)	567	0.00	0.008
(17)	Math score (z-score)	713	0.00	0.031
(18)	Digits score (z-score)	727	0.00	-0.005
(19)	Ravens score (z-score)	727	0.00	-0.087
Other				
(20)	Asset score (z-score)	705	0.00	-0.032
(21)	Married (0/1)	727	0.34	-0.008
(22)	Children (0/1)	727	0.32	-0.064
(23)	Close family works in Govt/GES/DA (0/1)	727	0.31	0.029
(24)	Urban (0/1)	689	0.68	-0.010
(25)	Top 10 Metro (0/1)	720	0.18	-0.004
(26)	Top 10 + District Capitals (0/1)	720	0.52	-0.028
	F-test statistic	362		1.188

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
Balanced baseline covariates are tested via OLS regressions for a sample of 3,600 individuals. Each row corresponds to such a regression. District x Trade Fixed Effects have been included. F-test statistic reported.

Table A.2: Balance of Baseline Characteristics by Treatment/Control - Females in Cosmetology

		(1) **N**	(2) **Control Mean**	(3) **Treatment**
Demographics				
(1)	Age (yrs)	1,194	23.05	-0.165
(2)	Years of schooling	1,158	7.47	-0.219
(3)	HH size (adults+children)	1,119	6.01	0.322
(4)	Mother: years of schooling	969	4.87	-0.891***
(5)	Father: years of schooling	820	7.42	-0.513
Labor				
(6)	Started an apprenticeship (0/1)	1,203	0.24	0.014
(7)	Working (0/1)	1,203	0.41	-0.012
(8)	Wage empl. (0/1)	1,203	0.05	-0.003
(9)	Self-empl. (0/1)	1,203	0.18	-0.004
(10)	Total hours (hrs)	1,203	9.55	-1.317
(11)	Wage empl. (hrs)	1,203	2.58	-0.609
(12)	Self-empl. (hrs)	1,203	6.96	-0.708
(13)	Total earnings (GHC)	1,203	10.94	-2.069
(14)	Wage empl. (GHC)	1,203	1.42	-0.328
(15)	Self-empl. (GHC)	1,203	7.68	-1.070
Skills				
(16)	Vocabulary score (z-score)	872	0.00	0.093
(17)	Math score (z-score)	1,148	0.00	0.041
(18)	Digits score (z-score)	1,200	0.00	-0.004
(19)	Ravens score (z-score)	1,198	0.00	0.018
Other				
(20)	Asset score (z-score)	1,145	0.00	0.005
(21)	Married (0/1)	1,203	0.27	-0.004
(22)	Children (0/1)	1,203	0.51	-0.043
(23)	Close family works in Govt/GES/DA (0/1)	1,203	0.31	-0.038
(24)	Urban (0/1)	1,144	0.80	0.018
(25)	Top 10 Metro (0/1)	1,199	0.15	0.001
(26)	Top 10 + District Capitals (0/1)	1,199	0.50	0.032
	F-test statistic	453		0.877

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
Balanced baseline covariates are tested via OLS regressions for a sample of 3,600 individuals. Each row corresponds to such a regression. District x Trade Fixed Effects have been included. F-test statistic reported.

Table A.3: Balance of Baseline Characteristics by Treatment/Control - Females in Garments

		(1) N	(2) Control Mean	(3) Treatment
Demographics				
(1)	Age (yrs)	1,400	22.94	-0.010
(2)	Years of schooling	1,364	6.90	0.111
(3)	HH size (adults+children)	1,348	6.90	-0.164
(4)	Mother: years of schooling	1,184	3.35	-0.138
(5)	Father: years of schooling	1,052	5.68	-0.151
Labor				
(6)	Started an apprenticeship (0/1)	1,410	0.22	-0.003
(7)	Working (0/1)	1,410	0.42	0.069***
(8)	Wage empl. (0/1)	1,410	0.04	-0.008
(9)	Self-empl. (0/1)	1,410	0.18	0.036*
(10)	Total hours (hrs)	1,410	7.50	1.999*
(11)	Wage empl. (hrs)	1,410	1.46	-0.014
(12)	Self-empl. (hrs)	1,410	6.04	2.013**
(13)	Total earnings (GHC)	1,410	9.36	2.915
(14)	Wage empl. (GHC)	1,410	1.65	-0.636
(15)	Self-empl. (GHC)	1,410	6.44	1.413
Skills				
(16)	Vocabulary score (z-score)	1,001	0.00	0.073
(17)	Math score (z-score)	1,340	0.00	-0.016
(18)	Digits score (z-score)	1,409	0.00	0.089*
(19)	Ravens score (z-score)	1,407	0.00	0.059
Other				
(20)	Asset score (z-score)	1,351	0.00	0.075*
(21)	Married (0/1)	1,410	0.36	-0.003
(22)	Children (0/1)	1,410	0.50	0.023
(23)	Close family works in Govt/GES/DA (0/1)	1,410	0.29	0.007
(24)	Urban (0/1)	1,347	0.78	-0.001
(25)	Top 10 Metro (0/1)	1,401	0.13	-0.004
(26)	Top 10 + District Capitals (0/1)	1,401	0.57	0.017
	F-test statistic	573		0.601

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
Balanced baseline covariates are tested via OLS regressions for a sample of 3,600 individuals. Each row corresponds to such a regression. District x Trade Fixed Effects have been included. F-test statistic reported.

Table A.4: Apprenticeship Characteristics (1/2)

	(1) **Entrance fee** (GHC)	(2) **Exit fee** (GHC)	(3) **Firm size** (#)	(4) **Satis-faction** (0/1)	(5) **Travel time** (min)
	Full Sample				
Treatment	-65.904***	-44.768***	-0.336	-0.011	-1.176
	(8.103)	(11.783)	(0.233)	(0.015)	(1.039)
Adjusted p-value	0.000	0.001	0.459	0.709	0.564
Mean Control	167.497	122.597	4.139	0.877	26.347
Observations	2,207	1,757	2,255	2,261	2,240
	Males in Construction				
Treatment	-36.400	-35.020	0.931	-0.044	4.359
	(29.132)	(29.617)	(1.310)	(0.038)	(3.139)
Adjusted p-value	0.822	0.822	0.943	0.822	0.792
Mean Control	128.908	82.938	4.637	0.934	22.614
Observations	457	311	464	467	448
	Females in Cosmetology				
Treatment	-63.325***	-42.028**	-0.842***	-0.022	-1.916
	(13.313)	(16.312)	(0.313)	(0.024)	(1.625)
Adjusted p-value	0.000	0.060	0.051	0.746	0.664
Mean Control	180.979	119.591	4.484	0.878	26.029
Observations	772	687	792	793	792
	Females in Garment-making				
Treatment	-73.258***	-50.324***	-0.169	0.019	-2.281
	(10.785)	(18.349)	(0.317)	(0.023)	(1.585)
Adjusted p-value	0.000	0.072	0.934	0.863	0.706
Mean Control	165.980	131.148	3.881	0.857	27.721
Observations	895	697	914	915	914
Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Conditional on having started an apprenticeship. Imbalanced baseline covariates included as controls.

Table A.5: Apprenticeship Characteristics (2/2)

	(1) **Toolkit** (0/1)	(2) **Practice materials** (0/1)	(3) **Written** (0/1)	(4) **Testimonial** (0/1)	(5) **Exam** (0/1)
	Full Sample				
Treatment	-0.055**	0.005	0.043***	-0.053	0.097***
	(0.022)	(0.022)	(0.016)	(0.034)	(0.033)
Adjusted p-value	0.080	0.849	0.047	0.436	0.024
Mean Control	0.480	0.602	0.129	0.592	0.550
Observations	2,261	2,261	2,261	980	980
	Males in Construction				
Treatment	-0.037	-0.042	0.065	-0.081	0.012
	(0.067)	(0.072)	(0.045)	(0.132)	(0.123)
Adjusted p-value	0.943	0.943	0.763	0.943	0.943
Mean Control	0.363	0.527	0.110	0.350	0.350
Observations	467	467	467	195	195
	Females in Cosmetology				
Treatment	-0.099***	0.020	0.074***	-0.040	0.099**
	(0.035)	(0.037)	(0.027)	(0.047)	(0.047)
Adjusted p-value	0.044	0.746	0.044	0.746	0.158
Mean Control	0.413	0.541	0.128	0.618	0.611
Observations	793	793	793	400	400
	Females in Garment-making				
Treatment	-0.034	0.005	0.005	-0.059	0.100**
	(0.033)	(0.031)	(0.023)	(0.050)	(0.049)
Adjusted p-value	0.841	0.967	0.967	0.809	0.327
Mean Control	0.567	0.664	0.138	0.622	0.545
Observations	915	915	915	346	346
Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Conditional on having started an apprenticeship. Imbalanced baseline covariates included as controls.

Table A.6: First Stage - Heterogeneity

	(1) **Started apprenticeship?** (0/1)	(2) **Completed apprenticeship?** (0/1)	(3) **Apprenticeship duration** (months)
		Assets	
Treatment	0.132***	0.098***	4.049***
	(0.017)	(0.017)	(0.740)
Poor (z-score)	-0.042***	-0.051***	-2.805***
	(0.015)	(0.013)	(0.637)
Treatment x Poor	0.062***	0.016	2.637***
	(0.017)	(0.016)	(0.747)
Adjusted p-value *Treatment*	0.000	0.000	0.000
Adjusted p-value *Poor*	0.005	0.000	0.000
Adjusted p-value *Interaction*	0.000	0.297	0.002
		Ability	
Treatment	0.133***	0.099***	4.121***
	(0.017)	(0.017)	(0.742)
Ability (z-score)	0.008	0.026**	0.301
	(0.012)	(0.011)	(0.529)
Treatment x Ability	0.002	-0.005	0.023
	(0.012)	(0.012)	(0.539)
Adjusted p-value *Treatment*	0.000	0.000	0.000
Adjusted p-value *Ability*	0.717	0.052	0.717
Adjusted p-value *Interaction*	0.979	0.969	0.979
		Network	
Treatment	0.136***	0.084***	4.339***
	(0.020)	(0.019)	(0.863)
Network (0/1)	0.021	-0.005	1.144
	(0.028)	(0.026)	(1.214)
Treatment x Network	-0.011	0.050	-0.825
	(0.035)	(0.035)	(1.600)
Adjusted p-value *Treatment*	0.000	0.000	0.000
Adjusted p-value *Network*	0.684	0.824	0.678
Adjusted p-value *Interaction*	0.817	0.327	0.817
Mean Control	0.626	0.249	18.608
Observations	3,270	3,270	3,270
Controls	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01; ** p<0.05; * p<0.1
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Estimation via OLS with treatment assignment as the independent variable. Imbalanced covariates included as controls. Asset and ability scores have been computed via PCA and then been standardized. Ability comprises vocabulary, math, ravens and digits. Network indicates whether apprentice has close family working in the government, GES or district assembly.

Table A.7: Labor Supply - Intensive Margin

	(1) **Working** (hrs)	(2) **Wage empl.** (hrs)	(3) **Self empl.** (hrs)	(4) **Own farm** (hrs)	(5) **App' ship** (hrs)	(6) **Unpaid work** (hrs)
				Full Sample		
Treatment	-1.19	-6.95***	5.01	-2.69**	3.21	0.13
	(3.82)	(2.60)	(3.21)	(1.22)	(2.50)	(1.79)
Adjusted p-value	0.755	0.038	0.301	0.098	0.356	0.943
Mean Control	117.25	28.24	44.76	9.48	23.19	11.97
Observations	3,270	3,270	3,270	3,270	3,270	3,270
				Males in Construction		
Treatment	6.21	-2.49	-9.87	-10.05*	25.50***	3.24
	(10.41)	(9.50)	(8.08)	(5.26)	(8.13)	(4.57)
Adjusted p-value	0.551	0.800	0.486	0.199	0.011	0.725
Mean Control	132.79	44.57	34.16	21.76	23.23	9.08
Observations	685	685	685	685	685	685
				Females in Cosmetology		
Treatment	-0.59	-9.96**	11.74**	-2.58*	-1.54	1.26
	(6.60)	(4.22)	(5.56)	(1.52)	(3.74)	(2.83)
Adjusted p-value	0.929	0.089	0.131	0.229	0.875	0.875
Mean Control	110.01	29.30	47.24	6.13	18.36	9.71
Observations	1,129	1,129	1,129	1,129	1,129	1,129
				Females in Garment-making		
Treatment	-4.91	-5.61	0.03	-0.72	3.75	-2.18
	(5.42)	(3.48)	(4.52)	(1.68)	(3.72)	(2.68)
Adjusted p-value	0.366	0.428	0.996	0.893	0.776	0.815
Mean Control	116.68	22.80	45.24	8.16	25.58	15.14
Observations	1,327	1,327	1,327	1,327	1,327	1,327
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parantheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$
P-values adjusted for multiple hypothesis testing provided. Method: Westfall and Young 1993. Outcome variable: Unconditional monthly hours. Total hours comprise wagejob, own business, own farm, apprenticeship and unpaid work. Imbalanced baseline covariates by strata included as controls.

Table A.8: Balance of Apprentice Characteristics for Match Randomization Sample

		MCP Math		MCP Profits		Apprentices Trained		Wage Bill	
		Mean "Control"	Match "Treatment"	Mean "Control"	Match "Treatment"	Mean "Control"	Match "Treatment"	Mean "Control"	Match "Treatment"
Demographics									
(1)	Age (yrs)	22.90	0.598	23.31	-0.232	23.25	-0.001	23.31	0.110
(2)	Years of schooling	7.30	0.545	7.84	-0.009	7.65	0.523	7.82	-0.368
(3)	HH size (adults+children)	8.02	-0.559	7.87	-0.041	7.89	-0.504	7.96	-0.310
(4)	Mother: years of schooling	3.14	-0.609	3.07	-0.352	3.01	-0.436	3.34	-2.104***
(5)	Father: years of schooling	5.10	0.541	5.17	-0.291	5.07	0.568	5.35	-0.631
Labor									
(6)	Started an apprenticeship (0/1)	0.29	-0.030	0.27	-0.021	0.25	0.092*	0.28	-0.025
(7)	Working (0/1)	0.49	0.031	0.50	0.040	0.52	-0.092	0.52	-0.054
(8)	Wage empl. (0/1)	0.05	0.032	0.06	-0.002	0.06	0.018	0.06	0.020
(9)	Self-empl. (0/1)	0.23	0.007	0.23	-0.019	0.23	-0.039	0.24	-0.065
(10)	Total hours (hrs)	10.13	0.743	10.71	-0.351	10.73	-0.900	11.28	-1.856
(11)	Wage empl. (hrs)	1.46	1.683	2.41	1.145	2.37	1.158	2.87	0.273
(12)	Self-empl. (hrs)	8.67	-0.940	8.31	-1.496	8.36	-2.058	8.41	-2.129
(13)	Total earnings (GHC)	19.45	-1.947	19.98	-5.349	17.77	3.665	19.87	-0.520
(14)	Wage empl. (GHC)	2.28	-0.482	2.06	0.431	2.03	2.090	2.31	2.071
(15)	Self-empl. (GHC)	11.55	1.483	10.89	0.002	12.04	-1.733	9.39	9.664
Skills									
(16)	Vocabulary score (z-score)	0.00	0.258**	0.00	0.060	0.00	-0.075	0.00	0.023
(17)	Math score (z-score)	0.00	-0.117	0.00	-0.033	0.00	-0.098	0.00	0.319***
(18)	Digits score (z-score)	0.00	-0.006	0.00	0.120	0.00	0.023	0.00	0.037
(19)	Ravens score (z-score)	0.00	0.076	0.00	-0.027	0.00	0.039	0.00	-0.053
Other									
(20)	Asset score (z-score)	0.00	0.003	0.00	-0.002	0.00	0.008	0.00	0.061
(21)	Married (0/1)	0.34	0.094*	0.36	0.004	0.35	0.044	0.35	0.061
(22)	Children (0/1)	0.48	-0.010	0.46	0.035	0.46	-0.024	0.47	0.014
(23)	Close family works in Govt/GES/DA (0/1)	0.29	0.007	0.33	-0.045	0.34	-0.106*	0.32	0.010
(24)	Urban (0/1)	0.70	0.029	0.71	0.033	0.71	0.014	0.73	0.000
(25)	Top 10 Metro (0/1)	0.15	0.013	0.14	0.015	0.15	-0.038*	0.15	-0.014
(26)	Top 10 + District Capitals (0/1)	0.53	0.024	0.52	0.095**	0.54	-0.031	0.54	-0.012
	F-test statistic	258	2.774	258	2.104	258	0.766	258	0.739
	Observations	567		567		567		567	

Robust standard errors in parantheses. *** p<0.01; ** p<0.05; * p<0.1
Balanced baseline covariates are tested via OLS regressions. Each row corresponds to such a regression.
District x Trade Fixed Effects have been included and standard errors are robust
F-test statistic reported.
"Control" to the remaining MCPs.